SPARK

THE TULLIE AND RICKEY FAMILIES
SPARK AWARDS FOR
INNOVATIONS IN IMMUNOLOGY

2018 – 2022

FIVE YEAR REPORT

April 2022

La Jolla Institute FOR IMMUNOLOGY | Life Without Disease.®

SPARK | THE TULLIE AND RICKEY FAMILIES SPARK AWARDS FOR INNOVATIONS IN IMMUNOLOGY

FIVE YEAR REPORT

TABLE OF CONTENTS

La Jolla Institute FOR IMMUNOLOGY | Life Without Disease.®

SPARK | THE TULLIE AND RICKEY FAMILIES SPARK AWARDS FOR INNOVATIONS IN IMMUNOLOGY

Five Years of Fueling the Next Generation of Researchers

By: Kelsey Dale
Tullie and Rickey Families SPARK Awards Program Manager

When the La Jolla Institute SPARK Awards program was first conceived in 2017, we had high hopes for the program but it was the first of its kind that we were aware of, and of course there is always a risk in charting a new course. We were filled with questions about whether it would work.

The program was developed to meet two complementary needs at the institute. First, young researchers often have fresh and innovative ideas, and their career advancement is dependent on them making their own contributions to the field of immunology, but they are often challenged with finding both the opportunity and funding to take their ideas from concept to solid data. And secondly, we had heard from our supporters that it could be difficult for them to feel like they were making an impact since many full-scale research projects or equipment require large amounts of funding and because it could take so long to see if their donation was contributing to progress or success. So, we were tasked with how to make our research more accessible and inspiring to donors. In thinking about these needs, we realized there was an opportunity to address both by connecting our supporters directly with our early-career scientists eager to explore their big ideas with the help of seed funding.

Our goal the first year was to fund at least two $25,000 awards to establish the program. We hoped that the program would attract the interest of our postdocs and also resonate with our supporters, but only time would tell. Almost immediately we saw there was indeed an appetite for this opportunity from our scientists when we received nearly 25 applications from the first request for proposals. We narrowed the pool down to four finalists that we then trained in lay-friendly communication skills, and prepared them to pitch their projects to prospective donors. Next, we needed to see if these scientists and their projects would appeal to donors.

We didn't have to wait long. In that first year, thanks to an early first commitment from Board Member François Ferré and his wife Magda Marquet, we moved into our fundraising campaign with momentum and secured enough funding for all four of the initial projects, and more. In fact, in 2018 we funded eight projects. This confirmed that there was indeed a desire for our supporters to connect with our individual scientists and fund specific, short-term projects. But even then, we

were left with outstanding questions. Would we be able to sustain this interest? How could we ensure this program wouldn't just fizzle out after that first year? After all, it might take a couple years before we could really see if the program was going to be successful in helping these scientists use the data they collected from their projects to attract follow-on funding for their project or compete for critical career-advancement grants and fellowships.

After sitting on the pitch review panel the first year in 2018, Tom Tullie saw the potential of the program, but also recognized the delicate position it was in. That was when he and his wife, Judy, came forward with an incredible $250,000 pledge to ensure the program could fund at least one award a year for the next 10 years. Not only this, but Tom reached out to his long-time friend Dave Rickey and invited him and his wife Brenda to learn about the program and get involved as well. Their joint commitments of $375,000 secured the foreseeable future for the program, and to honor their dedication the program, it was renamed *The Tullie and Rickey Families SPARK Awards for Innovations and Immunology*.

With the security of the program in place, we were able to blaze boldly ahead. In 2019, we continued to attract donors to join the Tullie and Rickey families in supporting the program, and we successfully funded another six SPARK Award projects. That summer LJI and Tom Tullie took the program to a national stage at the Classy Collaborative conference in Boston, to share with others in the non-profit space how we were able to successfully connect donors to our mission through the program. And in the fall of 2019, our first "SPARK Star," Dr. Rana Herro, advanced her career by securing a faculty position at Cincinnati Children's Hospital, which she attributed in part to the experience and skills she gained from her involvement with the SPARK program.

In the early years, there were additional unexpected benefits about the program that we learned from SPARK Award winner feedback about the most valuable aspects of the program. For example, we had underestimated how valuable and unique the training we were giving our postdocs was for building their own confidence and preparing them to not only effectively convey their ideas during the pitches, but also at interviews for faculty positions or writing grant applications for private foundations and individuals. And this skill also was valuable in helping our supporters feel a more personal connection with the Institute as SPARK winners have also taken on a true ambassador role for LJI, participating in donor tours and events to provide a personal perspective of the amazing work underway at LJI.

By 2020 the program truly became an established gem of the institute, featuring exciting new projects from LJI's best and brightest young researchers as well as inspiring even more donors to become involved. And between 2020 to 2021 an additional 16 projects received funding, including several projects aimed at helping us better understand the virus behind the global pandemic SARS-CoV-2. We also started to see in those years how the first couple of cohorts of SPARK winners were showing that the investment in them and their ideas was in fact resulting in scientific breakthroughs and career advancement. We started to see the first significant follow-on funding grants begin to materialize in 2021, when Dr. Marco Orecchioni and Dr. Abhijit Chakraborty each

won six-figure awards from The Conrad Prebys Foundation to build on the data they collected from their SPARK projects. And later that spring, Dr. Ian Mathews, together with his P.I. Dr. Sonia Sharma, secured a $4.2M R01 grant from the National Cancer Institute stemming from the data he produced with his SPARK award in 2018. Not only that, but in those two years, six more SPARK Stars from the 2018-2021 cohorts went on to start their own labs (Dr. Daniela Weiskopf, Dr. Holger Winkels, Dr. Huy Dinh, Dr. Cecilia Lindestam Arlehamn, Dr. Vipul Shukla and Dr. Michael Norris).

And now in 2022 our fifth year of awards, the program shows no sign of slowing down. Not only did we fund six new SPARK Award projects in January, but we've also had three more former winners, Dr. Greet Verstichel, Dr. Nicolas Thiault, and Dr. Julie Burel secure six-figure awards from private individuals and foundations to continue their work.

Looking back over the past five years SPARK project data has contributed to the publication of at least 12 manuscripts in leading scientific journals and 4 applications for patents demonstrating the innovative aspects of the projects being funded through this program. And as of writing this there are roughly 20 papers or grant applications in progress by SPARK Award winners indicating the number of publications and awards will continue to climb in the next five years.

We are so grateful to the approximately 200 supporters of this program who have helped us raise more than $1.1M for SPARK since 2017. And we are especially grateful for the investment and belief of the Tullie and Rickey families at such a critical time in this program's history. This partnership helped us confidently build this program to where it is today and has also allowed us to be nimble as we've piloted new aspects of the program to enhance it. For example, in 2021 we implemented a $5,000 bonus to the scientist with the top pitch to further incentivize our scientists to find ways to make their pitch really stand out to donors. And this year, we look forward to giving our 2022 winners an opportunity to compete for an additional $25,000 award after six months of progress, which would concentrate more award dollars behind the project showing the greatest promise of that group.

As we look to the future of the Tullie and Rickey SPARK Awards program, we still have questions about the program. However they are no longer about whether the program will survive or be successful, but instead we are excited to imagine just how successful it could be.

SPARK | THE TULLIE AND RICKEY FAMILIES SPARK AWARDS FOR INNOVATIONS IN IMMUNOLOGY

impact summary

5 years of awards

$1.1M
Raised in gifts and pledges

450
Gifts/Pledges

$835K
Awarded to early-career scientists exploring bold ideas

33
Awards

19
Different labs/cores

4 years of results*

$6.3M
Received in follow-on funding

7
Winners have started independent labs

13
Winners have advanced their career with a promotion or new position

12
Manuscripts published

20
Papers or grant applications in progress

* *Based on 27 completed SPARK projects 2018-21*

THE TULLIE AND RICKEY FAMILIES SPARK AWARDS FOR INNOVATIONS IN IMMUNOLOGY WINNER PROFILES (2018 – 2022)

(Alphabetical by scientist's last name)

La Jolla Institute FOR IMMUNOLOGY | Life Without Disease.®

SPARK | THE TULLIE AND RICKEY FAMILIES SPARK AWARDS FOR INNOVATIONS IN IMMUNOLOGY

Mehdi Benkahla, Ph.D.

Postdoctoral Researcher, La Jolla Institute for Immunology
2019 and 2021 SPARK Award Winner
Disease Impact: Type 1 Diabetes

2021 PROJECT: CAN A VIRUS TRIGGER TYPE 1 DIABETES?

Funded by: Donors to the Sierra Shapery Memorial SPARK Award

Type 1 diabetes (T1D) is an autoimmune disease causing the destruction of insulin-producing beta cells in the pancreatic islets. Immunopathogenesis of T1D is still unknown. There are lines of evidence suggesting that both environmental factors such as viruses and genetic susceptibility play an important role in triggering the disease. Several viruses, mainly from the enterovirus family, have been considered as potential causal agents for human T1D. However, there is still little evidence from the human pancreas indicating a link between viral infections and T1D development. The main interest of my SPARK project was to investigate the impact of viral infections on human pancreatic slices.

We obtained pancreatic tissue slices from non-diabetic donors provided by the Network for Pancreatic Organ Donors with Diabetes (nPOD). Human pancreatic slices from non-diabetic donors were infected with a coxsackievirus B3 (CVB3) genetically engineered to express a green fluorescent protein (GFP). After 3 days of infection, the GFP was detected by high resolution confocal microscopy in the insulin-producing beta cells. These infected beta cells were also expressing HLA-I, which is a molecule by which cells present an antigen to the immune system to trigger an immune response. Expression of HLA-I is also a hallmark of T1D. We have also observed that tissue-resident macrophages were infected by CVB3. Our 3D analysis shows that HLA-I is hyper-expressed in the CVB3-infected slices compared to the controls. Since the slices have tissue-resident immune cells, we wanted to study the impact of the infection on CD68+ macrophages and whether they become attracted to islets during CVB3 infection. When we quantified the number of macrophages in CVB3-infected and non-infected islets on day 3 post-infection, we found no differences in the number of macrophages attracted to pancreatic islets.

This project has helped implement a new model to study type 1 diabetes in the lab. Using the preliminary data generated during my SPARK project, we applied for follow-up funding from NIH in 2021 in the form of an R21 (estimated total award would be $503,250 for two years). We've also secured other sources of human pancreatic tissues from City of Hope and Prodolabs, as well as through a collaboration with the UC San Diego Moores Cancer Center. For future studies we plan to evaluate the impact of CVB3-GFP infection on beta cells and tissue-resident macrophages and whether the infection could lead to an immune destruction of beta cells when co-cultured with PBMCs. Tissue-resident macrophages will be characterized by single-cell RNA-seq and their interaction with PBMCs will be assessed.

2019 PROJECT: ROLE OF HUMAN HERPESVIRUS-6 INFECTION IN THE PATHOGENESIS OF TYPE 1 DIABETES

Funded by Robert and Rachel Perlmutter and Various Donors

Type 1 diabetes is an autoimmune disease where T cells destroy insulin-producing beta cells in the pancreatic islets. Scientists still don't understand the process that drives the immune system to target these cells. There is evidence suggesting that both genetic susceptibility and environmental factors such as viruses play an important role in triggering the disease. Scientists have investigated several viruses, mainly from the enterovirus family, as potential drivers of type 1 diabetes; however, little is known about the involvement of other viruses, such as herpesviruses. My project was specifically focused on determining whether human herpesvirus-6 (HHV-6) could play a role in the development of type 1 diabetes.

For this project, I used high-resolution confocal microscopy to detect HHV-6 in pancreas samples from diabetic and non-diabetic organ donors. I specifically focused on the HHV-6 gB protein, a protein that plays a critical role during membrane fusion and viral entry, as a marker for the presence or absence of HHV-6 in individual pancreatic islets. Our data showed that patients with type 1 diabetes have a significantly higher level of viral gB protein in the pancreas, compared to donors without diabetes. We detected the gB protein in the islets of six out of seven type 1 diabetes donors, compared with only one out of four non-diabetic donors, and two out of five donors with auto-antibodies.

We published the results of the study in the Journal of Autoimmunity in November 2019. In summary, this novel research has shown that donors with type 1 diabetes are more likely to have been infected by a virus like HHV-6. This suggests a viral infection may be a potential cause of the disease through an indirect mechanism. For example, maybe pancreata of type 1 patients are in general more susceptible to infections, possibly due to alterations in cellular pathways that could make them more vulnerable to harmful viruses.

SPARK STAR HIGHLIGHTS

Papers Published Related to Projects:

- HLA class I hyper-expression unmasks beta cells but not alpha cells to the immune system in pre-diabetes, Journal of Autoimmunity, May 2021
- Human herpesvirus-6 is present at higher levels in the pancreatic tissues of donors with type 1 diabetes, *Journal of Autoimmunity*, November 29, 2019

Other Highlights Related to Research or Career:

- Dr. Benkhala submitted a request to Network for Pancreatic Organ Donors with Diabetes (nPOD) with the hope of getting access to precious patient samples, this was approved in 2020. This allowed Dr. Benkhala and others in the von Herrath lab to use live human pancreatic slices to study viral infections.

SPARK | THE TULLIE AND RICKEY FAMILIES SPARK AWARDS FOR INNOVATIONS IN IMMUNOLOGY

Simon Brunel, Ph.D.

Postdoctoral Researcher, La Jolla Institute for Immunology
2021 SPARK Award Winner
Disease Impact: Autoimmune Diseases

CAN WE CURE AUTOIMMUNE DISEASES WITH A SINGLE THERAPY?

Funded by: The Rosemary Kraemer Raitt Foundation Trust, John and Cim Kraemer, and Various Donors

The goal of this project is to establish proof-of-principle that it is possible to reduce or prevent the development of spontaneous autoimmunity by rewiring the immune system to avoid destruction of molecules or cells that will lead to pathology. Genetic modifications of dendritic cells for in vivo applications have been performed in the past and are the subject of numerous clinical trials. If this therapeutic strategy works, it could represent a universal therapeutic approach for any autoimmune disease, as long as the molecules involved in the disease process are known.

Thus far, I have tested the techniques and technology we'll need to develop the best therapeutic cells. I was able to demonstrate that cells from bone marrow can differentiate into dendritic-like cells and migrate to the thymus, the organ where T cells develop. This discovery shows us that bone marrow cells can express the right molecular tools that allow them to migrate through the body—even after they are modified in the laboratory. Our follow-up research in a mouse model confirmed that cells expressing this particular protein, called CCR9, preferentially migrate to the thymus.

Thanks to this work, we now have the tools in place to analyze and track these modified cells. The next step is to confirm that the migration of those cells to the thymus can prevent or cure autoimmune disorders. The first model of autoimmune disorder will be type 1 diabetes in mice.

Unfortunately, because the COVID-19 pandemic caused supply shortages and high staff turnover in my lab, my project was delayed. I was granted a 6-month extension in January 2022. This spring, I've been developing the animal protocol with our Internal Review Board and Department of Laboratory Animal Care to get the necessary approvals to run my experiments with the mouse model. I've also

been working on stocking all the reagents and supplies needed for the experiments so I have everything in place once I'm able to start my experiments.

We hope this work will give us a window into treating type 1 diabetes and the mystery of how that disease develops in the first place. I look forward to providing my final report in the fall.

Estimated Project Completion Date: 07/31/22

SPARK | THE TULLIE AND RICKEY FAMILIES SPARK AWARDS FOR INNOVATIONS IN IMMUNOLOGY

Julie Burel, Ph.D.

Instructor, La Jolla Institute for Immunology
2019 SPARK Award Winner
Disease Impact: Infectious Disease

IMMUNE CROSSTALK MAY IMPROVE THE DIAGNOSIS OF INFECTIOUS DISEASES

Funded by: Larry and Tiki Spitcaufsky

There are many types of immune cells, and we're still figuring out how they work together. One interesting immune cell pairing are T cell: monocyte "doublets." For years, scientists have been observing these conjoined cell pairs in patient samples, but many regarded doublets as an accidental artifact of the flow cytometry process, and it was commonplace to "dump" those conjoined cells before gathering data. We wondered—could doublets actually be a sign that the immune system is fighting off a disease? To answer this question, we needed to overcome the technological hurdles that make it difficult to detect doublets in patient samples. We also needed to isolate their genetic signatures to learn more about what they're doing in the body.

My project showed that combining several cell imaging techniques allows us to visualize, for the first time, immune cell doublets in human peripheral blood. We focused on a particular type of cell doublets, the pairing of one patrolling cell (myeloid lineage, monocyte) and one fighting cell (lymphocyte lineage, T cell). We looked at how common these doublets are and how they are functioning in healthy individuals, as well as in the context of infection and vaccination. We found that doublet frequency fluctuates over time following tuberculosis treatment or Tdap vaccination. Doublet levels are also higher in patients with severe cases of dengue fever. With the publication of this research in eLife in June 2020, our group is now positioned as a major contributor in this completely novel field of research. At the same time, I addressed several technological challenges in the field. This work, published in the journal Cytometry Part A in May 2020, revealed that current single-cell technologies are not sensitive enough to reliably flag the presence of doublets within human cell suspensions such as human blood, which will "contaminate" single-cell studies and

potentially lead to wrong interpretations, such as the discovery of novel immune cell types with mixed lineage features. To help address this technology gap, I identified robust data analyses and experimental strategies to help other researchers distinguish between doublets and singlets, and thus avoid data misinterpretation. Finally, I launched an experiment to use RNA sequencing to study which types of T cells and monocytes tend to be present in doublets. This work may open the door to a fundamental understanding of doublet biology.

After seeing the results from the experiments funded by my SPARK award, my Principal Investigator, Bjoern Peters, Ph.D., committed additional funding to further this study. This allowed me to perform single-cell RNA sequencing data in T cells and monocytes in doublets vs. singlets isolated from a small number of patients with active tuberculosis.

As of March 2022, one of my projects is a direct continuation of my SPARK project. The main focus is to understand the molecular information contained within immune cell doublets present in human blood samples collected from patients with infectious diseases such as tuberculosis, or following vaccination. How and where do doublets form, which information do they contain, can this information be used as a biomarker of disease—or to gauge how well an individual is responding to a vaccine?

My other current project focuses on the characterization of immune cell populations circulating in the blood of individuals infected with Mycobacterium tuberculosis at various stages of disease, to gain a better understanding of what constitutes a good vs. bad immune response to the bacteria, in order to design more effective therapeutic and vaccine strategies.

SPARK STAR HIGHLIGHTS

Follow-on Funding Secured

- Dr. Burel won a $425,000 one-year grant from The Conrad Prebys Foundation in February 2022.

Papers Published Related to Projects:

- Circulating T cell-monocyte complexes are markers of immune perturbations, *eLifeSciences*, June 25, 2019
- The challenge of distinguishing cell-cell complexes from singlet cells in non-imaging flow cytometry and single-cell sorting, *Journal of Quantitative Cell Science*, April 18, 2020

Career Advancement:

- Dr. Burel was promoted to Instructor at LJI in January 2020.

SPARK | THE TULLIE AND RICKEY FAMILIES SPARK AWARDS FOR INNOVATIONS IN IMMUNOLOGY

Heather Callaway, Ph.D.

Postdoctoral Fellow, La Jolla Institute for Immunology
2022 SPARK Award Winner
Disease Impact: Rabies

WHAT IF WE COULD DESIGN LONGER-LASTING VACCINES?

Funded by: The Rosemary Kraemer Raitt Foundation Trust

Rabies virus is the source of an ongoing global pandemic that affects both humans and animals. Untreated rabies infection is nearly 100% lethal, and in spite of the existence of rabies vaccines, rabies continues to kill over 50,000 people every year and place a significant financial burden on the livestock industry. One reason rabies persists is because vaccine protection is short-lived. After about a year, anti-rabies antibodies decline and are no longer protective, necessitating annual repeat vaccinations to maintain immunity. Post-exposure treatment for rabies is also possible, but consists of multiple doses of rabies vaccine and expensive antiviral antibodies, making it unaffordable in low-income countries where most rabies deaths occur. To prevent these deaths, we need an effective, long-lasting rabies vaccine so vulnerable individuals only need to be vaccinated once.

Rabies vaccines work by generating antibodies against the rabies glycoprotein, the only protein on the surface of rabies virus. To develop a robust, long-lasting immune response, a vaccine needs to present multiple copies of identically shaped rabies glycoprotein trimers (three glycoproteins forming a functional unit) to the immune system. However, rabies glycoprotein is unusually heterogeneous and adopts multiple shapes, making it difficult for the immune system to "see" the right forms of rabies glycoprotein to generate long-lasting antibody-producing cells. In the Saphire Lab, I have engineered a stabilized version of the rabies glycoprotein to greatly reduce its heterogeneity. With SPARK Award funding, I plan to vaccinate mice with this newly stabilized protein and test how the immune system responds to stabilized rabies glycoprotein compared to existing vaccines containing heterogenous, unmodified rabies glycoprotein. I will determine whether the immune response is longer lasting after vaccination with stabilized glycoprotein and isolate anti-rabies antibodies that can be used both to develop more potent antiviral therapies

and to further engineer the rabies glycoprotein to improve vaccines. If successful, I will obtain preliminary data that will give valuable insight into the production of an improved, long-lasting rabies vaccine and potent antiviral antibodies, and allow me to apply for additional funding for larger scale testing.

Estimated Project Completion Date: 09/30/22

SPARK | THE TULLIE AND RICKEY FAMILIES SPARK AWARDS FOR INNOVATIONS IN IMMUNOLOGY

Abhijit Chakraborty, Ph.D.

Instructor, La Jolla Institute for Immunology
2020 SPARK Award Winner
Disease Impact: Cancer

SOLVING THE PUZZLE OF A SHATTERED CHROMOSOME

Funded by: François Ferré & Magda Marquet and Various Donors

Around 40 percent of cancer patients have tumors that show a “catastrophic” chromosomal rearrangement. This suggests that an understudied phenomenon called chromothripsis, a “shattering” of the chromosome, is a key factor in cancer progression.

For my SPARK project, I studied the role of chromothripsis in triggering acute lymphoblastic leukemia (ALL), which makes up 74% of childhood leukemia diagnoses, and is the most common childhood cancer. In 2021, there were more than 5,600 new cases of ALL diagnosed in the United States alone. I focused on genomic rearrangements involved in chromosome 21 (iAMP21), which are linked to the highest risk ALL cases. My goal was to shed light on the health consequences of genomic rearrangements at this site—and improve our general understanding of chromothripsis. I also aimed to devise new strategies for studying genomic rearrangements as a way to predict cancer growth and provide cancer patients with accurate prognoses.

My first step was to establish a collaboration with iAMP21 and ALL expert Dr. Anusha Preethi Ganesan, an Assistant Adjunct Professor and a hematologist/oncologist at Rady Children’s Hospital, San Diego. Working with Dr. Ganesan led to important cancer insights and gave us access to pediatric cancer samples. I then performed a differential gene expression analysis between samples from iAMP21 and non-iAMP21 patients. This work showed that 1,885 genes were significantly altered between these two groups. This was a fascinating finding, and led us to uncover potential clues as to how gene expression leads to rapid progression in pediatric blood cancer with rearrangements in iAMP21. Our analysis revealed that differentially and highly expressed genes in iAMP21 are involved in biological processes linked to blood formation, immune system development, myeloid cell differentiation, and other relevant functions. I also investigated

gene fusion in these samples, which gave us an estimate of the level of genomic stability in the tumors.

Ultimately, we devised new methods to understand 3D maps of the human genome. This research is critical for understanding chromothripsis, a devastating phenomenon in cells where shattered chromosomes can lead to especially aggressive cancer.

We know that genomic rearrangements at iAMP21 are devastating for patients; however, this project has given me reason for hope for future patients. Because iAMP21 rearrangements come with certain tell-tale gene alterations, there may be a way to expose specific vulnerabilities of cancer cells in these patients and develop targeted therapies.

Going forward, we will perform whole genome sequencing to identify structural variants at base-pair level resolution; Hi-C to understand the alterations in 3D organization; ATAC-seq to understand genome accessibility, and ChIP-seq to explore histone modifications in iAMP21 samples.

In 2021, I received a one-year $245,000 research grant from The Conrad Prebys Foundation to expand this line of investigation further. I'm working on this grant in collaboration with the Ay and Vijayanand Labs at LJI, and Rady Children's Hospital and Children's Oncology Group. This additional funding has been made possible by the initial SPARK funds, which allowed us to explore this area of research. My hope is that this investment by The Conrad Prebys Foundation will allow us to collect sufficient data to compete for additional substantial multi-year research funds.

In 2022, I expect to have sufficient data for an important proof-of-concept research paper, which will be the first of its kind in the studies of pediatric ALL cancers and specifically for iAMP21.

SPARK STAR HIGHLIGHTS

Follow-on Funding Secured

- Dr. Chakraborty won a $215,416 one-year grant from The Conrad Prebys Foundation in March 2021.

SPARK | THE TULLIE AND RICKEY FAMILIES SPARK AWARDS FOR INNOVATIONS IN IMMUNOLOGY

Huy Dinh, Ph.D.

Assistant Professor, University of Wisconsin-Madison
2019 SPARK Award Winner
Disease Impact: Cancer

NOVEL DNA BLOOD TEST FOR EARLY CANCER DIAGNOSIS

Funded by: Mark and Katie Bowles, Kathleen (Cini) and Ira Robb, and Various Donors

Doctors, patients and researchers desperately need ways to detect cancer earlier in the disease progression. One way to detect cancer may be to look for changes in neutrophils, the most abundant blood cell type. The goal of this project was to study if neutrophils and the cells they are derived from exhibit any special markers throughout cancer progression. To do so, we aimed to identify the molecular signatures of neutrophils in the blood of cancer patients. We hypothesized that neutrophils would change their molecular features after interacting with cancer cells, giving us a clue that the cancer was growing.

We compared protein and gene expression as well as DNA activity changes in neutrophils isolated from healthy individuals and from patients with melanoma. Our promising preliminary data gave us insights into why neutrophils are more abundant in the blood of cancer patients—and why they tend to suppress the immune system. The data plays an important part in our recent findings, published in the journal Immunity in August 2020, which brought us closer to determining important and cancer-related gene signatures and biological pathways in cancer patients. To the best of my knowledge, this is the first time this extent of data generation and analysis were done in the context of cancer and neutrophils in humans. I also believe that knowing how human neutrophils develop is especially relevant today because immature neutrophil levels are higher in both the blood and lungs of severe COVID-19 patients.

As of March 2022, I am continuing to work on research related to my SPARK project as I am studying neutrophils as a marker for early cancer detection and response to treatment.

SPARK STAR HIGHLIGHTS

Papers Published Related to Projects:

- Coexpression of CD71 and CD117 identifies an early unipotent neutrophil progenitor population in human bone marrow, *Immunity*, August 18, 2020

Career Advancement:

- In August 2020, Dr. Dinh began a faculty position at the University of Wisconsin-Madison.

SPARK | THE TULLIE AND RICKEY FAMILIES SPARK AWARDS FOR INNOVATIONS IN IMMUNOLOGY

Annie Elong Ngono, Ph.D.

Instructor, La Jolla Institute for Immunology
2021 SPARK Award Winner
Disease Impact: Infectious Disease

HOW DOES PRIOR EXPOSURE TO CLOSELY-RELATED VIRUSES CONTRIBUTE TO PROTECTION OR DISEASE DURING SUBSEQUENT INFECTION WITH A RELATED VIRUS?

Funded by: Barbara Donnell, Bill Passey & Maria Silva, and 2020 Various Donors

Dengue virus has been causing explosive outbreaks in numerous countries, increasing the number of endemic countries, impacting 3.6 billion people worldwide. The severe form of dengue has been known to kill thousands of children annually for decades. The proposed research was aimed at identifying how prior exposure to Japanese Encephalitis Virus (JEV) will modulate disease severity in patients with acute dengue infection. The long-term goal of this research is to explore the host-immune response of individuals with prior immunity to a closely related virus and define which components of the immune response induced protection or pathogenesis in order to design better vaccines. My lab's unique approach and framework explore a major concern of prior immunity in designing and testing vaccines against dengue. This is urgently required to solve the dengue problem that has been escalating worldwide for the past several decades. Moreover, the world is in urgent need for new insights to develop vaccines/therapeutics against dengue and emerging infectious diseases, including COVID-19. Many countries in the world face this problem of co-circulating viruses in the context of flaviviruses (JEV, DENV, ZIV) but also coronaviruses (SARS-CoV-2, HCoVs).

Although we were not able to draw any conclusions about a possible association between pre-existing immunity to JEV and dengue disease severity due to the small sample size, we were able to (1) detect and classify dengue infection in 2017 in Nepal; (2) determine and confirm the presence of prior-exposure to JEV in patients experiencing dengue infection; (3) establish and strengthen the collaboration between LJI and Nepal to successfully continue our research work (in terms of legal

documents and experiments); (4) substantially advance our understanding of cross-reactive immunity; (5) generate substantial preliminary data for grants including NIH opportunities; and (6) prepare the team in Nepal for the future dengue sample collection in 2022 (predicted outbreak year based on dengue cycle in Nepal). Nepal represents a unique model system to address many critical questions related to immunology controlling or exacerbating DENV, flaviviruses, and emerging infectious diseases.

The SPARK Award allowed me to realize how ambitious the project I proposed was and the different steps needed to start working on a new pathogen such as JEV. The SPARK Award set up a critical framework for this new research angle at LJI. The results generated during this project are important and will be used as preliminary data to apply for an NIH grant to work for the next 5 years on this project.

SPARK | THE TULLIE AND RICKEY FAMILIES SPARK AWARDS FOR INNOVATIONS IN IMMUNOLOGY

Nadine Hartmann, Ph.D.

Alliance Manager (Biopharma Services), Epic Sciences
2018 SPARK Award Winner
Disease Impact: Cancer

CAN THE ENVIRONMENT GIVE YOU CANCER?

Funded by: Rachel and Robert Perlmutter, The Thomas C. Ackerman Foundation and Various Donors

In the United States, every three minutes somebody gets diagnosed with blood cancer and every nine minutes, somebody dies from it, making it one of the most fatal cancer types. Yet, we still do not fully understand the causes and factors that contribute to the disease. Mutations in certain genes, such as TET2, are often found in leukemia patients, but there are healthy individuals who harbor these mutations without developing leukemia. Thus, secondary factors might be necessary to trigger the disease. Similar to humans, we found a mouse model of TET2 mutant mice, which develop several types of cancer in some animal facilities while staying healthy in others. Therefore, the aim of our study was to determine whether environmental influences like infections or chronic inflammation are involved in cancer development in these mice. The SPARK Award was used to purchase and maintain wildtype (WT) and TET2 mutant mice at the LJI facility, sequence their gut microbiome and perform pilot infectious experiments.

The first step of our project was to sequence the gut microbiome of wildtype mice and healthy TET2 mutant mice housed in LJI's facility to establish whether they would have similar microbiomes. Dr. Bana Jabri, at the University of Chicago, had recently published that their wildtype mice and sick TET2 mutant mice harbored a similar microbiome, so it was important for us to confirm whether the same was true for the mice in LJI's facility. Once this was established, we were then interested in comparing the difference in the gut bacteria between the healthy TET2 mutant mice at LJI and the sick TET2 mice at the University of Chicago. We found that our healthy TET2 mutant mice have a more diverse microbiome and lack certain bacteria that are known to promote gut inflammation. To further analyze if inflammation-causing bacteria can trigger cancer development in TET2 mutant mice, we colonized these and wildtype mice either with an acute inflammation-causing bacteria

species or chronic but mild inflammation-causing bacteria species and then compared them to control groups that received saline. Interestingly, only the TET2 mutant mice infected with the acute inflammation-causing bacteria showed elevated myeloid cell levels in their blood four weeks after infection. This suggests a strong inflammatory signal might be involved as a first step in triggering cell expansion during disease onset.

SPARK STAR HIGHLIGHTS

Career Advancement:

- Dr. Hartmann got a new position as a Non-Clinical Scientist at Intercept Pharmaceuticals in August 2019.
- Dr. Hartmann got a new position as an Alliance Manager at Epic Sciences in June 2021.

SPARK | THE TULLIE AND RICKEY FAMILIES SPARK AWARDS FOR INNOVATIONS IN IMMUNOLOGY

Rana Herro, Ph.D.

Assistant Professor, Cincinnati Children's Hospital Medical Center, University of Cincinnati
2018 SPARK Award Winner
Disease Impact: Allergy

PROBIOTICS OR ANTIBIOTICS TO CURE ALLERGY

Funded by: Larry and Tiki Spitcaufsky

Over the last decades, the incidence of chronic immune disorders such as food allergies, eczema, asthma, type 1 diabetes, and multiple sclerosis has skyrocketed. Through association studies in humans and experimental animal models, we know that chronic inflammatory disorders are linked to the microbiome in humans and that changes in microbiota (dysbiosis) can cause disease in animal experimentation models. This SPARK project was designed to profile the microbiome in patients during the "atopic march," a term referring to the typical progression of allergic diseases from skin (atopic dermatitis) to gut (food allergy) to lung (asthma) during early life. The overarching goal is to identify a "protective" versus "pathogenic" microbial signature that can be targeted in therapies to halt allergic diseases.

We sequenced stool samples of four atopic march patients matched with healthy controls and were able to see a striking 6.54-fold decrease in alpha diversity in the allergic patients as compared to healthy controls. This indicates a decrease in microbial diversity with predominant microbe species abundantly present in allergic stools compared to healthy ones. Regarding the skin, we observed a clear dysbiosis within the same atopic march individual (history of eczema, asthma) that had contact urticaria on the left arm but not on his right arm, which didn't develop an allergic reaction. Interestingly, the same dysbiosis was amplified in eczematous skin with three predominant microbes abundantly homing the skin. These results are very encouraging, as it clearly indicates a direct relationship between dysbiosis and allergy, observed in localized contact urticaria and amplified in eczema.

We had hoped to expand our human studies, where we could introduce pathogenic microbiomes into allergy-resistant rodents to study whether dysbiosis alone is sufficient to promote atopic diseases. And vice versa, where we would treat allergy-sensitive strains with antibiotics that suppress the pathogenic microbiome, or promote beneficial dysbiosis by supplementing their diet with probiotics to test for disease reversal when a healthy microbiome is restored. If successful, the follow up study would have confirmed the hypothesis that microbial dysbiosis can exacerbate if not cause atopic diseases, and whether developing selective therapies to change the microbiome can benefit patients with allergies. However, the pandemic disrupted this progress, and at this point we haven't been able to resume work on this project. But perhaps in the future if I were able to secure more funding for this, we could aim to form a library of microbial metabolites that could potentially serve as drug targets in human therapies to halt allergic diseases, with little to no side effects considering their prokaryotic origin.

SPARK STAR HIGHLIGHTS

Career Advancement:

- In March 2020, Dr. Herro started an Assistant Professor position (independent lab) at Cincinnati Children's Hospital Medical Center.

SPARK | THE TULLIE AND RICKEY FAMILIES SPARK AWARDS FOR INNOVATIONS IN IMMUNOLOGY

Sara Landeras Bueno, Ph.D.

Postdoctoral Researcher, La Jolla Institute for Immunology
2020 SPARK Award Winner
Disease Impact: Ebola

COMBATING EMERGING EBOLA VIRUS THREATS WITH AFFORDABLE CURES

Funded by Rachel and Bob Perlmutter, Raydene and Peter St.Clair and Various Donors

Ebola virus relies on a viral protein called the polymerase to infect cells. This protein makes copies of the viral genetic material needed to form new virus particles that go on to infect other cells or other people. Polymerase is also a potential target for antiviral drugs. To understand precisely where these drugs should attack the polymerase, scientists need to be able to see the protein structure at the atomic level. However, there are no atomic structures of polymerases, largely because these proteins are difficult to produce, purify and image with techniques such as cryo-electron microscopy (cryo-EM) or crystallography. I proposed purifying Ebola polymerase protein for use in structure work. Succeeding meant overcoming several major technical hurdles.

I designed a platform that allows the rapid production and screening of variants of Ebola virus polymerase, that are engineered, to be more stable and produce more protein. I modified a battery of ebolavirus polymerases from different key pathogenic species of Ebola virus (Zaire and Sudan), the related Marburg virus, the original strain from bats (Mayinga), as well as the non-pathogenic Reston virus with a tag that allows more specific purification of the protein. Interestingly, I found that the expression levels of the various ebolavirus polymerases differed significantly among virus strains. Sudan and Mayinga ebola viruses expressed the highest amount of polymerase complex in human cells. I also stabilized the polymerase complex by co-producing it in cells with another viral protein called VP35. I rapidly upscaled the purification of the most promising engineered candidates using liters of insect cells. It was important to use insect cells for this step because they offer a high level of protein expression.

Preliminary purifications of the polmerase-VP35 complex had low protein yield, but by modifying our purification method, we increased the amount of protein we could get by 50%. From among the purification strategies, I selected those that showed both better yield and purity of the

polymerase-VP35 complex. For the first time, I can visualize the polymerase-VP35 complex clearly in a SDS-PAGE gel, but there are other lower abundant proteins in the mixture. We hypothesize that these proteins are contributing to the stability of the complex. We are labeling the complex with gold in order to locate the polymerase-VP35 complex in the electron microscope grid.

I am now testing different sample preparation techniques to find conditions that will allow me to image this protein complex using cryo-EM, a high-resolution imaging technique that can shed light on the atomic-level details of viral proteins. I plan to further optimize the purification process and the stability of the Ebola virus polymerase-VP35 complex to pursue a high-resolution structure of the Ebola polymerase-VP35 complex. The interaction between the polymerase and VP35 is essential to fix the polymerase in one of its multiple functional states, but during this project, we learned that it is transient. Capturing this shifting interaction will affect the structural homogeneity required for getting high-resolution models by cryo-EM. We plan to increase the stability of the interaction through the binding with viral RNA and through a new artificial interface generated by self-complementing methods. We are also mapping the polymerase domains that serve as interaction partners for VP35. With this knowledge, we can engineer a construct that binds covalently both proteins and stabilizes their interaction. We are ready to start imaging the Ebola polymerase-VP35 complex using LJI's Titan Halo Electron Microscope. Once we get a low-resolution model of the complex, we will start the imaging of the Ebola polymerase-VP35 complex using LJI's Titan Krios Cryo-Electron Microscope. After processing the images, we aim to generate a three-dimensional model that will be further refined and validated.

SPARK | THE TULLIE AND RICKEY FAMILIES SPARK AWARDS FOR INNOVATIONS IN IMMUNOLOGY

Yuan Lin, Ph.D.

Scientist, Amgen
2018 SPARK Award Winner
Disease Impact: Cancer

NANOPARTICLES TO DELIVER CANCER IMMUNOTHERAPY?

Funded by: Various Donors

With solid tumor cancers, the challenge for the immune system is both to identify and then successfully attack the tumor. Recent cancer immunotherapies involving immune checkpoint inhibitors have had success in yanking away tumor cells' "invisibility cloak." However, one of the protective molecules deployed by tumors, adenosine, accumulates in such high concentrations that it can render the immunotherapy ineffective. This suggests that multiple immune suppression mechanisms co-exist and there might be a benefit to combining adenosine blockers (ARIs) with existing check-point immunotherapies. There are several ARI drugs in development, however the challenge remains to design a tool that delivers these drugs in a manner that allows them to counteract the high concentration of adenosine. This SPARK project aimed to determine if nanoparticles could be used as an effective delivery mechanism for ARIs.

For this project, we wanted to test our hypothesis that by using nanoparticles to deliver ARIs, we could more effectively deliver high amounts of ARIs to tumors to shrink and ultimately kill them. We used SPARK funding to combine a particular ARI (SCH58261) with a fluorescent dye and encapsulate them together in a biodegradable polymer to create drug-eluting nanoparticles, which, once injected into a tumor, result in a slow, localized drug release. We then used fluorescence microscopy and flow cytometry to visualize the fluorescent dye first in cell culture to help determine which immune cells take up nanoparticles, and then in vivo in tumor-bearing mice to determine its effectiveness in reducing tumor size. When we analyzed the uptake of nanoparticles in mouse immune cells and then human immune cells, we identified phagocytic macrophages as the primary target for nanoparticle uptake in both. We then moved on to investigating in vivo uptake of nanoparticles, and found that virtually all of the particles remained at the site of injection and that no fluorescence could be detected outside of the tumor, suggesting that

intratumoral injection of nanoparticles could be used to localize drugs within tumors to improve potency and reduce toxicity. Further studies found that our ARI-loaded nanoparticles could more effectively reduce tumor size in mice. The findings from this SPARK-funded experiment have great potential for clinical translation since encapsulation of ARI drugs into nanoparticles may prove to be safer and more effective than orally administered ARIs currently in clinical trials.

SPARK STAR HIGHLIGHTS

Career Advancement:

- In 2020, Dr. Lin got a new position as a full-time scientist at Amgen.

SPARK | THE TULLIE AND RICKEY FAMILIES SPARK AWARDS FOR INNOVATIONS IN IMMUNOLOGY

Cecilia Lindestam Arlehamn, Ph.D.

Research Assistant Professor, La Jolla Institute for Immunology

2020 SPARK Award Winner

Disease Impact: Infectious Disease, Tuberculosis

THE FIGHT AGAINST MYCOBACTERIA

Funded by: Various Donors

To control tuberculosis (TB), we need to understand the mechanisms of host resistance to the Mycobacterium tuberculosis bacterium (Mtb).

Non-tuberculous mycobacteria (NTM) are a group of over 140 mycobacterial species. For this project, I took a closer look at Mycobacterium avium complex (MAC), the NTM species that most frequently colonizes and infects humans. Some evidence suggests that MAC exposure varies in different geographic regions, which might explain the disparate outcomes of BCG vaccination (given where tuberculosis is common), as well as susceptibility to Mtb infection and TB. Despite their differences, MAC and Mtb are related and may have important structural similarities. These conserved “antigens” may induce similar T cell responses against MAC, TB and TB vaccines.

It has been traditionally hard to study these T cell responses due to a lack of reliable tools to measure MAC-specific immune responses in individuals. My initial goal was to identify reagents (T cell epitopes) to measure MAC-specific immune responses in the blood.

I found that MAC contains unique peptides and proteins not present in Mtb or other NTM, i.e. MAC-specific. However, T cell responses were low in all individuals with a history of MAC infection. Therefore, individuals with a history of MAC lung infection/disease do not respond with the same antigen-specific T cell responses seen in most patients with other infectious diseases. This finding suggests individuals with MAC disease may have other underlying immune deficiencies that predispose them to MAC infection. Importantly, my results further indicated that individuals with MAC infection do not have a general defect in the other arms of their immune response, only in their MAC-specific T cell response. This is because they mounted T cell responses against Epstein-Barr and cytomegalovirus comparable to other patient cohorts.

I also found that individuals with MAC disease have lower frequencies of a specific T cell subset known to fight mycobacteria (including NTMs and Mtb), which was also reflected in their specific gene signature. This adds evidence for a T cell deficiency in individuals with MAC disease, which explains the lack of conventional antigen-specific T cell responses. In fact, this could explain why these individuals were infected with MAC in the first place. Further studies will pinpoint the reason for this lack of a specific T cell subset.

In January 2020, I was promoted to Research Assistant Professor, and I now run my own lab. As of March 2022, I'm focusing on both tuberculosis and neurodegenerative disease research. For tuberculosis, I continue to be interested in understanding T cell responses in tuberculosis and what the targets are for that response. For neurodegenerative diseases, my lab is also working on T cell responses and their role in Parkinson's disease, Alzheimer's disease and ALS.

SPARK STAR HIGHLIGHTS

Papers Published Related to Projects:

- Peripheral blood T-cell deficiency and hyperinflammatory monocyte responses associate with MAC lung disease, *bioRxiv*, Under Review

Other Highlights Related to Research or Career:

- Dr. Lindestam Arlehamn published a study on Parkinson's Disease another focus area of hers, in 2022 as follows: Transcriptional analysis of peripheral memory T cells reveals Parkinson's disease-specific gene signatures, *Nature*, March 21, 2022

Career Advancement:

- Dr. Lindestam Arlehamn received a promotion to Research Assistant Professor at LJI in 2020, which means she now has her own lab.

SPARK | THE TULLIE AND RICKEY FAMILIES SPARK AWARDS FOR INNOVATIONS IN IMMUNOLOGY

Ian Mathews, Ph.D.

Visiting Scientist, La Jolla Institute for Immunology
Medical Student, UC San Diego
2018 SPARK Award Winner
Disease Impact: Cancer

PREDICTING PATIENT RESPONSES TO CANCER IMMUNOTHERAPY

Funded by: Gail Naughton

Immunotherapies are a novel class of cancer-targeting drugs that reawaken the immune system against a patient's own tumors. These treatments have shown remarkable clinical success treating melanoma and some of the most advanced tumors known with efficacy, never before shown against these forms of cancer. However, some patients either do not respond to therapy or get worse because of immunotherapy, developing autoimmune-like side effects. Immunotherapy can be miraculous for some and horrendous for others, underscoring the potency of this class of drug. Understanding which patients to give immunotherapy and how we can improve a patient's response to treatment are therefore important clinical questions. With this SPARK Award, we set out to measure the tens of thousands of molecules in the blood of cancer patients treated with immunotherapy, in the hopes that we could detect molecules that can predict who will respond to therapy, and potentially find signatures that suggest how we might improve patient response.

In melanoma patients treated with the immunotherapy ipilimumab (Yervoy), we found molecules that, when present in high levels in a patient's blood, could predict whether that patient would respond to therapy. In addition, we found molecules that either increased or decreased in patients' blood while on therapy, some of which were only changed in patients who developed autoimmune side effects. These latter molecules may be a window into understanding the still unknown mechanisms by which some patients develop severe side effects while other individuals do not. This study was an important first step in a project to ultimately develop a simple blood test for predicting and improving cancer patient response to immunotherapy.

As of March 2022, we are currently expanding the SPARK-funded study of immunotherapy response signatures in blood to R01-funded research into one such molecule with protective qualities in immune checkpoint blockade. This research focuses on predictive features in human blood for different inflammatory states, including cancer immunotherapy and COVID-19 outcome, and has IP around the protective molecule discovery made in cancer immunotherapy patients.

SPARK STAR HIGHLIGHTS

Follow-on Funding Secured:

- Dr. Mathews secured a $117,600, three-year F31 Predoctoral Fellowship from the National Cancer Institute in January 2019.
- Dr. Mathews, with his PI Dr. Sonia Sharma, secured a $4.2M, five-year R01 NCI Grant from the National Cancer Institute in March 2021 to build on results obtained from Dr. Mathews SPARK project.

Intellectual Property:

- LJI has filed 2 patents related to the IP surrounding this project.

Papers Published Related to Projects:

- Functional analysis of immune signature genes in Th1* memory cells links ISOC1 and pyrimidine metabolism to IFN-γ and IL-17 production, *The Journal of Immunology*, February 5, 2021
- Cellular sensing of extracurricular purine nucleosides triggers an innate IFN-β response, *Science Advances*, July 22, 2020
- Plans to submit a paper to *Nature Medicine* in summer 2022.

Career Advancement:

- Dr. Mathews was accepted to the UC San Diego School of Medicine in Fall 2020 for an M.D./Ph.D. program, after which he hopes to pursue a career in translational and clinical research around cancer prediction and immune-based modalities for cancer prevention.

SPARK | THE TULLIE AND RICKEY FAMILIES SPARK AWARDS FOR INNOVATIONS IN IMMUNOLOGY

Maria Inês Matias, Ph.D.

Postdoctoral Fellow, La Jolla Institute for Immunology
2022 SPARK Award Winner
Disease Impact: Cancer

WHAT IF WE COULD CONVERT PRO-TUMOR CELLS INTO TUMOR-FIGHTING CELLS?

Funded by: Larry and Tiki Spitcaufsky

Anti-tumor immunotherapies represent a promising approach for cancer treatment. These therapies are based on using the patient's own immune cells, equipping them with superior anti-tumor properties, expanding them in the laboratory and re-injecting them back into the patient. Despite the impressive successes that cancer immunotherapies have had, these approaches have also been met with disappointing roadblocks. One major obstacle is the recruitment and accumulation of suppressive cells that are favored in the metabolic environment of the tumor and which strongly inhibit the protective function of anti-tumor killer cells. In my previous research I have shown that it is possible to change the function of a cell by manipulating its nutrients and metabolites. Moreover, my host laboratory has demonstrated that these suppressive cells can be converted into protective killer cells. Together these advanced findings fueled the "spark" that led to my hypothesis that it should be possible to turn suppressor cells into potent killer cells by manipulating nutrients. Instead of blocking the anti-tumor immune response, these cells would greatly promote the efficacy of the immune system to eliminate the tumor and prevent the spread of the cancer cells. Therefore, I propose to test my hypothesis and design a strategy to convert blocking suppressive cells into anti-tumor killer cells by manipulating nutrients or the pathways they activate in the cell. My ultimate goal is to overcome a major roadblock in the fight against cancers, by designing a strategy that will take advantage of the tumor-seeking suppressor cells and turn them into potent killer cells that instead of obstructing and suppressing the anti-tumor response, greatly enhance the efficacy and success of protective immunity and medical interventions.

SPARK funding will make it possible for me to embark on this pioneering study and generate compelling data not only to prove my initial hypothesis but also to fuel my research career goals to fully develop this innovative concept into most successful immunotherapies that effectively cure and even prevent all cancers.

Estimated Project Completion Date: 09/30/22

SPARK | THE TULLIE AND RICKEY FAMILIES SPARK AWARDS FOR INNOVATIONS IN IMMUNOLOGY

Sara McArdle, Ph.D.

Staff Scientist, La Jolla Institute for Immunology
2019 SPARK Award Winner
Disease Impact: Cancer

SUPER-RESOLVING CANCER IMMUNITY

Funded by: Anthony Carr, The Ecke-Meyer Family Foundation, and Bill Passey and Maria Silva

Understanding how the immune system recognizes and responds to cancer is essential for developing treatments that augment our natural defenses. I aimed to develop a brand-new technique that allows for single-molecule imaging of human tumor biopsies, which has never been done before. I believe that better imaging will lead to new insights into how immune cells interact with cancer cells and fuel the discovery of new therapeutic targets.

My Tullie and Rickey Families SPARK Award made it possible for me to explore a recently developed super-resolution microscopy technique called DNA-PAINT. With this technology, we are closer to developing a method for super-resolution imaging of human tissue sections. Super-resolution microscopy lets us image the interactions between cancer cells and immune cells with unprecedented detail, down to the single-molecule level. There are currently multiple techniques for super-resolution imaging but most require intricate sample preparation protocols which limits its applications.

In contrast, DNA-PAINT is theoretically applicable to a wider range of samples, including tissue sections from patients' biopsies. I am adapting the existing published DNA-PAINT methods to perform multi-color imaging of clinical tissue biopsies. This will help scientists study the complex interactions between tumors and the immune system. I have also worked through the processes and pitfalls of DNA-PAINT imaging and taking on major technical challenges in image processing.

By investigating these issues, I have gained far more knowledge of super-resolution imaging. I am now better prepared to help any researchers who come to the LJI Microscopy Core to perform super-resolution imaging, whether through DNA-PAINT or a different technique. Now in 2022, I'm working on a project studying Ebola virus using the same microscope and the same technique as

my SPARK project. The knowledge I gained working on implementing DNA-PAINT for my SPARK project accelerated this work

SPARK STAR HIGHLIGHTS

Other Highlights Related to Research or Career:

- In 2019, Dr. McArdle was selected as one of 17 “Imaging Scientists,” and given a five-year grant from the Chan Zuckerberg Initiative. This award provides multi-year funding for experts in core facilities who make novel or complex technologies more accessible to life scientists to help accelerate their research.

SPARK | THE TULLIE AND RICKEY FAMILIES SPARK AWARDS FOR INNOVATIONS IN IMMUNOLOGY

Melanie McCauley, M.D.

Research Physician, Walter Reed Army Institute of Research
2018 SPARK Award Winner
Disease Impact: Dengue

FLAVIVIRAL INFECTIONS IN NEPAL

Funded by: Hunt and Laura Pettit and Various Donors

Dengue virus has circulated in tropical countries for decades. Now, the virus and its mosquito vector are moving into regions at higher elevations classically associated with cooler temperatures. This means millions more people are going to be at risk of this disease. As new populations become vulnerable to diseases like dengue virus for the first time, we have the opportunity to learn how a naïve immune system responds when an individual is first infected. In addition, we wanted to compare viruses from two separate climatic regions within Nepal to elucidate any genomic differences that may give dengue virus a selective advantage at cooler climate regions. Our goal was to gain valuable data on the immune system's response to primary and secondary infection that could be used to design a safe vaccine for dengue.

To get started, it was crucial to bring much-needed supplies and training to our partners in Nepal so that we could strengthen our clinical research protocol and train new team members in experimental design and implementation to increase the accuracy and speed of our data collection. I used the SPARK funding to travel to Nepal for this critical stage and bring the much-needed supplies to make the sample collection a success. I was also able to purchase and supply our collaborators in Nepal with electronic tablets that can securely store patient data and allow our team at La Jolla Institute to keep track of the collection in real time. While there, I trained the team to use these tablets; discussed procedures for the year's collection; and enrolled several patients who presented with suspected dengue illness in our study. One of the most important purchases the SPARK grant made possible was a temperature-controlled multipurpose centrifuge, which is critical to study human blood samples and preserve cell lines that are crucial to the study of the immune response to dengue virus infection. Another vital component to the study was providing access to very expensive diagnostic kits that help us look for dengue virus IgM antibody, IgG antibody and nonstructural protein 1 (NS1) antigens. SPARK funds allowed me to purchase

enough kits to analyze banked samples from 2016 and 2017. These diagnostic studies allowed me to distinguish between patients who have no disease, a primary infection, or a secondary infection. We now know that 27 of our samples are positive for NS1, and from those we have isolated dengue virus and are planning to do full genome sequencing. Further analysis of samples collected in Nepal demonstrated that people in Nepal exhibit immunity to one or multiple serotypes of dengue virus, supporting our hypothesis that these viruses are more widespread in the country than previously observed. We did not find evidence of Zika virus infections.

SPARK STAR HIGHLIGHTS

Other Highlights Related to Research or Career:

- Dr. McCauley's team at Walter Reed developed a SARS-CoV-2 vaccine based on nanoparticle technology and took the platform from animal models into clinical trials. She was an associate investigator on the phase 1 trial, and her team hopes to leverage this technology toward a pancoronavirus vaccine, effective against epidemic and pandemic-potential viruses like SARS1, MERS, and SARS-CoV-2 or future coronavirus outbreaks.

Other Highlights Related to Research or Career:

- Dr. McCauley got a new position as a research physician at The Henry M. Jackson Foundation for the Advancement of Military Medicine at Walter Reed Army Institute of Research.

SPARK | THE TULLIE AND RICKEY FAMILIES SPARK AWARDS FOR INNOVATIONS IN IMMUNOLOGY

Melissa Meyer, Ph.D.

Postdoctoral Fellow, La Jolla Institute for Immunology
2022 SPARK Award Winner
Disease Impact: Cancer

WHAT IF WE COULD TARGET PRO-TUMORAL IMMUNE CELLS WHILE LEAVING INFECTIOUS-DISEASE FIGHTING IMMUNE CELLS UNHARMED?

Funded by: The Tullie Family Foundation

Neutrophils are the body's first responders to infection and injury. They move swiftly through the blood to sites of infection to control bacteria and alert the immune system to the presence of invaders. In cancer, the neutrophil response becomes persistent and pathologic, often supporting cancer development. Chemotherapy, the most common type of cancer treatment, often leads to low white counts, or depletion of neutrophils, helping tumor regression but also leaving the patient susceptible to infection. I wonder if there is a way to target pro-tumoral neutrophils to support tumor regression while leaving anti-bacteria neutrophils intact to prevent infection in cancer patients. To target the differences between pro-tumoral and anti-bacterial neutrophils, I first need to understand their molecular differences. I will profile cell surface protein expression and coordinate intracellular signaling pathways by RNA expression in healthy individuals and melanoma patients. Understanding the intracellular signaling pathways will help us distinguish between the "good" antibacterial neutrophils and the "bad" pro-tumoral neutrophils. Not only can we distinguish between pro-tumoral and antibacterial neutrophils, but I will also identify differences in cell surface proteins between these populations. The location of cell surface proteins makes them optimal drug targets. So, once we have determined specific cell surface proteins for each population, we can use those particular proteins to target one population while leaving the other intact. Together, these data will connect specific functions of neutrophils subsets with accessible drug targets present on the surface of neutrophils in cancer patients.

Estimated Project Completion Date: 09/30/22

SPARK | THE TULLIE AND RICKEY FAMILIES SPARK AWARDS FOR INNOVATIONS IN IMMUNOLOGY

Mike Norris, Ph.D.

Assistant Professor, University of Toronto
2021 SPARK Award Winner
Disease Impact: Infectious Disease

WHAT CAN WE DO *NOW* TO PREVENT A FUTURE, EVEN DEADLIER PANDEMIC?

Funded by: The Tullie Family

In 2020, the global economy was shut down by a pandemic caused by a virus for which humanity has no drugs available. To ensure we avoid another similar disaster, we must develop antiviral drugs to combat the next viral pandemic. Health experts believe the next pandemic could be caused by a member of the paramyxovirus family. This sprawling family of viruses includes measles, which still infects over 7 million people every year; parainfluenza, which causes croup; multiple livestock pathogens that cost the industry >$100 million annually; and the fairly new but lethal Nipah virus, which kills 9 of 10 people infected. All paramyxoviruses are highly infectious. The infectiousness of measles, for example, is unmatched: a single cough in a room of 100 unvaccinated people will infect 90 of those 100 people. We have no therapies to treat paramyxovirus infections and a paramyxovirus pandemic will be catastrophic.

Despite their deadly potential, paramyxoviruses carry only 6 genes. This is all that is required to copy themselves thousands of times. A key part of the paramyxovirus machinery is the matrix protein, which acts as a critical field marshal that gathers and guides the assembly of viral components to build new viruses and release them from the infected cell. If we could develop an antiviral drug that targets the matrix protein, we could arrest viral assembly and in turn, halt virus spread.

I recently determined the high-resolution molecular structures of several different paramyxovirus matrix proteins, learned how they hijack the human cell membrane, and identified three key vulnerable sites shared by many paramyxoviruses. A drug targeting one of these sites could block all or many paramyxoviruses. LJI has filed a provisional patent application on this concept. I now seek to find the candidate molecules to protect and capitalize on these discoveries. In this project,

I will screen and identify small molecule drugs that bind into these sites to block the assembly and spread of new viruses. I will then test them against numerous paramyxoviruses relevant for human and livestock health.

One of the most important pieces of these viruses is the matrix protein, also known as M. In an infected cell, the virus manufactures thousands of copies of M as it begins making copies of itself. If we could develop an antiviral drug that targets the M protein interaction with the membrane lipid, we could stop infections in their tracks.

For this project, I built on my previous success in solving the 3-D structure of a paramyxovirus M protein in complexes with a membrane lipid. This structure became my “blueprint” to find drugs that can inhibit the membrane binding process. I used a technique called virtual drug screening to evaluate millions of drugs for their potential to bind M. Through this process, I narrowed down the list of over 3.2 million drugs to a few thousand.

As of March 2022, I am continuing research in the same vein of my Spark project. For my project, I computationally screened over 7 million unique compounds for their ability to bind pockets critical for activity of the viral “matrix” protein. These matrix proteins act as molecular scaffolds to drive formation and assembly of new viruses in the infected cell. Compounds that bind and inhibit these proteins will potently halt the formation and subsequent spread of new viruses. Based on the predicted affinity, scaffold diversity, and selectivity across binding sites, I identified a short-list of 100 high potential compounds. To test these compounds, I developed a novel budding assay for high-throughput screening of assembly inhibition. This assay was developed in cells expressing matrix proteins and allows us to identify compounds at BSL-2 that prevent matrix-driven assembly and budding. The assay is cellular, so initial hits will be cell permeable, directly interfere with matrix-driven assembly and budding (no deconvolution of target), and can easily be identified as overtly cytotoxic or not, allowing identification of hits tractable for advancement. My work is now focused on testing the identified compounds in this cellular assay and against live virus and initiating hit-to-lead optimization to create a second generation of compounds with increased specificity and potency.

SPARK STAR HIGHLIGHTS

Career Advancement:

- Dr. Norris accepted a faculty position at University of Toronto and will start his lab there in Spring 2022. His lab will seek to understand the structure and molecular assembly of pathogenic viruses, with special emphasis on paramyxoviruses and filoviruses. The three-dimensional molecular structures we reveal provide detailed information on the biological mechanisms driving viral assembly and will lay the foundation for the development of new, effective, and targeted therapeutic strategies.

SPARK | THE TULLIE AND RICKEY FAMILIES SPARK AWARDS FOR INNOVATIONS IN IMMUNOLOGY

Marco Orecchioni, Ph.D.

Instructor, La Jolla Institute for Immunology
2019 SPARK Award Winner
Disease Impact: Heart Disease

WE ARE WHAT WE EAT... MACROPHAGES CAN SMELL THIS AND DRIVE HEART DISEASE

Funded by: The Tullie Family Foundation

One in four Americans will die from the consequences of atherosclerosis, the buildup of plaques of fat and cholesterol in the arteries. The development of atherosclerosis is strongly affected by what we eat. The bacteria in our intestine eat with us, and some of the metabolites they produce when we eat a western style diet are associated with atherosclerosis progression. Our lab has discovered hundreds of olfactory receptors expressed in specific immune cells called macrophages (big eaters) isolated from atherosclerotic aortas of mice fed with a diet similar in composition to western food. Macrophages are important immune cells because they serve as the first line of defense, protecting the body from bacteria, viruses and cancer cells. They are also the primary contributors to atherosclerosis, the cause of heart attacks and strokes. I investigated how these olfactory receptors might activate macrophages in response to diet-induced volatile compounds.

Thanks to SPARK support, I confirmed the importance of olfactory receptors in mediating the function of macrophages in response to diet-induced volatile compounds. This finding was novel, and in fact, I've worked with LJI's Technology Development team to create an initial filing for intellectual property related to this discovery. The first screening of olfactory receptor function showed that seven olfactory receptors were able to induce inflammation. One particular receptor stood out: Olfr2, which acts as the receptor for octanal. Octanal, a diet-induced compound, can be detected in the blood of mice and humans—and can increase three times when the mice are set on a high-fat diet, similar to a typical western diet. Interestingly, we found that mice given high doses of octanal had a substantial increase in atherosclerosis plaque progression. I discovered that we could reduce inflammation and protect mice from atherosclerosis by blocking the receptor for octanal with a chemical inhibitor, or by modulating its expression with genomic editing techniques. These data give us a promising path for drug development to prevent and treat atherosclerosis-based cardiovascular diseases. This work also answered questions about the basic biology of olfactory receptors and the work in mice appears to be very relevant in humans. I discovered that

the human octanal receptor, called OR6A2, functions similarly to the receptor in mice, inducing inflammation in response to octanal. Finally, some olfactory receptors, including the receptor for octanal, are expressed in certain white blood cells in humans. These cells can be incorporated in atherosclerotic plaques.

The next step in my research is to study the functionality of the newly discovered olfactory receptors, expressed in human monocytes and macrophages. I will also expand the patient population of the study. Additionally, I'll be collaborating with a colleague at Duke University to identify the newly discovered olfactory receptors' agonists, which are the substances that initiate physiological responses when combined with a receptor. Previously established protocol will be applied through that process. Then, new monoclonal antibodies will be generated to detect the protein expression of the most promising receptors. I also plan to screen for small molecule compounds that inhibit OR6A2 through high-throughput screening, and have already identified the required libraries of small molecule compounds approved clinically and by the Food and Drug Administration. I'll then assess the efficacy of successful candidate molecules using flow cytometry.

The international patent related to my preliminary findings is still pending. Additionally, I am awaiting the final results of my National Institutes of Health (NIH) highly competitive Pathway to Independence Award (K99/R00) application. In March 2022, I was notified that I received a highly competitive score for this opportunity.

SPARK STAR HIGHLIGHTS

Follow-on Funding Secured

- Dr. Orecchioni received a $208,000 one-year award from The Conrad Prebys Foundation in March 2021.
- In March 2022, Dr. Orecchioni was notified that he won a $231,000 three-year fellowship from the American Heart Association.

Papers Published Related to Projects:

- Olfactory receptor-2 in vascular macrophages drives atherosclerosis by NLRP3-dependent IL-1 production, *Science*, January 13, 2022
- How the immune system shapes atherosclerosis: roles of innate and adaptive immunity, *Nature Reviews Immunology*, August 13, 2021
- Chapter 47: Flow Cytometry and Mass Cytometry for Measuring the Immune Cell Infiltrate in Atherosclerotic Arteries, *Methods in Molecular Biology: Atherosclerosis: Methods and Protocols*, January 1, 2022

Career Advancement:

- Dr. Orecchioni was promoted to an Instructor at LJI in January 2022.

SPARK | THE TULLIE AND RICKEY FAMILIES SPARK AWARDS FOR INNOVATIONS IN IMMUNOLOGY

Estefania Quesada-Masachs, M.D., Ph.D.

Postdoctoral Fellow, La Jolla Institute for Immunology
2020 and 2022 SPARK Award Winner
Disease Impact: Type 1 Diabetes

2022 PROJECT: WHAT IF A PERSON'S INSULIN-PRODUCING BETA CELLS ARE ORCHESTRATING THEIR OWN DEMISE?

Funded by: Larry and Tiki Spitcaufsky

Type 1 diabetes (T1D) Type 1 diabetes (T1D), an autoimmune disease that usually affects young people, is on the rise worldwide, and there is no cure to date. We still do not completely understand the underlying processes that drive the dysfunction and destruction of insulin-producing β cells in the pancreas leading to a lifelong dependency on insulin injections. It is key that we understand whether this disease is primarily triggered by the β cell itself or by the immune system to offer better therapeutic options.

One of the questions that has arisen is whether β cells are orchestrating their own demise by causing a direct response in the immune cells around them. Several discoveries have suggested this notion. Recently, we have demonstrated that β cells can express HLA class II in patients with T1D. HLA class II is a molecule that can be used by our own cells to "give orders" to our immune system. When an immune cell contacts HLA class II it responds to proteins that otherwise would probably be ignored. Whether β cells can express HLA class II in T1D has been a controversial topic for decades, and we have just settled this controversy recently. I have demonstrated that even healthy β cells in isolated human islets can express HLA class II when they are stressed. Based on these advancements, my hypothesis is that β cells can affect the behavior of immune T cells around them.

As part of an ongoing collaboration, I will obtain human pancreatic islets (from City of Hope), and human islet microtissues using an innovative 3D platform (from InSphero). T cells from compatible donors will be cultured with these islets. I will determine changes in islet function and structure and the magnitude of T cell infiltration under different conditions, using state-of-the art

microscopy systems. Subsequent experiments will seek to uncover the effect of those interactions on the immune T cells.

It is key to understand whether β cells influence the responses of T cells around them. This finding will be of major importance to our understanding of T1D pathogenesis, and it will drastically impact future treatment approaches.

Estimated Project Completion Date: 09/30/22

2020 PROJECT: EXPLORING TOXIC EFFECTS OF T1D PATIENT'S SERUM IN PANCREATIC B CELLS

Funded by Barbara Donnell and Denny Sanford

Patients with type 1 diabetes (T1D) develop the disease when the immune system damages their insulin-producing beta cells, leading to loss of beta cell function and, ultimately, irreversible insulin deficiency. This autoimmune attack of insulin-producing beta cells is thought to be primarily mediated by autoreactive immune cells. Older studies have suggested that the blood serum of patients with T1D may directly harm beta cells, and scientists have hypothesized that autoantibodies and other unknown players in blood serum may be partly responsible for disease progression. For this project, we investigated whether serum derived from patients with T1D could alter the function of healthy islets and induce a different response than serum from non-diabetic patients.

We found that islets incubated in serum from patients with T1D make more insulin in response to conditions of high and low glucose, compared to islets incubated in serum from non-diabetic patients. Furthermore, the increased insulin secretion observed in the islets cultured with serum from the T1D group was similar to that observed in islets incubated with proinflammatory molecules. This suggests that the serum of patients with T1D may be stressing these islets, leading them to secrete more insulin.

Despite using state-of-the-art microscopy techniques, we didn't observe differences in the presence of insulin, glucagon or inflammatory molecules, such as HLA-I, in the islets receiving the serum of T1D patients. To determine what may cause the altered islet insulin secretion, we next characterized the serum of the patients by measuring the levels of 97 different factors known to affect inflammation, immune-regulation, and cell growth. In our preliminary analysis, we found subtle differences in some of these molecules between the serum from T1D patients and the serum from non-diabetic patients, but not enough to justify the functional differences observed. Our more complex multivariate analysis is ongoing, which may yet uncover potential causes for these mysterious alterations in islet function.

SPARK STAR HIGHLIGHTS

Other Highlights Related to Research or Career:

- Dr. Quesada-Masachs Submitted a proposal to the Human Islet Research Network (HIRN) for the HIRN Catalyst Initiative in spring 2022.

SPARK | THE TULLIE AND RICKEY FAMILIES SPARK AWARDS FOR INNOVATIONS IN IMMUNOLOGY

Thomas Riffelmacher, Ph.D.

Instructor, La Jolla Institute for Immunology
2020 SPARK Award Winner
Disease Impact: Heart Disease, Autoimmune Disease

A NEW CELLULAR TARGET TO PREVENT TISSUE DAMAGE FOLLOWING HEART ATTACK

Funded by: The Rickey Family

I was always fascinated with the question of how the lack of oxygen would impact immune functions on a cellular level. Of course, it didn't take a pandemic to teach us how quickly we are in trouble when we can´t breathe—but what does the lack of oxygen mean for our immune cells within tissue? During a heart attack, a stroke and in sickle cell disease, blood vessels become clogged, but we don't understand, on a cellular level, the impact this hypoxia has on the inflammatory response that follows. Using single cell RNA sequencing, I sought to explore the mechanistic changes within innate immune cells in this context, to see if it offers cues to novel druggable targets.

The pandemic presented significant challenges to my project because my original plan relied on outpatient samples from patients who suffered heart attacks. Unfortunately, these samples were no longer collected during the pandemic. Also, since my original collaborator was in the U.K. and I was going to have to travel to retrieve the samples, that wasn't possible due to the travel bans between the United States and Europe. So I had to step back and think about what I could do, in the short time frame I had, to address this question with what is available more locally. Fortunately, thanks to a previous clinical study collaboration between LJI and the University of Wisconsin, I was able to get access to blood samples from sickle cell disease patients. This is a disease of chronic, recurring episodes of blood vessel occlusion, which causes inflammation of the lung. Simultaneously, I used a mouse model of sickle cell disease, harboring the human disease-causing mutation in the hemoglobin gene, for mechanistic follow-up experiments. With these samples, I was still able to explore the impact of vaso-occlusion and resulting hypoxia on immune cells, albeit in a different disease context.

I initially set up a single cell sorting strategy to isolate the Natural Killer T ("NKT") cells and optimize the conditions using the phlebotomy core at LJI for sourcing blood. We requested and received IRB approval to get healthy African American volunteers to donate blood at LJI as appropriate controls for sickle cell patients. I then sorted the relevant immune cells from over 20 sickle cell disease patients and controls. Originally, I proposed to focus only on NKT cells, but we realized, thanks to the capabilities of LJI's sequencing and bioinformatics teams, that we could add four more relevant innate lymphocyte populations. I am now analyzing the data from a total of 10,841 lymphocytes to test if we find evidence for altered inflammatory signaling. The first results were promising, indicating that innate lymphocytes were noticeably depleted in the blood of patients but not in the controls, meaning they may have been recruited into the lung tissue to mediate inflammation. There were several thousand genes differentially expressed, including the pathways that produce inflammatory cytokines. Having the humanized mouse model allowed me to study this in more detail within lung tissue, which of course we cannot get from patients. Indeed, inside the lungs, the NKT cells dramatically increased in numbers and contributed to inflammation. They were also metabolically different.

Due to the logistical challenges and need to pivot my project due to COVID-19, I had an extension to finish this through December 2021, and some of these results are still pending additional experiments and analysis. One of the challenges is to extract the T cell subset identities from the single cell transcriptomic data, which I am now achieving by interrogating gene expression signatures from the literature. Our bioinformatic experts at the LJI are crucial for this effort, which will be key to instruct further mechanistic experiments in sickle cell disease mice. One important conclusion we drew from our data is that depletion of NKT cells may be a promising therapeutic approach, and we are now testing this with a new NKT-depleting antibody we have generated.

SPARK STAR HIGHLIGHTS

Career Advancement:

- Dr. Riffelmacher was promoted to an Instructor at LJI in January 2022.

SPARK | THE TULLIE AND RICKEY FAMILIES SPARK AWARDS FOR INNOVATIONS IN IMMUNOLOGY

Artem Romanov, Ph.D.

Postdoctoral Researcher, La Jolla Institute for Immunology
2021 SPARK Award Winner
Disease Impact: Covid-19, Sex-Based Differences in Disease

WHY ARE CERTAIN VULNERABLE PATIENT POPULATIONS PRONE TO DEVELOPING SEVERE COVID DISEASE?

Funded by: Various Donors

The SARS-CoV-2 pandemic and COVID-19 disease quickly spread worldwide and significantly impacted almost every aspect of public health and daily life. Preliminary metabolic analysis had already revealed that increased levels of certain metabolites in the blood of COVID patients are associated with the development of severe disease. The goal of my project was to comprehensively characterize the blood metabolic profile of COVID disease in vulnerable patient populations, and based on this profile suggest new potential metabolic targets for treating severely affected COVID patients.

To measure the levels of circulating small molecule metabolites or chemicals in the blood of COVID-19 patients, I used a cutting-edge profiling technology called “high dimensional and high throughput liquid chromatography tandem mass spectrometry” or “rLC technology” to analyze blood samples from COVID-19 patients who were admitted to the hospital. These samples were run using the instrumentation available to our lab on a fee-per-service basis in the laboratory of our long-term collaborator Dr. Mohit Jain at UC San Diego. Once I had performed the measurements of metabolites in these patients, I performed statistical data analysis in order to identify those specific blood-borne metabolites associated with mild, or moderate/severe COVID-19 disease outcomes in the patients. I was also interested in seeing if there were any male-specific metabolites in the context of severe COVID-19 disease, because men are more vulnerable to severe disease worldwide. The analysis can be described as the following specific aims:

Aim 1: Find metabolites that are significantly associated with COVID-19 disease severity

Aim 2: Find metabolites that influence COVID-19 disease severity in men

I had to make some adjustments to the data analysis due to the skewed distribution of males and females in the dataset (70% male, 30% female). Basically, the moderate and severe disease categories were combined to achieve sufficient numbers for significance in each group. The results of the analysis showed that 27,875 Metabolites were detectable in most of the

samples (~95%). Surprisingly, the data analysis showed that only a minority of the metabolites

(e.g. 1 metabolite) was significantly associated with developing severe COVID-19 disease . In contrast, most metabolites were specifically associated with developing mild COVID-19. These results demonstrate that presence or absence of specific blood borne metabolites are predictive of

mild COVID-19 disease as opposed to predicting severe COVID-19 disease. This is surprising and different from my original objective of identifying severe COVID-19 metabolites. However, this is also more promising and potentially exciting because mild COVID-19

metabolites can be further developed as biomarkers of mild COVID-19 disease or as immune-stimulatory therapeutic compounds to prevent severe COVID-19 disease.

Regarding sex-specific metabolites, metabolites were specifically associated with the development of either severe or mild disease in males. Interestingly, these metabolites had different effects on the probability of developing severe COVID-19 disease in males and females. These results demonstrate that circulating metabolites are a good indicator of sex-specific differences in COVID-19 disease severity, and may be useful in developing sex-specific biomarkers of COVID-19 disease severity or therapies.

Based upon these exciting findings, I would like to validate and extend the initial observations I made during the course of this SPARK project by 1) expanding my metabolite analysis to additional patient cohorts to see if they replicate in other COVID-19 patients; 2) identifying the chemical nature of the mild COVID-19 metabolites and the male-specific COVID-19 metabolites; 3) exploring the use of these metabolites as biomarker predictors or modulators of COVID-19 disease outcomes. To do this, I will apply for R21 or R00/R99 funding from the NIH/NIAID, and also work with my PI Sonia Sharma to apply for a collaborative R01 grant funding from the NIH/NIAID. Additional patient cohorts can be obtained from key collaborations between myself, Dr. Sharma and Dr. Shresta at LJI.

SPARK | THE TULLIE AND RICKEY FAMILIES SPARK AWARDS FOR INNOVATIONS IN IMMUNOLOGY

Payel Roy, Ph.D.

Postdoctoral Fellow, La Jolla Institute for Immunology
2021 SPARK Award Winner
Disease Impact: Heart Disease

CAN WE DIAGNOSE EARLY SIGNS OF HEART DISEASE USING BLOOD TEST-BASED SCREENING STRATEGY?

Funded by: The Rickey Family

Current screening for heart disease is limited to traditional risk factors like cholesterol, smoking, hypertension and diabetes. However, it is becoming increasingly evident that disease progression is driven by immune cells, particularly T cells. My long-term goal is to develop assays that can detect as well as reveal the biological identities of harmful T cells, and my SPARK project was an important part of collecting data to inform this approach.

A T cell response is not a random event, but is directed towards specific molecules known as “antigens”. Normally, T cells respond only to foreign antigens, such as bacteria and viruses. But in case of heart disease, some rogue T cells recognize self-antigens, i.e., molecules that are present in our own bodies. In mouse models, it has been shown that apolipoprotein B, the protein that forms the backbone of low-density lipoprotein (LDL), is a major inducer of heart-disease related harmful immune responses. Therefore, I focused my attention on this protein and began to set up experimental strategies to examine whether human T cells similarly respond to this self-antigen. I optimized my immunological assays and achieved high levels of sensitivity that allowed detection of rogue T cells in the blood of these people. These T cells exhibited an inflammatory phenotype, concordant with their postulated role in triggering chronic inflammation that characterizes the long latent phase of heart disease. I used multiple positive and negative controls to ensure that my detection protocol is specific. I discovered that six specific regions, known as “epitopes,” in the human apolipoprotein B self-antigen have predominant roles in inducing the rogue T cells.

Having ascertained the identities of rogue T cell responses in human heart disease, I wanted to explore the potential of my immunological assays as a screening tool. To do this, I adopted two different strategies.

1. I collected blood samples from donors from different age groups (range 20–66 years). One part of the sample was sent to clinical labs at UC San Diego to measure lipid risk factors, such as triglyceride, total cholesterol, LDL-cholesterol, etc. The other part was used to run my immunological tests. None of these donors had a known condition of heart disease. I found that donors who had elevated levels of lipid risk factors also showed increased responses in my assays. Such heightened immunological responses were detectable even in donors who were in their 20s and 30s.
2. I obtained blood samples from patients undergoing coronary angiography at University of Virginia. They were divided into low and high severity groups based on their heart disease burden (standard angiographic scores). Rogue T cell responses were significantly higher in patients who were angiographically determined to have high cardiovascular disease burden.

These data suggest that my assays can detect heart-disease related inflammatory T cell responses in young and middle-aged people who have abnormal levels of blood lipids and cholesterol. Strong correlation exists between my assay scores and angiographic scores of heart disease. This signifies the tremendous potential of this workflow in evaluating early signs of the disease. This study is the first of its kind in the field of heart disease.

I hope to secure additional funding to continue this work so we can learn more about these rogue T cells so that we can develop our immune-based screening strategy. We have acquired more samples from patients recruited through Scripps, UCSD and University of Virginia. The goal is these studies will facilitate immunological assessment of cardiovascular risks and will shed light into the pathophysiological nature of autoimmune-related cardiovascular disease.

SPARK STAR HIGHLIGHTS

Papers Published Related to Projects:

- How the immune system shapes atherosclerosis: roles of innate and adaptive immunity, ***Nature Reviews Immunology***, August 13, 2021

Other Highlights Related to Research or Career:

- Dr. Roy is currently in the process of IP filings related to discoveries made as a result of her SPARK project data.

SPARK | THE TULLIE AND RICKEY FAMILIES SPARK AWARDS FOR INNOVATIONS IN IMMUNOLOGY

Priyanka Saminathan, Ph.D.

Postdoctoral Fellow, La Jolla Institute for Immunology
2022 SPARK Award Winner
Disease Impact: Covid-19, Sex-Based Differences in Disease

WHAT IF WE COULD UNCOVER THE BIOLOGICAL MECHANISM THAT MAKES MEN MORE VULNERABLE TO COVID-19?

Funded by: Bill Passey and Maria Silve and Various Donors

COVID-19 infection begins as a mild upper respiratory tract infection for most individuals and rapidly evolves into a severe respiratory distress syndrome, acute respiratory failure, coagulopathy, septic shock, and metabolic acidosis in a subset of patients. A recent metadata analysis of more than 3 million cases across the globe has shown that men are three times more likely to require intensive care therapy in comparison to their female counterparts. Several factors such as sex hormones and their ability to modulate an inflammatory response have been suggested as possible causes for this disparity. Increased interferon beta (IFNβ – an inflammatory cytokine) expression in females has been previously associated with antiviral effects with a more favorable clinical outcome during viral infection. In an attempt to unravel this mysterious phenomenon, our lab has been closely analyzing human metabolomic data from several COVID-19 patient cohorts. We observed that the ADA2 (adenosine deaminase 2) enzyme, which has been previously shown by the Sharma Lab to inhibit the production of IFNβ, has higher enzyme activity in the serum of male patients in comparison with their female counterparts. We hypothesize that the increased ADA2 activity therefore suppresses the production of a robust and early antiviral response in men, and therefore contributes to the previously described sex bias among COVID patients. By modeling COVID-19 infections in a humanized COVID-mouse model, we will monitor disease onset, severity, viral dissemination and lethality in the presence of increased ADA2 activity. Thus, through this current proposal we hope to validate the importance of ADA2 activity and the role it plays in the SARS-CoV-2 pathology.

Estimated Project Completion Date: 09/30/22

SPARK | THE TULLIE AND RICKEY FAMILIES SPARK AWARDS FOR INNOVATIONS IN IMMUNOLOGY

Gurupreet S. Sethi, Ph.D.

Postdoctoral Fellow, La Jolla Institute for Immunology
2022 SPARK Award Winner
Disease Impact: Asthma

WHAT IF WE COULD CURE ASTHMA BY ERASING THE IMMUNE SYSTEM'S MEMORY OF ALLERGENS?

Funded by: The Rickey Family

Allergic asthma is a complicated disease of the lung airways. It causes inflammation, swelling and mucus production, which makes breathing very difficult and painful. Basically, it is the result of a misguided attempt by our immune system to combat the entry of allergens into the body. Similar to an infectious disease, our immune system keeps the memory of all the previous allergen encounters within a cell type called T lymphocytes. We hypothesize that a very small population of these allergic memory T cells are long-lived and can persist in our lungs even after the removal of allergens. Interestingly, these cells can keep renewing themselves similar to stem cells. We suggest that these long-lasting immortal memory T cells of the lungs are responsible for the severity of disease upon subsequent allergen exposure and in many individuals maintain the disease from early childhood into late adulthood. We believe the main obstacle to an effective curative therapy is the need to eliminate this previous memory of our immune system. Thus, in this project we will try to discover the identity of these long-lasting memory cells and to uncover the factors and signaling pathways that make them immortal and then further use this information to eradicate these allergic memory cells from the lungs.

The SPARK funding will allow us to answer this question by conducting a series of experiments to generate a proof of concept of our hypothesis. For this, we will use a well-established murine model of severe persistent asthma where long-lasting memory cells in the lungs are induced upon allergen exposure. We will sort the lung cells and examine which genes are highly expressed in the long-lasting memory T cells. For this we are using a technique called

single-cell RNA sequencing. Overall, we will attempt to identify these long-lasting immortal memory stem T cells in the lungs through their genetic profiles and also discover what proteins they express that could be potential targets to develop an effective therapeutic approach for asthmatics.

Estimated Project Completion Date: 09/30/22

SPARK | THE TULLIE AND RICKEY FAMILIES SPARK AWARDS FOR INNOVATIONS IN IMMUNOLOGY

Vipul Shukla, Ph.D.

Assistant Professor, Feinberg School of Medicine, Northwestern University

2020 SPARK Award Winner
Disease Impact: Cancer

CRACKING THE FOLDED CODE IN OUR DNA

Funded by: The Tullie Family

My SPARK proposal was aimed at understanding the biological significance of unusual folded codes, known as G-folds in the genome. G-folds form in regions of DNA with a high abundance of "Gs," which is one of the four letters that make up DNA (A, T, G and C). For the last several decades, researchers have focused on understanding the mechanisms by which the linear sequence of letters in our DNA is read and interpreted within cells but we know very little about how cells decipher the information encoded in unusual DNA structures, such as G-folds.

To answer this question with my SPARK project, I focused on two important objectives. The first objective was to develop new technologies to detect G-folds in normal cells, as well as their altered formation in cancer cells. My second objective was to identify precisely how our cells read and interpret information encoded in G-folds.

We used a special chemical called NMM to detect G-folds in normal cells while also studying their perturbations during the development of cancer. We then used this newly developed NMM-based detection method to measure G-fold levels in normal cells and revealed a striking increase in the levels of G-fold structures in blood cancer cells that arise upon loss of certain DNA modifying proteins. Using unbiased genome-wide approaches, we also found that G-folds are a "fragility signature" in DNA and represent common sites in the genome where DNA frequently breaks. These DNA breaks are a hallmark of many different cancers.

Next, I set out to understand the mechanisms by which the information encoded in G-folds is relayed inside cells. I performed a mass-spectrometry screen to identify specific G-fold recognizing proteins. Since proteins are the key functional units in cells, this approach shed light on the biological functions of G-fold structures. We were able to identify several previously unknown G-

fold interacting proteins. As we hypothesized, many of the G-fold recognizing proteins that we identified through our screen are known to be essential for controlling key functional genomic states, thereby further endorsing the notion that G-folds represent a critical and underappreciated code in the genome.

In summary, these findings have implications towards our fundamental understanding of genome biology and represent a crucial step towards elucidating the biological functions of these enigmatic DNA structures.

SPARK STAR HIGHLIGHTS

Papers Published Related to the Project:

- TET deficiency perturbs mature B cell homeostasis and promotes oncogenesis associated with accumulation of G-quadruplex and R-loop structures, ***Nature Immunology,*** December 2021

Career Advancement:

- Dr. Shukla started his own lab at Northwestern University in Winter 2022. His lab's current research is focused on understanding the biology of alternative structural conformations in DNA. This is essentially a theme that emerged from Dr. Shukla's SPARK project. The lab members are applying state-of-the-art genetics, genomics, molecular biology and cell biology techniques to decipher the functions of cytosine modifications and structural conformations as alternative DNA codes in the genome. These studies have broad implications on many established paradigms in genome biology and will address fundamental questions related to origins of several different cancers with the ultimate goal of identifying vulnerabilities that could be therapeutically targeted.

SPARK | THE TULLIE AND RICKEY FAMILIES SPARK AWARDS FOR INNOVATIONS IN IMMUNOLOGY

Nicolas Thiault, Ph.D.

Postdoctoral Fellow, La Jolla Institute for Immunology
2021 SPARK Award Winner
Disease Impact: Cancer

WHAT IF THERE IS A HIGHLY EFFECTIVE ENGINEERED ANTI-TUMOR IMMUNE CELL THAT WOULD SAFELY KILL TUMOR CELLS AND LEAVE HEALTHY TISSUES INTACT?

Funded by: Larry and Tiki Spitcaufsky

Anticancer-immunotherapies have resulted in major breakthroughs in the treatment of cancer, especially with immune killer cells engineered to effectively target and kill tumors. Unfortunately, significant roadblocks remain, not in the least the killer cells themselves: They are either from cancer patients and function suboptimally, or are from healthy donors and pose the risk of also harming healthy tissues. Through my research, I discovered a potent killer cell, which will not attack healthy cells and therefore, I predict, will be the ideal cell to be engineered for superior and safe anticancer immunotherapy. My SPARK project will test this hypothesis and lay the foundation for translational research to effectively cure cancer.

Engineering chimeric antigen receptor (CAR) T cells represent one of the most promising immunotherapies against cancer. Unfortunately, CAR T cells are often associated with severe toxicity and a high failure rate in certain types of cancers. I'm aiming to apply the existing CAR T cell technology to new immune cells, discovered at La Jolla Institute for Immunology, called double negative T (DNT) cells. These cells harbor more effective killing capacity with lower toxic potential.

So far, I have focused on establishing methods to generate CAR DNT cells and their appropriate tumor targets. I also genetically modified tumor cells to allow for better recognition by engineered CAR DNT cells and provide the proof-of-concept for my hypothesis. These steps are essential for conducting upcoming in vitro and in vivo experiments. In fact, such characterization of the DNT cell biology has never been described in fundamental nor translational research and would be worth considering for future publications or grant submission.

Unfortunately, because of the COVID-19 pandemic and multiple reagent shortages, the in vitro experiments I needed to run to finish the project were delayed, so I was granted a 6-month extension in January 2022. I have now received all critical components for the necessary in vitro experiments to validate my protocols and address DNT efficiency against tumor cells. Simultaneously, I will confirm that our tumor model can grow in mice. Then I will address whether engineered DNT cells can safely kill tumors and cure cancer in vivo. I look forward to providing a final update this fall.

Estimated Project Completion Date: 07/31/22

SPARK STAR HIGHLIGHTS

Follow-on Funding Secured:

- Dr. Thiault, with his PI Dr. Hilde Cheroutre, secured a $350,000 grant from a private family in January 2022 to build on this work and see if the approach could be successful in the context of glioblastoma (brain cancer).

SPARK | THE TULLIE AND RICKEY FAMILIES SPARK AWARDS FOR INNOVATIONS IN IMMUNOLOGY

Greet Verstichel, M.D., Ph.D.

Postdoctoral Fellow, La Jolla Institute for Immunology
2020 SPARK Award Winner
Disease Impact: Autoimmune Diseases

FROM MOUSE TO MAN, TOWARDS A CURE FOR AUTOIMMUNE DISEASES

Funded by: Barbara Donnell, Bill Passey and Maria Silva, and Various Donors

More than 50 million Americans suffer from autoimmune diseases, and these numbers are rising at alarming rates. We know that patients' immune cells damage healthy organs in autoimmune disease, but scientists don't understand exactly why these immune cells go off course. The thymus is where the body sorts helpful T cells from harmful T cells. T cells that recognize healthy tissue should be educated not to cause autoimmune disease, but in some cases this education is disturbed and harmful autoreactive T cells are the result. We've learned a lot about this process by studying mouse models, but we need more tools to see T cell receptor (TCR) selection in action. For this project, I set up an organoid culture system for human T cells, starting from stem cells isolated from human thymus samples. My hope is to use this system to interfere during human T cell development and assess the functional profile and TCR repertoire of mature T cells.

We used an organoid culture system for human T cells, starting from stem cells, and tested if we could mimic the genetic models of mouse T cell generation by adding pharmaceutical compounds to the culture. We used single cell RNA sequencing to compare the nature of the generated T cells with those from normally developed T cells. In addition, we wanted to get an idea of the ability of the cells to recognize a human body's own tissue and organs. For this purpose, we obtained the genetic sequences of the key structure of the T cell that recognizes its target.

We found that the use of a pharmaceutical compound (2-DG) during human T cell organoid culture mimics the genetic defect we observed in mouse models. This means that it is likely that similar mechanisms underlie the generation of functional T cells in humans and in mice. Identifying how we can manipulate these processes will significantly boost our knowledge of the pathways leading to autoimmunity and will aid in the design of effective new strategies to cure or even prevent autoimmunity.

My Principal Investigator, Hilde Cheroutre, Ph.D., and I will continue to build on this project together, thanks to a $400,000 grant from the Praespero Foundation. With this support, I will build on these data and perform in-depth analyses of the TCR sequences we found. I will need to score the self-reactivity computationally and answer the question of whether inhibiting glycolysis during beta-selection alters the recognition of self-antigens. When I have this data, I will work with Dr. Cheroutre to add it to the paper I hope to prepare and submit to journals in 2022. Additionally, if the hypothesis proves right, this will be important preliminary data to submit an R01 grant to the National Institutes of Health on this project.

SPARK STAR HIGHLIGHTS

Follow-on Funding Secured:

- Dr. Verstichel, with her PI Dr. Hilde Cheroutre, is building on her SPARK project with a $400,000 grant from the Praespero Foundation made in July 2021.

SPARK | THE TULLIE AND RICKEY FAMILIES SPARK AWARDS FOR INNOVATIONS IN IMMUNOLOGY

Daniela Weiskopf, Ph.D.

Research Assistant Professor, La Jolla Institute for Immunology
2018 SPARK Award Winner
Disease Impact: Dengue

DEVELOPMENT OF A DIAGNOSTIC TEST TO PREDICT DISEASE SEVERITY AFTER INFECTION WITH DENGUE VIRUS

Funded by: Francois Ferre and Magda Marquet

Dengue virus (DENV) is the most significant mosquito-borne viral disease in humans. The virus, which actually has four serotypes or strains (DENV1-4), causes a wide spectrum of clinical disease ranging from asymptomatic infection, undifferentiated fever, to dengue fever and dengue hemorrhagic fever (DHF). Currently there is no vaccine or specific treatment available to prevent or treat dengue, which poses a serious risk since as many as 400 million dengue infections occur worldwide annually, of which 96 million manifest clinically. The most serious reaction to dengue, DHF, occurs in a minority of patients and is characterized by bleeding and plasma leakage. Plasma leakage may lead to shock and can result in death if not managed appropriately in a timely fashion. Treatment is largely supportive in nature and relies on diagnosis by a highly trained physician. Interestingly, only a minority of patients will actually progress to severe dengue but currently physicians are unable to identify those patients reliably, requiring careful attention to fluid management and proactive treatment of hemorrhage of most patients. To be able to predict plasma leakage on the basis of a simple blood test is thus a key issue in the study and management of dengue fever. My SPARK project aimed to collect preliminary data to support the development of a diagnostic test that can predict disease severity in dengue patients.

For my SPARK project, I studied blood samples from dengue patients with different disease severity outcomes and defined transcriptional profiles of genes that are up- or down-regulated in these individuals. Interestingly, we identified a panel of genes that clearly predicted the development of bleeding and plasma leakage in this patient subset. We are now in the process to further validate these genes in more detail in an independent study cohort as well as develop a diagnostic test based on establishing short DNA segments, so called primers, that can be used to amplify these

genes of interest in basic clinical settings. This diagnostic test to assess genes associated with disease progression to plasma leakage has the potential to identify patients at risk for severe disease and will be invaluable in the management of dengue virus infection and associated clinical disease. Our findings are expected to make an important, innovative, and original contribution to advancing solutions relevant to military and public health and ultimately lead to an improved outcome for the general public.

As of March 2022, after my focus shifted to SARS-CoV-2 over the past two years, I'm reviving my efforts to understand the immunity against dengue. I have received funding under a P01 grant together with UC Berkeley and the University of North Carolina. With this grant, we are planning to study people that have experienced asymptomatic infection with DENV and compare them with people with severe infections. Importantly, we are studying the immune response before a person gets infected, which is possible by collecting longitudinal samples and following them over a couple of years. Likewise, we are collecting DENV vaccinated individuals and following them over a couple of years. We will then compare breakthrough infections to samples from people that have not been infected.

SPARK STAR HIGHLIGHTS

Follow-on Funding Secured:

- Dr. Weiskopf secured $100,000 from the Human Immune Profiling Consortium in 2019.
- Dr Weiskopf secured $100,000 from the Human Immune Profiling Consortium again in 2020.
- Dr. Weiskopf was also awarded a $274,350 P01 grant from NIAID at National Institutes of Health in 2021.

Papers Published Related to Projects:

- A population of CD4+CD8+ double-positive T cells associated with risk of plasma leakage in dengue viral Infection, *Viruses*, January 5, 2022

Other Highlights Related to Research or Career:

- Dr. Weiskopf was awarded the 2021 Junior Principal Investigator Award from the organization Austrian Scientists & Scholars in North America in September 2021.
- Dr. Weiskopf has been instrumental in the research related to Covid-19 over the course of the pandemic, as follows:
    - SARS-CoV-2 infection generates tissue-localized immunological memory in humans, *Science Immunology*, November 19, 2021
    - Low-dose mRNA-1273 COVID-19 vaccine generates durable memory enhanced by cross reactive T cells, *Science*, October 22, 2021
    - Impact Of SARS-CoV-2 Variants on the total CD4+ and CD8+ T cell reactivity in infected or vaccinated individuals, *Cell Reports Medicine*, July 20, 2021
    - SARS-CoV-2 human T cell epitopes: adaptive immune response against COVID-19, *Cell Host Microbe,* July 14, 2021

- Immunological memory to SARS-CoV-2 assessed for up to 8 months after infection, *Science*, February 5, 2021
- Targets of T cell responses to SARS-CoV-2 coronavirus in humans with COVID-19 disease and unexposed individuals, *Cell*, May 20, 2020

Career Advancement:

- Dr. Weiskopf received a promotion to Research Assistant Professor at LJI in January 2020, which means she now has her own lab. Her focus is on infectious diseases such as dengue fever and COVID-19.

SPARK | THE TULLIE AND RICKEY FAMILIES SPARK AWARDS FOR INNOVATIONS IN IMMUNOLOGY

Holger Winkels, Ph.D.

Neven DuMont Associate Professor for Experimental Cardiology of Aging, University Hospital Cologne (Cologne, Germany)
2018 SPARK Award Winner
Disease Impact: Heart Disease

NEW IMMUNE CELLS IN ATHEROSCLEROSIS

Funded by: San Diego Advisor of the Year Awards Program

The immune system plays a pivotal role in the progression of atherosclerosis, the underlying cause of fatal cardiovascular events, including myocardial infarction and stroke. In previous work, I analyzed the immune cell landscape in aortas of atherosclerotic mice at different stages of disease. During the analysis of these data sets I identified a new aortic immune cell population, which appears to consist of T cell progenitors. T cell progenitors, so called thymocytes (CD4CD8 cells), are thought to only reside in the thymus. Finding them in the aorta may indicate that the thymus, where T cells develop, might have a leak allowing progenitors to escape. This could be potentially dangerous, as uncontrolled T cells responses can lead to autoimmunity. The SPARK award was used to further characterize and describe these new aortic immune cells.

I first screened 16 organs of non-atherosclerotic mice and was able to confirm the presence of CD4CD8 cells not only in the thymus as expected, but also the mediastinal fat and the aortic arch. Using state-of-the-art 3D whole mount fluorescence microscopy of tissue blocks containing the thymus, mediastinal fat, and the aortic arch I confirmed the presence of these cells in all three tissues. Interestingly, CD4CD8 cells in the arch were localized in clusters, whereas fat CD4CD8 cells were aligning along the vasculature in fat tissue. This points to an organization of these cells and argues against a random distribution. A significant amount of my SPARK Award was used to gain further insight into the functionality and identity of these cells. To do so, I isolated CD4CD8 cells from the thymus, mediastinal fat, and aortic arch by flow cytometry and subjected those cells to an extremely low input sequencing protocol to assess their transcriptome. I could demonstrate that CD4CD8 cells isolated from the mediastinal fat and aortic arch could develop into functional T cells in a tissue culture experiment. However, in exploring their gene expression landscape, I found that extra-thymic CD4CD8 cells seem to be impaired in forming particular T cell responses. Further

transcriptomic analysis revealed that extra-thymic CD4CD8 cells express factors, which are important for the maintenance of neurons. A functional innervation of the vasculature, among others, is essential to sense, integrate, and regulate blood pressure. Dysregulated blood pressure can lead to heart disease. Extra-thymic CD4CD8 cells might thus nurse nerves in vessels, a hypothesis I will test in future experiments.

As of March 2022, my current group studies underlying immunological mechanisms of cardiovascular diseases. We are exploring novel mechanisms of macrophage activation in myocardial infarction and aortic aneurysm formation, two fatal diseases. In addition, we study T cell responses towards auto antigens in atherosclerosis, the underlying cause of myocardial infarction.

SPARK STAR HIGHLIGHTS

Papers Published Related to the Project:

- Thymus-derived CF4+CD8+ cells reside in mediastinal adipose tissue and the aortic arch, *Journal of Immunology*, November 5, 2021

Career Advancement:

- Dr. Winkels got a position as Assistant Professor (independent lab) in the Cardiology Department of the Medical Faculty at the University of Cologne (Germany).

SPARK | THE TULLIE AND RICKEY FAMILIES SPARK AWARDS FOR INNOVATIONS IN IMMUNOLOGY

Hui Zhi, Ph.D.

Postdoctoral Fellow, La Jolla Institute for Immunology
2021 SPARK Award Winner
Disease Impact: Infectious Disease, Cancer

CAN WE HARNESS THE HIDDEN POWERS OF OUR OWN TISSUE-RESIDENT IMMUNE CELLS AGAINST VIRAL INFECTIONS?

Funded by: Jim Isaacs

Viral infections are a huge health and economic threat to society, exemplified by the current COVID-19 pandemic. In turn, despite recent immunotherapy approaches to treating cancer, new strategies are still badly needed. Here we propose to study a recently identified "new" immune cell type called ILC1. We hypothesize ILC1 are critical players in controlling infection and tumors, and we will use state of the art 'omics' techniques available at LJI to characterize their immune arsenal. This study will establish a solid foundation to harness the power of ILC1 in humans for fighting against viral infection and even cancer.

Type 1 innate lymphoid cells (ILC1) are among the first responders when a viral infection is trying to gain a foothold in the body. ILC1 cells reside in the tissues and run constant immunosurveillance to spot and eradicate any danger. My goal in this new study is to harness single-cell RNA sequencing techniques to identify the most powerful proteins ILC1 cells use to combat viruses.

Most likely, ILC1 cells don't use their full complement of "weapons," unless they encounter stressed cells, possibly due to the constant expression of CD200R on these cells (an anti-inflammatory receptor that controls the activation of ILC1 cells). Therefore, in my experiments thus far, I've stimulated these cells by treating mice with either pro-inflammatory reagents or antibodies that block CD200R signaling, which did in turn further activate ILC1 cells. After these stimulations, the liver and salivary gland ILC1 cells were purified and sent for single-cell RNA sequencing for further analysis. Currently I'm analyzing the sequencing data to identify effector proteins of ILC1 cells uniquely expressed under stimulation. From these data, we plan to further test which of those proteins are most powerful for treating viral infection and cancers.

Estimated Project Completion Date: 07/31/22

PRESS RELEASES AND ARTICLES FEATURING THE TULLIE AND RICKEY FAMILIES SPARK AWARD WINNERS

(Chronological, newest to oldest.)

La Jolla Institute FOR IMMUNOLOGY | Life Without Disease.®

The Conrad Prebys Foundation grants more than $1.5 million to support critical infrastructure—and a fascinating branch of immunology

Grants to LJI scientists will open up new areas of study against infectious disease and more

Madeline McCurry-Schmidt
Science Writer
February 22, 2022

Share this article:

New support to the LJI Flow Cytometry Core will strengthen research across San Diego. Photo courtesy La Jolla Institute for Immunology

LA JOLLA, CA—The Conrad Prebys Foundation has given more than $1.17 million to support La Jolla Institute for Immunology's (LJI) Flow Cytometry Core and more than $415,000 to fuel infectious disease research led by LJI Instructor Julie Burel, Ph.D.

"This support allows LJI scientists to buy essential equipment and pursue fascinating early career research," says LJI President and CEO Erica Ollmann Saphire, Ph.D. "We are proud to strengthen our relationship with The Conrad Prebys Foundation as we lead immunology research here in San Diego."

The Conrad Prebys Foundation works to improve quality of life—primarily in San Diego County—through the distribution of grants to support medical research, healthcare, visual and performing arts, and other charitable causes. These new grants carry on founder Conrad Prebys' legacy of supporting truly innovative research in the health sciences.

The new funding to LJI's Flow Cytometry Core will allow the Institute to purchase new cutting-edge equipment for research into diseases such as cancers, COVID-19 and heart disease.

"The Foundation is pleased to support this world-renowned institute," says Erin Decker, Director of Grantmaking at The Conrad Prebys Foundation. "LJI is driving understanding and progress across a wide range of diseases. The Institute's vanguard flow cytometry core and its top team of technicians are poised for future breakthroughs, and this funding means LJI scientists, and the entire San Diego research community, will have access to enhanced equipment."

Flow cytometry is a method for sorting individual immune cells and analyzing their roles in fighting disease or contributing to autoimmunity. Skilled technicians in the Flow Cytometry Core at LJI carry out experiments for LJI scientists and offer services to nearby research collaborators at UC San Diego, Moores Cancer Center, the Sanford Burnham Prebys Medical Discovery Institute, Scripps Research and more.

Julie Burel, Ph.D.

The Conrad Prebys Foundation's support will allow Burel to build on the surprising findings of her 2019 Tullie and Rickey Families SPARK Award for Innovations in Immunology. Burel, a member of the Peters Lab, has pioneered the study of doublets. Doublets occur when two immune cells are stuck together. For decades, researchers thought doublets were an accidental artifact of certain laboratory procedures commonly used for analyzing clinical samples. Burel's work, in collaboration with the LJI Flow Cytometry Core, has shown that some doublets do have a role in the body—they may even be a sign that the immune system is fighting an infection.

Burel's research so far suggests doublets do not form randomly, but instead are enriched for certain immune cell subsets with unique gene expression patterns, suggesting a previously uncharacterized form of communication between immune cells.

"Doublets are found at increased frequency in the blood of individuals with diseases such as tuberculosis, or dengue fever," says Burel.

Detecting doublets and analyzing their content could show researchers how well an individual is fighting an infection or responding to a vaccine.

"Dr. Burel and her exciting early stage research is a wonderful illustration of the type of groundbreaking work that Conrad Prebys would have been fascinated with and eager to help move forward," says Decker.

"Mr. Prebys was a visionary who had a long and successful career of discovering and advancing pioneering ideas and the people behind them," adds Decker. "The Foundation identified a gap in the ability for younger investigators to get significant independent funding for their innovative ideas, so this new grant represents an opportunity for us to play a role in furthering scientific breakthroughs and careers."

New research gets to the roots of type 1 diabetes

Madeline McCurry-Schmidt
Science Writer
February 18, 2022

Share this article:

A Q&A with Estefania Quesada-Masachs, M.D., Ph.D.

The immune system wants to attack pathogens and cancers—not a person's own tissues. To do this, the immune system needs to communicate with all the other cells in the body.

Healthy beta cells in the pancreas can talk to the immune system by displaying special markers—little molecular flags—to tell other cells "I'm supposed to be here!" Some of these markers are HLA molecules. Along with other proteins, HLA molecules can influence the immune response, for instance, by stopping immune cells from attacking the body's own tissues.

Diabetes researcher Estefania Quesada-Masachs, M.D., Ph.D.

La Jolla Institute for Immunology (LJI) Postdoctoral Fellow Estefania Quesada-Masachs, M.D., Ph.D., a member of the von Herrath Lab, is investigating cellular communication in the pancreas. She wants to know whether insulin-producing beta cells in the pancreas express markers that trigger T cells to mistakenly kill them. Her work brings us closer to understanding the roots of type 1 diabetes.

In this Q&A, Quesada-Masachs shares her most recent findings, which were published in the February issue of *Diabetologia*.

Q: What big question were you trying to answer with this study?

A: I wanted to figure out whether beta cells in the pancreas are expressing HLA class II in patients with type 1 diabetes. These beta cells are the ones that die in type 1 diabetes.

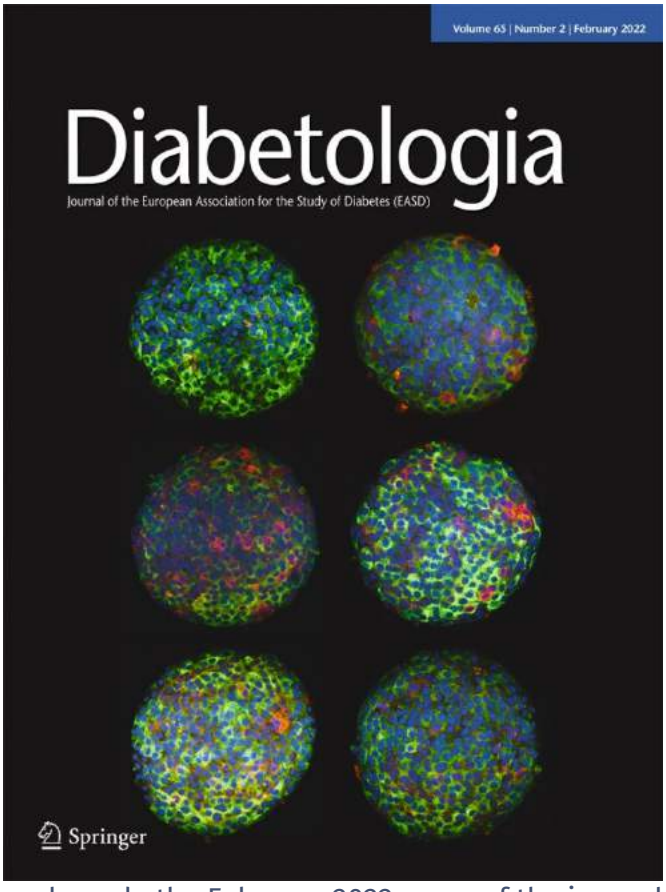
Figuring out whether these cells express these proteins, is important to understand how they communicate with CD4+ T cells.

We looked at pancreatic tissue section slides from the Network for Pancreatic Organ donors with Diabetes (nPOD), and we analyzed samples from healthy controls and from patients with type 1 diabetes. We found that approximately 28 percent of beta cells in patients with type 1 diabetes have upregulated HLA class II in the pancreas.

Q: So are these HLA class II molecules signaling that a beta cell is in danger?

A: We don't know whether expressing HLA class II is detrimental to beta cells or a last cry for help. It could be both. These could just be dysfunctional cells, or they could have a lot of stress around them. They might be trying to regulate the cells around them to say "don't attack me."

We don't really know why the beta cells do it. We know it's probably not favorable for them in the end because they are still attacked in type 1 diabetes.



This new research made the February 2022 cover of the journal Diabetologia

Q: Your paper calls this area of diabetes research a "controversy." Why is that?

A: HLA class II expression in beta cells was first described in the 1980s. But then in the late 80s and 90s, several groups tried to replicate these findings and were unsuccessful.

So part of the scientific community thought that beta cells were not able to express HLA class II and part of the community thought they were able to express HLA class II in certain conditions. For that reason it was controversial.

It was hard to replicate this research because it is difficult to process pancreas and obtain high-quality tissues. We can do this well nowadays thanks to nPOD. Also, the expression of HLA class II by the beta cells is very heterogeneous even within the pancreas of the same individual. Additionally, the antibodies we could use as tools, our microscopy technology, and image analysis software were not as good in the 90s. So much of the work was manual. We have semi-automated software and much better tools now.

In 2019, when my study was already underway, another research group from University of Massachusetts did transcriptomic analysis to uncover an upregulation of HLA class II molecules and class II antigen presentation pathway components in pancreatic beta cells. This was the only study on the topic after 20 years of silence. Our study here at LJI was ongoing, and we decided to keep working since we were using a very different approach than them.

Having two papers arriving at the same conclusion, in our opinion, closes the debate.

Q: What was different about your study?

A: For this study I did immunofluorescence and two kinds of analyses with the nPOD samples. I did whole tissue analysis using a machine learning approach, and this was the part where my colleagues in the LJI Microscopy Core, Sara McArdle and Zbigniew Mikulski, helped me. And then I worked with Bill Kiosses in the Microscopy Core, who randomly selected thirty islets for us to look at using confocal imaging.

Islets are the areas of the pancreas where insulin-producing beta cells live. I had thousands of islets in the whole tissue samples, and these are precious samples, so I wasn't satisfied looking at just 30 islets. So Zbigniew and Sara helped me apply innovations to the project so I could finally analyze more than 7,000 islets.

We also used 3D human cultures with human islets—not animal cells or tissue lines—but actual human islets and human islet microtissues. I demonstrated that stimulating even healthy islets with different concentrations of cytokines can lead beta cells to express HLA class II. That's when these cells receive inflammatory stimuli, because cytokines are inflammatory molecules. We also did RNA studies and found there was an upregulation in the expression of HLA class II RNA-wise in those "inflamed" islets.

So we settled that controversy, and I'm happy the field can move forward.

Q: What are the next steps for you?

A: The next question is whether the beta cells are using this HLA class II expression to communicate with CD4+ T cells? Are they inducing any sort of change in the T cells?

This is what I'm trying to answer with my recent Tullie and Rickey Families SPARK Award project. I want to give a yes or no answer to whether this communication is happening.

Learn more about the study in Diabetologia.

Macrophages in the artery wall 'smell' their surroundings

Immune cells can sniff out octanal in blood, triggering dangerous inflammation and atherosclerosis

Madeline McCurry-Schmidt
Science Writer
January 13, 2022

Share this article:

LA JOLLA, CA—An artery is not like a nose. Or is it?

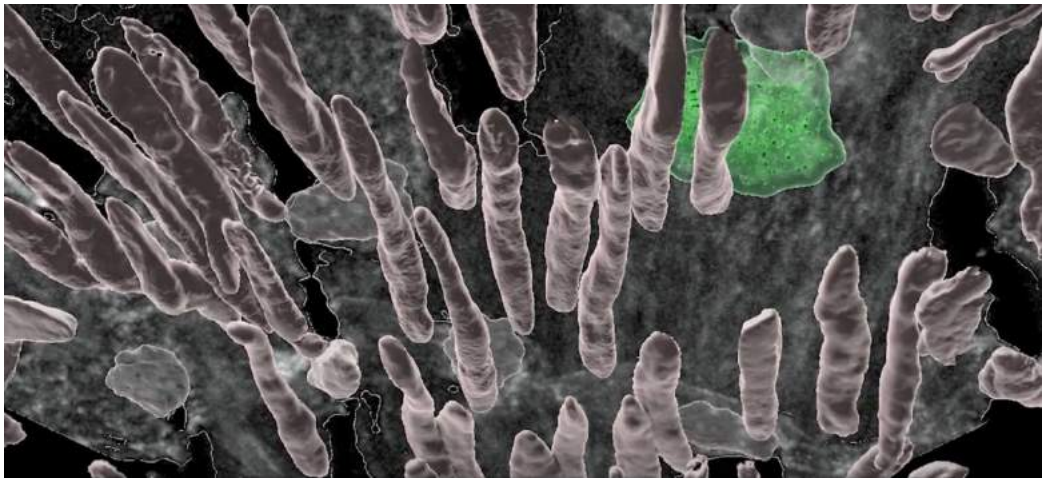
Scientists at La Jolla Institute for Immunology (LJI) have discovered that immune cells in arteries can "sniff" out their surroundings and cause inflammation.

"Smelly molecules can be pro-inflammatory," says study leader LJI Professor Klaus Ley, M.D., a member of the LJI Center for Autoimmunity and Inflammation.

The new study, published in *Science*, shows that this inflammation can lead to cardiovascular disease and atherosclerosis in mice. The researchers reversed this inflammation by blocking immune cells called macrophages from sensing a compound called octanal.

Everyone has a small amount of octanal in their blood, but LJI scientists have shown that people with markers of cardiovascular disease, such as high LDL cholesterol, also have higher levels of octanal. This extra octanal can end up in blood due to diet or a phenomenon in cells called oxidative stress.

The human nose is already good at smelling octanal. Ley describes it as a warmed-over chicken kind of odor. "Like chicken that is not so nice anymore," he says.



Macrophages (shown in green) express olfactory receptors that can detect octanal, which is increased in individuals with markers of cardiovascular disease. (Image courtesy of Matt Ellenbogen, La Jolla Institute for Immunology.)

A 2019 study spearheaded by LJI Scientific Associate Sara McArdle, Ph.D., was the first to show that macrophages in blood vessel walls also have some of the olfactory receptors needed to "smell" molecules. In 2020, LJI scientists were the first to report that these macrophages can sense octanal, thanks to an olfactory receptor called OR6A2.

"Macrophages are some of the most important cells in our immune system," says study first author Marco Orecchioni, Ph.D., an instructor at LJI. "They are constantly checking for signals. We could say they 'sniff' their environment and respond."

The new study is the first to show precisely how sniffing out octanal can boost inflammation in the arteries.

Orecchioni tested the effects of injecting octanal into normal "wild type" mice and into mice where the gene for the mouse macrophage receptor Olfr2 (which corresponds to OR6A2 in humans) was deleted. By comparing these mouse groups, Orecchioni found that inflammation gets much worse as the Olfr2 receptor senses octanal. Over time, the arteries even begin to develop the lesions seen in atherosclerosis.

The researchers then used a molecule called citral (which has a lemon-like odor), known to block this mouse olfactory receptor, and saw that inflammation went down. By making macrophages blind to octanal, they reversed the disease progression.

Ley and Orecchioni think it may be possible to block OR6A2 in humans too. "These receptors are very well known as drug targets," says Ley. "In fact, most drugs on the market today act on this type of receptor, called a GPCR."

Orecchioni's work was supported by La Jolla Institute's Tullie and Rickey Families SPARK Award, which makes it possible for early career LJI scientists to lead "high risk, high reward" projects as they launch an independent research career.

"This project came from a chance discovery of these olfactory receptors," says Orecchioni. "Thanks to SPARK, I was able to explore these more in mice and look at human cells."

Orecchioni was then able to follow up on his preliminary SPARK-funded findings thanks to a generous grant from The Conrad Prebys Foundation.

The researchers are now investigating the jobs of other olfactory receptors found on macrophages. They are also examining how OR6A2 operates in humans.

Orecchioni plans to continue studying atherosclerosis, but he's curious whether olfactory receptors may play a role in metabolic diseases, such as type 2 diabetes. "This study is just the first hint of something new," he says. "It's opened up years of research ahead of us."

Additional authors of the study, "Olfactory receptor-2 in vascular macrophages drives atherosclerosis by NLRP3-dependent IL-1 production," include Kouji Kobiyama, Holger Winkels, Yanal Ghosheh, Sara McArdle, Zbigniew Mikulski, William B. Kiosses, Zhichao Fan, Lai Wen, Yunmin Jung, Payel Roy, Amal J. Ali, Yukiko Miyamoto, Matthew Mangan, Jeffrey Makings, Zhihao Wang, Angela Denn, Jenifer Vallejo, Michaela Owens, Christopher P. Durant, Simon Braumann, Navid Mader, Lin Li, Hiroaki Matsunami, Lars Eckmann, Eicke Latz, Zeneng Wang, and Stanley L. Hazen.

The study was supported by the American Heart Association (grants AHA18POST34060251, 18CDA34110426), the Tullie and Rickey Families SPARK Awards Program, The Conrad Prebys Foundation, Deutsche Forschungsgemeinschaft (GZ WI 4811/1-1), the Chan Zuckerberg Initiative, the National Institutes of Health (grants R01 HL 145454, DK 120515, HL103866, HL126827, DK106000, HL 130819, HL 115232, 145241, HL088093, RR027366, S10OD021831), and Kyowa Kirin.

DOI: 10.1126/science.abg3067

Strange DNA structures may drive cancer development

LJI researchers shed light on the role of TET enzymes in genomic stability and cancers

Madeline McCurry-Schmidt
Science Writer
December 22, 2021

LA JOLLA, CA—Scientists at the La Jolla Institute for Immunology (LJI) have uncovered how loss of TET enzymes can lead to B cell lymphoma. Their research, published in Nature Immunology, could potentially open opportunities for designing drug treatment strategies to target malignant cells in many cancers.

The new research was led by LJI Professor Anjana Rao, Ph.D., in the LJI Center for Cancer Immunotherapy, and experiments were spearheaded by LJI Instructor Vipul Shukla, Ph.D., (soon to be an Assistant Professor at Northwestern University) and UC San Diego Graduate Student Daniela Samaniego-Castruita.

The new research helps scientists finally link two dangerous phenomena in cancer cells.

In previous studies, scientists spotted mutations that cause TET enzymes to lose their function in many patients with blood cancers and solid cancers. Researchers have also found that genomic instability, such as double-stranded breaks in the DNA code, are a common feature in cancer cells

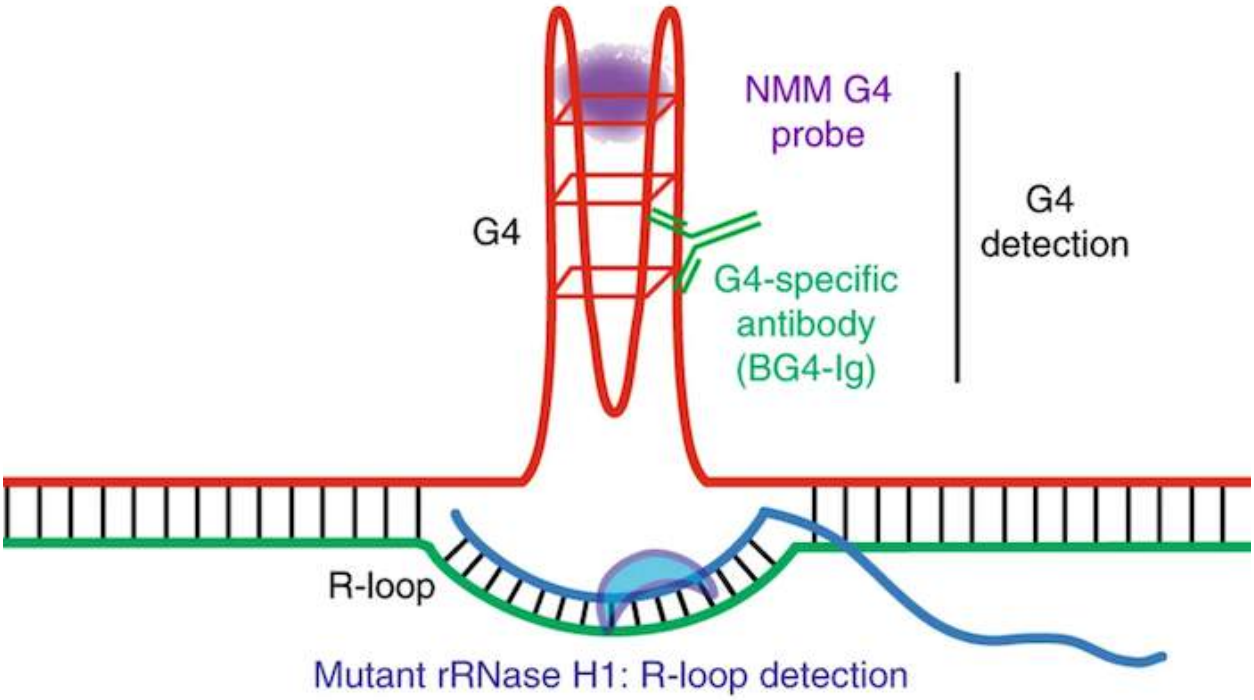
In this project, scientists explored one potential way in which TET deficiency is connected to genomic instability

“This study provides insights about an important question in the field,” says Shukla.

By studying a mouse model of lymphoma, the researchers found that deleting TET2 and TET3 enzymes in mature B cells had huge consequences for B-cell homeostasis. “The TET-deficient mice developed lymphoma, and we observed an increase in marks associated with genomic instability, such as double strand breaks,” says Samaniego-Castruita.

The team then performed genomic analysis for clues to what was happening at the molecular level. They saw that without TET2 and TET3, the DNA became riddled with unusual DNA structures called G-quadruplexes and R-loops.

DNA normally has two strands running parallel to each other, like two rails of a ladder. R-loops appear when a third rail, made of RNA, slips in and forces a gap between the two DNA rails. G-quadruplexes act like knots on the DNA rails. Both R-loops and G-quadruplexes make it hard for the original two rails of DNA to "unzip" as the cell tries to read the DNA code and keep the cell working properly.



Diagrammatic representation of a G-quadruplex (G4) with an associated R-loop structure, illustrating the reagents used for detection of G-quadruplexes and R-loops. (Courtesy Shukla et al.)

Shulka and Samaniego-Castruita examined these DNA structures in depth thanks to funding from the La Jolla Institute's Tullie and Rickey Families SPARK Awards for Innovations in Immunology. "These structures represent sites in the DNA that are much more fragile than other regions," says Shukla. "With this study, we found that TET enzymes are perhaps related to the regulation of these structures, which could in turn explain one mechanism for acquisition of genomic instability in the absence of TET enzymes."

When it comes to B cell malignancies, G-quadruplexes and R-loops appear to be a missing link between tell-tale TET mutations and dangerous genomic instability.

So if G-quadruplexes and R-loops are causing problems, is there a way to stop them from forming?

Shukla and Samaniego-Castruita observed that DNMT1 was upregulated in TET-deficient B cells. DNMT1 is a key enzyme responsible for maintaining marks on DNA called "DNA methylation." DNA methylation is an important regulatory mark in the genome, and is normally removed through the activity of TET enzymes.

Without TET enzymes, the normal give-and-take of DNA methylation marks was broken. So in their next experiment, the scientists also deleted the Dnmt1 gene in TET-deficient B cells in mice to test if levels of G quadruplexes and R-loops could be altered upon removal of DNMT1 protein.

Indeed, deleting DNMT1 was associated with a striking delay in the development of aggressive B-cell lymphomas. Deleting DNMT1 was also associated with decreased levels of G-quadruplexes and R-loops, says Samaniego-Castruita.

The researchers emphasize that regulating G-quadruplexes and R-loops may be just one way TET enzymes control genomic stability. There's also more work to be done to uncover the precise steps that lead TET-deficient cells to accumulate these enigmatic structures in DNA. Someday, the team hopes to devise strategies through which G-quadruplexes and R-loops could be targeted to help cancer patients.

As the Rao Lab continues investigating the effects of TET enzymes, Shukla will be joining the faculty of the Department of Cell and Developmental Biology at Northwestern University this winter. He plans to launch his own laboratory focused on studying alternative structural conformations in DNA.

The study, "TET deficiency perturbs mature B cell homeostasis and promotes oncogenesis associated with accumulation of G-quadruplex and R-loop structures," was supported by the National Institute of Health (grants S10OD016262, S10RR027366, DP2-NS105576, S10OD025052, 2500 S10OD016262, R35 CA210043, R01 AI109842, AI128589 and K99/R00 CA248835); the NIH

National Cancer Institute (grant CA248835); a Leukemia and Lymphoma Society Postdoctoral Fellowship (grant 5463-18); CONACYT/ UCMEXUS fellowships; research funds from LLS grant 5463-18; and the The Tullie and Rickey Families SPARK Awards for Innovations in Immunology Program.

The researchers used resources of the Advanced Light Source, which is a DOE Office of Science User Facility under contract no. DE-AC02-05CH11231.

Additional study authors include Zhen Dong, Edahi Gonzalez-Avalos, Qingqing Yan, and Kavitha Sarma.

DOI: 10.1038/s41590-021-01087-w

Why we give

Raydene and Peter St. Clair

Behind every "eureka!" science headline is a story of perseverance. At La Jolla Institute for Immunology (LJI), donor support sustains many long-term studies and funds "high risk, high reward" projects from many of immunology's brightest minds.

LJI is where supporters Raydene and Peter St. Clair really get to see the nuts and bolts behind scientific breakthroughs. The San Diego philanthropists started attending the LJI "Life Without Disease" lecture series a few years ago and were immediately enthralled.

"I could feel the excitement," says Raydene. "What the scientists were doing had 'tomorrow' applications. I could see immediately how these studies could lead to new drugs or treatments for diseases like cancer."

Peter was struck by how the culture at LJI fueled research breakthroughs. "What really makes LJI stand out is the focus on collaboration," says Peter. "LJI is both an example and a beacon to the other professional groups around the world, demonstrating on a daily basis what scientists are able to accomplish by working together."

Peter is a retired leader in banking and real estate finance. Raydene is a trained nurse and teacher who worked as a special education administrator. This duo knows about long work hours—and still, the work ethic at LJI has impressed them.

At one evening lecture, Peter noted how the scientists spoke passionately about their science and then headed right back up to their laboratories to keep working. The St. Clairs also appreciate how experts at LJI have dedicated time to speaking publicly during the COVID-19 pandemic.

The St. Clairs support LJI researchers through regular donations, including funding for early career researchers via The Tullie and Rickey Families SPARK Awards for Innovations in Immunology.

"It may sound trite to say this, but giving to LJI gets you a big bang for your buck," says Raydene. "People sharing ideas at LJI results in exponential growth once these projects launch."

The St. Clairs are eager for more people to see the excellent effort of scientists of LJI as they research cancer immunotherapies, diabetes, viral infections, and more.

"That is an important part of the process," says Peter. "New support comes through friends talking to friends and letting them come and listen—and see for themselves the value of the incredible science underway at LJI."

SPARK

THE TULLIE AND RICKEY FAMILIES SPARK AWARDS FOR INNOVATIONS IN IMMUNOLOGY

Left to right: Michael Norris, Ph.D., Simon Brunel, Ph.D., Annie Elong Ngono, Ph.D., Jacqueline Tullie, Tom Tullie, Judy Tullie, Mehdi Benkahla, Ph.D., Payel Roy, Ph.D., Artem Romanov, Ph.D., Nicolas Thiault, Ph.D.

The Tullie and Rickey Families **SPARK** Awards for Innovations in Immunology 2021 Awardee Celebration

La Jolla Institute for Immunology welcomed guests to an outdoor reception on June 29 in celebration of the brilliant winners of the 2021 Tullie and Rickey Families SPARK Awards who were announced earlier this winter. The reception was a chance for LJI to recognize supporters of the Tullie and Rickey Families SPARK Awards program and honor the newest class of SPARK Award winners. Attendees had the opportunity to hear from former SPARK winners about their research, recent achievements, and the program's impact.

Left to right: Abhijit Chakraborty, Ph.D, Marco Orecchioni, Ph.D., Nicolas Thiault, Ph.D., Joani Nelson

The immune cell with two faces

Madeline McCurry-Schmidt
Science Writer
September 7, 2021

Hui Zhi, Ph.D., is one of the few scientists in the world with his eye on a very strange immune cell.

Zhi is a postdoctoral researcher at La Jolla Institute for Immunology (LJI). He recently launched new research into mysterious immune cells called ILC1s, thanks to $25,000 in funding from The Tullie and Rickey Families SPARK Awards for Innovations in Immunology. Zhi's specific SPARK Award was funded by long-time LJI supporter James Isaacs, Jr.

LJI Postdoctoral Researcher Hui Zhi, Ph.D.

Some immune cells go after viruses guns-a-blazing. Others follow that old piece of advice: *Speak softly and carry a big stick*.

ILC1s are cells with a hidden talent: they can quickly turn into killers. These cells look like mild-mannered innate lymphoid cells, which usually send out signaling molecules when they spot a potential threat. Scientists long thought that these cells were important players, sure, but they didn't lead the charge against disease.

LJI scientists have shown that ILC1s are actually waiting—and watching. Recent studies led by LJI Professor Chris Benedict, Ph.D., revealed that ILC1s can actually seek out and destroy infected cells. This quick response can stop pathogens such as cytomegalovirus from replicating in the body.

The problem is that no one knows exactly how ILC1s do it.

"A small population of ILC1s may be functioning differently. These cells might make a big difference in fighting off viruses," says Zhi.

Zhi believes ILC1s have been overlooked for too long. He wants to build on the findings from the Benedict Lab and figure out exactly how ILC1s fight pathogens. The trick is to catch ILC1s in action. Zhi thinks this work could give us tools to harness ILC1s and better fight viruses such as SARS-CoV-2 and even cancers.

As Zhi explains, ILC1s are normally part of the flock of mild-mannered "tissue resident" cells that hang out in tissues like the salivary glands and liver. Researchers have studied how ILC1s behave under normal conditions, but they haven't captured how the cells shift to respond to disease.

Zhi is just the person to investigate how ILC1s work. In fact, he has studied both sides of the pathogen vs. immune system battle. As a Ph.D. student at Texas A&M University, Zhi focused on the microbiome and how pathogens attack the body. He went on to study gut bacteria and the microbiome at the University of California, San Diego, where he discovered a passion for immunology.

"When I saw the opening in the Benedict Lab, I had to go for it," says Zhi. "This lab is perfect for me because we study the pathology of viruses, but we also lead research into the immune system's defenses."

For the new study, Zhi will study ILC1s that have either been exposed to a viral challenge or inflammatory molecules. Zhi thinks these conditions will trigger the ILC1s to jump into their disease-fighting Hulk mode.

"We want to see if they start or generate new functions in response to the stimuli," he says.

Zhi will work with LJI's Flow Cytometry Core to separate the ILC1s from other immune cells in the tissue samples. He will then work with the LJI High-Throughput Sequencing Core to perform single-cell DNA sequencing and get a clear picture of which genes are expressed in the ILC1s.

"We need to know when and where these genes are expressed in different cells in the ILC1 population," says Zhi.

Thanks to donor support, Zhi may uncover ways to use ILC1s to help patients fight off viruses, cancers and more. "This is a starting point for me to get into the next phase of my career," says Zhi.

Preparing T cells for the long haul

New LJI research could keep cancer-fighting T cells from burning out before the job is done

Madeline McCurry-Schmidt
Science Writer
July 19, 2021

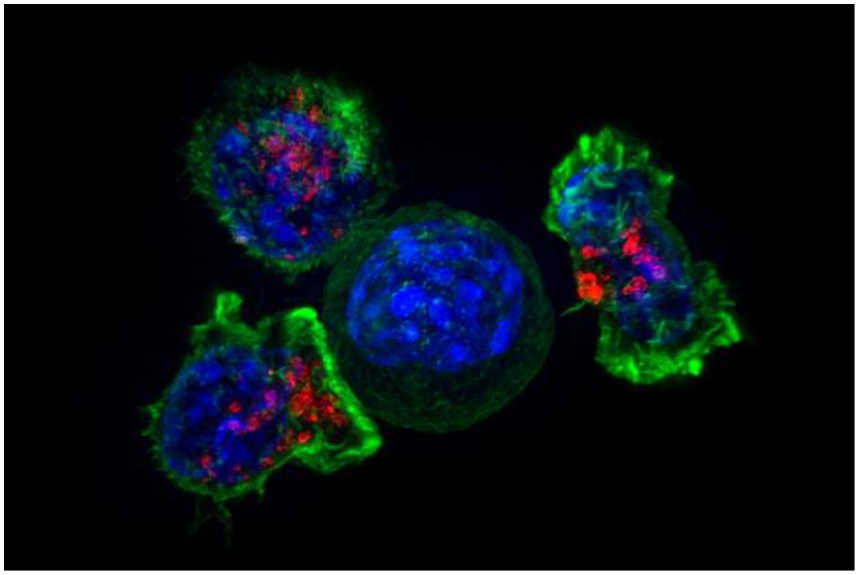
Share this article:

LA JOLLA—Fighting a tumor is a marathon, not a sprint. For cancer-fighting T cells, the race is sometimes just too long, and the T cells quit fighting. Researchers even have a name for this phenomenon: T cell exhaustion.

In a new *Nature Immunology* study, researchers at La Jolla Institute for Immunology (LJI) report that T cells can be engineered to clear tumors without succumbing to T cell exhaustion.

"The idea is to give the cells a little bit of armor against the exhaustion program," says LJI Professor Patrick Hogan, Ph.D. "The cells can go into the tumor to do their job, and then they can stick around as memory cells."

This research builds on a decades-long collaboration between Hogan and LJI Professor Anjana Rao, Ph.D. Their work has shown the key role of proteins called transcription factors in the cellular pathway that triggers T cell exhaustion.



Prolonged overactivity of killer T cells summoned to a tumor site can render them useless, a cellular state immunologists call "exhaustion." Image: Superresolution image of a group of killer T cells (green and red) surrounding a cancer cell (blue). Courtesy of National Institutes of Health.

This work is important because T cell exhaustion continues to plague even the most cutting-edge cancer immunotherapies.

With CAR T therapies, for example, researchers take T cells from a cancer patient and "arm" them by altering the expression of genes that aid in the cancer fight. Researchers make more of these special T cells, which then go back into the patient. CAR T therapies are different from immunotherapies, which aim to activate the patient's existing T cell population.

With both approaches, T cell exhaustion rears its ugly head. "Many people have tried to use CAR T therapies to kill solid tumors, but it's been impossible because the T cells become exhausted," says study co-first author Hyungseok Seo, Ph.D., a former postdoctoral fellow in the Rao Lab who is currently working at Novartis.

The new study addresses this problem by giving T cells the ability to fight exhaustion itself.

To accomplish this, the researchers screened T cells to uncover which transcription factors could boost a T cell's "effector" program, an important step in readying T cells to kill cancer cells.

This screening process led the researchers to BATF, a transcription factor that they found cooperates with another transcription factor called IRF4 to counter the T cell exhaustion program.

In mouse melanoma and colorectal carcinoma tumor models, altering CAR T cells to also overexpress BATF led to tumor clearance without prompting T cell exhaustion. The CAR T therapy worked against solid tumors.

"BATF and IRF4 are cooperating to make T cells better," says Seo.

Further testing showed that while IRF4 is important, it shouldn't be overexpressed to the same degree as BATF. For maximum effect, BATF was overexpressed around 20 times more than in normal cells.

Encouragingly, some altered T cells also stuck around and became memory T cells. This is important because T cell exhaustion often prevents T cells from mounting a strong memory response to recurrent cancers.

“We didn’t just increase the ability of T cells to fight exhaustion—we increased the ability of cells to fight tumors,” says study co-first author Edahí González-Avalos, a graduate student in the Rao Lab who led the bioinformatic analysis for the project.

Hogan thinks overexpressing BATF could be a promising approach for improving CAR T therapies and for tackling some hard-to-treat cancer types, such as pancreatic ductal carcinoma. These types of cancers are known as “immunologically cold” because they don’t spark a strong anti-cancer response from the immune system. T cells don’t fight them in force.

Other laboratories have been exploring ways to make these cold tumors “hot,” so they will attract T cells. The LJI team thinks a promising strategy would combine those approaches with targeting transcription factors to render T cells exhaustion-proof.

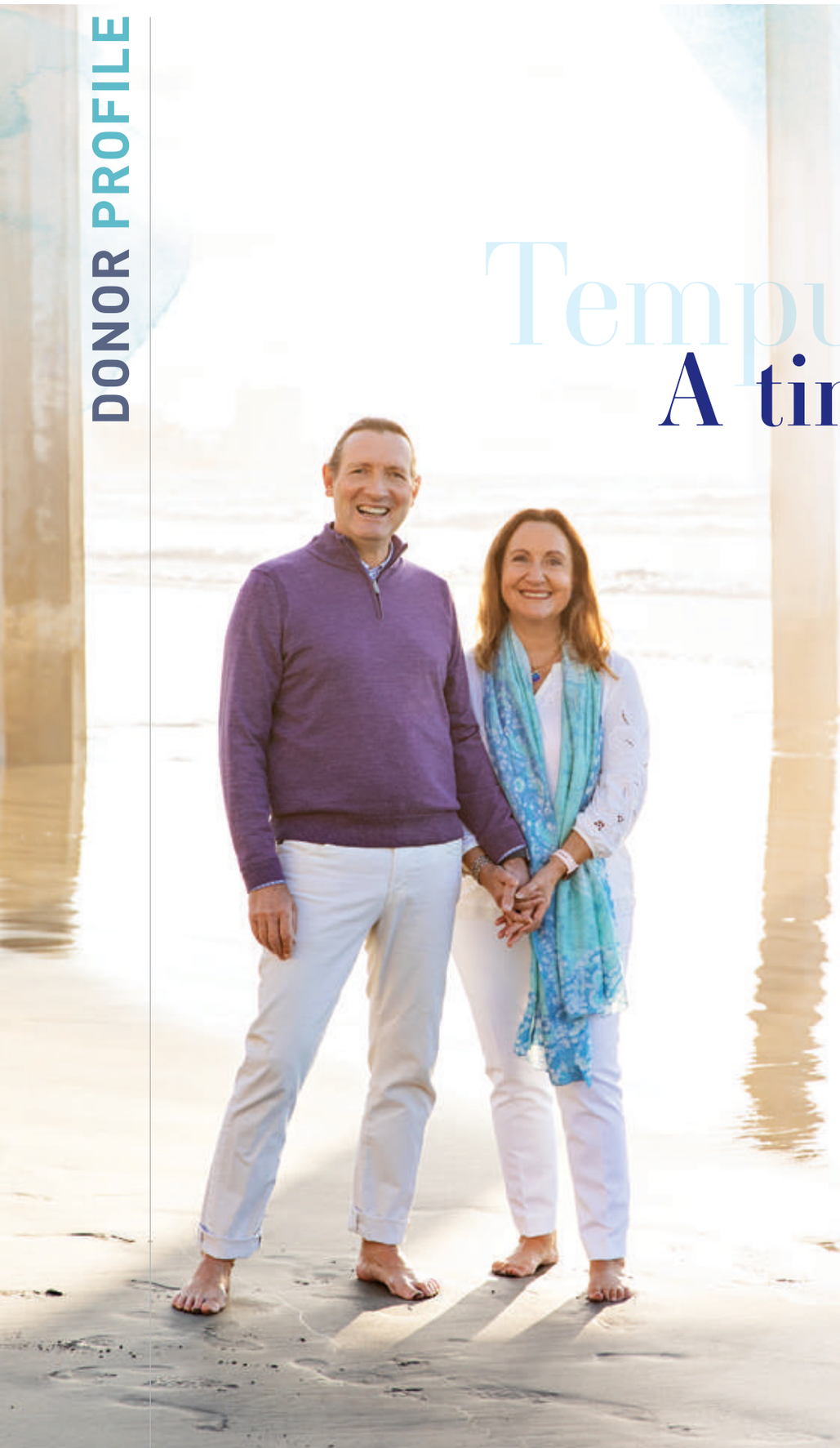
“We wouldn’t necessarily need a transgenic approach to do this,” says Hogan. “Maybe even an oral drug molecule could do it, if you knew what transcriptional pathways you wanted to go after.”

The researchers emphasize that BATF is just one of many transcription factors that may prove important to manipulate when countering T cell exhaustion.

“We’re going to keep looking for answers,” adds González-Avalos.

The study, “BATF and IRF4 cooperate to counter exhaustion in tumor-infiltrating CAR T cells,” was supported by the National Institutes of Health (grants AI109842, AI040127, DE028227, RR027366, OD018499); a AACR-Genentech Immuno-oncology Research Fellowship 18-40-18-SEO; a Donald J. Gogel Cancer Research Institute Irvington Fellowship; a University of California Institute for Mexico and the United States (UC-MEXUS) and El Consejo Nacional de Ciencia y Tecnología (UCMEXUS/CONACYT) predoctoral fellowship; NIH T32 predoctoral training grant in the UCSD Cardiovascular Bioengineering Training Program; Cancer Research Institute Irvington Fellowship; an Independent Investigator Fund (La Jolla Institute/Kyowa Kirin) and a Career Transition Award from the National Cancer Institute (K22CA241290).

Additional authors of the study include Wade Zhang, Payal Ramchandani, Chao Yang and Chan-Wang J Lio. DOI: 10.1038/s41590-021-00964-8



Tempus dare

A time to give

"How can I help?" That was the life-saving question this year, as people around the world stepped up to volunteer for clinical trials and staff vaccination sites.

"How can I help?" That was the question supporters of La Jolla Institute for Immunology (LJI) asked as the Institute took on COVID-19 research.

"The extraordinary progress our researchers made in addressing the global pandemic was a direct result of private philanthropy," says Christopher Lee, LJI's Chief Advancement Officer. "Critical early funding from individuals and foundations, both locally and nationally, allowed for immediate work in studying the virus at LJI—as federal funding was being organized and eventually distributed."

"Without our donors, our highly impactful work—now cited and used around the globe—would have been limited if not impossible," says Lee.

San Diego biotech entrepreneurs François Ferré, Ph.D., and Magda Marquet, Ph.D., knew supporting LJI would be critical for advancing

> "Philanthropy can help researchers jump into more adventurous science. That's really the core of why we need strong philanthropy at LJI at all times."
> – François Ferré, Ph.D.

our understanding of COVID-19. Scientists at LJI dedicated themselves early on to long days and nights studying how immune cells and antibodies could fight SARS-CoV-2.

"We really connected with the fact that LJI was in the weeds, trying to make new COVID-19 research happen," says Dr. Ferré, who also serves on LJI's Board of Directors.

Drs. Ferré and Marquet knew scientists would be hard-pressed to find funding for such rapid shifts in research. After listening to a presentation from LJI Professor Erica Ollmann Saphire, Ph.D., the couple made a gift to LJI that allowed the Saphire Lab to purchase two incubators essential for COVID-19 antibody therapeutic research.

"François and Magda are real difference makers here at LJI," Lee says. "We know that when urgent action is required, LJI can always count on these two remarkable philanthropists. We are deeply honored to be associated with them."

As LJI continues to shed light on COVID-19, many are now wondering how to build on the public's new grasp of medical research. LJI has been a leader in immune system research for more than 30 years, and now the public is in a position to understand how insights into immune cells can open the door to better vaccines, cancer therapies, autoimmune disease treatments, and more.

"People have realized what LJI is doing—in a much more practical way," says Dr. Marquet.

"We're at a point in medicine right now where it is so important to truly understand the functioning of the immune system," Dr. Ferré adds. "I cannot think of a better cause right now."

Drs. Ferré and Marquet took an interest in LJI long before the pandemic. Dr. Ferré joined the LJI board in 2017, and the two have funded early career research projects through the Tullie and Rickey Families SPARK Awards for Innovations in Immunology.

The couple appreciates that LJI is open about how gifts are used. Following their donation in 2020, they received a personal letter from Dr. Saphire. "She explained in detail where the money was going, describing why this incubator was so important for the research she was doing," says Dr. Ferré.

"LJI recognizes the importance of donors, and that goes a long way," says Dr. Marquet.

They also recognize the importance of efficiency at LJI. "Being an entrepreneur myself, I really like the Institute's approach," says Dr. Ferré. "It's not encumbered by layers of bureaucracy."

Dr. Ferré says he's looking forward to where the Institute goes from here, especially in the area of autoimmune disease. "One of the big crises of our immune system is inflammation going berserk," he says, adding that with donor support, researchers at LJI will be in a good position to quickly follow up on promising findings—for any disease.

"Philanthropy can help researchers jump into more adventurous science," Dr. Ferré says. "That's really the core of why we need strong philanthropy at LJI at all times." •

Alba Grifoni, Ph.D.

Speaking Up for Science

Most scientists don't expect to end up in the news. Alba Grifoni, Ph.D., certainly didn't. As a postdoctoral researcher (later promoted to instructor) at La Jolla Institute for Immunology (LJI), Dr. Grifoni had always worked on diseases that were global killers—but not headline news.

Then came COVID-19. In the early months of 2020, Dr. Grifoni worked closely with LJI Professor Alessandro Sette, Dr. Bio. Sci., to publish the first clues to how T cells target SARS-CoV-2. Their work inspired scientists around the world to consider how T cells, and not just antibodies, could drive immunity to the virus.

For many, the work gave hope that a COVID-19 vaccine would be possible. *The New York Times*, *Washington Post*, and many others began publishing updates from the research team.

"I just kept blinking my eyes. Am I reading this? Am I dreaming this? It was a shock," says Dr. Grifoni. "They were acknowledging that our field of study was important in the fight against COVID."

Dr. Grifoni comes from a family of medical doctors. As a girl growing up in Italy, she assumed she would go into medicine too. Yet as she took college courses, she started to wonder if a career as a doctor, where she'd see patients one by one, was the right fit. She started to consider studying immunology instead.

"What if I could do research that doesn't just help one person in the moment but could help way more people—even after I'm not around anymore?" she says. "That was the path I wanted to follow."

Dr. Grifoni had also grown up with several role models in the sciences. One was Rita Levi-Montalcini, M.D., the Italian neurobiologist who received a Nobel Prize in 1986 for the co-discovery of important nerve growth factors. "She showed me that being a woman should not hinder my aspirations," says Dr. Grifoni.

The others were Dr. Sette and LJI Professor Bjoern Peters, Ph.D. The pair leads cutting-edge research in epitope identification. Put simply,

they pinpoint the parts of a pathogen that provoke an immune response. “Everyone in immunology knows about Sette and Peters,” she says.

Dr. Grifoni earned her Ph.D. in immunology in 2014. Then she saw a job listing for a postdoctoral position in the Sette and Peters labs. “I thought, ‘You know what—I’m just going to apply for the postdoc. I’m going to give it a shot.’”

At LJI, Dr. Grifoni spearheads research that combines immunology and genetics. She uses bioinformatics tools to analyze viral genomes. This research can reveal which viral epitopes might be targeted by the immune system. Dr. Grifoni can then work with clinical samples to see exactly which potential targets are most important for fighting off the virus.

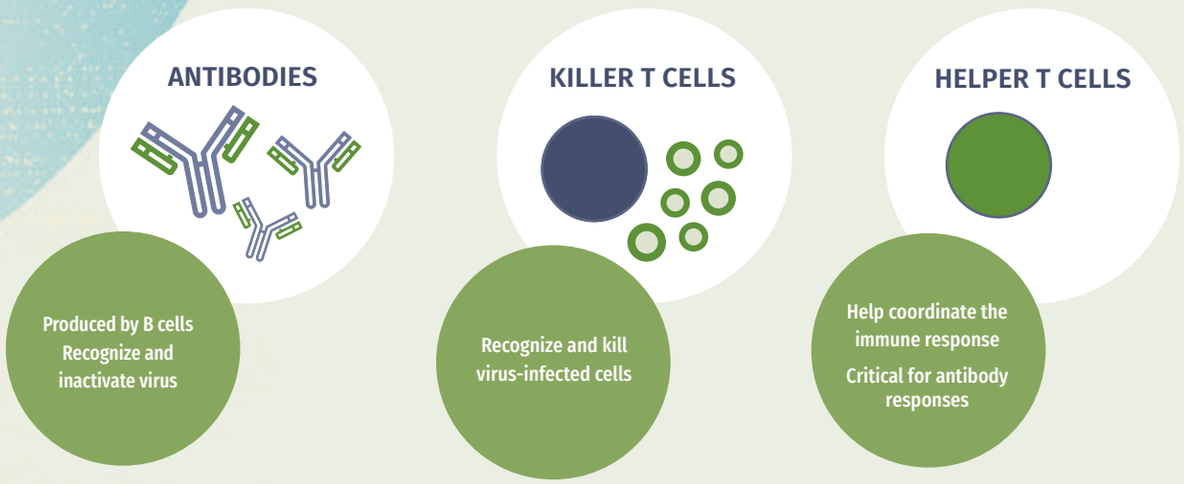
Her work has shed light on HIV, hepatitis B and C, Ebola virus, Zika, dengue, yellow fever, and many other pathogens. “I’ve lost count,” says Dr. Grifoni. “But the focus is always the same. I want to know how human T cells respond to viral infections.”

The hardest part of the job? “That I need to wait to see the data!” she says with a laugh.

“Some steps of my research require patience,” she adds. “You need to get the technical steps figured out. But once that’s set up, I want to go fast because I want to see the results.”

“What if I could do research that doesn’t just help one person in the moment but could help way more people—even after I’m not around anymore?”

WHAT KIND OF IMMUNITY IS IMPORTANT AGAINST COVID-19?



Dr. Grifoni's persistence in figuring out key SARS-CoV-2 T cell epitopes was recognized last fall when she received the Embassy of Italy Award to Honor Young Italian Researchers for Research to Fight COVID-19 from the Italian Scientists and Scholars in North America Foundation (ISSNAF).

As an ISSNAF winner, Dr. Grifoni had the chance to present her work virtually to leading Italian scientists and speak with the Italian media.

"It meant a lot to have my science recognized in the U.S. and Italy," says Dr. Grifoni. "It gave me the opportunity to share my research not only with other scientists in the field but with the general Italian population, in my own language, and deliver important information for them to understand the impact of SARS-CoV-2 and the importance of vaccination."

Dr. Grifoni is the kind of researcher who can easily explain her work to a non-scientific crowd, and that task has gotten even easier during the pandemic. "Now, when you talk to the average person about T cells, they know what those are," says Dr. Grifoni.

Going forward, Dr. Grifoni wants to help people understand the same cutting-edge research techniques and sense of urgency we've had with COVID-19 are needed to fight other global threats, such as dengue and chikungunya. She says researchers need to better understand how T cells recognize these viruses. They need to know how to balance T cell responses with antibody responses, and how to spark the right kind of immune responses with a vaccine.

"I find the need to simplify the science when speaking with a wide audience, but sometimes it would be good for people to really understand the complexity of the research," says Dr. Grifoni.

While much more work is to be done, Dr. Grifoni has already achieved what she'd hoped for. Her research has affected many other people and inspired new lines of research.

"I've had a great passion to guide me," she says. Thinking back to her university days, she remembers feeling discouraged, that she wouldn't make big contributions to science. Dr. Grifoni has advice for young scientists today.

"To them I'd say, 'Be brave. Throw yourself into the sky. No matter where you end up, you will definitely enjoy the journey—and if your work can also help others, that will be the best reward you can aim for.'" •

Melanoma meets its match

Madeline McCurry-Schmidt
Science Writer
April 7, 2021

Share this article:

How Ian Mathews used Tullie and Rickey Families SPARK Awards funding to make cancer therapy advances and launch a career in medicine

Ian Mathews, Ph.D., may be the youngest La Jolla Institute for Immunology (LJI) scientist to ever win funding through The Tullie and Rickey Families SPARK Awards for Innovations in Immunology. He applied for seed funding through the program as a graduate student, back when his doctorate thesis was still in the fledgling stage.

Ian Mathews, Ph.D.

Three years later, Mathews has a Ph.D. from the University of California, San Diego, under his belt and he's enrolled in med school. Mathews' work has opened the door to major cancer research—and new funding.

"These data were foundational to my career in a few ways, both concrete ways and more philosophical," says Mathews.

The $25,000 grant from the donor-driven Tullie and Rickey Families SPARK Awards program led Mathews to receive a highly competitive NIH fellowship in 2019. More recently, the National Cancer Institute also awarded $4.2 million in additional research funding for LJI-led follow-up studies, based on the data Mathews produced for his SPARK project.

In fact, his mentor, LJI Associate Professor Sonia Sharma, Ph.D., has done the math. She found that the SPARK award led to about a 165-fold return on investment.

"With young scientists and new scientists, the focus needs to be on potential, potential, potential," adds Sharma. "They need to have their chance to perform, to grow, to succeed."

How it started

Except for Mathews, Tullie and Rickey Families SPARK Awards recipients have all been postdoctoral researchers who have already earned a Ph.D. and are in the midst of launching a research career. Mathews says he didn't expect to be even considered for funding.

"Personally, I was hoping just to have a chance to practice my writing, but was fortunate to be considered and eventually supported," says Mathews.

Indeed, LJI Board Member Gail K. Naughton, the donor behind his award, saw something promising in his cancer research proposal.

With funding from the Tullie and Rickey Families SPARK program, Mathews hoped to develop a way to predict how patients with melanoma would respond to cancer immunotherapy. Harnessing the immune system to fight cancers has proven tough. Cancer immunotherapy can save lives, but some patients see no change and can suffer from terrible side effects. Mathews planned to focus on how molecules called metabolites could help doctors spot the patients who could benefit.

"Ian thinks big," says Sharma. "He likes to focus on the larger question and then narrow in on the details. With his SPARK project, he knew he wanted to take a highly translational approach towards pinpointing the somewhat mysterious metabolites that control cancer immunotherapy response."

Launching this study meant Mathews had to build up new skills, fast.

"He knew from the get-go that he wanted to use a humans-first approach, which is very much in line with our lab's overall vision. So, it was a matter of gathering the right clinical collaborators, who could provide us with samples and outcomes data for cancer immunotherapy patients, all the while Ian was learning fundamental mass spectrometry-based metabolomics techniques to be able to study the cohorts."

Once all that was done, Mathews had to determine how to validate his initial discovery in humans by using *in vivo* models.

When SPARK serves as a stepping stone

Within a year of receiving the Tullie and Rickey Families SPARK Award, Mathews and his colleagues in the Sharma Lab found that some melanoma patients indeed had high levels of metabolites in their blood that could serve as a good predictor for whether they would respond to therapy.

The data from Mathews' project opened the door to his NIH fellowship. The new funding from the National Cancer Institute (an R01 award to Sharma) will allow scientists to build on Mathews' findings and foster scientific collaboration between UC San Diego, Cedars-Sinai and the Salk Institute.

"With the R01 award we can focus on developing these immunotherapy-associated metabolites as novel cancer therapeutics," Sharma says.

Last summer, Mathews defended his thesis (pandemic-style via Zoom). As he readies his study manuscript for publication, Mathews is also tackling his life-long goal: med school.

"Direct patient care is my north star in research, it gives my life's endeavors meaning and guides what scientific inquiries I find interesting and worthwhile," says Mathews.

"Actually, it occurs to me that SPARK can have a similar effect," he says. "Innovative, narrow, translationally-driven proposals which incentivize broadly disseminable science communication skills: SPARK can teach you a lot about how to be an investigator with clinical outcomes on the mind."

The Conrad Prebys Foundation grants more than $450,000 for studies into childhood cancer, cardiovascular disease

New funding will support high-risk, high-reward projects led by two early-career scientists

Madeline McCurry-Schmidt
Science Writer
March 23, 2021

Share this article:

LA JOLLA—The Conrad Prebys Foundation has awarded more than $450,000 in funding to support two early-career researchers at La Jolla Institute for Immunology (LJI). LJI Instructor Abhijit Chakraborty, Ph.D., will use the funding to investigate how genomic instability can lead to childhood cancer. LJI Postdoctoral Fellow Marco Orecchioni, Ph.D., will lead a project to uncover how human immune cells may trigger atherosclerosis, the cause of heart attacks and strokes.

"We are honored at La Jolla Institute to be in the first round of recipients of the generosity of The Conrad Prebys Foundation," says Mitchell Kronenberg, Ph.D., President and Chief Scientific Officer of LJI. "Conrad Prebys was a visionary in so many ways, and his deep understanding of the power of fundamental research to address a wide range of health conditions affecting humanity has been incredibly impactful. With these two projects funded by the foundation, LJI hopes to build on Conrad's legacy of propelling great scientific ideas forward and as a result, helping many people suffering from disease."

Abhijit Chakraborty, Ph.D., (left) and Marco Orecchioni, Ph.D., (right).

The Conrad Prebys Foundation works to improve quality of life—primarily in San Diego County—through the distribution of grants to support medical research, healthcare, higher education, and other charitable causes. These new grants carry on founder Conrad Prebys' legacy of supporting truly innovative research in the health sciences.

"The past year has highlighted the imperative behind-the-scenes work that researchers are conducting to ensure we have the insights, treatments and medicine we need to keep communities healthy," said Erin Decker, director of grantmaking at The Conrad Prebys Foundation. "The foundation is honored to have a role in ensuring these meaningful research projects have the resources needed to fuel future solutions in medical care."

For LJI, another exciting aspect of these grants from The Conrad Prebys Foundation is that they are enabling expanded research based on data that Orecchioni and Chakraborty were able to collect thanks to seed-investment grants they won from LJI's Tullie and Rickey Families SPARK Awards for Innovations in Immunology in 2019 and 2020 respectively. The mission of the Tullie and Rickey Families SPARK Awards program, is to provide $25,000 in private funding for LJI's young investigators so they can explore novel ideas and hopefully generate promising data that will lead to larger awards that could advance the research and help further their careers.

"Not only are we grateful as an Institute to receive support from the The Conrad Prebys Foundation," shares Dr. Kronenberg, "I'm also particularly encouraged that these grants represent the realization of the vision of the Tullie and Rickey Families SPARK Awards program, which is to launch bold new ideas of our young researchers. Our donors to the Tullie and Rickey SPARK Awards program can feel proud to see these examples of their support being leveraged ten-fold for both exciting projects thanks to these new awards from The Conrad Prebys Foundation."

Chakraborty, a member of the Ferhat Ay Lab at LJI, is focused on uncovering how cancer-causing genomic rearrangements make their way into the genome and alter cellular functions. In some cases a person's genetic material can break apart—become "completely shattered," he says—and reform in a random fashion, which leads to dangerous rearrangements (a phenomenon called chromothripsis) in the already unstable cancer genome. With the help from his SPARK funding, he has developed methods to identify these significant changes in the human genome.

He now intends to study how such catastrophic rearrangements alter the three-dimensional physical structures of chromosomes and thus affects cellular functions in cancer patients. He will use methods he's developed to study the effects of iAMP21, a chromosomal rearrangement that leads to a rare but lethal subtype of childhood cancers.

"This is really one of the worst of all the cancers, and we are in the dark," says Chakraborty.

For the new project, Chakraborty will analyze clinical samples from young cancer patients with iAMP21 conditions and compare those to control samples. This work will shed light not only on where cancer-causing structural variations lurk in the genome but how those rearrangements alter the gene expression pattern leading to the oncogenesis.

Orecchioni, a member of the Klaus Ley Lab at LJI, will pursue new research into how immune cells may drive the development of atherosclerosis, the condition that leads to blocked arteries and causes heart attacks and strokes. He has found that immune cells called macrophages can actually express an olfactory receptor—the same type of receptor that in the nasal passages allows us to smell odors. He has shown that this receptor in mice, called Olfr2, can sense a marker correlated with increased cholesterol in the body and trigger the inflammation seen in atherosclerosis. Blocking the receptor in mice helped stop the disease progression.

"Mice help us understand how things work, but now we need to see if this is the same in humans," Orecchioni says.

In addition to studying macrophages in human samples, Orecchioni will investigate whether other types of immune cells also express olfactory receptors. He is especially interested in the possibility that immune cells called monocytes could express olfactory receptors and contribute to inflammation. By seeing which cells patrol the body to sense changes in diet, Orecchioni will be a step closer to understanding how new therapies could address the root causes of atherosclerosis.

The new funding from The Conrad Prebys Foundation puts both early-career scientists in a position to make major advances in their fields.

Chakraborty says he can't wait to get started, especially after a year working outside the lab due to social distancing. "I'm very fortunate," he says.

Orecchioni says LJI is the perfect place to take on these kinds of high-risk, high-reward projects. "We have fantastic scientists and cutting-edge facilities," he says. "This environment actually pushes you to try new things."

Researchers discover new particle in the blood of septic patients

LJI scientists get first glimpse of how mysterious particles break off of immune cells

Madeline McCurry-Schmidt
Science Writer
December 4, 2020

Share this article:

LA JOLLA—Researchers at La Jolla Institute for Immunology (LJI) have found that people with sepsis have never-before-seen particles in their blood. The scientists are the first to show that these particles, called elongated neutrophil-derived structures (ENDS), break off of immune cells and change their shape as they course through the body.

“We actually found a new particle in the human body that had never been described before,” explains LJI Instructor Alex Marki, M.D., who served as first author of the study. “That’s not something that happens every day.”

The research, published December 4, 2020 in the *Journal of Experimental Medicine* shows the importance of understanding how immune cells change over the course of a disease.

“ENDS are not normal—they are not detectable in healthy people or mice,” says LJI Professor Klaus Ley, M.D., who served as senior author of the study. “But ENDS are very high in sepsis, and I would not be surprised if they were high in other inflammatory diseases.”

The beginning of the ENDS

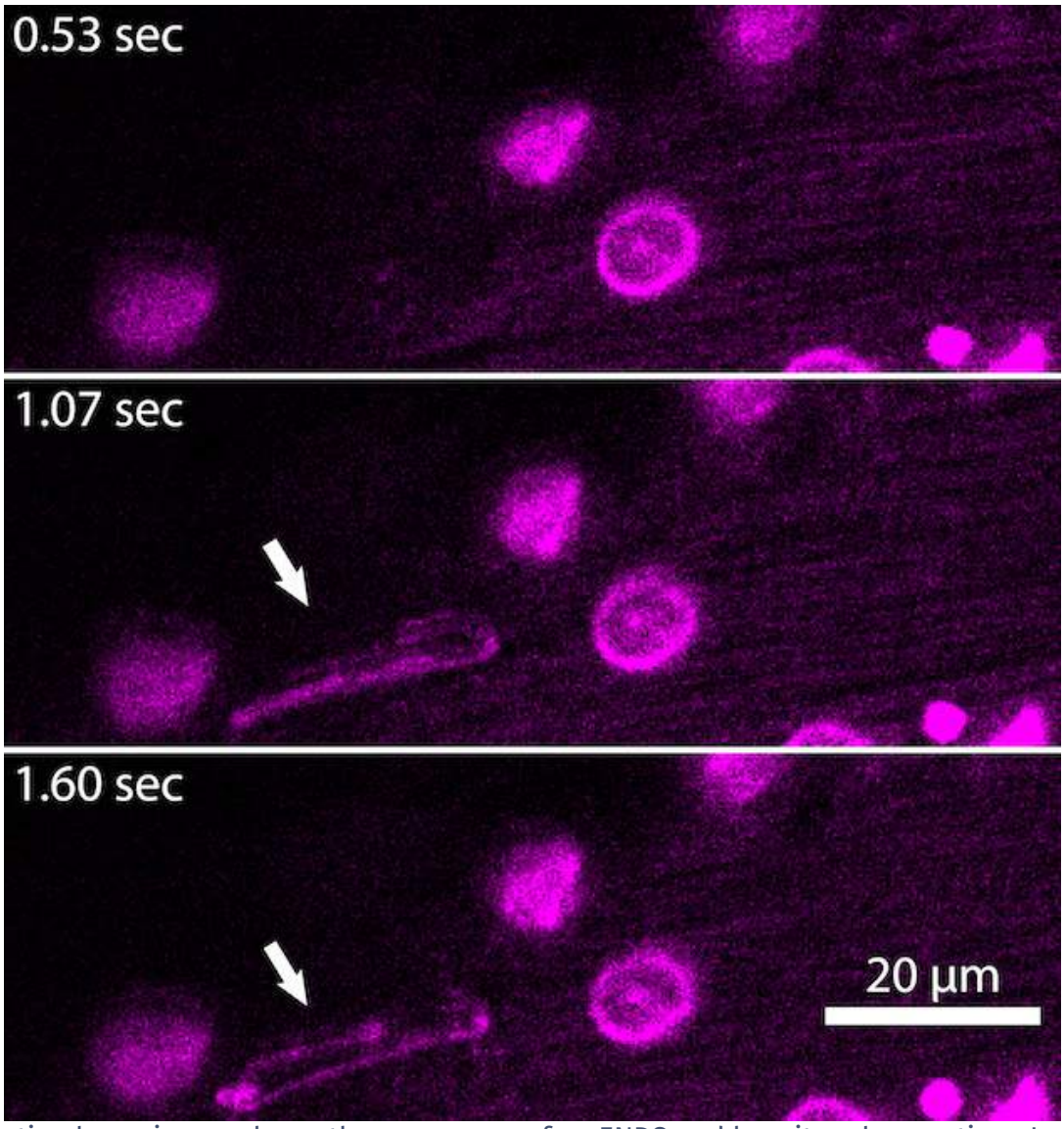
The discovery of ENDS started with an odd observation.

Marki was studying neutrophils, a kind of immune cell that moves through the bloodstream and slips into tissues to fight infections. At the time, he was studying living mice to confirm the presence of tubes called tethers. These tethers are attached to neutrophils as they roll on the blood vessel wall.

During these experiments Marki noticed long, thin objects of neutrophil origin sticking to the vessel wall. Since no such structure was described in the scientific literature, the team had to come up with a name for them. The initial lab slang name “sausages” was eventually replaced by the elongated neutrophil-derived structures or ENDS.

Desperate for learning more about these new objects, the LJI team developed a series of new techniques to study how ENDS form and degrade—and to detect them in human and mouse blood plasma.

Thanks to sophisticated imaging techniques, the LJI team figured out that tethers *become* ENDS. As the neutrophils flop and roll along, their tethers get longer and longer. Eventually the tethers become too thin—just 150 nanometers (around 1/500th the width of a human hair). Then they break in the middle. Part of the tether stays with the neutrophil, but the broken fragment flies away in the bloodstream, off to form an ENDS.



This timelapse image shows the appearanc of an ENDS and how it curls over time. Image courtesy of Alex Marki, M.D., Ley Laboratory, La Jolla Institute for Immunology.

The researchers showed that these ENDS curl against the vessel wall until they get a rounded shape. It's likely that the ENDS stay intact for a while, but not for long. Without any life-sustaining organelles inside, the ENDS begin to die. In fact, the researchers found that the ENDS secrete tell-tale signaling molecules that promote inflammation.

Compared with healthy subjects, the researchers showed that ENDS are around 100-fold more detectable in septic patients.

What this means for sepsis

Sepsis can occur when the immune system overreacts to an infection by flooding the body with dangerous chemicals. Instead of just fighting the infection, these chemicals trigger organ damage as they course through the bloodstream. The mortality rate for septic "shock" is 30 percent.

"Once you're in the hospital, sepsis is the most common cause of death," Ley says.

Ley and Marki are still not sure why ENDS form in patients with sepsis. To learn more, Marki hopes to collect more patient samples to track ENDS formation and frequency over time. "I'd like to study blood from several time points from each patient—to see the dynamics of how ENDS change," he says.

Ley says it is theoretically possible that ENDS could one day serve as a biomarker for early sepsis detection, but it is currently impossible to detect them in a clinical setting. "Right now, the assay is not practical because it takes specialized instrumentation," says Ley.

Rather than serving as a diagnostic, Ley thinks studying ENDS could reveal secrets to how the immune system evolved. He's curious to learn how the process to form ENDS evolved—and why.

"Neutrophils are very soft cells that can deform to reach almost any place in the body," says Ley. "So one hypothesis I have is that ENDS might be the price you pay for having such a soft cell—that if you pull too hard, it falls apart."

This study, titled “Elongated neutrophil-derived structures are blood-borne microparticles formed by rolling neutrophils during sepsis,” was supported by The National Institutes of Health (grants P01 HL078784, P01 HL151433, R01HL145454, S10OD021831), the American Heart Association (17POST33410940, 18POST34060251, 16POST31160014, 18CDA34110426) and a DFG award (GZ WI 4811/1-1).

Marki is a recipient of the Tullie and Rickey Families SPARK Award at La Jolla Institute for Immunology.

Additional study authors included Konrad Buscher, Cristina Lorenzini, Matthew Meyer, Ryosuke Saigusa, Zhichao Fan, Yi-Ting Yeh, Nadine Hartmann, Jennifer M. Dan, William Bill Kiosses, Gregory J. Golden, Rajee Ganesan, Holger Winkels, Marco Orecchioni, Sara McArdle, Zbigniew Mikulski, Yoav Altman, Jack Bui, Mitchell Kronenberg, Shu Chien, Jeffrey D. Esko, Victor Nizet, David Smalley and Johannes Roth.

DOI: 10.1084/jem.20200551

Fighting Zika? Call in the T cells

Stopping emerging viruses could depend on getting T cells and antibodies to work together

Madeline McCurry-Schmidt
Science Writer
November 4, 2020

Share this article:

LA JOLLA—Where *Aedes* mosquitoes fly, Zika virus may not be far behind. Although the explosive 2015–2016 Zika epidemics in the Americas are behind us, Zika may re-emerge, and "in many countries, Zika may be spreading in silence," says Sujan Shresta, Ph.D., a professor at La Jolla Institute for Immunology (LJI). "We need to develop effective vaccines."

In a new *Science Advances* study, Shresta and her colleagues at LJI report that the immune system's T cells have the power to prevent Zika infection in mice. This finding suggests that effective Zika vaccines need to activate T cells to work alongside antibodies.

"If we combine T cells and antibodies, we have even stronger protection and longer-term protection," says Annie Elong Ngono, Ph.D., a postdoctoral fellow at LJI and first author of the new study.

Zika virus cases are usually mild, but the virus can cause serious congenital malformations in infants and neurological complications in adults and children. Since Zika made headlines in 2016, when cases of the virus peaked in the Americas, researchers have developed more than 40 Zika vaccine candidates. The vast majority of these vaccines are designed to prompt the body to make antibodies that target one specific protein on the virus.

Left: Aedes mosquito. Right: Zika virus particles (red) Credit: NIAID

Unfortunately, there is a drawback to this neutralizing antibody approach. In many parts of the world, Zika virus spreads alongside related mosquito-borne viruses, such as dengue. Scientists have found that the presence of anti-Zika antibodies can make a subsequent case of dengue much, much worse. In a 2018 study, Shresta's lab showed that newborn mouse pups harboring anti-Zika antibodies were more vulnerable to death from dengue exposure than mice that lacked anti-Zika antibodies.

Theoretically, similar cases of "antibody-dependent enhancement" may lead to cases where lingering anti-Zika antibodies in a patient will actually make it *easier* for dengue to enter host cells—leading to especially devastating consequences in pregnant patients. This means that a Zika vaccine that prompts only antibody production may be risky in areas where both dengue and Zika are common. Luckily, the immune system can make more than antibodies.

For the new study, Shresta and Elong Ngono tested an experimental Zika vaccine in a mouse model. The vaccine was designed to elicit the arm of the immune system that makes T cells. The mice were given the vaccine, given a second vaccine boost four weeks later, and then exposed to Zika six weeks after that.

The team found that the vaccine could induce a strong immunity against a potentially lethal Zika virus infection by inducing mainly CD8+ T cells, also called "killer" T cells, against the virus. The vaccine also prevented Zika transmission through the placenta from mother to fetus in pregnant mice.

This vaccine approach was even more effective when combined with a vaccine candidate that induced neutralizing antibodies. "We found that it is better to have a vaccine that induces both T cells and antibodies than either one alone," says Elong Ngono.

The new research also shows the importance of targeting more than one viral protein when fighting flaviviruses, the group of viruses that include Zika, dengue, yellow fever and Japanese encephalitis. By getting T cells and antibodies to recognize key sites on these related viruses, researchers may be closer to developing a "pan-flavivirus" vaccine to protect people in areas where several of these diseases are common.

“We think this approach can be used against other infectious diseases,” Elong Ngono says. For example, recent research from LJI scientists suggests that COVID-19 vaccines may also need to elicit T cells to work alongside antibodies.

“Now the challenge is finding how best to elicit appropriately balanced antibody and T cell responses,” says Shresta. “We also don’t know how durable the vaccine protection is—if it’s fairly short, we want to figure out how to enhance it.”

The study, titled “CD8+ T cells mediate protection against Zika virus induced by an NS3-based vaccine,” was supported by the National Institutes of Health (grants AI116813, AI140063 and NS106387) and a grant from the Chiba-UCSD Center for Mucosal Immunology, Allergy, and Vaccine Development.

Additional study authors include Thasneem Syed, Anh-Viet Nguyen, Jose Angel Regla-Nava, Mercylia Susantono, Darina Spasova, Allison Aguilar, Melissa West, Jessica Sparks, Andrew Gonzalez, Emilie Branche, Jason L. DeHart, Jerel Boyd Vega, Priya Prakash Karmali, Padmanabh Chivukula, Kurt Kamrud, Parinaz Aliahmad and Nathaniel Wang.

Doi: http://doi.org/10.1126/sciadv.eabb2154

Why I Give: Barbara Donnell on the importance of endurance—and hope

Madeline McCurry-Schmidt
Science Writer
October 14, 2020

For Barbara Donnell, La Jolla Institute for Immunology (LJI) has always been a place for inspiration and collaboration. The San Diego philanthropist and fundraiser has enjoyed attending events, learning about research, and getting acquainted with staff and scientists over the years.

Together with the Juvenile Diabetes Research Institute (JDRF), Donnell has organized the annual “Meet the Scientists” event at LJI, at which kids diagnosed with diabetes and their families can meet diabetes researchers and learn about the latest science. “It’s an all-day educational event,” says Donnell. “It’s very interactive and uplifting for the kids and families as well as the scientists.”

The COVID-19 pandemic changed everything. “Meet the Scientists” was cancelled. The research funding world was reeling. With her experience in philanthropy, Donnell knew how hard it would be for researchers to get new projects off the ground.

Donnell has been a steady supporter of LJI labs, and she had given additional funding this year to LJI's SPARK Program for early career investigators. When she heard that LJI would be running the international Coronavirus Immunotherapy Consortium, she decided to make an additional gift to the institute.

"Every bit counts," says Donnell. "I just felt compelled to break my budget and give an additional gift of stock."

As the pandemic continues, Donnell is encouraged by LJI's "Live from the Laboratory" webinar series, which highlights LJI teams and their progress toward COVID-19 therapies and vaccines.

Donnell also understands the importance of every penny raised for scientific research. She began working with JDRF in 2001, when her daughter was diagnosed with type 1 diabetes. One of Donnell's roles was to act as a bridge between the foundation and San Diego's research institutes. That work brought her to LJI.

"I just fell in love with the place and the scientists," Donnell says. "La Jolla Institute shines for me."

Donnell is confident researchers will develop new therapies and vaccines for COVID-19. It will just take time. "I've been on a quest for a type 1 diabetes cure for years, so I understand that you have to have so much endurance and patience."

"I know how dedicated, how hardworking, how passionate these scientists are, and I know they are pushing the envelope. They are going way beyond," says Donnell. "La Jolla Institute researchers are in this, heart, and soul. And because I know that, I have hope."

Blocking nerve signals to the pancreas halts type 1 diabetes onset in mice

New LJI discovery could help explain patterns of cell death in many autoimmune diseases

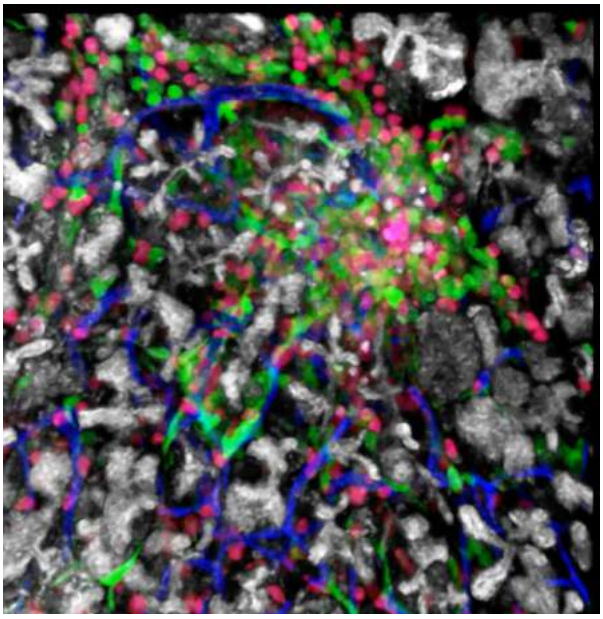
August 25, 2020

Share this article:

LA JOLLA—Your pancreas is a little sweet potato-shaped organ that sits snug behind your stomach. The pancreas is studded with islets, the cell clusters that house insulin-producing beta cells. In people with type 1 diabetes, the body's own immune cells head for the islets and start attacking the beta cells. No one knows exactly what triggers this attack.

One clue may lie in the pattern of beta cell death. Many beta cells are killed off in big patches while other beta cells are mysteriously untouched. Something seems to be drawing immune cells to attack specific groups of beta cells while ignoring others.

In a new *Science Advances* study, researchers at La Jolla Institute for Immunology (LJI) report that the nervous system may be driving this patchy cell die-off. Their new findings in a mouse model suggest that blocking nerve signals to the pancreas could stop patients from ever developing type 1 diabetes.



"It's astonishing that this process may be stoppable through neuronal influence," says LJI Professor Matthias von Herrath, M.D., who served as the study's senior author.

The von Herrath Lab has been working to uncover the cause of type 1 diabetes. Although there are environmental and genetic risk factors for the disease, type 1 diabetes often seemingly strikes at random. Over the years, researchers have sought an explanation for the observed patchy pattern of cell death. One theory has been that these patches have differences in blood flow or they have been damaged by a virus that might be sparking an immune attack.

But recently, researchers have been exploring a new field called neuroimmunology, which is the idea that nerve signals can affect immune cells. Could nerve signals drive immune cells to attack certain areas of the pancreas?

“We thought that could explain a lot,” says study first author Gustaf Christoffersson, Ph.D., a former LJI postdoctoral researcher now at the University Uppsala, Sweden.

To test this theory, the researchers used a mouse model that can be experimentally induced to have beta cell death. They “denervated” the mice, either surgically or through use of a neurotoxin or a pharmacological agent, to block most of the sympathetic nerve signals to the pancreas. The researchers then used LJI’s world-class imaging facility to track the pattern of beta cell death in living mice.

The team found that blocking the nerve signals protected mice from beta cell death, compared with no effect in mice given no treatment and mice given only beta blockers. Without innervation, it was like the pancreas had gone dark and immune cells were unable to find their targets

“We were pretty surprised to see that these nerve blockers led to pretty significant differences in the onset of diabetes,” says Christoffersson.

More work needs to be done before this method can be tested in people. Von Herrath explains that doctors would first need a reliable way to identify patients at risk of type 1 diabetes onset. Once these patients are identified, von Herrath believes they could be treated either through electrostimulation or drugs to block nerve signals. There are also non-surgical, intravascular methods for blocking nerve signals.

The new discovery might explain much more than the patchiness seen in type 1 diabetes. Several autoimmune diseases share this same patchiness—but in a symmetrical pattern. For example, the skin condition vitiligo causes skin to lose its pigment, often in symmetrical

areas across the faces and hands. Arthritis also tends to strike symmetrically, with inflammation in both knee, elbow or wrist joints.

The new study suggests that these areas may be innervated by nerves that branch out symmetrically through the body.

“This symmetry is very striking, and it’s been almost impossible to explain,” says von Herrath.

Von Herrath thinks breakthroughs in neuroimmunology could have broad implications for explaining why the body turns against its own organs in many autoimmune diseases. Going forward, he and Christoffersson hope to investigate the cellular mechanisms that connect the nervous system and type 1 diabetes.

The study, “Interference with pancreatic sympathetic signaling halts the onset of diabetes in mice,” was supported by the National Institutes of Health (R01 AI092453), the Swedish Research Council (2014-06840 and 2018-02314), the Swedish Society for Medical Research, the Göran Gustafsson Foundation, and the Science for Life Laboratory.

The study was co-authored by Sowbarnika S. Ratliff of LJI.

LJI team gets first-ever look at a rare but vital stem cell in humans

New understanding of immune cell origins may offer cancer and COVID-19 drug targets

Madeline McCurry-Schmidt
Science Writer
August 18, 2020

LA JOLLA—Neutrophils are the warriors of the immune system. They are always ready to spring to action to help heal injuries or fight off disease. Unless, that is, something goes wrong in their developmental process.

Immature neutrophils aren't all warriors—they can be dangerous turncoats. High levels of immature neutrophils in the bloodstream can be a tell-tale sign of cancer and may even be a biomarker for COVID-19.

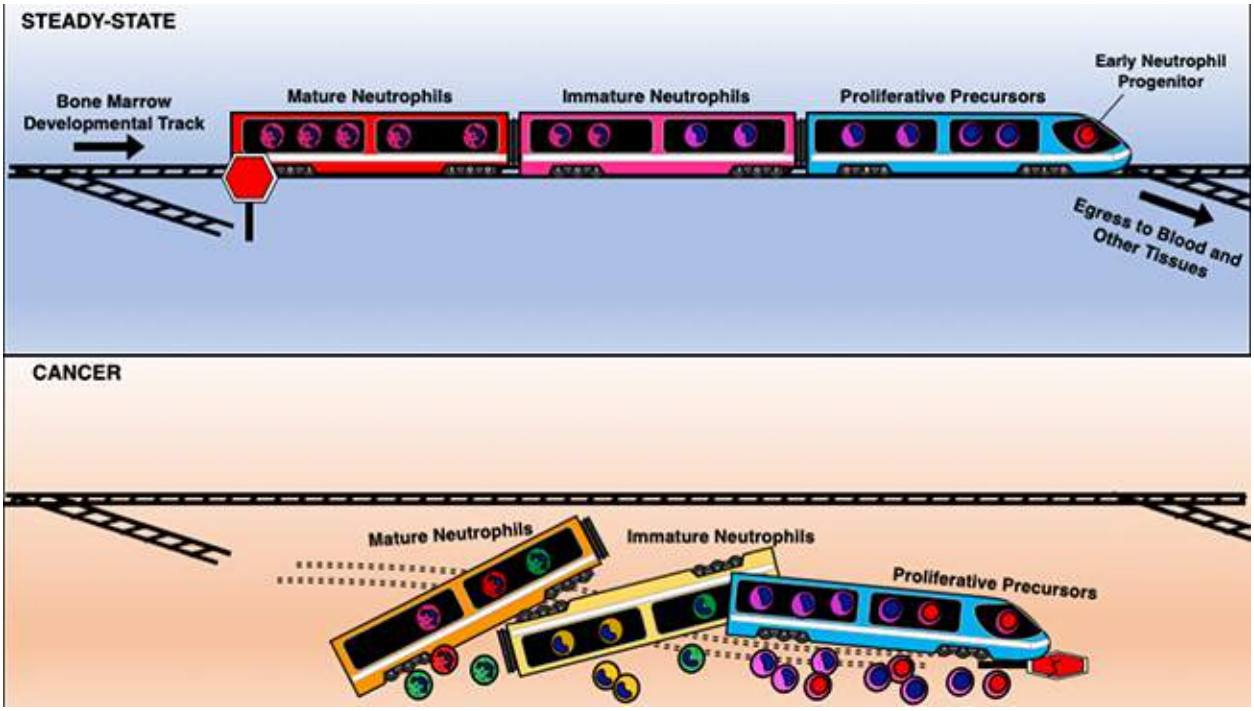
Now scientists at La Jolla Institute for Immunology (LJI) have tracked down the rare stem cells that generate neutrophils in human bone marrow. This research, published August 18, 2020, in Immunity, gives researchers a potential path for intervening in diseases where neutrophil development goes awry.

"We have identified the stem cells that are the early origins of neutrophils, the most abundant blood cell type in humans," says Huy Dinh, Ph.D., a former LJI postdoctoral associate who recently moved to a faculty position at The University of Wisconsin–Madison. Dinh led the study with LJI Professor Catherine C. Hedrick, Ph.D. "Knowing how human neutrophils develop is especially relevant today because immature neutrophils have been found to be elevated in both the blood and lungs of severe COVID-19 patients."

Despite their importance, neutrophils have proven very hard to study. They don't hold up well outside the body, and the stem cells that make them are even harder to investigate because they only live in bone marrow.

In 2018, the Hedrick Lab reported the discovery of a group of "progenitor" stem cells that give rise to mature neutrophils. These progenitors' sole job was to generate neutrophils, yet they appeared to also promote tumor growth. The researchers believed that detecting these progenitors could give doctors a better way to catch early cancer cases. But first, the team needed to know a lot more about neutrophil development.

The new research revealed a progenitor cell type that exists even earlier in human neutrophil development. Dinh, a past SPARK Award recipient, together with Tobias Eggert, Ph.D., a LJI visiting scientist and Melissa Meyer, Ph.D., a LJI postdoc, who served as the co-first authors in the study, spearheaded the effort to use a tool called cytometry by time-of-flight (CyTOF) to distinguish these rare cells from other types of immune progenitor cells. This work also made it possible for the researchers to identify more specific protein markers on this early progenitor cell surface.



Cancer can derail neutrophil development. Image courtesy of Daniel Araujo, La Jolla Institute for Immunology.

The discovery of these protein markers was important because until now, scientists have used only a few of markers to track neutrophils over time. The new study gives scientists specific markers for tracking neutrophil development from day one.

The researchers also found that cases of skin and lung cancers are often accompanied by a flood of immature neutrophils including the early progenitor cells into the bloodstream. These immature neutrophils change as they interact with tumor cells, though the researchers aren't sure yet how these changes affect cancer progression.

Dinh likens the stages of neutrophil development to the cars on a train. The early progenitors are like the train engine, keeping everything going smoothly along the track to maturity. Cancer shakes everything up, and immature neutrophils jump off the track before they reach maturity. "It's like the train is falling apart," Dinh says.

Neutrophil development has been in the news recently due to the COVID-19 pandemic, as studies have shown immature neutrophils are also more abundant in some patients with COVID-19. Dinh and Hedrick think perhaps the threat of the virus prompts the body to churn out neutrophils too quickly, again forcing immature cells off the track to maturity.

“We need to study this phenomenon further to see if these neutrophils can be tied to case prognosis or if they can be a drug target for COVID-19,” says Dinh.

The researchers hope to continue their work to discover the exact mechanisms that stop neutrophils from reaching maturity. “Knowing the earliest cell that gives rise to neutrophils is really critical for trying to target and control these cells,” says Hedrick. “But we don’t know exactly how to do that yet.”

The study, titled, “Coexpression of CD71 and CD117 identifies an early unipotent neutrophil progenitor population in human bone marrow,” was supported by the Tullie and Rickey Families SPARK Awards program and the National Institute of Health (grants 2T32AR064194, U01 CA224766, R01 CA202987, R01 HL134236, P01 HL136275.)

Additional study authors include Tobias Eggert, Melissa Meyer, Yanfang Zhu, Claire Olingy, Ryan Llewellyn, Runpei Wu.

DOI: 10.1016/j.immuni.2020.07.017

Researchers track down metabolic enzyme that protects against inflammation

LJI study shows new link between metabolism and the immune system

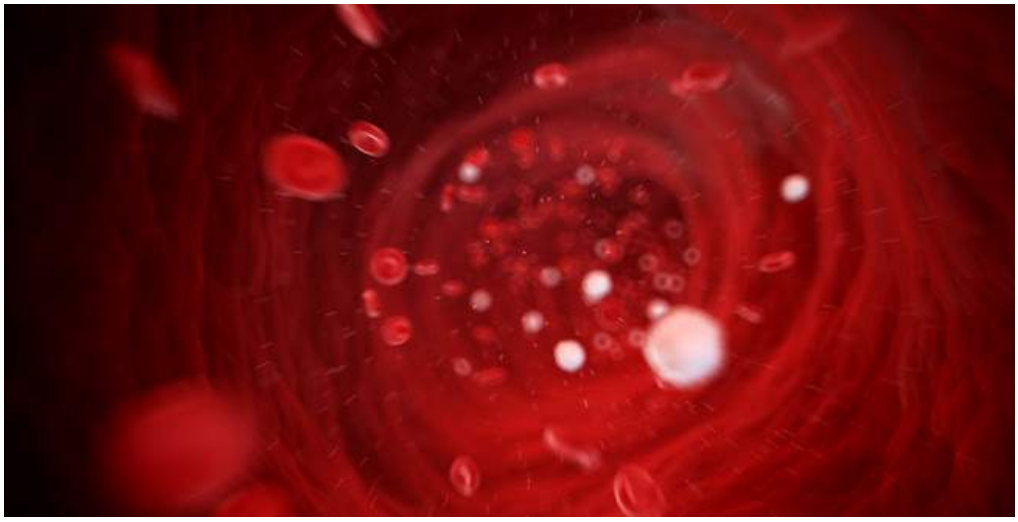
Madeline McCurry-Schmidt
Science Writer
July 17, 2020

LA JOLLA, CA—Scrape your knee, and you'll see some red puffiness appear around the injury. This is inflammation, and it is driven by the immune system.

Inflammation around a scrape is minor, but inflammation around joints—due to diseases like rheumatoid arthritis—can be incredibly painful. And inflammation in and around the blood vessels can prove deadly.

Now a new study from scientists at the La Jolla Institute for Immunology (LJI) shows that a previously poorly understood enzyme actually inhibits inflammation in blood vessels. The research offers a potential path to treating DADA2, an inflammatory blood vessel disease in children that is similar to Kawasaki disease.

"This is all pretty uncharted because for a long time, this was a very misunderstood enzyme," says Sonia Sharma, Ph.D., associate professor at LJI and senior author of the new study in *Science Advances*. "But we showed that the metabolic activity of this enzyme, ADA2, is very important for restraining the immune system in both vascular cells and immune cells."



DADA2 vasculitis, which leads to massive inflammation of small and medium blood vessels, is caused by a mutation in the enzyme ADA2.

Linking a mysterious enzyme to inflammation

Sharma's lab is dedicated to uncovering the earliest cellular and molecular origins of inflammation. Her work has led her to focus on stromal cells, which can initiate inflammation as a result of cellular damage, transformation or infection. This group of cells includes the endothelial cells that line your vascular system. Because endothelial cells are in constant contact with your blood, they are in a good position to raise the alarm when they detect viruses, bacteria, tumor cells or tissue damage. This kind of early detection is called an innate immune response, and it forms the foundation of downstream inflammation.

"We think the stromal cells' capacity for sensing early pathogenic threats is likely very important for initially coordinating and developing the magnitude, duration and quality of the downstream immune response," says Sharma, "and ultimately determining whether the immune response is protective or pathogenic in nature."

Sharma and her colleagues looked at a group of uncharacterized human disease genes linked to lupus-like autoimmune diseases that all cause multi-organ systemic inflammation or vasculitis. All the diseases were linked to single-gene mutations, but in most cases no one had uncovered exactly how each mutation affects the innate immune system. For the new study, the researchers asked if any of the disease-linked genes play a role in how stromal cells respond to threats by initiating an innate immune response.

Their search led them to an enzyme called ADA2, which is mutated in ADA2 deficiency disease (DADA2), and whose loss stimulated a robust innate immune response in ADA2-deficient endothelial cells and in white blood cells called monocytes.

The researchers were intrigued. Although ADA2 was implicated as a bio-marker for infection and inflammation, ADA2 is not thought of as a functional player in the immune system. Instead, this protein with metabolic enzyme activity was considered mostly redundant with its cousin, ADA1, which breaks down immune-modulatory molecules called purine nucleosides. Sharma says ADA2 was misunderstood.

"For a long time, ADA2 was ignored—no one thought it did anything," says Sharma. "What we know now is that both ADA1 and ADA2 play a much more nuanced role in regulating the bio-activity of purine nucleosides, which are powerful signaling molecules for the immune system."

In fact, two 2014 studies linked genetic loss of ADA2 to the inflammatory condition DADA2, which is a childhood-onset vasculitis syndrome. The new study gave Sharma the impetus to further investigate ADA2's functional role in inflammation.

Sharma and her colleagues showed that ADA2 enzyme activity naturally inhibits the innate immune response by restraining the production of a cytokine called type 1 interferon beta. Although type 1 interferon beta is a protective molecule against viral infections and cancer, too much type 1 interferon beta causes harmful inflammation, as seen in DADA2 and lupus.

The new finding from Sharma's lab is the first one to show that the metabolic enzyme activity of ADA2 can play a role in regulating vascular inflammation. "Loss of ADA2 drives metabolic dysregulation which in turn drives immune dysregulation and inflammation, so what we ended up identifying was new immuno-metabolic axis," says Sharma.

At last ADA2's role in disease made sense.

A new path for improving human health

Sharma says that a better understanding of ADA2 and purine nucleoside metabolism may open the door to new therapies for treating DADA2 and other types of multi-organ systemic inflammation. She believes that targeting the roots of these diseases will likely require a gene therapy approach or bone marrow transplants to reconstitute ADA2 in patients who cannot make the enzyme themselves.

"I think our study suggests the simplest treatment must be pursued to correct the metabolic dysregulation—we must reconstitute this particular enzyme activity in patients, ultimately," Sharma says.

Going forward, Sharma plans to further study ADA2 function in human cells and develop *in vivo* models of ADA2 deficiency. She also sees a link between ADA2 research and COVID-19 studies.

Sharma's lab is currently studying how the novel coronavirus may infect stromal endothelial cells and other cells involved in stimulating the body's innate immune response. They are hoping to shed light on how the viral pathogenesis is linked to vascular complications and the recently described Multisystem Inflammatory Syndrome in Children (MIS-C), which has similarities to Kawasaki disease and DADA2. By tracking ADA2 activity and purine nucleoside

levels in the blood of COVID-19 patients, Sharma's lab may be able to see if the virus really is targeting the vascular system.

"We're studying COVID patients now to see if ADA2 activity or purine nucleosides change," says Sharma.

The study, titled "Cellular sensing of extracellular purine nucleosides triggers an innate IFNb response," was supported by the National Institutes of Health (grants R01CA199376, U01DE028227, S10OD020025, R01ES027595, RO1 AI101423, T32GM007752 and F31CA236405), the Vasculitis Foundation and a SPARK award from the LJI Board of Directors.

Additional authors of the study included Rekha Dhanwani, Mariko Takahashi, Ian T. Mathews, Camille Lenzi, Artem Romanov, Jeramie D. Watrous, Bartijn Pieters, Catherine C. Hedrick, Chris A. Benedict, Joel Linden, Roland Nilsson, Mohit Jain.

DOI: https://doi.org/10.1126/sciadv.aba3688

WE ARE WHAT WE EAT

Postdoc's exploration of olfactory receptor function could lead to therapies to prevent atherosclerosis.

When 16-year-old Marco Orecchioni, Ph.D., abandoned his dream of becoming a chef to pursue science, little did he know that years later food would play a major role is his quest to counter the pro-inflammatory effect poor diets have in triggering atherosclerosis.

Dr. Orecchioni, a postdoctoral fellow in the lab of professor Klaus Ley, M.D., at the La Jolla Institute for Immunology (LJI), is a rising star at the Institute, having already notched a number of key research successes in a little over two years since he arrived in La Jolla after receiving his Ph.D. in his native Sardinia, an Italian island in the Mediterranean.

One of the signs that Dr. Orecchioni is an outstanding young talent came last year when he was awarded one of the Institute's Tullie and Rickey Families SPARK Awards for Innovation in Immunology. SPARK awards provide $25,000 of flexible start-up funding so that young scientists just beginning their careers can launch their ideas and generate enough data to attract scientific grants enabling them to take their projects to the next level.

Dr. Orecchioni's research expanded on the Institute's discovery that olfactory receptors are not restricted to our nose but can be found in other parts of the body in the form of macrophages, a type of white blood cell.

MARCO ORECCHIONI, PH.D.

"At first, we couldn't understand what role these olfactory receptors could play outside the nose," Dr. Orecchioni says. "But then we discovered that mice fed cholesterol-rich diets expressed hundreds of these olfactory receptors on macrophages isolated from their atherosclerotic arteries.

"That's when I really began to get excited because I was able to demonstrate that vascular macrophages can sense certain odorants in response to which they induce a pro-inflammatory response. It still has to be proven in humans, but my theory is that these olfactory receptors expressed in macrophages just might be one of the missing links between high fat diets and the progression of atherosclerosis."

Another indication Dr. Orecchioni is on a fast track is that his SPARK project, which normally would remain a side effort to work on in his spare time, has now become part of the Ley lab's main research focus (see Q&A with Dr. Ley, page 18).

While Dr. Orecchioni is amazed at his success at LJI, he traces its origins to his decision in high school not to follow in the footsteps of his brother—who was a chef—after he took his first biology class.

"I remember studying a cell and realizing how complicated and fascinating it was and how I really wanted to understand everything I could about it," Dr. Orecchioni recalls. "And then when I began to study immunology that's when my real passion began. I was intrigued because the immune system is not only one of the most intricate and complex systems in the body, there's a delicate balance between its amazing power to protect us against disease and its destructive ability to turn on us and harm our bodies through autoimmune diseases."

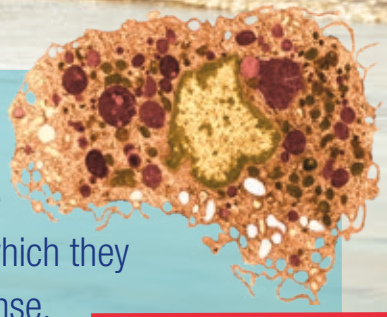
And Dr. Orecchioni is eager to tackle the autoimmunity issue head on. He knows his research on olfactory receptors is in the early stages, but he's optimistic enough to believe that within a few years it will lead to clinical trials with humans and ultimately development of a drug that would block the pro-inflammatory effect of the macrophages.

"One of the reasons I love being a scientist is that my research has the potential to help people," he says. "Nothing would be more exciting for me than to make a contribution to the health of millions around the world by coming up with a way to protect the human body against cardiovascular disease, which is still the leading killer around the world."

There is an even more personal reason Dr. Orecchioni wants his research to succeed.

"I'd like to help my own family, which has a bad history of cardiovascular disease. My grandfather died of a heart attack at 54 and my uncle has heart disease. My mother has high blood pressure and high cholesterol and I'm always telling her to eat this and not that. Everyone in the family knows that I'm working on this and they're very happy for my success and very proud," Dr. Orecchioni adds, laughing, "but they want me to hurry up!" •

"One of the reasons I love being a scientist is that my research has the potential to help people."



I was able to demonstrate that vascular macrophages can sense certain odorants in response to which they induce a pro-inflammatory response.

La Jolla Institute Board of Directors elects semiconductor industry pioneer

Former AMCC Chair and CEO David M. Rickey joins Institute Board

March 2, 2020

Share this article:

LA JOLLA, CA – The La Jolla Institute for Immunology is pleased to announce that David M. Rickey, former chairman and CEO of Applied Micro Circuits Corporation (AMCC), has been elected to its Board of Directors.

“We’re delighted to welcome David Rickey to our board,” said Mitchell Kronenberg, Ph.D., La Jolla Institute president and chief scientific officer. “Dave has enjoyed a legendary career in the telecommunications and semiconductor industries, including building AMCC from a struggling firm into one of the powerhouses of San Diego. His vast experience in business and high tech—which ranges from creating startups to running major corporations—will be invaluable to the Institute as we increasingly move our research discoveries into the clinical and private sector environment.”

Rickey, who is based in San Diego, was born in Geneva, a small town in western New York where his father worked as a chemist and his mother was a nurse and homemaker. He graduated summa cum laude with a BS in mathematics from Marietta College in Ohio. He also received a BS in metallurgy and materials science from Columbia University and a masters in materials science from Stanford University. Rickey’s first job out of school in 1981 was with IBM where he worked as a semiconductor engineer for four years.

In 1985, Rickey moved to San Diego where he was hired by the semiconductor division of Northern Telecom, the Canadian

telecommunications giant. By the time he left Nortel eight years later, he was director of operations and engineering and was running the company's semiconductor manufacturing operations in San Diego and Ottawa. He then left Nortel to join one of the company's customers, a small struggling company called Applied Micro Circuits Corporation, where he was vice president of operations for 20 months before he took the same job with NextGen in San Jose. He was there for only eight months when the chairman of AMCC asked him to return to the company as CEO. When Rickey rejoined AMCC the company was worth approximately $30 million, but over the next nine years, he and his team turned AMCC around and led the company to a top valuation of over $30 billion.

Since 2000, Rickey has served on a number of boards, including Netlist, AMI Semiconductor, Agility Communications, Silicon Wave, and Cytori Therapeutics, where he is currently board chair. He recently served as interim CEO and consultant to Calient Technologies.

Rickey has also been an active philanthropist, primarily through the Brenda and Dave Rickey Foundation he created with his wife. The foundation donated the funds to build a planetarium at Rickey's alma mater, Marietta College; provided the lead gift for construction of the Boys and Girls Club in his native Geneva, N.Y., and is helping finance the renovation of that city's YMCA.

Rickey's initial involvement with La Jolla Institute came last year when LJI board member Tom Tullie, Rickey's COO at AMCC, invited the couple join forces to create The Tullie and Rickey Families SPARK Awards for Innovation in Immunology at the Institute. The SPARK Awards provide flexible start-up funding that enables LJI's early career investigators to act on their promising projects for bold new approaches to diagnoses, treatments and possibly cures for diseases that afflict millions today.

"Brenda and I were thrilled to become involved with the Institute through the SPARK Awards," Rickey said. "We're inspired by the talent and drive of these young investigators and we're privileged to be able to provide them with a steppingstone to launch their research careers. We believe they are the future leaders of science and based on the quality and breadth of their projects, we have no doubt their research is not only going to be successful, it will ultimately have a major impact on improving public health for millions around the world."

From left to right: Judy and Tom Tullie, Dave and Brenda Rickey

Donors endow SPARK Program to fund innovative research by Institute's young scientists

Some of the most exciting and potentially life-saving immunology research never sees the light of day because it lies only in the imaginations of young scientists who have no path to develop their innovative ideas.

Two San Diego high tech executives with deep entrepreneurial experience are determined to provide that path. Tom Tullie and Dave Rickey, who worked together as chief operating officer and chairman/CEO, respectively, of Applied Micro Circuits Corporation, have made substantial donations to endow La Jolla Institute for Immunology's two-year old SPARK program over the next decade.

The Tullie and Rickey Families SPARK Awards for Innovation in Immunology provides flexible start-up funding so that young scientists just beginning their careers can launch their ideas and generate enough data to attract scientific grants, enabling them to take their projects to the next level.

Tullie says he and his wife, Judy, are thrilled to support a program they believe will lead to groundbreaking science.

"I've been an entrepreneur and involved in innovation my entire career, but this is the most exciting program I've seen because the ideas these young scientists have are not only fascinating, they have the potential to generate breakthrough discoveries that will someday save lives," says Tullie, who also serves on the Institute's board.

Each year, the SPARK program receives dozens of proposals from Institute scientists. A panel that includes Institute scientists, board members, and top business leaders reviews the proposals, selects the finalists, and has each of them present their projects in person. Four award winners are selected and each are provided $25,000 in start-up funding.

The projects now under way range from using nanoparticles to deliver cancer immunotherapy and developing a universal molecular diagnostic test for infectious diseases, to using pro/antibiotics to treat and cure allergies and targeting the microbiome to prevent the infection and chronic inflammation that cause cancer.

For the Rickeys, involvement in the SPARK program was spurred by the couple's belief in the Institute's mission itself.

"Brenda and I love the idea of 'life without disease' because so many of our friends and family have suffered or died from cancer, heart problems, dementia, and other ailments," Rickey says. "A lot of non-profits don't have clear goals, but the Institute's mission is concise while being extremely bold and ambitious. It may not be fully realized in our lifetime, but we believe with the amazing talent of their scientists, the Institute's mission is achievable."

Both Tullie and Rickey have helped attract other donors to join forces with them on SPARK, bringing the total amount of the fund to approximately $1 million.

"A lot of other young scientists have proposed really worthy projects that are potential home runs, but we just couldn't fund them all," Tullie says. "We are hoping others will join us and provide additional seed funding to advance some more of these promising ideas."

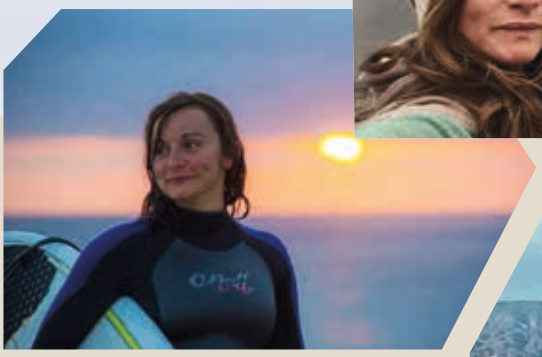
A SCIENTIST ABROAD

LJI postdoctoral researcher

Julie Burel

is changing the way we think about

INFECTIOUS DISEASE



When La Jolla Institute (LJI) postdoctoral researcher Julie Burel, Ph.D., isn't in the lab, she's adventuring. She's gone diving in Australia and hiked all over the world. A recent trip to Colombia included a four-day high-altitude trek in the Paramo, a unique alpine tundra ecosystem in the Northern Andes.

But it was a vacation to the South Pacific nation of Vanuatu that really stuck with Dr. Burel. "It's a fantastic place," she says. Her backpacking group stopped in a village, and Dr. Burel—who is originally from France—found the people there spoke a little French. Dr. Burel quickly made friends with a local teacher, and her group was welcomed by the chief.

Walking through the village, Dr. Burel saw what looked like a grave. A child had recently died of malaria, she was told.

Dr. Burel's research focuses on better understanding immune responses to infectious diseases to improve diagnosis, treatment, and vaccine efficiency. She stared at the grave. *This is why I'm doing what I'm doing*, she thought.

Although she is still early in her career, Dr. Burel's work has already forced immunologists to look more closely at how the body responds to infections. Earlier this year, Dr. Burel and her mentor, Bjoern Peters, Ph.D., Professor in the Vaccine Discovery Division, published a study in the journal *eLife* showing that monocytes and T cells, two key cell players in the immune system, appear more likely to stick together as "doublets" when the body is facing a threat, such as tuberculosis or dengue fever.

Dr. Burel thinks that looking for this signature in the blood could help doctors more quickly diagnose deadly diseases, and highlight which patients are more likely to face severe disease progression or treatment failure. She says LJI is uniquely set up to encourage her to pursue these kinds of unexpected discoveries. It's a place where collaboration makes every project stronger. For example, Dr. Burel's research often takes her down to LJI's Flow Cytometry Core, where the team seems just as excited as her to see new data.

"You have a sense of community because of the Institute's small size," says Dr. Burel. "Everything you need is on site. There are strong core facilities, and you have a lot of communication between the administration teams and the research teams. That's really important."

LJI's generous donor community has fueled this work. In 2019, Dr. Burel received $25,000 in unrestricted research funding through the Tullie and Rickey

"LJI lets you really focus 100 percent on doing research. I think that's what makes this place really successful."

UP & COMING

Why are some people able to fight off infections while others are not?

Samples collected from geographically diverse populations provide insight into human immune diversity.

GOAL

To isolate white blood cells from samples and analyze which specific genes are turned off and on.

These gene patterns can reveal exactly how the immune system has reacted to the disease.

Families SPARK Awards for Innovation in Immunology. The SPARK Program allows researchers to pursue bold ideas that traditional funding agencies often see as too risky.

LJI has found that this research often leads to unexpected breakthroughs, such as Dr. Burel's "doublet" cell findings, so donors have stepped in to make SPARK funding available to early career scientists.

The SPARK Program was Dr. Burel's first experience in earning grant money. She said she learned a lot from presenting her work to non-scientists during the award selection process. "I think communicating with the public is really important, and scientists don't do it enough," Dr. Burel says.

Dr. Burel's next goal is to dive into one of the biggest questions in immunology: Why are some people able to fight off infections while others are not? To answer this question, researchers need to know more about the mechanisms driving immune system variability. Dr. Burel wants to know exactly what goes wrong in some patients. Which immune cells aren't working?

"That's very important for us to know to do more personalized medicine," Dr. Burel says.

Malaria is a good example of this variability. Most people can control the disease and may just experience a bad fever. But the disease can prove deadly for pregnant women and children, who may not have as strong an immune system.

Tuberculosis is another example. Most people can control the disease in an asymptomatic stage, called latent tuberculosis. But in a small fraction of individuals, mostly those with a weakened immune system, such as children and HIV-seropositive people, the tuberculosis bacteria can develop into a deadly lung disease. Dr. Burel wants to know what to test for in the blood to predict how the body will react to the disease.

"We still don't know how it works in most cases," she says. "That is why we can't have one therapeutic approach that works for everyone."

Studying individual variability comes with the challenge of processing many, many human samples. These days, Dr. Burel spends hours hunkered down in a tissue culture room, working with samples from patients with tuberculosis. Her goal is to isolate white blood cells from the samples and analyze which specific genes are turned off and on. These gene patterns can reveal exactly how the immune system has reacted to the disease. She's handling samples right now from patients in Peru, Sri Lanka, and South Africa. "It's quite geographically diverse, which is important if we want to understand human variability at a global level," Dr. Burel says. She's also working hard to strengthen her skills in bioinformatics, a field that will help her better analyze the large datasets that come from her research.

"LJI lets you really focus 100 percent on doing research," says Dr. Burel. "I think that's what makes this place really successful."

As she forges ahead, Dr. Burel knows there are lives on the line. As she meets people during her travels, they often ask what she does back home.

"Whenever I tell them, they say, 'this is good, this is helping.'" •

No cell is an island

LJI team finds that 'doublet' immune cells are much more common—and more important in disease—than previously thought

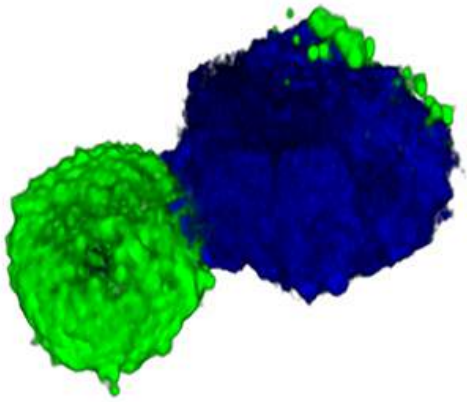
Madeline McCurry-Schmidt
Science Writer
June 24, 2019

Share this article:

LA JOLLA, CA – Sorting through individual immune cells is a handy way to see how the body responds to disease. For years, researchers around the world have used a technique called flow cytometry to separate different types of immune cells, such as T cells, B cells and monocytes. But occasionally they'd see two cells stuck together as a "doublet." Many regarded doublets as an artefact of the flow cytometry process, and it was common place to "dump" those conjoined cells before gathering data.

That was until scientists at the La Jolla Institute of Immunology (LJI) took a closer look at doublets and uncovered their link to disease. In a new study, published on June 25, 2019, in the journal eLife, the researchers report that higher levels of doublets can be found in people with severe cases of tuberculosis or dengue fever.

This finding shows that doublets are not simply technical artefacts of the flow cytometry process. Instead of ignoring these doublets, scientists could study them to predict disease progression and even vaccine effectiveness.



A representative 3D-reconstructed image of a doublet of a T cell (CD3, green) and a monocyte (CD14, blue). Image courtesy of Dr. Yunmin Jung, La Jolla Institute for Immunology.

“It’s an important lesson learned,” says Bjoern Peters, Ph.D., a professor in the Vaccine Discovery Division at LJI and senior author of the new study. “By studying these cells and how they communicate, we might learn more about the immune system.”

The presence of doublets also explains cases where an immune cell appears to show gene expression of two distinct types of immune cells. Peters and his team have studied thousands of these odd “double expressors” and shown that these are actually doublets. “You look closer, and really it’s two cells,” says Peters.

The team was inspired to investigate doublets when they ran into a puzzling group of cells in blood samples from patients with tuberculosis. Their flow cytometry experiments revealed immune cells in these patients the characteristics of both monocyte cells and T cells. The scientists saw the same strange cells in patients with dengue fever.

Julie Burel, Ph.D., postdoctoral research fellow at LJI, realized that high levels of these double expressors correlated with a higher risk of disease progression in the patients. Her follow-up experiments showed that these cells were in reality doublets made up of a monocyte and a T cell stuck together.

The researchers then started looking for these doublets in more samples. They found that while recently infected patients tend to have more doublets, doublets appear to be surprisingly common in everyone. “Every blood sample we’ve analyzed contains doublets,” says Burel, who served as first author of the new study.
So why did it take so long to find them?

“Instrument sensitivity is a lot higher now, so we’re able to see things no one could see 10 years ago,” Peters says.

The presence of doublets makes complete sense if you consider the role of T cells in the body, Peters explains. “T cells are meant to monitor what’s going on in other cells, and detect and even kill infected cells —so it makes sense to study T cells that are found in complexes with other cells,” says Peters.

In tuberculosis and dengue fever, monocytes become reservoirs for disease, so T cells would be more likely to come in contact with monocytes and form doublets.

Peters says doublets are likely relevant in many other diseases. After all, infection prompts immune cells to communicate and interact more closely. That could mean they are more likely to form doublets.

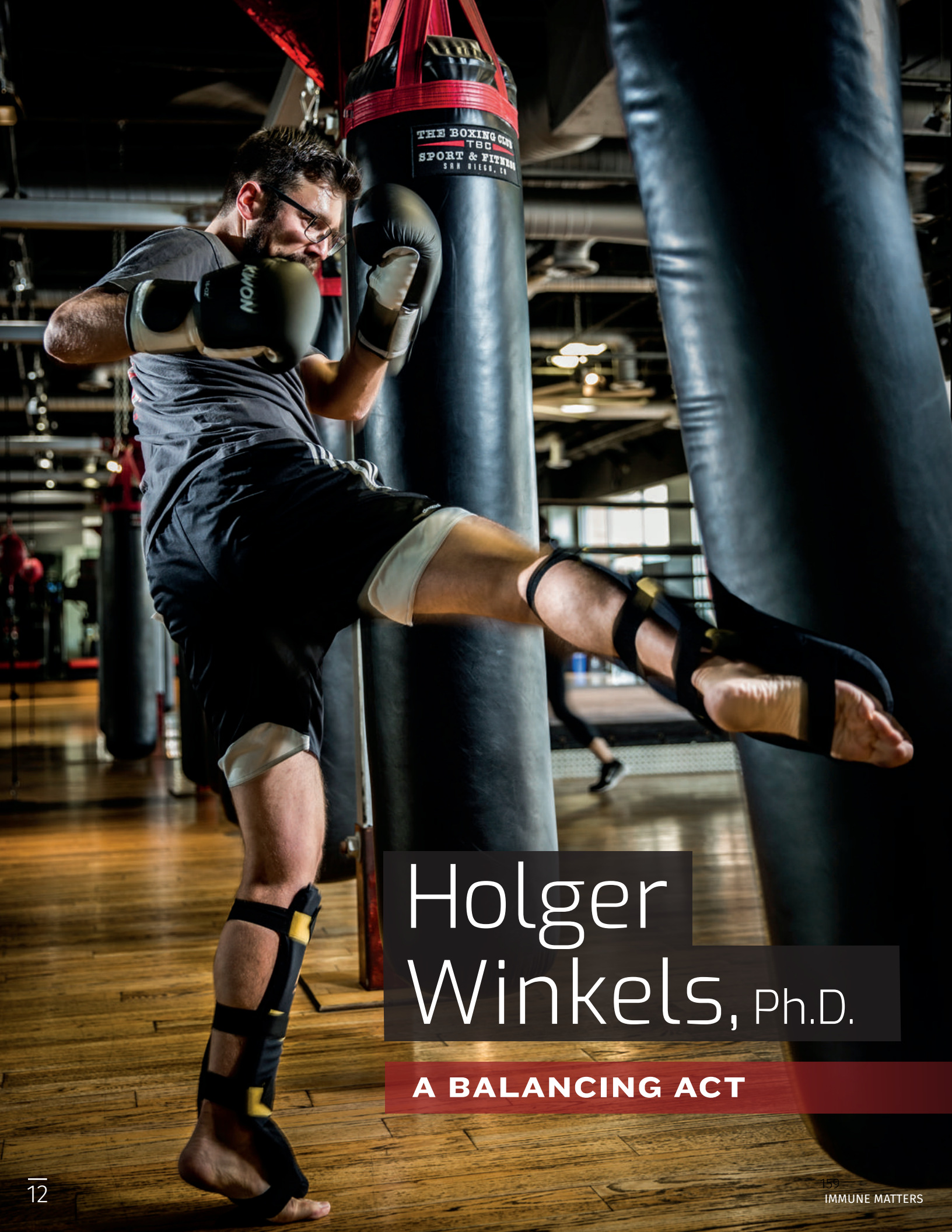
“The immune system is a very complicated, social network,” says Burel. “Flow cytometry is meant to break cells apart, but actually, doublets matter.”

As Peters likes to put it, “no cell is an island.” The new research has inspired his lab to look beyond the single-cell analysis usually done with flow cytometry and instead gather more data on doublets.

One question is whether measuring doublets could allow researchers to detect disease before a patient even shows the symptoms. This could be critical for stopping fast-moving diseases such as dengue. Doublets could also indicate when a vaccine is sparking an effective immune response, which could help speed up vaccine research.
“Doublets could be a biomarker for any immune perturbation,” Burel says.

Full citation:
Julie G. Burel, Mikhail Pomaznoy, Cecilia S. Lindestam Arlehamn, Daniela Weiskopf, Ricardo da Silva Antunes, Yunmin Jung, Mariana Babor, Veronique Schulten, Grégory Seumois, Jason A. Greenbaum, Sunil Premawansa, Gayani Premawansa, Ananda Wijewickrama, Dhammika Vidanagama, Bandu Gunasena, Rashmi Tippalagama, Aruna D. deSilva, Robert H. Gilman, Mayuko Saito, Randy Taplitz, Klaus Ley, Pandurangan Vijayanand, Alessandro Sette and Bjoern Peters. https://doi.org/10.7554/eLife.46045. eLife. 2019

DOI: 10.7554/eLife.46045



Holger Winkels, Ph.D.

A BALANCING ACT

Whether in the lab or the dojo, Holger Winkels, Ph.D., spends a lot of time thinking about balance. As a devoted martial arts practitioner with many years of jiu jitsu, krav maga, and muay thai under his belt, Dr. Winkels reads his opponents, finding an opening to push them off balance. As a postdoctoral researcher in the lab of Professor Klaus Ley, M.D., he studies the immune system to figure out how to achieve the perfect balance between pro-inflammatory and anti-inflammatory T cells to prevent atherosclerosis.

Born and raised in Germany, Dr. Winkels had always been interested in science. As a high school student he hesitated between chemistry and biology when an internship at Bayer propelled him toward biology. "I actually got to synthesize aspirin and it was fascinating," remembers Dr. Winkels. But it left him uninspired—unlike his first immunology lecture, "I immediately knew immunology was it!" Dr. Winkels' course was set and he never wavered.

During his diploma (undergraduate) thesis, Dr. Winkels studied T cell homeostasis in atherosclerosis before he dug deeper into T cell activation during his graduate studies in the lab of Professor Esther Lutgens, M.D., Ph.D., and Professor Norbert Gerdes, Ph.D., at the Institute for Cardiovascular Prevention in Munich. "I got into my current field very early on," he says. The added advantage of being able to split his time between Munich, Amsterdam, and Boston, Mass., "was fantastic," he says and adds with a laugh, "when I first arrived in Boston, I wasn't even sure I had landed in the U.S. since a lot of the announcements were in Spanish." He quickly fell in love with the city and Americans' cheerful and positive attitude. "It was the first thing I noticed and it was very refreshing," he says.

After finishing his graduate studies in 2016, the newly minted Ph.D. joined the lab of Dr. Ley. "If you are in immunology, you can't do much better than La Jolla Institute," he states as one of the reasons. The other one was Dr. Ley.

"Klaus has an outstanding reputation and is very well known in Germany," he says. Dr. Winkels' extensive expertise in T cell biology in the context of atherosclerosis was the perfect match for Dr. Ley's heart vaccine project. "We are looking at the question why a certain subset of T cells, so called regulatory T cells, is unstable in cardiovascular disease," he says. "The balance is skewed and our goal is to almost surgically dissect the immune system and find the dials that we can turn to stabilize regulatory T cells."

In an unexpected twist, Dr. Winkels' experiments turned up hitherto unknown immune cells that seem to play a role in atherosclerosis. Dr. Winkels successfully competed for a SPARK grant to follow up on his finding.

When he is not in the lab, Dr. Winkels is trying to find his own balance out in nature, at a music festival, a local bar, or in the muay thai boxing gym. Probably even more than the physical aspects, he values the mental workout martial arts provide him. "It forces me to focus and really be in the moment," he says. "It quiets my mind and gives me a feeling of peace." •

Lori and Paul Thiel present Dr. Holger Winkels with a check on behalf of the Advisor of the Year Awards Program, which will fund his successful SPARK proposal. From left to right: Dr. Mitchell Kronenberg, Dr. Winkels, Lori and Paul Thiel.

» READ MORE ON DR. LEY'S HEART VACCINE | PAGE 5

WE ARE LOOKING AT THE QUESTION WHY A CERTAIN SUBSET OF T CELLS, SO CALLED REGULATORY T CELLS, IS UNSTABLE IN CARDIOVASCULAR DISEASE.

Advisor of the Year presents check to SPARK program

March 15, 2018

Share this article:

LA JOLLA-On behalf of the San Diego Advisor of the Year Awards program, Paul and Lori Thiel presented a check for $25,000 to the SPARK program at La Jolla Institute for Immunology (LJI) in an awards ceremony held on February 26, 2018.

The newly created SPARK program was designed to help young researchers act on promising ideas. Without a proof-of-concept to convince granting agencies to fund a bold research proposal, many daring ideas are never put into action. The SPARK program overcomes the hurdle by providing flexible start-up funding to help young researchers close the gap between their imagination and that first solid set of data that allows them to attract external funding.

Inspired by the enthusiasm and dedication of the young researchers who applied for funding through the SPARK program, Lori and Paul Thiel topped off the proceeds from last year's Advisor of the Year Award event with a generous contribution to support the work of postdoctoral researcher Holger Winkels, Ph.D., who studies the role of the immune system in atherosclerosis.

From left to right: Trever Acers, Victor Ramshauer, Leo Kjiln, Judy Susser-Travis, John Kraemer, David Crean, Cim Kraemer, Steve LaSala, Lori Thiel, Mitchell Kronenberg, Paul Thiel, Holger Winkels, David Duval, Lamar Rutherford, Channing Hamlet, Clark Jordan, Bill Super

“The reason Paul and I decided to direct these funds toward heart disease research is because we felt it touches on all our lives in one way or another,” said Lori Thiel, an internationally experienced communications strategist, who oversees the administration and strategic planning of the Advisor of the Year.

Arterial plaque build-up, the main driver of heart attacks and strokes, is initiated and perpetuated by immune cells accumulating at sites of vessel tissue damage. Dr. Winkels is particularly interested in understanding how the interactions of immune cells within atherosclerotic plaques may lead to fatal events.

During the course of his work, Dr. Winkels uncovered two previously unknown B cell subtypes as well as non-traditional T cells, which captured his imagination. “I am incredibly grateful for the support and the vote of confidence,” said Winkels. “We know that immune cells play a crucial role in heart disease but there is still a lot we don’t know and the SPARK grants allows me to follow up on some exciting observation.”

“One reason La Jolla Institute has always resonated with us is that in addition to the spectacular work they do, you can actually make a tangible difference here,” said Paul Thiel, a wealth advisor who helped create Advisor of the Year, a not-for-profit organization that hosts an annual awards event to recognize and celebrate local advisors who provide extraordinary service to owners of middle market companies.

“By supporting Holger’s work we hope to not only drive heart disease research forward but also direct attention to something of substance in an era when civic engagement is often overlooked,” added Paul Thiel, who also co-chairs La Jolla Institute’s Planned Giving Advisory Council.

PAPERS PUBLISHED WITH DATA GENERATED BY THE TULLIE AND RICKEY FAMILIES SPARK AWARD WINNERS

(Alphabetical by scientist's last name)

La Jolla Institute FOR IMMUNOLOGY | Life Without Disease.®

Journal of Autoimmunity xxx (xxxx) xxxx

Contents lists available at ScienceDirect

Journal of Autoimmunity

journal homepage: www.elsevier.com/locate/jautimm

Human herpesvirus-6 is present at higher levels in the pancreatic tissues of donors with type 1 diabetes

Somayeh Sabouri[a,1], Mehdi A. Benkahla[a], William B. Kiosses[b], Teresa Rodriguez-Calvo[a,c], Jose Zapardiel-Gonzalo[a,c], Ericka Castillo[a], Matthias G. von Herrath[a,*]

[a] Type 1 Diabetes Center, La Jolla Institute for Immunology, La Jolla, CA, 92037, USA
[b] Core Microscopy, La Jolla Institute for Immunology, La Jolla, CA, 92037, USA
[c] Helmholtz Zentrum München, Helmholtz Diabetes Center, Institute of Diabetes Research Heidemannstrasse, 1, 80939, München, Germany

ARTICLE INFO

Keywords:
Human herpesvirus-6
Type 1 diabetes
Environmental factors

ABSTRACT

Human herpesvirus-6 (HHV-6) is a ubiquitous pathogen associated with nervous and endocrine autoimmune disorders. The aim of this study was to investigate the presence of HHV-6 in pancreatic tissue sections from non-diabetic, auto-antibody positive (AAB+), and donors with type 1 diabetes (T1D) and explore whether there is any association between HHV-6 and MHC class I hyperexpression and CD8 T cell infiltration.

HHV-6 DNA was detected by PCR and its protein was examined by indirect immunofluorescence assay followed by imaging using high-resolution confocal microscopy. Viral DNA (U67) was found in most pancreata of non-diabetic (3 out of 4), AAB+ (3 out of 5) and T1D donors (6 out of 7). Interestingly, HHV-6 glycoprotein B (gB) was more expressed in islets and exocrine pancreas of donors with T1D. However, gB expression was not directly associated with other pathologies. Out of 20 islets with high gB expression, only 3 islets (15%) showed MHC class I hyperexpression. Furthermore, no correlation was found between gB expression and CD8 T cell infiltration on a per-islet basis in any of the groups.

Our observations indicate that HHV-6 DNA and protein are present in the pancreas of non-diabetic subjects but gB expression is higher in the pancreas of donors with T1D. The possible role of HHV-6 as a contributory factor for T1D should therefore be further investigated.

1. Introduction

Type 1 diabetes (T1D) is characterized by the immune-mediated destruction of insulin-producing beta cells (β cells) in the pancreas. It is thought that an interaction between predisposing genes and environmental factors such as viral infections may trigger the disease. Several viruses, mainly from the *Enterovirus* genus have been considered as potential causal agents for human T1D [1]. However, little is known about the involvement of other viruses such as herpesviruses.

Following primary infection, herpesviruses remain in a latent state in various tissues of the human host and become reactivated later in life [2], particularly in patients with severe immunosuppression. Human Herpesvirus-6 (HHV-6), a *Beta-herpesvirus* subfamily member, is a ubiquitous virus associated with roseola or *exanthema subitum* [3]. Some reports have implicated HHV-6 in several autoimmune diseases such as multiple sclerosis, autoimmune connective tissue diseases, and Hashimoto's thyroiditis [2]. HHV-6 preferentially replicates in activated T cells but can also infect natural killer cells and several non-immune cells, and it is known to cause immunosuppression [2]. It is also commonly present in salivary glands, thyroid, liver, kidney, and therefore potentially present in pancreas.

HHV-6 utilizes envelope glycoproteins such as glycoprotein B (gB), an abundantly expressed glycoprotein, for membrane fusion and viral entry [4]. Previous reports have suggested the involvement of HHV-6 in fulminant T1D [5,6]. However, the contribution of viral triggers in T1D development has been controversial and there is no documented association between roseola infection and T1D. Thus, here we attempted to explore the possible role of HHV-6 on disease pathogenesis and its presence in the pancreas. In line with this idea, the presence of HHV-6 in the pancreas has been examined recently by Ericsson et al. [7], where the authors detected HHV6-B DNA in pancreatic islets from brain-dead diabetic and non-diabetic donors by PCR assay [7]. In our study, in addition to PCR, we have used high-resolution confocal microscopy on a per-islet basis to investigate the presence or absence of the HHV-6 gB

* Corresponding author. La Jolla Institute for Immunology, 9420 Athena Circle, La Jolla, CA, 92037, USA.
E-mail address: matthias@lji.org (M.G. von Herrath).
[1] Sanofi, Translational Medicine & Early Development, 1 Mountain Road, Framingham, MA, 01701, USA

https://doi.org/10.1016/j.jaut.2019.102378
Received 6 August 2019; Received in revised form 25 November 2019; Accepted 29 November 2019

protein in pancreatic tissues. We also examined whether HHV-6 infection is associated with MHC class I expression and CD8 T cell infiltration, which are key hallmarks of T1D. Our results indicate that HHV-6 is not associated with MHC class I expression or CD8 infiltration. However, the virus is more frequently present in diabetic pancreata than in subjects without diabetes, suggesting that diabetic pancreata might be more susceptible to acute and persistent viral infections.

2. Materials and methods

2.1. Subjects

Human pancreata were collected from cadaveric organ donors via nPOD through accredited organ procurement organizations that are authorized to serve all regions across the United States. Six μm sections from formalin-fixed paraffin-embedded (FFPE) tissue samples from the head (PH), body (PB) and tail region (PT) were obtained from non-diabetic (n = 4) donors, non-diabetic autoantibody positive donors (n = 5) and donors with T1D (n = 7). Donor information is summarized in Supplementary Table 1.

The experimental procedures of this study were performed in accordance with relevant guidelines and regulations. La Jolla Institute for Immunology Institutional Review Board approved all experimental procedures (protocol number DI3-054-1112). Informed consent from donor families was obtained by Organ Recovery Partners at the nPOD (https://www.jdrfnpod.org/for-partners/organ-recovery-partners/).

2.2. Indirect immunofluorescence (IF)

To determine the expression of HHV-6 gB, MHC class I, and CD8 T cells, pancreas sections were subjected to a standard triple IF staining protocol. After deparaffinization and rehydration in descending ethanol concentrations, sections were exposed to heat-based citrate antigen retrieval for 20 min. Sections were blocked with 10% goat serum for 1 h. HHV-6 protein was detected using a monoclonal mouse antibody H-AR-2 (10 mg/ml) directed against gp110 glycoprotein of both HHV-6A and HHV-6B (kindly provided by Janos Luka, Bioworld Consulting Laboratories; 1:2000) incubated overnight at 4 °C followed by an Alexa Fluor 647 conjugated goat anti-mouse IgG antibody (Life Technologies, Grand Island NY; 1:1000) incubated at room temperature for 30 min. Staining for insulin and glucagon was performed subsequently at room temperature for 1 h using the following antibodies: polyclonal guinea pig anti-insulin (Dako, Carpinteria CA; 1:600) and polyclonal rabbit anti-glucagon (Dako, Carpinteria CA; 1:2000). Detection was performed at room temperature for 30 min using polyclonal goat anti-guinea pig IgG Alexa Fluor 488 (Life Technologies; 1:1000) and polyclonal goat anti-rabbit IgG Alexa Fluor 555 (Life Technologies; 1:1000), respectively. The specificity of the HHV-6 gB antibody was confirmed by using HSB-2 cell lines (NIH AIDS Reagent Program) [8] infected or uninfected with HHV-6 as a positive and negative control, respectively (Fig. 1a).

For CD8 and MHC class I staining, a combination of insulin, CD8, and MHC class I antibodies was used in a consecutive section to the one stained for gB. Deparaffinization, rehydration, heat-induced antigen retrieval, and blocking steps were similar to the gB staining as mentioned above. CD8 staining was performed by using a polyclonal rabbit antibody (ab4055, Abcam, Cambridge MA; 1:400) incubated overnight at 4 °C followed by an Alexa Fluor 555 conjugated goat anti-rabbit IgG antibody (Life Technologies, Grand Island NY; 1:1000) incubated at room temperature for 30 min. Staining for insulin and MHC class I was performed subsequently at room temperature for 1 h using the polyclonal guinea pig anti-insulin (Dako, Carpinteria CA; 1:600) and monoclonal mouse anti-human anti-HLA class I ABC (Abcam, ab70328; 1:200). Detection was performed at room temperature for 30 min using polyclonal goat anti-guinea pig IgG Alexa Fluor 488 (Life Technologies; 1:1000) for insulin and polyclonal goat anti-mouse IgG Alexa Fluor 647 (Life Technologies; 1:1000) for MHC class I. For both sets of the staining (insulin/glucagon/gB or insulin/CD8/MHC class I), after applying the secondary antibodies, sections were washed with distilled water, and counterstained with Hoechst 33342 (Life Technologies; 1:200) for 20 min. Sections were then mounted with Prolong Gold anti-fade medium (Life technologies).

2.3. Image acquisition

Z stacks of images for insulin/glucagon/gB staining and consecutive sections for insulin/CD8/MHC class I staining were acquired using a laser scanning Olympus FV10i confocal microscope and a 60x/1.3NA objective. Images were acquired at Nyquist resolution (using a step size of 0.5 μm and optimal image frame size at 1024 × 1024), obtained as multi-paneled stacks that overlapped at 10%, and were automatically stitched using the FV10-ASW 4.2 Olympus software. These multi-stitched images were flattened using Fiji software (LOCI, Madison, WI) at maximum intensity and then imported into Image-Pro Premier software (Media Cybernetics, Rockville, MD) for further processing. All 8-bit images were acquired using the full dynamic intensity range (0–256) that was set based on the population of cells, which displayed the highest signal. The same instrument settings (based on laser power and gain at the detector) were also used to acquire all islets from the various donors. Similarly, they were used to obtain secondary antibody alone images and these were in turn used to define thresholds of real signal above background.

2.4. Exocrine tissue imaging

To determine the presence or absence of gB in the exocrine tissue, the slides were scanned by a Zeiss Axio Scan Z.1 slide scanner (Carl Zeiss microscopy) using an Orca Flash 4.0 v2 fluorescent camera with 20x objective.

2.5. gB quantification analysis

Endocrine analysis: In Image-Pro Premier, up to 10 islets per case were randomly analyzed for levels of fluorescent signal above background. All images received the same thresholding in order to define the real signal above background signal. However, donor 6081 displayed a different background and the threshold was selected manually for this case. Periphery was defined as the exocrine area around the islet in the captured image.

Exocrine analysis: In order to make sure that the analysis excluded the islets, both images labeled for insulin and glucagon were used to make masks of the islets and to subtract those islet regions from the images labeled with gB only, therefore enabling us to analyze exocrine gB only in Image-Pro Premier (See image preparation in supplementary methods). This was accomplished using macros created in image J (NIH-FIJI-J). The images were imported into Image-Pro Premier and then, the smart segmentation feature was used to highlight, detect and auto outline all the gB positive zones in all the images. The analyzed data was sent into a data collector and then to excel in an automated fashion through custom-made macros using automate and batch processing modules within Image-Pro Premier project explorer. All techniques, image acquisition and quantification were approved by histology and microscopy specialists at the La Jolla Institute for Immunology in order to guarantee the appropriate experimental setup and analysis for this study. Information regarding the number of cases and islets that were analyzed for gB analysis is summarized in Supplementary Table 2a.

2.6. MHC class I analysis

For MHC class I analysis, consecutive sections were used to find the same islets as those stained for gB (up to 10 per donor). For MHC class I quantification, similar image acquisition and quantification method as

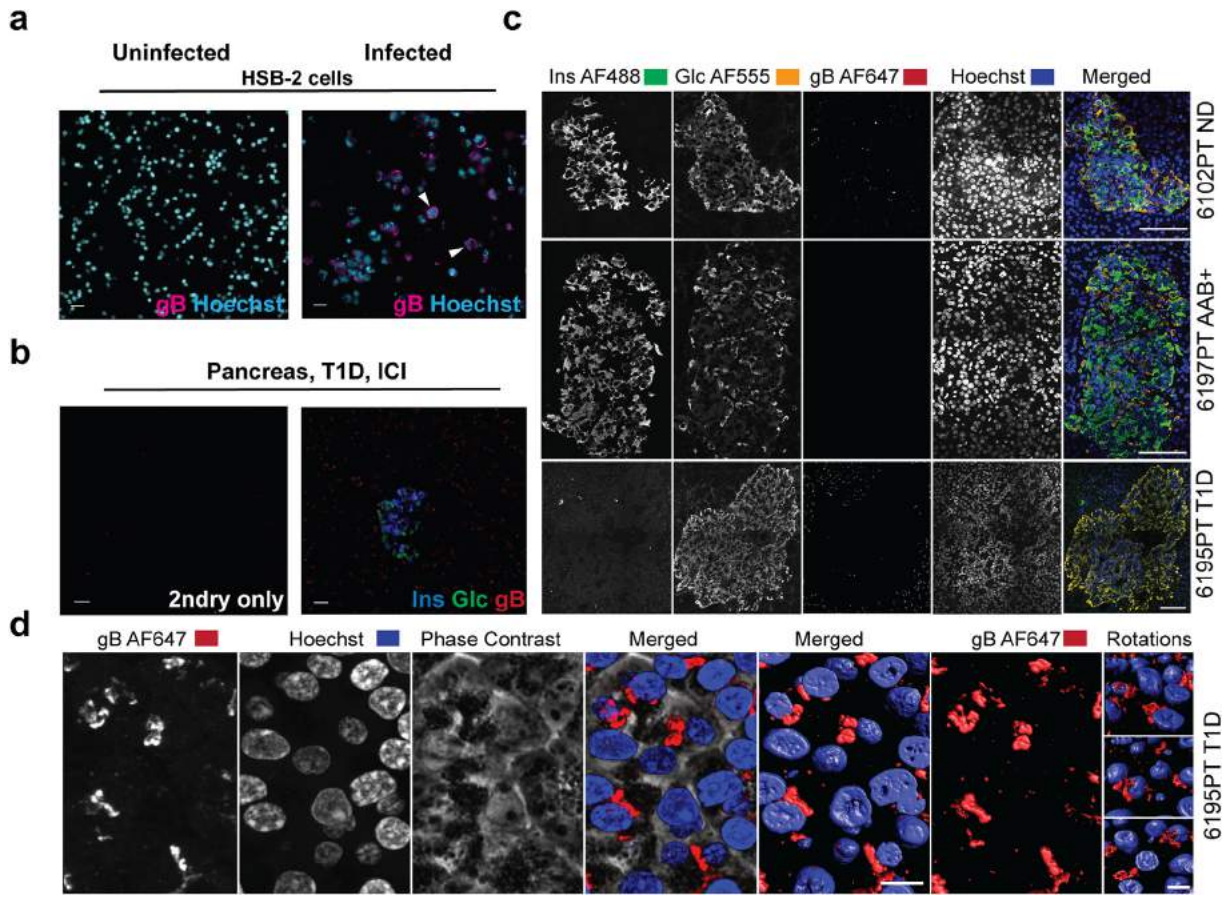


Fig. 1. Validation of the anti-gB antibody using HHV-6-infected cells and pancreas sections. (**a**) HSB-2 cells uninfected (left panel) or HHV-6-infected (right panel) were stained with anti-HHV-6 gB (H-AR-2) antibody. Goat anti-mouse AF647 was used as a secondary antibody and Hoechst was used for nuclear staining. Arrowheads show gB positive staining (*magenta*). (**b**) Representative triple staining of insulin (*blue*), glucagon (*green*), and gB (*red*) in the absence or presence of the primary antibodies (pancreatic tail region, case 6212). (**c**) Representative phenotypes of gB expression in a non-diabetic donor, AAB + donor and a donor with T1D. Sections were stained for insulin (AF488), glucagon (AF555), gB (AF647), and Hoechst for the nuclear staining. Last panel (right) shows the merged channels. Case number and the corresponding group are indicated on the right side of each panel. (**d**) Representative image of gB expression in high magnification (AF647) in a donor with T1D, case 6195, is shown. Fluorescently labeled samples were optically sectioned and reconstructed in 3 dimensions using a high-resolution microscope (Zeiss 780 LSCM) and Imaris software. T1D: Type 1 diabetes; ICI: Insulin-Containing Islets. Scale bars for (**a**) and (**b**): 20 μm, (**c**): 100 μm, (**d**): 10 μm. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

gB analysis was used. Information regarding the number of cases and islets analyzed for MHC class I is summarized in Supplementary Table 2b.

2.7. CD8 T cell analysis

In order to assess the number of CD8 T cells inside the islets, consecutive sections to the one stained for gB were used to find the same islets (up to 10 islets per case) (Supplementary Table 2c). CD8 positive T cells were manually counted on the sections that were scanned by the Zeiss Axio Scan Z.1 scanner using an Orca Flash 4.0 v2 fluorescent camera with 20x objective. Due to the different background levels between sections, contrast stretching was applied to the tissue sections in order to maximize the signal. To calculate the CD8 positive T cell density inside each islet, the number of CD8 positive T cells in each islet was divided by the islet size.

2.8. DNA extraction from pancreatic FFPE tissue sections

DNA was extracted from the consecutive FFPE sections to those stained by IF. Depending on the tissue size, 3 to 5 sections were used for DNA extraction, which was performed using the QIAmp DNA FFPE Tissue protocol (Qiagen). Briefly, pancreatic tissue sections were scraped using a razor blade. Paraffin was dissolved and removed using xylene followed by two times ethanol wash. Consequently, samples were lysed under denaturing conditions with Proteinase K. Formalin crosslinking was reversed by heat incubation. Finally, DNA was washed and purified.

2.9. Polymerase chain reaction

The extracted DNA was used to perform a PCR to amplify the U67 gene of both HHV-6A and HHV-6B. External primers [7] forward (GCGTTTTCAGTGTGTAGTTCGGCAG) and reverse (TGGCCGCATTCGT ACAGATACGGAGG) generated a 520 bp fragment.

PCR products were used to perform a nested PCR using internal primers. Forward (GCTAGAACGTATTTGCTGCAGAACG) and reverse (ATCCGAAACAACTGTCTGACTGGCA) primers generated a 258 bp fragment.

The amplified nested PCR products were analyzed by electrophoresis on a 2% agarose gel containing GelRed™ (Biotium). A 100 bp DNA ladder (Qiagen) was used as a molecular mass marker.

2.10. Statistical analysis

Data are presented as mean ± SD and analyzed using one-way ANOVA followed by Tukey's multiple comparison test. Analyses were performed using GraphPad Prism version 7 (GraphPad Software, San

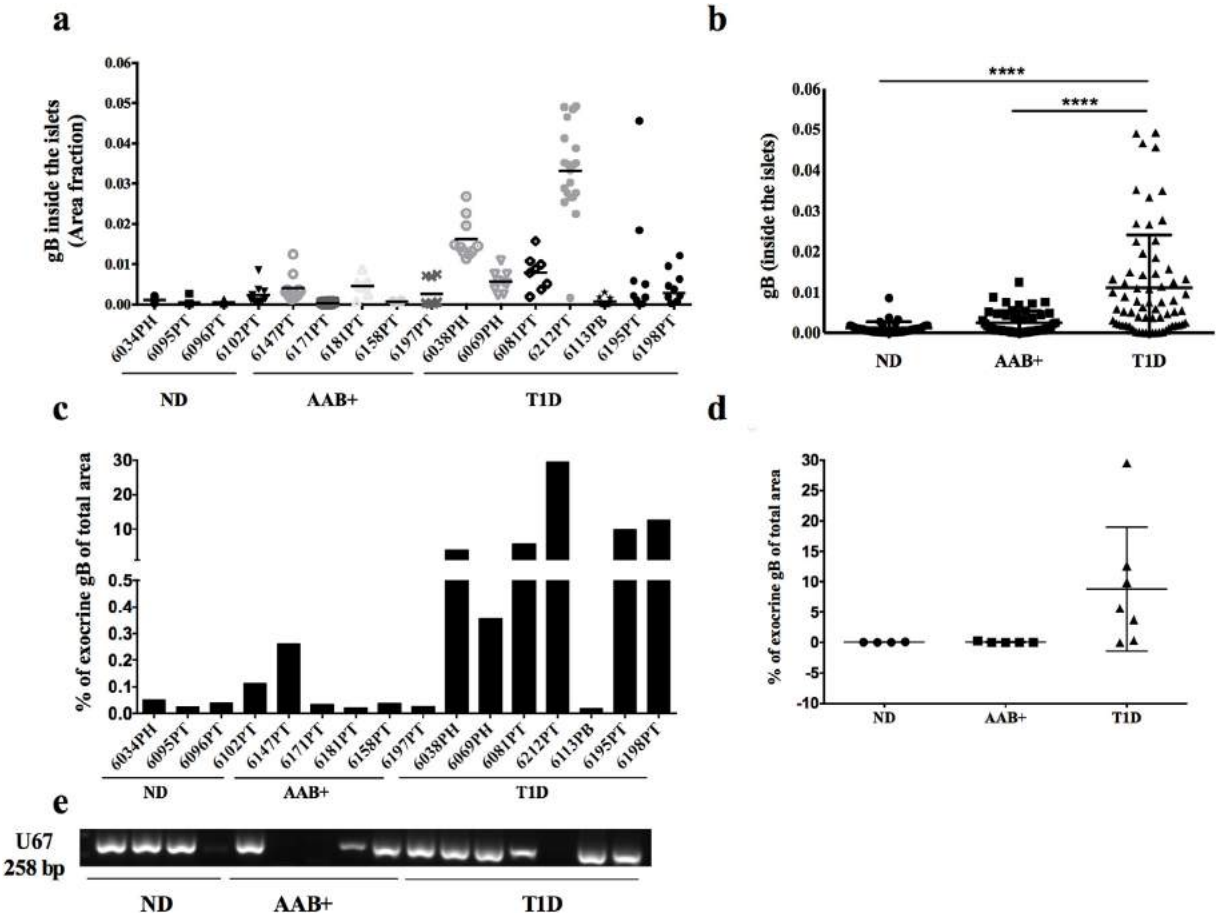


Fig. 2. gB is more frequently detected in the pancreas of T1D donors. Quantification of gB expression inside the islets (a, b) or in the exocrine (c, d) from non-diabetic donors (n = 4), AAB + donors (n = 5) and donors with T1D (n = 7). Each dot represents an islet and is presented as the positive fraction of the total islet area (a, b) or percentage of gB positive of total area (c). For each case, up to 10 islets were quantified. In (d), each dot represents a donor. (e) DNA was extracted from ND, AAB + and donors with T1D, and U67 viral gene (258bp) was detected by nested PCR. ND: Non diabetic; AAB +: Auto-antibody positive; T1D: Type 1 diabetes.

Diego, CA). A value of $P < 0.05$ was considered significant. The P values were corrected for multiple comparisons using the Benjamini and Hochberg method.

3. Results

3.1. gB protein of HHV-6 is detected more frequently in the pancreas of donors with type 1 diabetes than auto-antibody positive donors or donors without diabetes

HHV-6 gB antibody has been validated on infected and uninfected HSB-2 cells (materials and methods) (Fig. 1a) as well as on a pancreas section (Fig. 1b). Pancreas sections from 16 donors were assessed for the presence of gB protein by IF. Representative images of the triple staining of insulin (blue), glucagon (green) and gB (red) are presented in Fig. 1c and a higher magnification of gB expression is presented in Fig. 1d. A significantly higher level of gB expression was found in the islets of donors with T1D compared with AAB + donors and donors without diabetes (Fig. 2a and b). gB was detected in the islets of 6 out of 7 T1D donors compared to 1 out of 4 non-diabetic cases and 2 out of 5 AAB + donors (Fig. 2a). Consistent with high intra-islet expression of gB in the islets in donors with T1D, a higher expression level of gB protein in the periphery of the islets was also observed (Supplementary Fig. 1). In order to find out if gB expression was restricted to the islets and their surroundings, the presence of gB in the exocrine tissue was investigated. We observed that gB was more expressed in the exocrine tissue in donors with diabetes compared to non-diabetic donors (Fig. 2c–d). Indeed, gB was detectable in the exocrine compartment of all donors with T1D except case 6113. Also, gB was expressed at low levels in case 6034 and 6102 in the non-diabetic group. In AAB + donors, gB was only detected in the exocrine tissue of case 6147 (Fig. 2c–d).

3.2. HHV-6 DNA is detected in the pancreas of both donors with and without diabetes

Genomic DNA was extracted from the pancreas sections of non-diabetic donors, AAB + donors, and donors with diabetes. In order to detect HHV-6 genomic DNA, the samples were amplified by a nested PCR using specific primers. U67 gene, a major capsid protein of both HHV6-A and HHV6-B was present in 3 out of 4 non-diabetic donors, 3 out of 5 AAB + donors, and 6 out of 7 donors with T1D (Fig. 2e).

3.3. Intra-islet gB does not correlate with MHC class I expression or CD8 T cell infiltration in donors with type 1 diabetes

We analyzed MHC class I expression on consecutive sections to those stained for gB (Supplementary Table 2b). A representative image of MHC class I staining and gB staining on the consecutive section is shown in Fig. 3a, top and bottom panel, respectively. MHC class I protein level in the pancreatic islets was quantified similarly to the gB quantification as described above. In order to find out if there is a correlation between the expression of gB and MHC class I in the islets, data based on the area fraction was used for both gB and MHC class I.

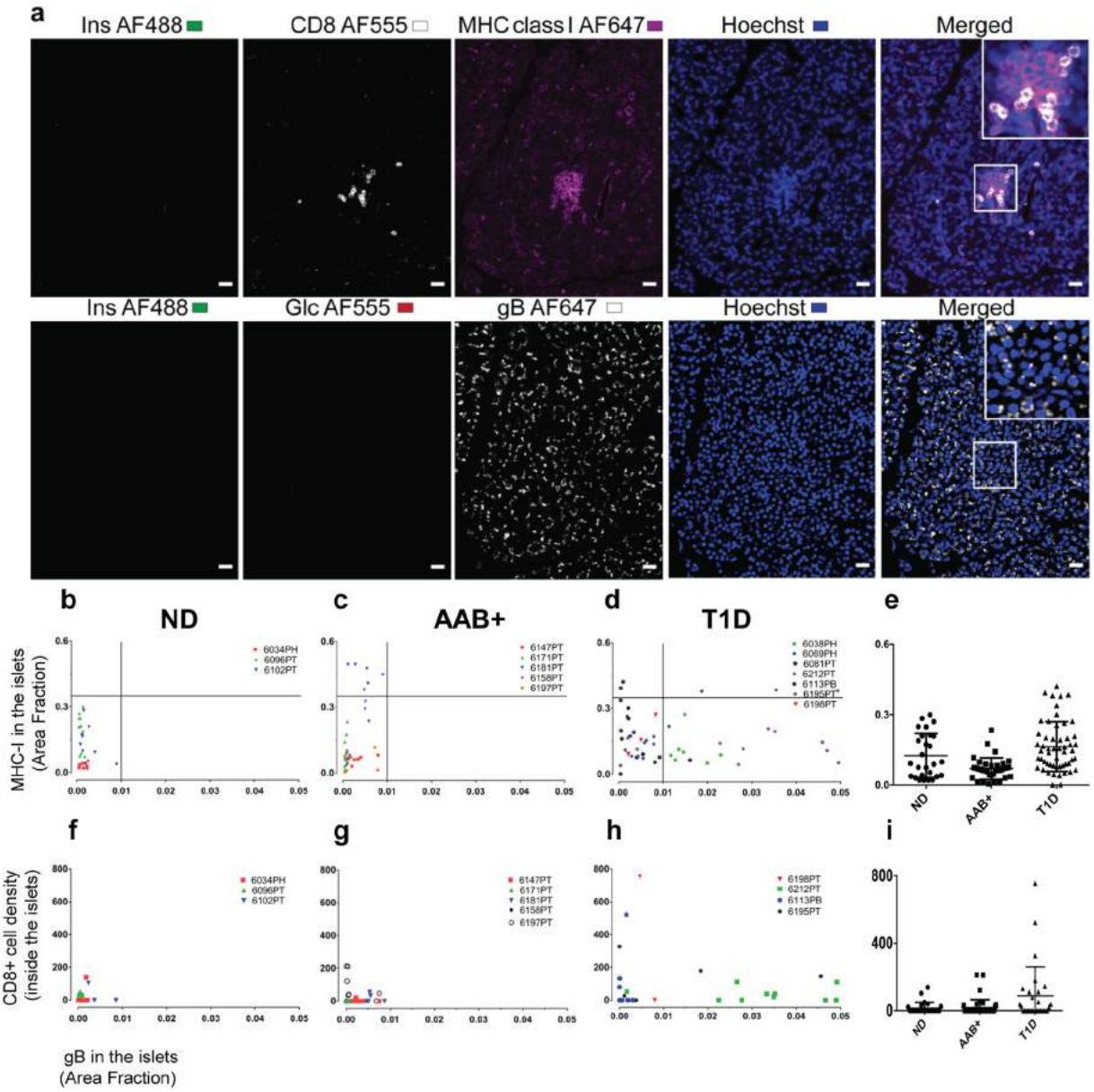


Fig. 3. Islet MHC class I expression and CD8 T cells infiltration is not associated with gB expression. Representative images from case 6195 (T1D) positive for CD8, MHC class I, and gB staining (a). Top panel: section stained with insulin (green), CD8 (white), MHC class I (magenta), and Hoechst (blue). Bottom panel: the consecutive section was stained with insulin (green), glucagon (red), gB (white), and Hoechst (blue). Last column in each panel (right) shows the merged channels and inserts show magnification of the area inside the box. Correlation between gB protein level and MHC class I expression inside the islets is presented based on the area fraction in (b) non-diabetic donors (n = 3), (c) AAB + donors (n = 5), (d) donors with T1D (n = 7) and (e) in groups. For each case up to 10 islets were quantified. Each dot represents an islet and each color represents an individual donor. Correlation between gB protein level and CD8 T cells density inside the islets in each group are presented (f–i). Each dot represents an islet and each color represents an individual donor. CD8 counting inside the islets was manually performed from scanned images and the CD8 density was presented as total CD8 number divided by the islet size. For ease of presentation, this value was multiplied by 10 × 6. . (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

However, no direct association between the presence of gB and MHC class I expression was found regardless of the diabetic status (Fig. 3b–d).

We performed a CD8 staining in combination with MHC class I staining on consecutive slides to that of gB staining, which enabled us to locate the same islets in both sections and perform a CD8 T cell correlation analysis (see Methods for more details). A representative image for CD8 staining and gB staining on the consecutive section is presented in Fig. 3a, top and bottom panel, respectively. Although we observed an increase in the number of $CD8^+$ T cells in donors with T1D (Fig. 3f–i), we could not find any association between the gB expression level and CD8 T cell numbers inside the islets on a per-islet basis (Fig. 3f–h and Supplementary Table 2c).

4. Discussion

Epidemiological and serological studies have suggested that several viruses including rotaviruses, retroviruses and other picornaviruses may have a role as triggers for the development of T1D [1]. However, due to the complexity of the human pancreas and the lack of available samples, there is limited information about the presence of other viruses such as HHV-6 in the pancreas. This is the first study in which the association of a viral infection with single islet pathology and not merely disease status is investigated.

The current study is different in many aspects from those of other investigators who reported the presence of HHV-6 in pancreata of healthy and T1D donors. In the present study, HHV-6 gB protein has been quantified and its possible association with MHC class I or CD8 T cell infiltration has been investigated.

Several considerations in the present report are noteworthy. Using high-resolution confocal microscopy, we observe a higher level of gB protein in the pancreatic islets from donors with T1D compared to non-diabetic donors. Similar results were obtained in the periphery of the islets as well as in the exocrine tissue of both diabetic and non-diabetic subjects. Furthermore, using nested PCR, the U67 viral gene has been detected in 75% of ND donors, 60% in AAB + donors and 85.7% in T1D donors. We have focused on U67, which is a late gene and expresses a major capsid protein of both HHV6-A and HHV6-B. The detection of U67 gene along with HHV-6 gB protein suggests an active viral infection. Our data are further supported by the detection of the viral receptor OX40 in the pancreas (data not shown). In accordance with our results, the presence of HHV-6 DNA in the pancreas from brain-dead organ donors with or without T1D has been recently reported [7]. However, in the study by Skog et al. there was a lack of concordance between PCR and IF.

According to the fertile field hypothesis, certain viruses may infect the pancreas, thus creating a fertile field for preexisting islet autoimmunity, resulting in its augmentation, increased β cell destruction, and ultimately leading to T1D development [10]. Our results indicate that HHV-6 infection might play such an indirect role in fostering the autoimmune process and may still act as a general pathogenic factor enhancing the development of the disease.

Hyperexpression of MHC class I has been reported in mouse models of T1D [11] as well as in the islets of patients with recent-onset T1D [12,13]. In agreement with previous studies [9], we observed a higher level of MHC class I expression in donors with T1D than in those without diabetes. However, there was no correlation between gB and MHC class I expression on a per-islet basis. The absence of correlation could be explained by the fact that HHV-6 downregulates the expression of MHC class I on the surface and within infected cells [14].

HHV-6 infects both CD4 and CD8 T cells *in vivo* [2]. Although we observed a higher number of $CD8^+$ T cells in islets of donors with T1D, we were unable to find any correlation between gB protein level and CD8 T cell infiltration on a per-islet basis (Fig. 3f–h).

These results indicate that, while there might be a link between gB expression and T1D, HHV-6 is not strongly associated with pancreas pathology during T1D (or in pre-diabetes) on a per-islet basis. Diabetic pancreata might be more susceptible to viral infections in general, indirectly contributing to disease pathogenesis. It is also possible that the higher expression level of gB seen in the pancreata from donors with T1D could be associated with exocrine subclinical pancreatitis [15] which often accompanies the disease, and not directly with the islet pathology.

Our findings are novel and indicate that although it is unlikely that HHV-6 is a primary cause for T1D and is not associated with MHC class I expression or CD8 infiltration, yet the virus is more frequently present in diabetic pancreata, which is similar to what has been reported for enterovirus infections [1,16].

The conclusion that one can observe increased frequency of infection by a virus like HHV-6 in type 1 diabetes pancreata without a direct association with islet pathology is novel and points towards a more indirect mechanism of potential causality. Maybe, pancreata of type 1 patients are in general more susceptible to infections, possibly due to alterations in the interferon pathway [17], which could result in more general deleterious effects of many different viruses during type 1 pathogenesis.

Our findings should prompt further studies to understand why certain viral infections occur more frequently in the diabetic pancreas and how they might trigger and/or aid in the development of the disease.

CRediT authorship contribution statement

Somayeh Sabouri: Data curation, Formal analysis, Writing - original draft. **Mehdi A. Benkahla:** Data curation, Formal analysis, Writing - review & editing. **William B. Kiosses:** Formal analysis, Software. **Teresa Rodriguez-Calvo:** Data curation. **Jose Zapardiel-Gonzalo:** Formal analysis. **Ericka Castillo:** Data curation. **Matthias G. von Herrath:** Funding acquisition, Supervision, Formal analysis, Writing - review & editing.

Declaration of Competing interests

The authors declare no competing financial interests.

Acknowledgements

The authors would like to thank Zbigniew Mikulski and Yasaman Lajevardiat from the La Jolla Institute for Immunology for help with the image acquisition and analysis and technical assistance, respectively. The following reagent was obtained through the NIH AIDS Reagent Program, Division of AIDS, NIAID, NIH: HSB2 HHV6A (GS) from Dr. Dharam Ablashi. We would like to acknowledge Dr. Kristin Loomis of the HHV-6 Foundation for her helpful comments.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jaut.2019.102378.

Funding

This research was performed with the support of nPOD, a collaborative type 1 diabetes research project sponsored by the Juvenile Diabetes Research Foundation International. Organ Procurement Organizations, partnering with nPOD to provide research resources, are listed at www.jdrfnpod.org/our-partners.php. This study was also supported by the National Institutes of Health/National Institute of Allergy and Infectious Diseases grant #U01AI102370-08.

S.S. was supported by the NIDDK of the NIH under Award number T32DK007494. M.A.B. was supported by a SPARK award from La Jolla Institute for Immunology.

References

[1] T. Rodriguez-Calvo, S. Sabouri, F. Anquetil, M.G. von Herrath, The viral paradigm in type 1 diabetes: who are the main suspects? Autoimmun. Rev. 15 (2016) 964–969, https://doi.org/10.1016/j.autrev.2016.07.019.

[2] H. Agut, P. Bonnafous, A. Gautheret-Dejean, Update on infections with human herpesviruses 6A, 6B, and 7, Med. Maladies Infect. 47 (2017) 83–91, https://doi.org/10.1016/j.medmal.2016.09.004.

[3] K. Yamanishi, T. Okuno, K. Shiraki, M. Takahashi, T. Kondo, Y. Asano, T. Kurata, Identification of human herpesvirus-6 as a causal agent for exanthem subitum, Lancet 1 (1988) 1065–1067, https://doi.org/10.1016/s0140-6736(88)91893-4.

[4] Y. Tanaka, T. Suenaga, M. Matsumoto, T. Seya, H. Arase, Herpesvirus 6 glycoproteins B (gB), gH, gL, and gQ are necessary and sufficient for cell-to-cell fusion, J. Virol. 87 (2013) 10900–10903, https://doi.org/10.1128/JVI.01427-13.

[5] A. Imagawa, T. Hanafusa, Fulminant type 1 diabetes–an important subtype in East Asia, Diabetes Metab. Res. Rev. 27 (2011) 959–964, https://doi.org/10.1002/dmrr.1236.

[6] S. Yoneda, A. Imagawa, K. Fukui, S. Uno, J. Kozawa, M. Sakai, T. Yumioka, H. Iwahashi, I. Shimomura, A histological study of fulminant type 1 diabetes mellitus related to human cytomegalovirus reactivation, J. Clin. Endocrinol. Metab. 102 (2017) 2394–2400, https://doi.org/10.1210/jc.2016-4029.

[7] M. Ericsson, O. Skog, Presence of human herpesvirus 6B in the pancreas of subjects with and without type 1 diabetes, Pancreas 46 (2017) 1341–1346, https://doi.org/10.1097/MPA.0000000000000927.

[8] D.V. Ablashi, N. Balachandran, S.F. Josephs, C.L. Hung, G.R. Krueger, B. Kramarsky, S.Z. Salahuddin, R.C. Gallo, Genomic polymorphism, growth properties, and immunologic variations in human herpesvirus-6 isolates, Virology 184 (1991) 545–552.

[9] S.J. Richardson, T. Rodriguez-Calvo, I.C. Gerling, C.E. Mathews, J.S. Kaddis, M.A. Russell, M. Zeissler, P. Leete, L. Krogvold, K. Dahl-Jørgensen, M. von Herrath,

A. Pugliese, M.A. Atkinson, N.G. Morgan, Islet cell hyperexpression of HLA class I antigens: a defining feature in type 1 diabetes, Diabetologia 59 (2016) 2448–2458, https://doi.org/10.1007/s00125-016-4067-4.

[10] M.G. von Herrath, R.S. Fujinami, J.L. Whitton, Microorganisms and autoimmunity: making the barren field fertile? Nat. Rev. Microbiol. 1 (2003) 151–157, https://doi.org/10.1038/nrmicro754.

[11] H.E. Thomas, J.L. Parker, R.D. Schreiber, T.W. Kay, IFN-gamma action on pancreatic beta cells causes class I MHC upregulation but not diabetes, J. Clin. Investig. 102 (1998) 1249–1257, https://doi.org/10.1172/JCI2899.

[12] A.K. Foulis, M.A. Farquharson, R. Hardman, Aberrant expression of class II major histocompatibility complex molecules by B cells and hyperexpression of class I major histocompatibility complex molecules by insulin containing islets in type 1 (insulin-dependent) diabetes mellitus, Diabetologia 30 (1987) 333–343.

[13] K.T. Coppieters, F. Dotta, N. Amirian, P.D. Campbell, T.W.H. Kay, M.A. Atkinson, B.O. Roep, M.G. von Herrath, Demonstration of islet-autoreactive CD8 T cells in insulitic lesions from recent onset and long-term type 1 diabetes patients, J. Exp. Med. 209 (2012) 51–60, https://doi.org/10.1084/jem.20111187.

[14] M. Ota, S. Serada, T. Naka, Y. Mori, MHC class I molecules are incorporated into human herpesvirus-6 viral particles and released into the extracellular environment, Microbiol. Immunol. 58 (2014) 119–125, https://doi.org/10.1111/1348-0421.12121.

[15] I. Capua, A. Mercalli, M.S. Pizzuto, A. Romero-Tejeda, S. Kasloff, C. De Battisti, F. Bonfante, L.V. Patrono, E. Vicenzi, V. Zappulli, V. Lampasona, A. Stefani, C. Doglioni, C. Terregino, G. Cattoli, L. Piemonti, Influenza A viruses grow in human pancreatic cells and cause pancreatitis and diabetes in an animal model, J. Virol. 87 (2013) 597–610, https://doi.org/10.1128/JVI.00714-12.

[16] H. Hyöty, K.W. Taylor, The role of viruses in human diabetes, Diabetologia 45 (2002) 1353–1361, https://doi.org/10.1007/s00125-002-0852-3.

[17] S.J. Richardson, N.G. Morgan, Enteroviral infections in the pathogenesis of type 1 diabetes: new insights for therapeutic intervention, Curr. Opin. Pharmacol. 43 (2018) 11–19, https://doi.org/10.1016/j.coph.2018.07.006.

Journal of Autoimmunity 119 (2021) 102628

Contents lists available at ScienceDirect

Journal of Autoimmunity

journal homepage: www.elsevier.com/locate/jautimm

HLA class I hyper-expression unmasks beta cells but not alpha cells to the immune system in pre-diabetes

Mehdi A. Benkahla [a], Somayeh Sabouri [a,1], William B. Kiosses [a], Sakthi Rajendran [a], Estefania Quesada-Masachs [a], Matthias G. von Herrath [a,b,*]

[a] La Jolla Institute for Immunology, 9420 Athena Circle, La Jolla, 92037, CA, USA
[b] Global Chief Medical Officer, Novo Nordisk, Denmark

ARTICLE INFO

Keywords:
HLA class I
Type 1 diabetes
Beta cells
Alpha cells
nPOD

ABSTRACT

Human leukocyte antigens of class-I (HLA-I) molecules are hyper-expressed in insulin-containing islets (ICI) of type 1 diabetic (T1D) donors. This study investigated the HLA-I expression in autoantibody positive (AAB+) donors and defined its intra-islet and intracellular localization as well as proximity to infiltrating CD8 T cells with high-resolution confocal microscopy. We found HLA-I hyper-expression had already occurred prior to clinical diagnosis of T1D in islets of AAB+ donors. Interestingly, throughout all stages of disease, HLA-I was mostly expressed by alpha cells. Hyper-expression in AAB+ and T1D donors was associated with intra-cellular accumulation in the Golgi. Proximity analysis showed a moderate but significant correlation between HLA-I and infiltrating CD8 T cells only in ICI of T1D donors, but not in AAB+ donors. These observations not only demonstrate a very early, islet-intrinsic immune-independent increase of HLA-I during diabetes pathogenesis, but also point towards a role for alpha cells in T1D.

1. Introduction

Type 1 diabetes (T1D) is an autoimmune disease characterized by increased blood glucose levels likely due to the destruction of insulin-producing beta cells resulting in lifelong dependence on exogenous insulin. This immune attack is mediated by CD8 T cells recognizing beta cell peptides presented by human leukocyte antigen of class I (HLA-I) [1].

Islet proteins are cleaved into peptides in the proteasome. Peptides are then transported from the cytosol into the endoplasmic reticulum (ER) by the transporter associated with antigen processing (TAP). In the ER, peptides are loaded into HLA-I molecules. The newly formed peptide-HLA-I complex is then discharged in vesicles that migrate to the cell surface through the Golgi apparatus. At the surface, these peptide-HLA-I complexes may be recognized by CD8 T cells [2].

T1D is associated with the presence of autoantibodies prior to disease onset. The expansion of autoreactive B cells results in autoantibodies targeting insulin (IAA), glutamate decarboxylase 2 (GAD65), insulinoma-associated protein 2 (IA-2) and zinc transporter 8 (ZnT8) [3]. Ziegler et al. showed that children with multiple autoantibodies have greater risk of developing the disease than children with single autoantibody [4].

Previous studies [5–9] have shown that HLA-I is hyper-expressed in the insulin-containing islets (ICI) of T1D donors. However, it is not known whether HLA-I hyper-expression is the cause or consequence of T1D pathogenesis. To the best of our knowledge, quantification of HLA-I hyper-expression has not been reported in the islets of AAB+ donors. We therefore characterized the expression and intracellular localization of HLA-I in single and double AAB+ donors as well as in non-diabetic and T1D donors.

Our new findings here show that HLA-I is hyper-expressed intracellularly and accumulates significantly in the Golgi compartments in the AAB+ and T1D donors. Also, HLA-I is mostly expressed on alpha cells, irrespective of disease status. Moreover, a moderate positive correlation between HLA-I expression and the number of CD8 T cells was only observed in the ICIs of T1D donors, but not prior to clinical diabetes in AAB+ donors. These results point towards an early, intrinsic defect of beta and alpha cells in T1D that ultimately might lead to immune recognition of autoantigenic peptides expressed by heightened HLA-I on beta cells.

* Corresponding author. Global Chief Medical Officer, Novo Nordisk, Denmark.
E-mail address: Matthias@lji.org (M.G. von Herrath).
[1] Current address: Sanofi, Translational Medicine & Early development, 1 Mountain Road, Farmingham, MA, 01701, USA.

https://doi.org/10.1016/j.jaut.2021.102628
Received 10 December 2020; Received in revised form 17 February 2021; Accepted 25 February 2021
Available online 9 March 2021

2. Materials and methods

2.1. Subjects

Human pancreata collected from cadaveric organ donors were obtained through Network for Pancreatic Organ Donors (nPOD). Six μm thick sections from formalin-fixed paraffin embedded (FFPE) tissue samples from non-diabetic (ND, n = 10), autoantibody positive donors (AAB+, n = 10) and donors with T1D (n = 12) were stained. Donor information is summarized in supplementary Table 1. The experimental procedures of this study were performed in accordance with relevant guidelines and regulations. La Jolla Institute for Immunology Institutional Review Board approved all experimental procedures (protocol number DI3-054-1112). Informed consent from donor families was obtained by Organ Recovery Partners at the nPOD (https://www.jdrfnpod.org/for-partners/organ-recovery-partners/).

2.2. Immunofluorescence staining

2.2.1. Surface staining

To examine HLA-I (A, B, C) and CD8 T cells expression, pancreatic sections were subject to an immunofluorescence staining as previously described [10]. Briefly, after deparaffinization and rehydration in descending ethanol concentrations, sections were exposed to heat-based citrate antigen retrieval for 20 min. Slides were then incubated with monoclonal mouse anti-human-HLA class I ABC (Abcam, ab70328; 1:200) for 1 h at room temperature (RT) followed by detection using polyclonal goat anti-mouse IgG-Alexa Fluor 647 (Life Technologies, 1:1000) for 30 min at RT. Polyclonal rabbit anti-CD8 antibody (ab4055, Abcam, Cambridge MA; 1:400) was incubated overnight at 4 °C, followed by goat anti-rabbit IgG-Alexa Fluor 555 (Life Technologies, 1:1000) for 30 min at RT. Staining for insulin (polyclonal guinea pig anti-insulin. Dako, Carpinteria CA; 1:600) and glucagon (polyclonal rabbit anti-glucagon. Dako, Carpinteria CA; 1:2000) was performed at room temperature for 1 h. Detection was performed at room temperature for 30 min using polyclonal goat anti-guinea pig IgG Alexa Fluor 488 (Life Technologies; 1:1000) and polyclonal goat anti-rabbit IgG Alexa Fluor 555 (Life Technologies; 1:1000), respectively.

2.2.2. Intracellular staining

GM130 was stained using the same protocol with slight modifications. Sections were blocked/permeabilized with 10% goat serum in 0.3% Triton-X-100 PBS for 1 h at RT. Monoclonal rabbit anti-GM130 (EP892Y) antibody, (Abcam ab52649; 1:50) was incubated overnight at 4 °C, followed by an Alexa Fluor 647 conjugated goat anti-rabbit IgG antibody (Life Technologies; 1:1000) incubated at RT for 1 h in the dark.

2.3. Image acquisition and analysis

All 3D images were acquired with a Zeiss laser scanning confocal microscope (LSCM)780 using a 63x (1.4na) objective. All Image stacks (on average 12 slices) were acquired with Nyquist resolution parameters using a 0.3 μm step size and optimal frame size of 2644 × 2644. All 8-bit images were acquired using the full dynamic intensity range (0–256) that was determined with the population of cells having the brightest signals. Samples stained for secondary antibodies and isotypic controls alone were used to define thresholds of real signal above background and non-specific signal. Based on these positive and negative controls, we developed an acquisition strategy that used the same laser power and detector signal amplification (digital gain) settings across all donors. Images were further processed using either Zen Pro (Zeiss) for colocalization or Image Pro Premier (Media Cybernetics) for CD8 T cells proximity analysis. Briefly, maximum intensity projections (MIPs) were generated in Zen and then processed using the Zen colocalization module. Here, regions of interest (ROIs) were drawn around each individual islet based on insulin and glucagon staining. With the inputs of the previously defined minimum thresholds of two fluorescent channels at a time, the software automatically calculated pixel intensity spatial overlap coefficients between them using Mander's coefficient. Mander's overlap coefficient is described by MOC = $\sum$i (Ri x Gi)$\sum$iRi2 x $\sum$iGi2 where Ri is the intensity of the first (red) fluorophore in an individual pixel, whereas Gi is the corresponding intensity for the second (green) fluorophore in the same pixel [11]. Channel signal comparisons included HLA-I vs insulin, HLA-I vs glucagon and HLA-I vs GM130.

MIPs for CD8 T cell distance localization analysis were created in Image J and further processed with Image Pro Premier. In Image Pro Premier, fluorescently labeled CD8 T cells were automatically detected and outlined as ROIs. These original ROIs were then used to generate expanded concentric radial circles of the original outlined signal. Four x 100 pixel expanded concentric circles were created as shown in Fig. 4b (defined as 100, 200, 300 and 400). These outlined radial circles were then overlaid onto the HLA-I signal images to extract the location, total area and signal intensity of HLA-I that resides within each annulus of concentric circles (See Fig. 4a).

2.4. Statistical analysis

Statistical analysis was performed using GraphPad Prism version 8 (GraphPad software, San Diego, CA). Kruskal Wallis was used to calculate p values and Dunn's multiple comparisons test used as post-hoc test and considered significant if $p < 0.05$. Correlations were evaluated and considered significant if $p < 0.05$ and a Pearson's rank correlation coefficient (r) > 0.70.

3. Results

3.1. HLA class I is hyperexpressed in the islets of autoantibody positive donors

Pancreas sections from 32 organ donors and a total of 802 islets were examined for the presence of HLA-I by indirect immunofluorescence. Consistent with previous studies, HLA-I was hyper-expressed in the islets of T1D donors and specifically in the insulin-containing islets (ICI) (Fig. 1a). Interestingly, HLA-I was already hyper-expressed in the islets of AAB+ donors (Fig. 1a). ICIs from T1D donors (20.05 ± 12.19%) showed the highest proportion of HLA-I positive area followed by AAB+ donors (12.63 ± 9.34%), IDIs from T1D donors (12.29 ± 6.29%) and ND donors (9.69 ± 5.55%) (Fig. 1b).

The positive area for HLA-I in the ICIs from T1D donors ranged from 1.77 to 54.59% with 70 islets (out of 142 islets) above the 20% threshold. The positive area for HLA-I in the islets of AAB + donors ranged from 0.6 to 42.66% with 52 islets (out of 223 islets) above the 20% threshold. While the positive area for HLA-I for ND donors and IDIs from T1D donors ranged from 0.62 to 26.01% and 0.84–38.41% respectively, only 10 islets and 19 islets were above the threshold for the ND donors and the IDIs from the T1D donors (Fig. 1b).

HLA-I was hyper-expressed in 52 out of 224 islets (23%) in AAB+ donors (Fig. 1b). A heterogenous expression of HLA-I was observed inter and intra-groups as well as within the same pancreas, where some islets were hyper-expressing HLA-I and others exhibiting a normal HLA-I expression (Fig. 1c). HLA-I hyper-expression was observed not only in recent onset of the disease but also in the islets of donors with T1D for 20 years (nPOD cases 6038, 6040 and 6195) (Fig. 1c).

3.2. Majority of HLA class I is expressed on alpha cells

We analyzed the distribution of HLA-I in the islets of ND, AAB+ and T1D donors. The same sections previously analyzed for the quantification of HLA-I were analyzed using Mander's overlap analysis, to quantify the degree of colocalization of HLA-I with insulin and HLA-I with glucagon. Expression of HLA-I was significantly higher in the alpha cells irrespective of disease status (Fig. 2a). In IDIs, alpha cells contributed to

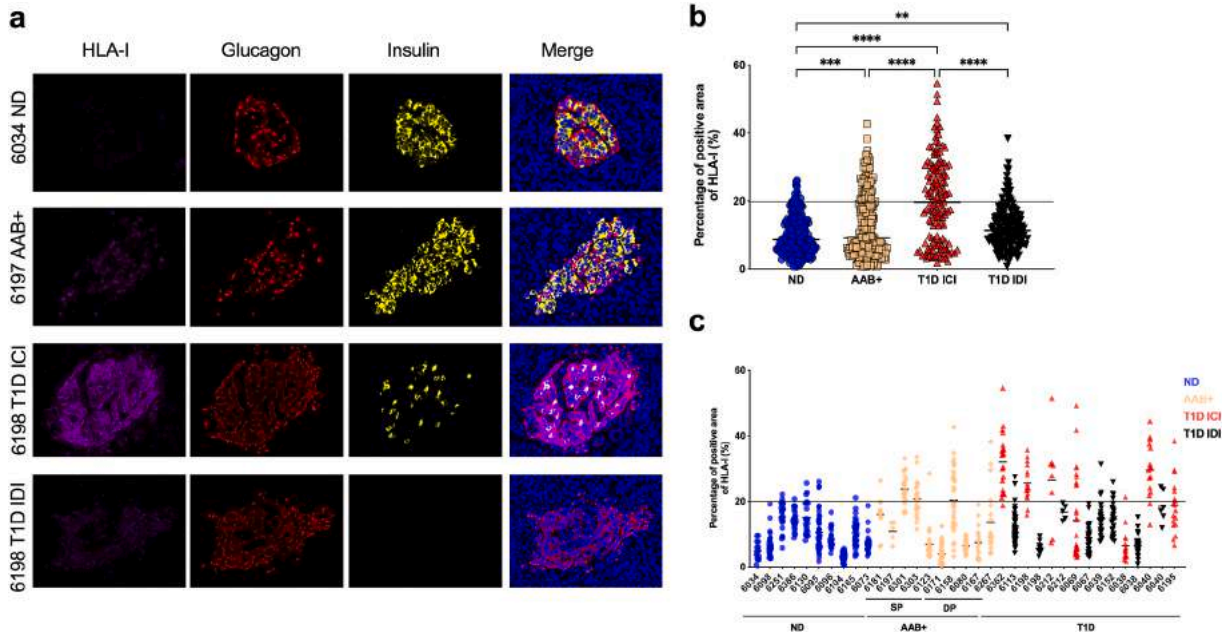


Fig. 1. HLA class I is hyper-expressed in the islets of T1D as well as in AAB+ donors. FFPE sections of human pancreata were stained with anti-HLA-I (magenta), anti-insulin (green) and anti-glucagon (red) and Hoechst (blue). **(a)** Representative images of an islet from a non-diabetic (ND) donor (nPOD 6034), an autoantibody positive (AAB+) donor (nPOD 6197), an insulin-containing islet (ICI) of a T1D donor (nPOD 6198) and an insulin-deficient islet (IDI) of a T1D donor (nPOD 6198) are shown. Images were acquired using a confocal microscope LSM 780 with a 63 × 1.4na objective. HLA-I was quantified using Zen software. **(b)** The percentage of total islet area positive for HLA-I was measured in at least 20 islets per section from non-diabetic donors (n = 10), single autoantibody positive (n = 5), double autoantibody positive (n = 5) and type 1 diabetic donors (n = 12). **(c)** Percentage of total islet area positive for HLA-I presented by case. Each data point represents an islet. An arbitrary threshold of 20% was set above which HLA-I was considered hyper-expressed. ***p = 0.0004 ****p < 0.0001.

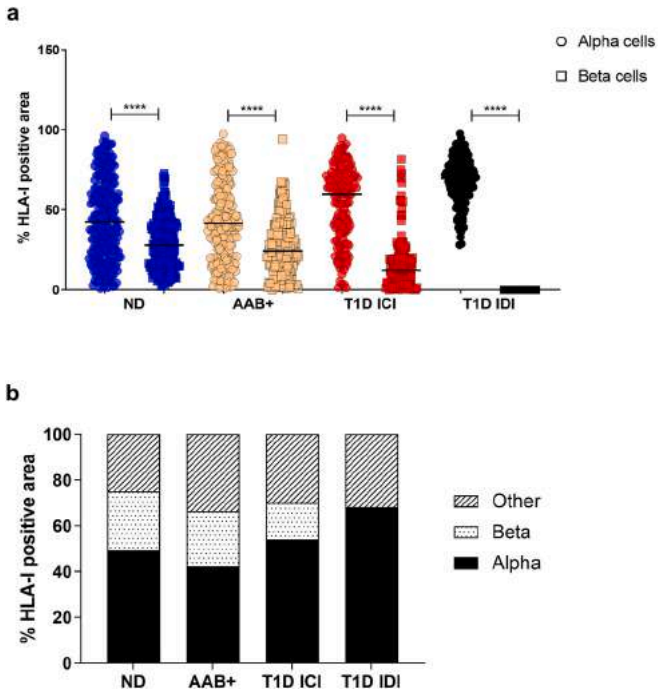


Fig. 2. HLA class I is mainly expressed on alpha cells irrespective of disease status. **(a)** Percentage of colocalization of HLA-I and glucagon (alpha cells-circle) and HLA-I and insulin (beta cells - square) quantified using the Mander's overlap coefficient using Zen software. Each data point represents an islet. **(b)** Mean percentage of HLA-I positive area per cell type. ND: non-diabetic; AAB+: autoantibody positive; T1D: type 1 diabetes; ICI: insulin-containing islet; IDI: insulin-deficient islet. ****p < 0.0001.

the highest level of HLA-I expression (68.56 ± 14.30%), which is expected due to the lack of beta cells. Alpha cells in ND, AAB+ and ICIs of T1D donors had a level of distribution of HLA-I of 43.89 ± 25.45%, 43.35 ± 25.44% and 53.51 ± 23.80%, respectively.

Beta cells in the islets represented the smallest fraction of cells expressing HLA-I in AAB+ and ICIs of T1D donors with a mean percentage of HLA-I expression of 26.79 ± 17.60% and 16.23 ± 16.51%, respectively (Fig. 2b).

3.3. HLA class I is localized in the golgi compartment in AAB+ and T1D donors

Next, the intracellular localization of HLA-I was examined. Pancreas section from 1 ND donor, 1 single AAB+ donor, 1 double AAB+ donor and 1 T1D donor were stained for GM130, a peripheral cytoplasmic protein tightly bound to the Golgi membrane, HLA-I, insulin and glucagon. The Mander's overlap colocalization analysis showed that GM130 and HLA-I were overlapping in the AAB+ and T1D donors with a percentage of colocalization of 7.10 ± 6.56% in nPOD 6181 AAB+, 5.96 ± 5.29% in nPOD 6197 AAB+ and 10.95 ± 11.76% in nPOD 6212 T1D. The non-diabetic donor nPOD 6034 showed the least percentage of colocalization (0.88 ± 0.62%, Fig. 3).

3.4. HLA class I is hyper-expressed in the proximity of CD8 T cells in T1D but not autoantibody positive donors

CD8 T cells are one of the main perpetrators of the destruction of beta cells through HLA-I [12]. Hence, we examined if the presence of infiltrating CD8 T cells could affect the expression of HLA-I.

Pancreatic sections from 2 ND donors, 1 single AAB+ donor, 1 double AAB+ donor and 2 T1D donors were stained with HLA-I, CD8 and insulin, to examine the expression of HLA-I in the proximity of CD8 T cells. HLA-I expression was invariable between the radial circles surrounding the CD8 T cells irrespective of disease status, but it was higher in the AAB+ and ICIs of T1D donors (Fig. 4a). When examining the

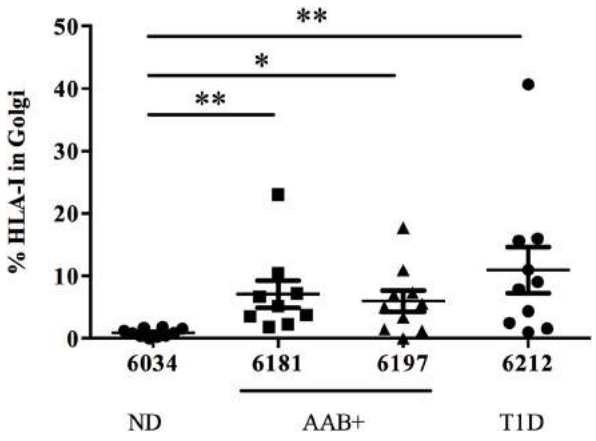


Fig. 3. HLA class I colocalizes with Golgi marker GM130 in the islets of AAB+ and T1D donors. Percentage of colocalization of HLA-I and GM130 quantified using Mander's overlap coefficient in the islets from non-diabetic donor (nPOD 6034), autoantibody positive donors (nPOD 6181 and 6197) and a type 1 diabetic donor (nPOD 6212). Colocalization analysis was performed in Zen software. ND: Non-diabetic; AAB+: autoantibody positive; T1D: type 1 diabetes. *$p < 0.05$ **$p < 0.01$.

whole HLA-I positive area surrounding CD8 T cells (the sum of radial circles), it was significantly higher in the T1D ICIs (Fig. 4b and c) compared to the other groups. Data was spread out over a wide range of values with a mean level of HLA-I area of 1597 ± 1469 μm^2 (Fig. 4c). The levels of HLA-I in the proximity of CD8 T cells ranged from 0 to 4599.7 μm^2 with a mean of 817.8 ± 991.9 μm^2 and it was not significantly different between ND donors and IDIs from T1D donors with 631.1 ± 487.7 μm^2 and 412.5 ± 550.5 μm^2 respectively (Fig. 4c). Then, we performed a correlation analysis between the number of CD8 T cells and the expression level of HLA-I surrounding those cells. We found a moderate positive correlation between HLA-I and the number of CD8 T cells in the ICIs of T1D with a rho of 0.64 and a p value of 0.016 (Fig. 4d). However, the correlation between HLA-I and the number of CD8 T cells in AAB+ donors was weak with a rho of 0.35 and did not achieve statistical significance with a p value of 0.051.

4. Discussion

In agreement with previous studies, HLA-I hyper-expression was observed in insulin-containing islets of T1D donors, confirming that HLA-I hyper-expression is an important pathological feature of type 1 diabetes [13]. Skog et al. have reported an intense staining of HLA-I in the islets of AAB+ donors, however, its expression was not characterized nor quantified [8].

The current study is different and extends these observations in many aspects: In addition to HLA-I quantification in the islets of ND, AAB+ and T1D donors, its intracellular localization and its possible association with CD8 T cell infiltration have been investigated. Using high-resolution confocal microscopy, we observed a hyper-expression of HLA-I in the insulin-containing islets of T1D donors as well as in the islets of AAB+ donors before the clinical disease onset. In normal conditions, HLA-I molecules are expressed on all cells and are regulated by transcription factors and its expression can be modulated by different stimuli, such as interferon gamma induced in response to viral infections. It has been reported that there was no correlation between HLA-I hyper-expression and enteroviral VP1 [9,14] and human herpesvirus 6 gB protein in pancreatic islets of T1D donors [10]. Thus, HLA-I upregulation seems to be pointing towards an intrinsic defect with beta and alpha cells preceding and maybe causing T1D.

An interesting new finding in the present study is that the majority of HLA-I is expressed on alpha cells irrespective of disease status. Although expressing higher levels of HLA-I, alpha cells are spared by the immune system compared to beta cells, which are systematically targeted and destroyed. Alpha cells also express other T1D autoantigens such as ZnT8,

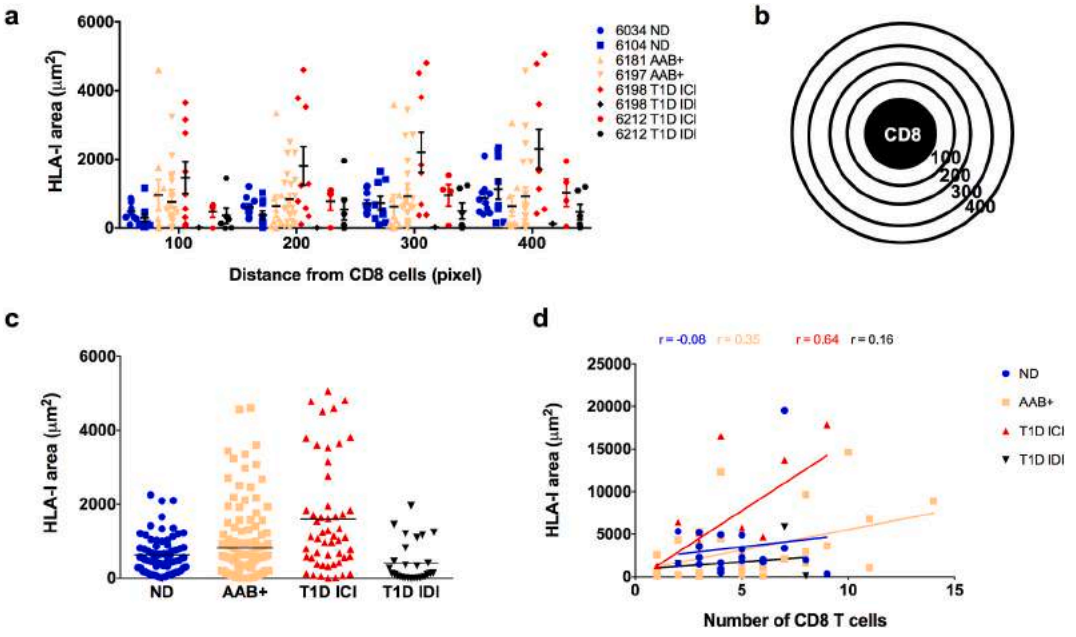


Fig. 4. Proximity analysis of HLA class I in the surroundings of CD8 T cells. Pancreatic sections from organ donors were stained for HLA-I, CD8, insulin and glucagon. CD8 T cells were detected in the pancreatic tissue sections and a proximity analysis was performed to detect the HLA-I expression in the surroundings of CD8 T cells. (a) Quantification of HLA-I using ImagePro premier in the radials surrounding the CD8 T cells, 100 being the closest to the CD8 cell and 400 the furthest presented by case. Each data point represents the level of HLA-I in a radial circle. (b) Schematic representation of CD8 T cell and the radial circles from 100 to 400 used for the proximity analysis. (c) Quantification of HLA-I presented as the sum of the radials from 100 to 400 and presented by group. Each data point represents an islet. Each data point represents the level of HLA-I in a radial circle (d) Correlation between the number of CD8 T cells per field of view and the total expression of HLA-I in the surroundings of those CD8 T cells. Each data point represents a field of view. ND: non-diabetic; AAB+: autoantibody positive; T1D: type 1 diabetes; ICI: insulin-containing islet; IDI: insulin-deficient islet.

IA-2 and Chromogranin A, however insulin is exclusively expressed by beta cells, implying that the immune attack of beta cells is related to the presence or presentation of insulin as such [15]. This might explain why the beta cells are more susceptible to the immune-mediated destruction. Moreover, Brissova et al. have shown that alpha cells from human pancreatic islets obtained from recent-onset T1D donors were functionally impaired and had altered gene expression, which could affect antigen presentation [16].

Furthermore, we have found that HLA-I accumulates in the Golgi apparatus of the islets of AAB+ and T1D donors. This might be explained by HLA-I being hyper-expressed in these donors, resulting in increased intracellular detection. Another possibility could be that ER stress caused by viral infections, environmental stresses or NO [17] could lead to accumulation of unfolded or misfolded proteins in the lumen of the ER [18]. The latter is further supported by RNA-sequencing analysis performed on alpha cells from T1D donors [16], where an increased expression of genes involved in the unfolded protein response (UPR) was observed. UPR is activated in response to the accumulation of unfolded or misfolded proteins in the ER.

Beta cells are targeted by the CD8 T cell-mediated immune destruction through recognition of autoantigenic peptides associated with HLA-I. HLA-I expression in beta cells is not required for the initiation of insulitis, however it is essential for the destruction of beta cells and incidence of diabetes [19]. In order to investigate the possible impact of CD8 T cells on HLA-I expression, we performed a proximity analysis of HLA-I in the surroundings of CD8 T cells. HLA-I expression was higher in the surroundings of CD8 T cells in AAB+ and T1D donors, but had normal levels in the ND and IDIs from T1D donors. The limitation of this study is that CD8 T cells are very dynamic, and move around very rapidly as seen in the mouse pancreas by 2-photon imaging [20]. Hence, the proximity analysis may not reflect a true status of CD8 T cell accumulation, although, we have seen a moderate positive correlation between HLA-I expression and the numbers of CD8 T cells in the insulin-containing islets from T1D donors.

In conclusion, our findings have shown that HLA-I is hyper-expressed and accumulates in the Golgi compartment of the islets of AAB+ and T1D donors. Despite being spared by the immune system, alpha cells are the main cell types expressing HLA-I irrespective of disease status. Also, HLA-I is hyper-expressed in the proximity of CD8 T cells in T1D but not AAB+ donors. Taken altogether, our observations points towards an early intrinsic defect in beta and alpha cells that precedes and possibly precipitates immune cell infiltration.

Declaration of competing interest

The authors declare no competing interest.

Acknowledgments

We thank the microscopy core at the La Jolla Institute for Immunology.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jaut.2021.102628.

Funding

This research was performed with the support of nPOD, a collaborative type 1 diabetes research project sponsored by the Juvenile Diabetes Research Foundation International. Organ Procurement Organizations, partnering with nPOD to provide research resources, are listed at www.jdrfnpod.org/our-partners.php. This study was also supported by the National Institutes of Health/National Institute of Allergy and Infectious Diseases grant #U01AI102370-08. M.A.B. was supported by the Tullie and Rickey SPARK program for innovation in Immunology award from La Jolla Institute for Immunology. S.S. was supported by the NIDDK of the NIH under Award number T32DK007494. EQM was supported by a postdoctoral grant of Fundación Alfonso Martín Escudero.

Author contributions

Mehdi. A. Benkahla: Data curation, Formal analysis, Writing – original draft, Writing – review & editing. Somayeh. Sabouri: Data curation, Writing – review & editing. William. B. Kiosses: Data curation, Formal analysis, Software, Writing – review & editing. Sakthi. Rajendran: Formal analysis, Writing – review & editing. Estefania. Quesada-Masachs: Formal analysis, Writing – review & editing. Matthias. G. von Herrath: Funding acquisition, Supervision, Formal analysis, Writing – review & editing.

Data and materials availability

Data are available upon reasonable request.

References

[1] A. Katsarou, S. Gudbjörnsdottir, A. Rawshani, D. Dabelea, E. Bonifacio, B. J. Anderson, L.M. Jacobsen, D.A. Schatz, Å. Lernmark, Type 1 diabetes mellitus, Nat Rev Dis Primers 3 (2017) 17016.
[2] P. Leone, E.-C. Shin, F. Perosa, A. Vacca, F. Dammacco, V. Racanelli, MHC class I antigen processing and presenting machinery: organization, function, and defects in tumor cells, J. Natl. Cancer Inst. 105 (2013) 1172–1187.
[3] J.P. Krischer, K.F. Lynch, D.A. Schatz, J. Ilonen, Å. Lernmark, W.A. Hagopian, M. J. Rewers, J.-X. She, O.G. Simell, J. Toppari, A.-G. Ziegler, B. Akolkar, E. Bonifacio, TEDDY Study Group, The 6 year incidence of diabetes-associated autoantibodies in genetically at-risk children: the TEDDY study, Diabetologia 58 (2015) 980–987.
[4] A.G. Ziegler, M. Rewers, O. Simell, T. Simell, J. Lempainen, A. Steck, C. Winkler, J. Ilonen, R. Veijola, M. Knip, E. Bonifacio, G.S. Eisenbarth, Seroconversion to multiple islet autoantibodies and risk of progression to diabetes in children, J. Am. Med. Assoc. 309 (2013) 2473–2479.
[5] G.F. Bottazzo, B.M. Dean, J.M. McNally, E.H. MacKay, P.G. Swift, D.R. Gamble, In situ characterization of autoimmune phenomena and expression of HLA molecules in the pancreas in diabetic insulitis, N. Engl. J. Med. 313 (1985) 353–360.
[6] A.K. Foulis, M.A. Farquharson, R. Hardman, Aberrant expression of class II major histocompatibility complex molecules by B cells and hyperexpression of class I major histocompatibility complex molecules by insulin containing islets in type 1 (insulin-dependent) diabetes mellitus, Diabetologia 30 (1987) 333–343.
[7] R. Pujol-Borrell, I. Todd, M. Doshi, D. Gray, M. Feldmann, G.F. Bottazzo, Differential expression and regulation of MHC products in the endocrine and exocrine cells of the human pancreas, Clin. Exp. Immunol. 65 (1986) 128–139.
[8] O. Skog, S. Korsgren, A. Wiberg, A. Danielsson, B. Edwin, T. Buanes, L. Krogvold, O. Korsgren, K. Dahl-Jørgensen, Expression of human leukocyte antigen class I in endocrine and exocrine pancreatic tissue at onset of type 1 diabetes, Am. J. Pathol. 185 (2015) 129–138.
[9] S.J. Richardson, T. Rodriguez-Calvo, I.C. Gerling, C.E. Mathews, J.S. Kaddis, M. A. Russell, M. Zeissler, P. Leete, L. Krogvold, K. Dahl-Jørgensen, M. von Herrath, A. Pugliese, M.A. Atkinson, N.G. Morgan, Islet cell hyperexpression of HLA class I antigens: a defining feature in type 1 diabetes, Diabetologia 59 (2016) 2448–2458.
[10] S. Sabouri, M.A. Benkahla, W.B. Kiosses, T. Rodriguez-Calvo, J. Zapardiel-Gonzalo, E. Castillo, M.G. von Herrath, Human herpesvirus-6 is present at higher levels in the pancreatic tissues of donors with type 1 diabetes, J. Autoimmun. 107 (2020) 102378.
[11] J. Adler, I. Parmryd, Quantifying colocalization by correlation: the Pearson correlation coefficient is superior to the Mander's overlap coefficient, Cytometry 77A (2010) 733–742.
[12] I. Gomez-Tourino, S. Arif, M. Eichmann, M. Peakman, T cells in type 1 diabetes: instructors, regulators and effectors: a comprehensive review, J. Autoimmun. 66 (2016) 7–16.
[13] S.J. Richardson, A. Willcox, A.J. Bone, A.K. Foulis, N.G. Morgan, The prevalence of enteroviral capsid protein vp1 immunostaining in pancreatic islets in human type 1 diabetes, Diabetologia 52 (2009) 1143–1151.
[14] L. Krogvold, B. Edwin, T. Buanes, G. Frisk, O. Skog, M. Anagandula, O. Korsgren, D. Undlien, M.C. Eike, S.J. Richardson, P. Leete, N.G. Morgan, S. Oikarinen, M. Oikarinen, J.E. Laiho, H. Hyöty, J. Ludvigsson, K.F. Hanssen, K. Dahl-Jørgensen, Detection of a low-grade enteroviral infection in the islets of langerhans of living patients newly diagnosed with type 1 diabetes, Diabetes 64 (2015) 1682–1687.
[15] M. Nakayama, Insulin as a key autoantigen in the development of type 1 diabetes, Diabetes Metab Res Rev 27 (2011) 773–777.
[16] M. Brissova, R. Haliyur, D. Saunders, S. Shrestha, C. Dai, D.M. Blodgett, R. Bottino, M. Campbell-Thompson, R. Aramandla, G. Poffenberger, J. Lindner, F.C. Pan, M. G. von Herrath, D.L. Greiner, L.D. Shultz, M. Sanyoura, L.H. Philipson, M. Atkinson, D.M. Harlan, S.E. Levy, N. Prasad, R. Stein, A.C. Powers, α cell

function and gene expression are compromised in type 1 diabetes, Cell Rep. 22 (2018) 2667–2676.
[17] H.U. Kataoka, H. Noguchi, ER stress and β-cell pathogenesis of type 1 and type 2 diabetes and islet transplantation, Cell Med. 5 (2013) 53–57.
[18] C.Y. Liu, R.J. Kaufman, The unfolded protein response, J. Cell Sci. 116 (2003) 1861–1862.
[19] E.E. Hamilton-Williams, S.E. Palmer, B. Charlton, R.M. Slattery, Beta cell MHC class I is a late requirement for diabetes, Proc. Natl. Acad. Sci. Unit. States Am. 100 (2003) 6688–6693.
[20] K. Coppieters, N. Amirian, M. von Herrath, Intravital imaging of CTLs killing islet cells in diabetic mice, J. Clin. Invest. 122 (2012) 119–131.

eLIFE elifesciences.org

RESEARCH ARTICLE

Circulating T cell-monocyte complexes are markers of immune perturbations

Julie G Burel[1], Mikhail Pomaznoy[1], Cecilia S Lindestam Arlehamn[1], Daniela Weiskopf[1], Ricardo da Silva Antunes[1], Yunmin Jung[2], Mariana Babor[1], Veronique Schulten[1], Gregory Seumois[1], Jason A Greenbaum[3], Sunil Premawansa[4], Gayani Premawansa[5], Ananda Wijewickrama[6], Dhammika Vidanagama[7], Bandu Gunasena[8], Rashmi Tippalagama[9†], Aruna D deSilva[1,9†], Robert H Gilman[10,11], Mayuko Saito[12], Randy Taplitz[13], Klaus Ley[2,14], Pandurangan Vijayanand[1,15], Alessandro Sette[1,15], Bjoern Peters[1,15]*

[1]Division of Vaccine Discovery, La Jolla Institute for Immunology, La Jolla, United States; [2]Division of Inflammation Biology, La Jolla Institute for Immunology, La Jolla, United States; [3]Bioinformatics core, La Jolla Institute for Immunology, La Jolla, United States; [4]Department of Zoology and Environment Sciences, Science Faculty, University of Colombo, Colombo, Sri Lanka; [5]North Colombo Teaching Hospital, Ragama, Sri Lanka; [6]National Institute of Infectious Diseases, Gothatuwa, Sri Lanka; [7]National Tuberculosis Reference Laboratory, Welisara, Sri Lanka; [8]National Hospital for Respiratory Diseases, Welisara, Sri Lanka; [9]Genetech Research Institute, Colombo, Sri Lanka; [10]Johns Hopkins School of Public Health, Baltimore, United States; [11]Universidad Peruana Cayetano Heredia, Lima, Peru; [12]Department of Virology, Tohoku University Graduate School of Medicine, Sendai, Japan; [13]Division of Infectious Diseases and Global Public Health, University of California, San Diego, La Jolla, United States; [14]Department of Bioengineering, University of California, San Diego, La Jolla, United States; [15]Department of Medicine, University of California, San Diego, La Jolla, United States

***For correspondence:** bpeters@lji.org

Present address: †Dept of Paraclinical Sciences, Faculty of Medicine, General Sir John Kotelawala Defence University, Ratmalana, Sri Lanka

Competing interests: The authors declare that no competing interests exist.

Funding: See page 18

Received: 13 February 2019
Accepted: 31 May 2019
Published: 25 June 2019

Reviewing editor: Jos WM van der Meer, Radboud University Medical Centre, Netherlands

Abstract Our results highlight for the first time that a significant proportion of cell doublets in flow cytometry, previously believed to be the result of technical artifacts and thus ignored in data acquisition and analysis, are the result of biological interaction between immune cells. In particular, we show that cell:cell doublets pairing a T cell and a monocyte can be directly isolated from human blood, and high resolution microscopy shows polarized distribution of LFA1/ICAM1 in many doublets, suggesting in vivo formation. Intriguingly, T cell-monocyte complex frequency and phenotype fluctuate with the onset of immune perturbations such as infection or immunization, reflecting expected polarization of immune responses. Overall these data suggest that cell doublets reflecting T cell-monocyte in vivo immune interactions can be detected in human blood and that the common approach in flow cytometry to avoid studying cell:cell complexes should be re-visited.

DOI: https://doi.org/10.7554/eLife.46045.001

Introduction

Communication between immune cells is a major component of immune responses, either directly through cell-cell contacts or indirectly through the secretion of messenger molecules such as cytokines. In particular, the physical interaction between T cells and antigen-presenting cells (APCs) is critical for the initiation of immune responses. APCs such as monocytes can take up debris from the

extracellular environment, and will display fragments of it on their surface to T cells, which can identify potentially harmful, non-self antigens. There is paucity of data regarding T cell-APCs interactions in humans in vivo, but they appear to be highly diverse in terms of structure, length and function, depending on the nature and degree of maturation of the T cell and APC (**Friedl and Storim, 2004**).

Despite the importance of interactions between immune cells, many experimental techniques in immunology specifically avoid studying cell:cell complexes, in particular for the analysis of clinical samples obtained ex vivo. The most notable example for this is in flow cytometry, in which cells are labeled with a panel of fluorochrome-conjugated antibodies, and each cell is then individually hit by a laser and its corresponding fluorescence emission spectra recorded. In this process, doublets (a pair of two cells) are routinely observed but are believed to be the results of technical artifacts due to ex vivo sample manipulation and are thus usually discarded, or ignored in data analysis (**Kudernatsch et al., 2013**).

Blood is the most readily accessible sample in humans with high immune cell content. We and others have shown circulating immune cells contain critical information that can be used for diagnostic-, prognostic- and mechanistic understanding of a given disease or immune perturbation (**Bongen et al., 2018**; **Burel et al., 2018**; **Grifoni et al., 2018**; **Roy Chowdhury et al., 2018**; **Zak et al., 2016**). Thus, whereas blood does not fully reflect what is occurring in tissues, it contains relevant information from circulating immune cells that have been either directly impacted by the perturbation, or indirectly through cell contact with tissue-resident cells in the affected compartment, including lymphoid organs.

However, the presence of dual-cell complexes (and their content) has never been studied in the peripheral blood and in the context of immune perturbations. Monocytes are a subtype of phagocytes present in high abundance in the peripheral blood, which play a critical role in both innate and adaptive immunity (**Jakubzick et al., 2017**). In particular, monocytes have the capacity to differentiate into highly specialized APCs such as macrophages or myeloid DCs (**Sprangers et al., 2016**). More recently, it has been highlighted that they might directly function as APCs and thus contribute to adaptive immune responses (**Jakubzick et al., 2017**; **Randolph et al., 2008**).

We recently identified a gene signature in memory CD4+ T cells circulating in the peripheral blood that distinguishes individuals with latent TB infection (LTBI) from uninfected individuals (**Burel et al., 2018**). Surprisingly, this dataset also led to the discovery of a group of monocyte-associated genes co-expressed in memory CD4+ T cells whose expression is highly variable across individuals. We ultimately traced this signature to a population of CD3+CD14+ cells that are not single cells but T cell:monocyte complexes present in the blood and that can be detected following immune perturbations such as disease or vaccination. The frequency and T cell phenotypes of these complexes appear to be associated with the nature of pathogen or vaccine. Thus, studying these complexes promises to provide insights into the impact of immune perturbation on APCs, T cells and their interactions.

Results

Unexpected detection of monocyte gene expression in CD4+ memory T cells from human subjects

We initially set out to investigate the inter-individual variability of gene expression within sorted memory CD4+ T cells from our previously characterized cohort of individuals with latent tuberculosis infection (LTBI) and uninfected controls (**Burel et al., 2018**). Within the 100 most variable genes, we identified a set of 22 genes that were highly co-expressed with each other (22-var set, **Figure 1A**, **Figure 1—source data 1**). Strikingly, many of the genes contained within the 22-var set were previously described as being highly expressed in classical monocytes (and to a lower extent non-classical monocytes) but not in T cells (**Figure 1B**, **Schmiedel et al., 2018**). In particular, the 22-var set contained the commonly used monocyte lineage marker CD14, the lysozyme LYZ and the S100 calcium binding proteins S100A8 and S100A9, which are known to be extremely abundant in monocytes (**Figure 1B**). By examining the flow cytometry data that were acquired during cell sorting and applying our memory CD4+ T cell gating strategy (**Figure 1—figure supplement 1A**), we identified that indeed there was a subpopulation within sorted memory CD4+ T cells that stained positive for CD14 (**Figure 1C**). More importantly, the proportion of memory CD4+ T cells that were CD14+ was

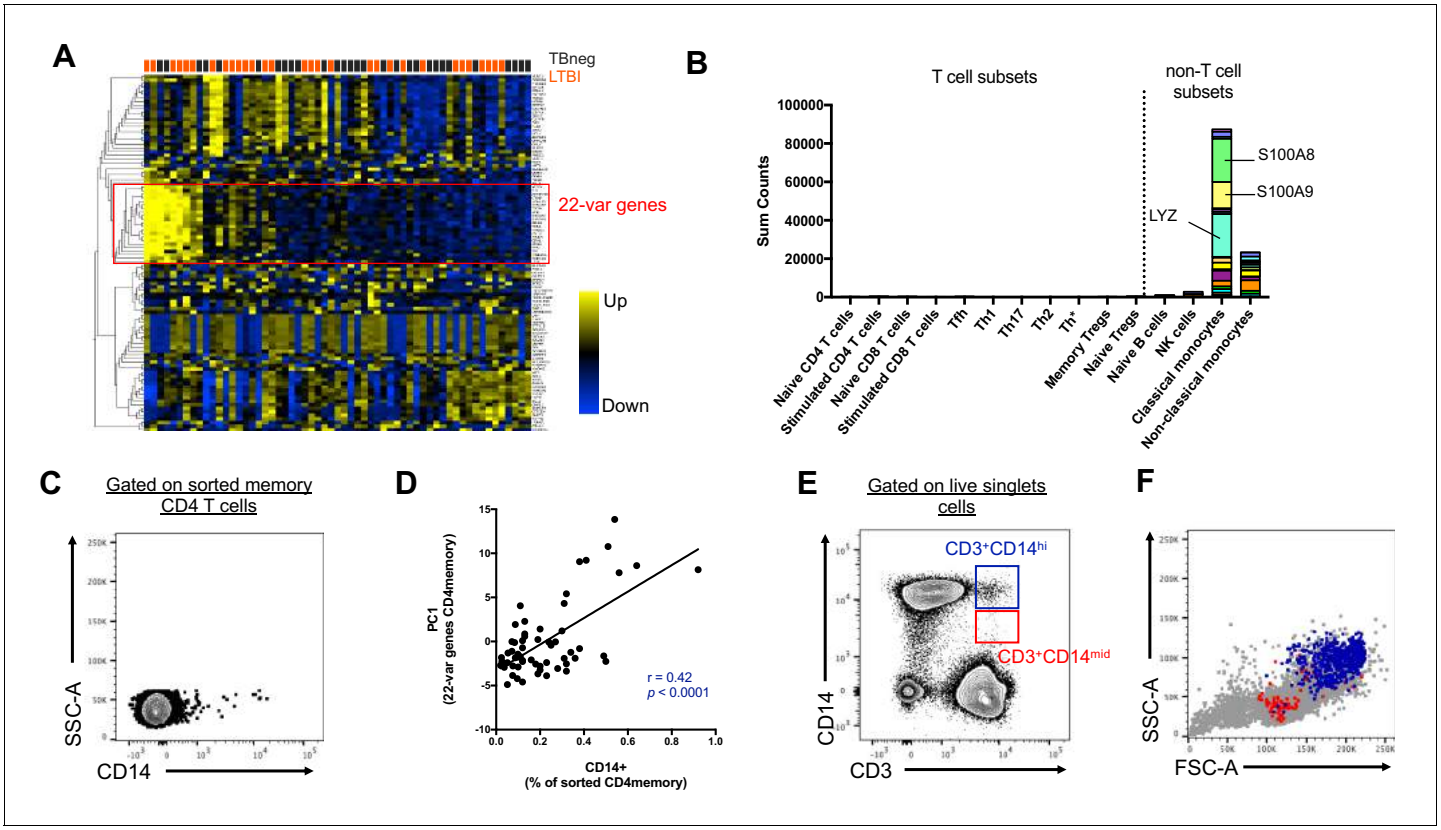


Figure 1. Two cell populations expressing both T cell (CD3) and monocyte (CD14) surface markers exist in the live singlet cell population of PBMC from human subjects. (**A**) The top 100 most variable genes in memory CD4+ T cells across TB uninfected (TBneg) and LTBI infected subjects. (**B**) Immune cell type specific expression of the 22-var genes identified in A). Every bar consists of stacked sub-bars showing the TPM normalized expression of every gene in corresponding cell type. Expression of genes for the blood cell types shown were taken from the DICE database (***Schmiedel et al., 2018***, http://dice-database.org/). (**C**) Detection of CD14+ events within sorted CD4+ memory T cells and (D) non-parametric spearman correlation between their frequency and the PC1 from the 22-var genes. (**E**) Gated on 'singlet total live cells', two populations of CD3+CD14+ cells can be identified based on the level of expression of CD14. (**F**) Based on FSC and SSC parameters, CD3+CD14hi cells are contained within the monocyte gate, whereas CD3 +CD14mid cells are contained within the lymphocyte gate. Data were derived from 30 LTBI subjects and 29 TB uninfected control subjects.

DOI: https://doi.org/10.7554/eLife.46045.002

The following source data and figure supplements are available for figure 1:

Source data 1. Non-parametric spearman correlation between the 22-var genes in memory CD4+ T cells from human subjects.

DOI: https://doi.org/10.7554/eLife.46045.005

Figure supplement 1. Identification of a CD14 +population within memory CD4 T cells.

DOI: https://doi.org/10.7554/eLife.46045.003

Figure supplement 2. Gating strategy to identify CD3+CD14+ cells.

DOI: https://doi.org/10.7554/eLife.46045.004

positively correlated with the 22-var set expression (spearman correlation coefficient r = 0.42, p<0.0001, ***Figure 1D***), suggesting that this cell subset is responsible for the expression of the monocyte-associated genes identified in ***Figure 1A***. The CD14+ memory CD4+ T cell population has similar forward and side scatter (FSC/SSC) values to other memory CD4+ T cells and was thus sorted along with conventional CD14- memory CD4+ T cells (***Figure 1—figure supplement 1B***). In particular, there was no indication that CD14+ memory CD4+ T cells were the product of a technical artifact, such as dead cells or a compensation issue.

Distinct CD3+CD14+ cell populations are present in the monocyte vs. the lymphocyte size gate

To further investigate the origin of the CD14+ T cell population, we analyzed our flow cytometry data, this time not restricting to the compartment of sorted memory T cells, but looking at all cells.

When gating on live FSC/SSC (including both monocytes and lymphocytes) singlet cells (***Figure 1—figure supplement 2***), two populations of CD3+CD14+ could be readily identified: CD3+CD14hi cells and CD3+CD14mid cells (***Figure 1E***). CD3+CD14hi cells were predominantly contained within the monocyte size gate, whereas CD3+CD14mid cells were contained within the lymphocyte size gate (***Figure 1F***).

CD3+CD14+ cells consist of T cells bound to monocytes or monocyte debris

To better understand the nature of CD3+CD14+ cells, we aimed to visualize the distribution of their markers using imaging flow cytometry. Live events were divided into monocytes (CD3-CD14+), T cells (CD3+CD14-), CD3+CD14hi cells, and CD3+CD14mid cells (***Figure 2A***), and a random gallery of images was captured for each population. As expected, monocytes and T cells contained exclusively single cells that expressed either CD14 (monocytes) or CD3 (T cells), respectively (***Figure 2B***, *first and second panel*). To our surprise, CD3+CD14hi cells contained predominantly two cells, sometimes even three cells, but no single cells (***Figure 2B***, *third panel*). The doublets (or triplets) always contained at least one CD14+ cell, and one CD3+ cell (***Figure 2B***, *third panel*). CD3+CD14mid cells contained predominantly single CD3+ cells, but also some doublets of one CD3+ cell and one CD14 + cell, but with CD14 expression lower than average monocytes (***Figure 2B***, *fourth panel*). The majority of CD3+ T cell singlets in the CD3+CD14mid population, but not in the CD3+CD14 T cell population, contained CD14+ particles, often seen at the periphery of the CD3+ T cell membrane (***Figure 2B–C***). Looking more closely at the CD14+ particles contained within the CD3+ CD14 mid population using confocal microscopy, they were found to have size and shape similar to cell debris (***Figure 2D***). To confirm our initial observation, we repeated the experiment with multiple individuals, and compared for each cell population the aspect ratio and area from the brightfield parameter collected with the image stream. Doublets are known to present a larger area but reduced aspect ratio, when compared to single cells. Thus, their overall ratio (area vs aspect ratio) is greater than in single cells. As expected, the area vs aspect ratio was significantly higher for CD3+CD14hi cells and CD3+CD14mid cells compared to single monocytes and T cells, and events in these two cell populations were found predominantly in the 'doublet gate' (***Figure 2E–F***). CD3+CD14hi cells also had a significantly higher ratio compared to CD3+CD14mid cells (***Figure 2F***).

Taken together, these results demonstrate that CD3+CD14hi cells are tightly bound T cell:monocyte complexes, in such strong interaction that sample processing and flow cytometry acquisition did not break them apart. Conversely, the CD3+CD14mid population appears to predominantly consist of single CD3+ T cells with attached CD14+ cell debris. This conclusion is further supported by CD3+CD14hi complexes being found in the monocyte size gate, whereas CD3+CD14mid cells were falling into the lymphocyte size gate (***Figure 1F***).

For the remaining of the manuscript, we refer to T cell:monocyte complexes as the CD3+CD14 + population gated from live singlets cells, as represented in ***Figure 1—figure supplement 2***.

T cell:monocyte complexes are not the result of ex vivo sample manipulation

Next, we sought to determine whether the physical association of T cells and monocytes within the T cell:monocyte complexes was the result of random cellular proximity or non-specific antibody staining during ex vivo sample manipulation, or if the complexes are originally present in peripheral blood. We could readily detect T cell:monocyte complexes in freshly isolated PBMC, and at similar frequencies as the same samples after cryopreservation (***Figure 2G***). In another set of samples, using red blood cell (RBC) magnetic depletion (and thus minimal sample manipulation), we could successfully identify T cell:monocyte complexes directly from whole blood at frequencies matching the same sample after PBMC isolation (***Figure 2—figure supplement 1A***). We also assessed the effect of the anti-coagulant used for blood collection, and found similar frequencies of T cell:monocyte complexes using either heparin or EDTA, in both fresh and frozen PBMC (***Figure 2—figure supplement 1B***). More strikingly, in a small healthy population bled longitudinally one week apart, their frequency was variable between individuals, but highly stable over time within each individual (non-parametric spearman correlation r = 1 and r = 0.9 for fresh and frozen PBMC, respectively, ***Figure 2—figure supplement 1C***). Finally, to rule out that non-specific binding of antibody conjugates

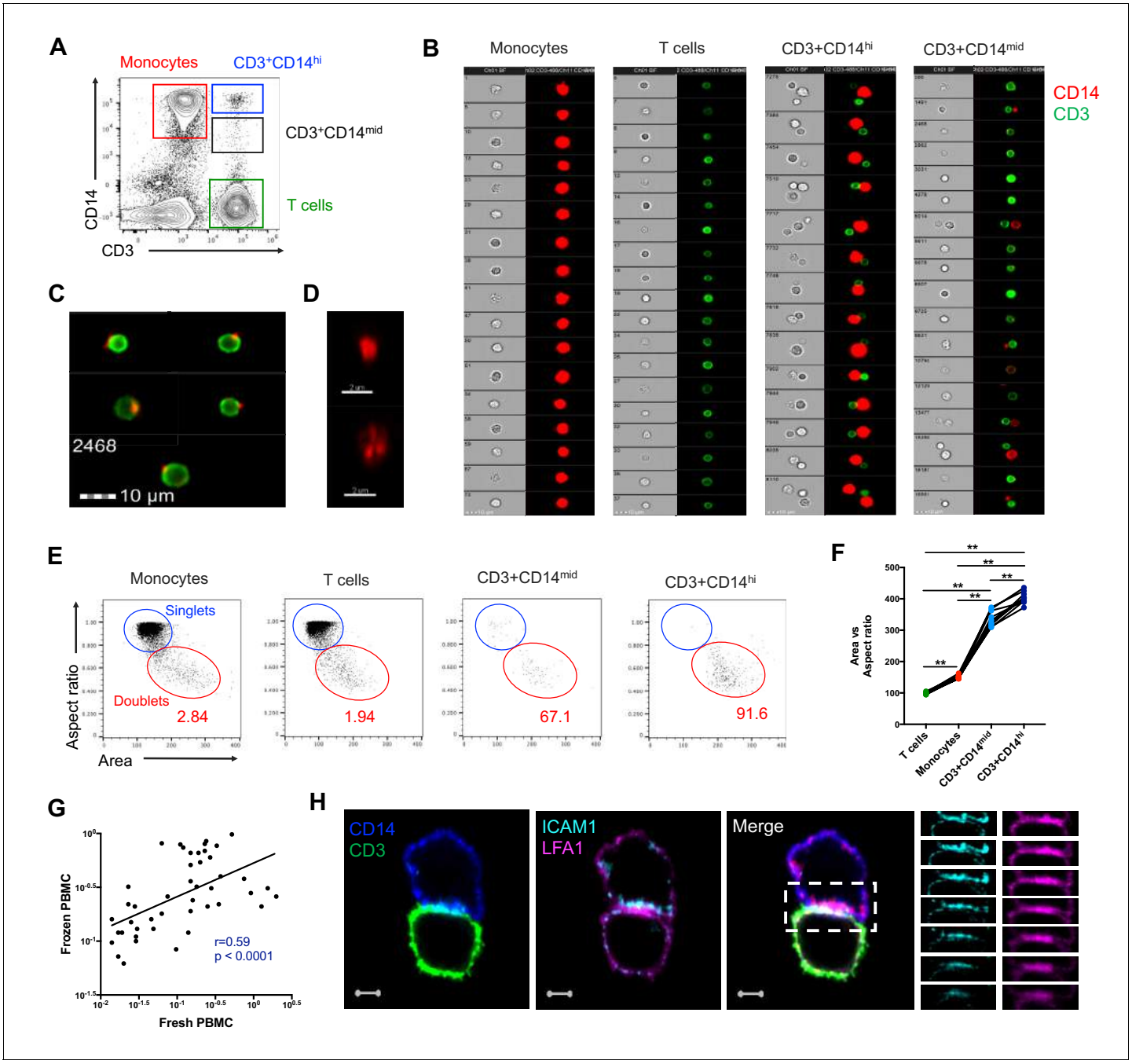


Figure 2. CD3+CD14+ cells are tightly bound T cell:monocyte complexes that represent in vivo association. (**A**) Gating strategy and (**B**) random gallery of events for monocytes (CD14+CD3), T cells (CD3+CD14-), CD3+CD14hi cells and CD3+CD14mid cells determined by imaging flow cytometry (ImageStreamX, MkII Amnis Amnis). CD14+ cell debris were identified within CD3+CD14mid cells (**C**) by imaging flow cytometry and (**D**) confocal microscopy after bulk population cell sorting. (**E**) Plots and (F) Ratio of Aspect ratio vs Area of the brightfield parameter for monocytes (CD14+CD3-), T cells (CD3+CD14-), CD3+CD14hi cells and CD3+ CD14 mid cells, determined by imaging flow cytometry. (**G**) Non-parametric Spearman correlation of the frequency of T cell:monocyte complexes in paired fresh PBMC vs cryopreserved PBMC derived from 45 blood draws of healthy subjects. T cell: monocyte complexes were defined as the CD3+CD14+ cell population gated from live singlets as represented in ***Figure 1—figure supplement 2***. (**H**) Single z-plan (0 µm) images (*left*) and z-plane stacks (*right*) of the region marked (dashed rectangle) from one sorted CD3+ CD14+ T cell:monocyte complex displaying accumulation of LFA1 and ICAM1 at the interface. Images show expression of CD14 (blue), CD3 (green), ICAM1(Cyan), and LFA1 (Magenta). Relative z-positions are indicated on the right, and scale bars represent 2 µm. Imaging flow cytometry data was derived from 10 subjects across three independent experiments and microscopy data was representative of the analysis of n = 105 CD3+CD14+ complexes isolated from three subjects across three independent experiments.

DOI: https://doi.org/10.7554/eLife.46045.006

Figure 2 continued on next page

Figure 2 continued

The following figure supplements are available for figure 2:

Figure supplement 1. Technical variations in sample preparation do not impact the frequency of T cell:monocyte complexes.

DOI: https://doi.org/10.7554/eLife.46045.007

Figure supplement 2. Accumulation of CD3, LFA1 and ICAM1 at the interface of a T cell:monocyte complex.

DOI: https://doi.org/10.7554/eLife.46045.008

Figure supplement 3. Effect of physical and chemical sample manipulation on the frequency of T cell:monocyte complexes.

DOI: https://doi.org/10.7554/eLife.46045.009

used for flow cytometry staining could be the result of complex formation (via Fc-receptors, known to be highly abundant in the surface of monocytes), we compared the frequency of T cell:monocyte complexes obtained from cryopreserved PBMC in the presence or absence of a Fc-receptor blocking reagent. The frequency of T cell:monocyte complexes frequencies was unchanged when blocking Fc-receptor binding of conjugated antibodies (***Figure 2—figure supplement 1D***).

Taken together these data rule out that ex vivo manipulation of the blood sample could be responsible for T cell:monocyte complexes formation and thus suggest their presence in vivo in peripheral blood.

T cell:monocyte complexes show increased expression of adhesion molecules at their interface

During T cell recognition of epitopes on APCs such as monocytes, the two cells are known to form an 'immune synapse' at their contact point, which is stabilized by key adhesion molecules such as LFA1 on the T cell, and ICAM1 on the APC (***Dustin, 2014***). Upon interaction these two molecules undergo a drastic redistribution by focusing almost exclusively at the cell:cell point of contact, thus forming a 'ring' that can be visualized (***Wabnitz and Samstag, 2016***). To identify candidate immunological synapses in T cell:monocyte complexes, we used high resolution Airyscan images of sorted doublets (see ***Figure 1—figure supplement 2*** for sorting strategy). Almost a third (thirty out of 105, 29%) of doublets analyzed from three different individuals displayed accumulation and polarization of ICAM1 and LFA1 at their interfaces (***Figure 2H***). The percentage of polarized doublets ranged from 17% to 67% between the subjects. In seven doublets, CD3 also accumulated together with LFA1 (***Figure 2—figure supplement 2***). However, we did not find well developed, classical immunological synapses, defined by central accumulation of CD3 and LFA1 exclusion from central region of a synapse (***Monks et al., 1998***; ***Thauland and Parker, 2010***). Overall, this suggests that a significant fraction of the detected T cell:monocyte complexes utilizes adhesion markers associated with T cell:APC synapse formation to stabilize their interaction, but they do not appear to undergo active TCR signaling at the moment of capture and acquisition.

To assess the strength of the adhesion between the T cell and the monocyte within complexes, we explored various in vitro conditions in which we attempted to break down their interaction. As seen in ***Figure 2—figure supplement 3***, in all three individuals tested, the frequency of T cell:monocyte complexes was reduced after vigorous pipetting up/down, and the strongest 'destructive effect' was observed with mild sonication. Incubation with RBC lysis buffer disrupted the T cell:monocyte complexes in one out of three individuals, while addition of high concentration of anti-chelating agent EDTA had no effect on their frequency. Thus, it appears that it is possible to disrupt T cell:monocyte complexes with physical methods. Conversely, we tried to promote their in vitro formation by stimulating PBMC for two hours to several days, with various concentrations of highly antigenic stimuli (such as LPS, PHA, SEB or live BCG) without success, which might be due to issues with monocytes attaching to the culture plates (data not shown).

Conventional flow cytometry parameters could not differentiate between T cells and monocytes in a complex versus not in a complex

So far, our identification of T cell:monocyte complexes has been solely relying on the co-expression of CD3 and CD14 within cells falling into the live singlet gate (***Figure 1—figure supplement 2***). In order to fine-tune their detection, we investigated whether we could identify some additional flow cytometry parameters that could separate T cells and monocytes in a complex vs. not in a complex.

The area (A), height (H) and width (W) of the peak from forward and side scatter parameters are routinely used in flow cytometry to identify doublets. A non-linear staining between any 2D combination of these three parameters is indicative of cell aggregates. The density plot of CD3+CD14+ cells overlapped with both CD3+ cells and especially CD14+ cells for any combination of A, H and W for both FSC and SSC parameters (***Figure 3A***). An additional parameter often used to 'clean up' the gating of cells within biological samples is the use of CD45-SSC gating (***Harrington et al., 2012***). We found that the frequency of T cell:monocyte complexes was not affected by applying an initial CD45-SSC gate filtering (***Figure 3B***, see gating in ***Figure 3—figure supplement 1***). In conclusion, we could not find any parameter from conventional flow cytometry that could separate singlets and T cell:monocyte complexes with sufficient resolution.

No difference in T cell and monocyte canonical marker expression was identified between T cells and monocytes in a complex versus not in a complex

In parallel, we analyzed the expression of various markers known to be exclusively expressed by either monocytes or T cells within T cell:monocyte complexes from healthy individuals. Monocytes in a complex (CD14+CD3+) vs. not in a complex (CD14+CD3-) had similar expression for monocyte canonical markers: CD33, CD36, CD64 and CD163 (***Figure 3C***). Similarly, no difference in the levels of expression of T cell canonical markers, CD2, CD5, CD7 and CD27 was observed between T cells in a complex (CD3+CD14+) vs. not in a complex (CD3+CD14-) (***Figure 3D***). Additionally, both CD4 and CD8 T cell subsets could be found in association with a monocyte (***Figure 3E***), as well as naïve and memory phenotypes (***Figure 3F***). Thus, there was no obvious T cell or monocyte marker that could differentiate between monocytes and T cells present in a complex vs. not in a complex, suggesting the ability to form complexes is a general property of all CD14 +monocytes and T cell subsets.

The frequency of T cell:monocyte complexes varies in the context of diverse immune perturbations

Next, we thought to examine whether the formation of T cell:monocyte complexes is affected by immune perturbations. In order to accurately assess and compare the frequency of complexes between cells of different types across different donor cohorts, we need to consider that their frequency is dependent on the abundance of its two components. Indeed, in healthy subjects, where we expect constant affinity between T cells and monocytes, we observed that the frequency of CD3+CD14+ cells is a linear function of the product of singlet monocyte and T cell frequencies (***Figure 4A***). To correct for this, we elected to express the abundance of T cell:monocytes complexes as a constant of association Ka, where similarly to a constant of chemical complex association, the frequency of T cell:monocyte complexes is divided by the product of the frequency of both T cells and monocytes (***Figure 4B***). As T cell and monocyte frequencies in the blood can fluctuate greatly during immune perturbations, the Ka is a more accurate readout of the likelihood of T cell:monocyte complex formation as opposed to raw frequencies, the latter being biased towards the overall abundance of each subset forming the complex.

We first investigated the T cell:monocyte Ka in the context of two diseases where monocytes are known to be important, namely active tuberculosis (TB) infection and dengue fever. In the case of TB, although macrophages are known to be the primary target for *Mycobacterium tuberculosis* (Mtb) infection and replication, monocytes can also be infected and contribute to the inflammatory response (***Srivastava et al., 2014***). In active TB subjects, we found a significant decrease in T cell:monocyte Ka at 2 months post treatment (***Figure 4C***). At the time of diagnosis, some subjects displayed a Ka much higher than any uninfected or LTBI individuals, but because of the high heterogeneity within the active TB cohort, these differences did not reach statistical significance (***Figure 4—figure supplement 1***). Dengue virus predominantly infects monocytes in the peripheral blood (***Kou et al., 2008***), and circulating monocyte infection and activation is increased in dengue hemorrhagic fever (the more severe form of dengue fever) (***Durbin et al., 2008***). In subjects with acute dengue fever from Sri Lanka, patients that developed hemorrhagic fever had higher T cell:monocyte Ka upon hospitalization compared to healthy, previously infected subjects (blood bank donors seropositive for dengue antibodies) (***Figure 4D***). In contrast, patients with a less severe form of acute

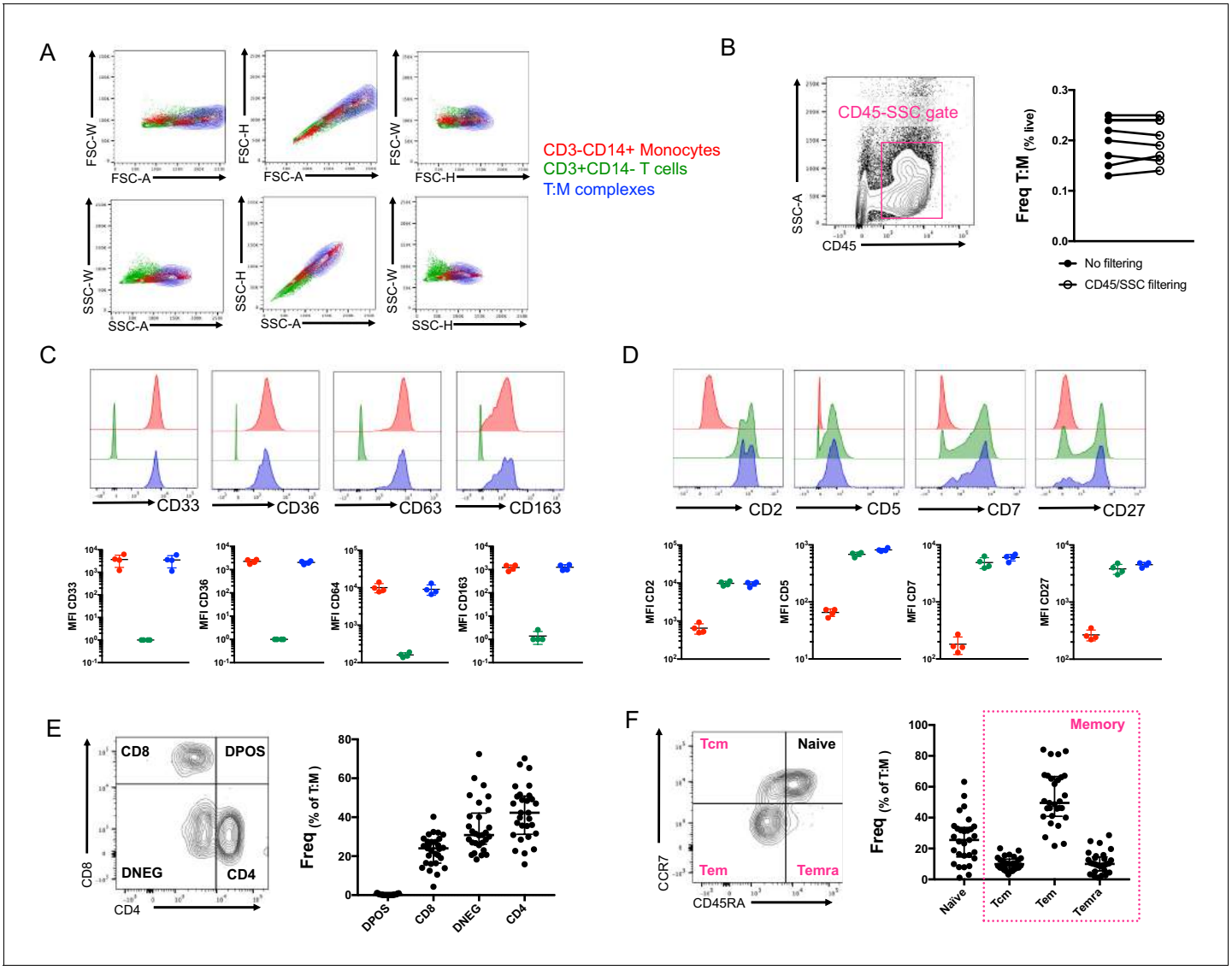


Figure 3. Conventional flow cytometry parameters and expression of T cell/monocyte canonical markers cannot differentiate between T cells and monocytes in a complex vs. not in a complex. (**A**) 2D density plots of A, H and W from FSC and SSC parameters for CD3-CD14+ Monocytes (red), CD3+CD14- T cells (green) and CD3+CD14+ T cell:monocyte complexes (T:M, blue). Representative staining of one healthy individual. (**B**) Frequency of T cell:monocyte complexes cells with or without addition of CD45-SSC filtering gate (see Figure 3 – figure supplement 1 for gating strategy). Expression of canonical markers for (**C**) monocytes and (**D**) T cells in CD3-CD14+ Monocytes (red), CD3+CD14- T cells (green) and CD3+CD14+ T cell:monocyte complexes (T:M, blue). (**E**) Expression of CD4 and CD8 and division into T cell subsets within T cell:monocyte complexes. (**F**) Expression of CD45RA and CCR7 and division into naïve, central memory (Tcm), effector memory (Tem) and effector memory re-expressing CD45RA (Temra) subsets within T cell:monocyte complexes. Data derived from frozen PBMC of n=30 (**A, E, F**), n=8 (**B**) and n=4 (**C, D**) healthy individuals. Unless otherwise stated, T cell:monocyte complexes were defined as the CD3+CD14+ cell population gated from live singlets as represented in ***Figure 1—figure supplement 2***.

DOI: https://doi.org/10.7554/eLife.46045.010

The following figure supplement is available for figure 3:

Figure supplement 1. Gating strategy to identify CD3+CD14+ cells with or without a CD45-SSC gate filtering.

DOI: https://doi.org/10.7554/eLife.46045.011

dengue infection showed no significant difference in T cell:monocyte Ka compared to healthy, previously infected donors (***Figure 4D***).

To assess whether vaccination also impacted the formation of T cell:monocyte complexes, we obtained samples from healthy adults that received the tetanus, diphtheria and pertussis (Tdap) booster vaccination. We indeed observed a significantly higher T cell:monocyte Ka at three days

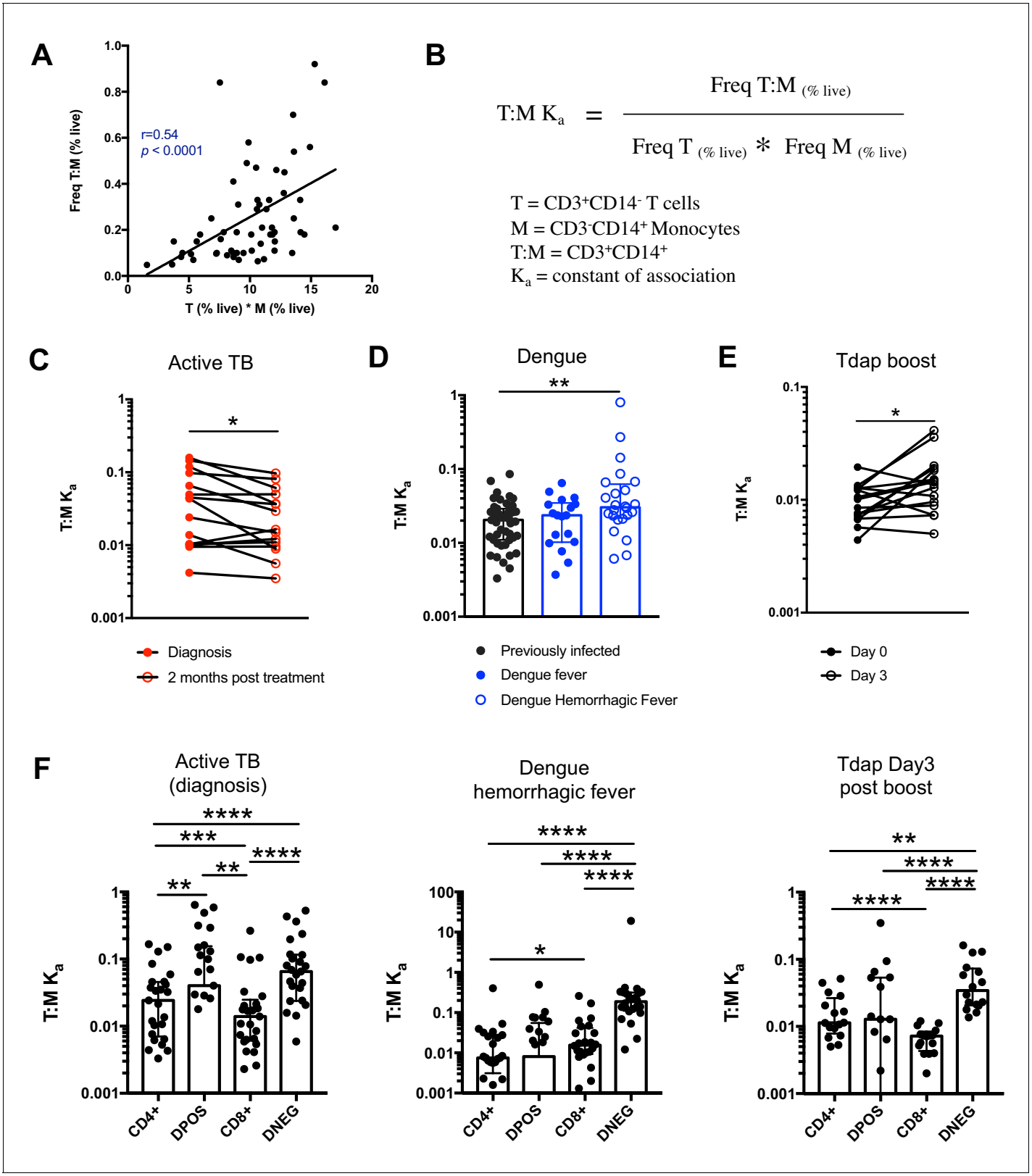


Figure 4. The constant of association Ka between monocytes and T cells (and T cell subsets) varies with the presence and nature of immune perturbations. (**A**) Non-parametric spearman correlation between the frequency of T cell:monocyte complexes and the product of singlet T cells and monocyte frequencies in healthy subjects (n = 59). (**B**) Formula for the calculation of the T cell:monocyte constant of association Ka. T cell:monocyte complexes constant of association Ka in (C) active TB subjects at diagnosis and 2 months post treatment (n = 15), (**D**) individuals with acute dengue

Figure 4 continued on next page

Figure 4 continued

fever (n = 18), acute dengue hemorrhagic fever (n = 24) or previously infected (n = 47) and (E) previously vaccinated healthy adults (n = 16) before and three days post boost with Tdap vaccine, calculated as explained in B). (**F**) The constant of association Ka between monocytes and T cell subsets in active TB subjects at diagnosis (n = 25), individuals with acute dengue hemorrhagic fever (n = 24) and previously vaccinated healthy adults three days post boost with Tdap vaccine (n = 16), calculated as explained in B). Statistical differences over time and across cell populations within subjects were determined using the non-parametric paired Wilcoxon test; other statistical differences were determined using the non-parametric Mann-Whitney test; *, p<0.05; **, p<0.01; ***, p<0.01; ****, p<0.0001. Plots represent individual data points, median and interquartile range across all subjects within each cohort. Raw frequencies of T cell:monocyte complexes for the different disease cohorts are available on Figure 4—figure supplement 4. T cell:monocyte complexes were defined as the CD3+CD14+ cell population gated from live singlets as represented in ***Figure 1—figure supplement 2***. CD4 and CD8 subsets within T cell:monocyte complexes were defined as presented in ***Figure 3E***.

DOI: https://doi.org/10.7554/eLife.46045.012

The following figure supplements are available for figure 4:

Figure supplement 1. T cell:monocyte constant of association Ka in subjects with active TB, latent TB or TB uninfected individuals.

DOI: https://doi.org/10.7554/eLife.46045.013

Figure supplement 2. T cell:monocyte constant of association Ka fluctuates as a function of time following Tdap boost administration.

DOI: https://doi.org/10.7554/eLife.46045.014

Figure supplement 3. Comparison of constant of association Ka between monocytes and T cell subsets across different immune perturbations.

DOI: https://doi.org/10.7554/eLife.46045.015

Figure supplement 4. Frequencies of T cell:monocyte complexes in different immune perturbation models.

DOI: https://doi.org/10.7554/eLife.46045.016

post boost compared to baseline (***Figure 4E***), but no significant changes at one, seven or fourteen days post boost (***Figure 4—figure supplement 2***). Taken together, these data confirm that circulating T cell:monocyte complexes can be found directly ex vivo in different immune perturbations, and their likelihood of formation is associated with clinical parameters such as disease severity, and they fluctuate as a function of time post treatment and post vaccination.

T cells with different phenotypes are found in T cell:monocyte complexes dependent on the nature of the immune perturbation

Finally, we reasoned that if immune perturbations increase the formation of T cell:monocyte complexes, then the nature of the T cells contained in the complexes could provide insights into which T cells are actively communicating with monocytes in vivo. In particular, the T cell subsets that will associate with an APC for the different perturbations studied above are expected to be distinct, and thus their likelihood to form a complex with a monocyte might differ too. The Tdap vaccine contains exclusively protein antigens and is known to elicit predominantly memory CD4+ T cell responses (***da Silva Antunes et al., 2018***). Mtb is a bacterial pathogen known to trigger strong CD4 + responses (***Lindestam Arlehamn et al., 2016***) as opposed to dengue virus, which is a viral antigen and thus expected to elicit CD8+ responses.

Similarly to global T cell:monocyte complexes (***Figure 4C–E***), we calculated the constant of association Ka with monocytes for each CD4/CD8 T cell subset individually. In subjects with active TB, the Ka between monocytes and CD4+CD8+ (DPOS) T cells or CD4+ T cells was significantly higher than for CD8+ T cells (***Figure 4F***) and both DPOS and CD4+ T cell:monocyte complexes had higher Ka in active TB compared to dengue hemorrhagic fever (***Figure 4—figure supplement 3***). Dengue hemorrhagic fever showed a higher T cell:monocyte Ka for CD8+ over CD4+ cells whereas Tdap day three post boost showed the opposite, with highest Ka for CD4+ over CD8+ cells (***Figure 4F***). The CD8+ T cell:monocyte Ka was also higher in Dengue and active TB compared to Tdap boost (***Figure 4—figure supplement 3***). Thus, the magnitude of Ka in CD4+ vs CD8+ T cell subsets matched what is expected based on the nature of immune perturbation. Interestingly, for all three immune perturbations studied the highest Ka with monocytes across all T cell subsets was for CD4-CD8- (DNEG) T cells (***Figure 4F***), and this effect was most pronounced in dengue (***Figure 4—figure supplement 3***). These cells could constitute gamma-delta T cells that are known to be strongly activated in the peripheral blood during acute dengue fever (***Tsai et al., 2015***).

In summary, these data indicate that the T cell subsets that are preferentially associated with monocytes differ from their individual frequencies in PBMC, and follow different patterns in the three

systems studied, further supporting the notion that these complexes are not the result of random association, and are specific to the nature of the immune perturbation.

Discussion

The unexpected detection of monocyte genes expressed in cells sorted for memory T cell markers led to the discovery that a population of CD3+CD14+ cells exist within the 'live singlet' events gate and that these cells are T cells that are tightly associated with monocytes, and less frequently, with monocyte-derived debris. Their presence in freshly isolated cells and the fact that a significant fraction of the complexes showed enriched expression for LFA1/ICAM1 adhesion molecules at their interface, suggest that they are not the product of random association of cells during processing, but represent interactions that occurred in vivo prior to the blood draw. The frequency of T cell:monocyte complexes fluctuated over time in the onset of immune perturbations such as following TB treatment or Tdap boost immunization and correlated with clinical parameters such as disease severity in the case of dengue fever. Furthermore, the T cell subset in preferential association within the monocyte in a complex varies in function of the nature of the immune perturbation.

Our initial observation of the presence of monocyte genes within the transcriptome of T cells was focused on memory CD4 T cells. This cell population was elected since our study aimed to define novel immune signatures associated with Mtb-specific CD4 T cells (***Burel et al., 2018***), which are expected to almost exclusively fall into the memory compartment in the context of latent TB infection (***Lindestam Arlehamn et al., 2013***). However, since we have found that both CD4 and CD8 T cells, from both memory and naïve phenotype, can be found in a complex with a monocyte, we think similar results would have been obtained with other sorted T cell populations. We have since detected the expression of monocytes-associated genes in several other T cell subsets, including memory CD8 T cells, and total CD4 T cells (unpublished observations).

Intact T cell:monocyte complexes were almost exclusively found in the top area of the FSC/SSC 2D plot, and were associated with high CD14 expression. In contrast, T cells with monocyte debris were associated with FSC/SSC values similar to regular non-complex T cells and an intermediate CD14 expression. This cell population might be the result of T cell:monocyte complexes from which the monocyte was disrupted during sample preparation or flow cytometry acquisition. Alternatively, these CD3+CD14mid cells could be the result of plasma membrane fragments exchange from monocytes to T cells following interaction. This phenomenon, known as trogocytosis, has been described to occur during cellular encounters between several immune cell types, including monocytes and T cells (***Daubeuf et al., 2010***; ***HoWangYin et al., 2011***).

Taken together, our results suggest circulating CD3+CD14+ complexes appear to be the result of in vivo interaction between T cells and monocytes. The origin and location of the complexes' formation is still unknown. These interactions might be occurring directly in the blood. Alternatively, it is possible that T cell:monocyte complex formation does not initially occur in peripheral blood, but rather in tissues or draining lymph nodes, and these complexes are then excavated into the peripheral circulation. The most studied physical interaction between T cells and monocytes is the formation of immune synapses. We found that about a third of complexes displayed LFA1/ICAM1 mediated interaction similarly to immune synapses, but no CD3 polarization. The immune synapse formation is a highly diverse event in terms of length and structure (***Friedl and Storim, 2004***), so it is possible that not all detected complexes are at the same stage in the interaction. In some complexes, the nature (and structure) of the architectural molecules forming the cell:cell contact might differ from traditional immune synapses, too. Studying the nature and physical properties of these interactions could provide insights into how T cells and monocytes can physically interact. Additionally, because monocytes are not the only cell type known to associate with T cells, we think the ability to form complexes with T cells should not be restricted to monocytes, but could apply more broadly to any APC. Thus, it is possible that other types of complexes pairing a T cell and other APCs such as B cells or dendritic cells can be found in the peripheral blood.

Increased immune cell:cell interactions might not necessarily always correlate with onset of immune perturbations. Nevertheless, our preliminary data suggest that determining the constant of association Ka of the T cell:monocyte (and likely more broadly any T cell:APC) complexes can indicate the presence of an immune perturbation to both clinicians and immunologists. We think that the Ka, rather than the frequency of live cells, is a more relevant parameter to measure the

occurrence of T cell:monocyte complexes (and thus the in vivo affinity between T cells and monocytes) since it corrects for random/aleatory association that is directly dependent on the T cell subset or monocyte abundance. For instance, DNEG T cells are in lower abundance than CD4+ or CD8+ T cells in peripheral blood, and thus DNEG T cells in a complex with a monocyte are found at a lower frequency compared to CD4+ or CD8+ T cell:monocyte complexes (***Figure 4—figure supplement 4D***). However, interestingly, their Ka is consistently much higher than CD4+ or CD8+ T cells across all disease cohorts analyzed (***Figure 4F***), suggesting a higher affinity of DNEG cells to monocytes. This information would have been missed if only frequencies were considered. In dengue infected subjects, a higher T cell:monocyte Ka at time of admission was associated with dengue hemorrhagic fever, the more severe form of disease. The distinction between hemorrhagic vs. non-hemorrhagic fever may become clear only days into hospitalization, so the ability to discriminate these two groups of individuals at the time of admission has potential diagnostic value. In the case of active TB, subjects presented a very high variability at diagnosis that might reflect the diverse spectrum associated with the disease (***Pai et al., 2016***), but for all subjects a significant decrease in T cell:monocyte Ka was observed upon treatment. This could thus be a tool to monitor treatment success and predict potential relapses. It will of course be necessary to run prospective trials to irrefutably demonstrate that the likelihood of association between T cells and monocytes have predictive power with regard to dengue disease severity or over the course of TB treatment. Additionally, the T cell:monocyte Ka was increased three days following Tdap booster vaccination. Therefore, in vaccine trials, it could be examined as an early readout to gage how well the immune system has responded to the vaccine. Finally, in apparently 'healthy' populations, or those with diffuse symptoms, an unusually high T cell:monocyte Ka in an individual could be used as an indicator of a yet to be determined immune perturbation.

Beyond detecting abnormal frequencies of T cell:monocyte complexes, characterizing the T cells and monocytes in these complexes might provide insights into the nature of immune perturbation and subsequent immune response based on which complexes were formed. Our data suggest that there are drastic differences in terms of T cell subsets in the complexes. As aforementioned, despite their lower frequency over CD4+ and CD8+ T cells in the peripheral blood, DNEG T cells show a clear increased association with monocytes. Gamma-delta T cells constitute the majority of circulating DNEG T cells in humans, and LFA1 dependent crosstalk between gamma-delta T cells and monocytes has been shown to be important in the context of bacterial infections (***Eberl et al., 2009***), which might be also generalized to viral infections. Thus, the DNEG T cell:monocyte complexes might well represent a novel type of interaction between T cells and monocytes, not necessarily involving classical alpha-beta T cells or involving the formation of 'traditional' immune synapses. Aside from the enrichment for DNEG cells in T cell:monocyte complexes, we also observed a relatively high Ka for DPOS cells in T cell:monocyte complexes in all samples analyzed, despite their very low abundance amongst T cells. Circulating DPOS T cells have been described in the context of several infections, in particular from viruses. They are associated with enhanced effector functions such as proliferation, cytotoxicity and cytokine production (***Kitchen et al., 2004***; ***Nascimbeni et al., 2004***). DPOS cells might thus represent a specific subset of T cells with enhanced surveillance and cell:cell communication functions, and thus have higher affinity for APCs hence higher likelihood to be found in a complex with a monocyte. Finally, we also found that the CD4 vs CD8 phenotype of the T cell present in complexes depends on the nature of the immune perturbation studied, and reflects the expected polarization of immune responses. Thus, looking for additional characteristics from T cells and monocytes present in the complexes, such as the expression of tissue homing markers, specific TCRs and their transcriptomic profile might provide further information about the fundamental mechanisms underlying immune responses to a specific perturbation.

Why were T cell:monocyte complexes not detected and excluded in flow cytometry based on gating strategies to avoid doublets? Surprisingly, all usual parameters (pulse Area (A), Height (H) and Width (W) from forward and side scatter) looked identical between T cell:monocyte complexes and singlet T cells or monocytes. The only parameter that could readily distinguish between intact CD3+CD14hi complexes and single T cells or monocytes was the brightfield area parameter from the imaging flow cytometer, which is a feature absent in non-imaging flow cytometry. Thus, it seems that gating approaches and parameters available in conventional flow cytometry are not sufficient to completely discriminate tightly bound cell pairs from individual cells.

Given that T cell:monocyte complexes are not excluded by conventional flow cytometry gating strategies, why were they not reported previously? Examining our own past studies, a major reason is that lineage markers for T cells (CD3), B cells (CD19) and monocytes (CD14) are routinely used to remove cells not of interest in a given experiment by adding them to a 'dump channel'. For example, most of our CD4+ T cell studies have CD8, CD19 and CD14, and dead cell markers combined in the same channel (**Arlehamn et al., 2014**; **Burel et al., 2018**). Other groups studying for example CD14 + monocytes are likely to add CD3 to their dump channel. This means that complexes of cells that have two conflicting lineage markers such as CD3 and CD14 will often be removed from datasets early in the gating strategy. Additionally, the detection of complexes by flow cytometry is not straightforward. In our hands, we have found that conventional flow analyzers give low frequency of complexes and poor reproducibility in repeat runs. This is opposed to cell sorters, presumably due to differences in their fluidics systems, which puts less stress on cells and does not disrupt complexes as much. Both the routine exclusion of cell populations positive for two conflicting lineage markers and the challenges to reproduce such cell populations on different platforms has likely contributed to them not being reported.

Moreover, even if a panel allows for the detection of complexes, and there is a stable assay used to show their presence, there is an assumption in the field that detection of complexes is a result of experimental artifacts. For example, we found a report of double positive CD3+CD34+ cells detected by flow cytometry in human bone marrow, which followed up this finding and found them to be doublets using microscopy imaging. The authors concluded that these complexes are the product of random association and should be ignored (**Kudernatsch et al., 2013**). Their conclusion may well be true for their study, but it highlights a common conception in the field of cytometry that pairs of cells have to be artifacts. Another study described CD3+CD20+ singlets cells observed by flow cytometry as doublets of T cells and B cells, and also concluded them to be a technical artifact, in the sense that these cells are not singlets double expressing CD3 and CD20 (**Henry et al., 2010**). In this case however, authors pointed out that 'Whether the formation of these doublets is an artifact occurring during staining or is a physiologic process remains to be determined' (**Henry et al., 2010**).

We ourselves assumed for a long time that we might have an artifact finding, but given the persistent association of T cell:monocyte complexes frequency and phenotype with clinically and physiologically relevant parameters, we came to a new conclusion: cells are meant to interact with other cells. Thus, detecting and characterizing complexes of cells isolated from tissues and bodily fluids, can provide powerful insights into cell:cell communication events that are missed when studying cells as singlets only.

Materials and methods

Key resources table

Reagent type (species) or resource	Designation	Source or reference	Identifiers	Additional information
Antibody	CCR7-PerCpCy5.5; clone G043H7; mouse monoclonal	Biolegend	Cat# 353220	1:50 (4 µl per test)
Antibody	CD2-BV421; clone RPA-2.10; mouse monoclonal	Biolegend	Cat# 300229	1:66 (3 µl per test)
Antibody	CD3-AF700; clone UCHT1; mouse monoclonal	BD pharmigen	Cat# 557943	1:66 (3 µl per test)
Antibody	CD3-AF488; clone UCHT1; mouse monoclonal	Biolegend	Cat# 300415	1:200 (1 µl per test)
Antibody	CD4-APCeF780; clone RPA-T4; mouse monoclonal	eBiosciences	Cat# 47-0049-42	1:200 (1 µl per test)
Antibody	CD5-APCCy7; clone L17F12; mouse monoclonal	Biolegend	Cat# 364009	1:66 (3 µl per test)

Continued on next page

Continued

Reagent type (species) or resource	Designation	Source or reference	Identifiers	Additional information
Antibody	CD7-APC; clone CD7-6B7; mouse monoclonal	Biolegend	Cat# 343107	1:66 (3 µl per test)
Antibody	CD8a-BV650; clone RPA-T8; mouse monoclonal	Biolegend	Cat# 301042	1:200 (1 µl per test)
Antibody	CD14-APC; clone 61D3; mouse monoclonal	Tonbo biosciences	Cat# 20–0149 T100	1:200 (1 µl per test)
Antibody	CD14-AF594; clone HCD14	Biolegend	Cat# 325630	1:200 (1 µl per test)
Antibody	CD14-AF647; clone 63D3	Biolegend	Cat# 367128	1:200 (1 µl per test)
Antibody	CD14-BV421; clone HCD14	Biolegend	Cat# 325628	1:200 (1 µl per test)
Antibody	CD14-PE; clone 61D3; mouse monoclonal	eBioscience	Cat# 12-0149-42	1:200 (1 µl per test)
Antibody	CD27-BV650; clone O323	Biolegend	Cat# 302827	1:100 (2 µl per test)
Antibody	CD33-APC; clone WM53	Biolegend	Cat# 303407	1:200 (1 µl per test)
Antibody	CD36-APCCy7; clone 5–271	Biolegend	Cat# 336213	1:100 (2 µl per test)
Antibody	CD45-PerCpCy5.5; clone HI30; mouse monoclonal	Tonbo biosciences	Cat# 65–0459 T100	1:66 (3 µl per test)
Antibody	CD45RA-eF450; clone HI100; mouse monoclonal	eBiosciences	Cat# 48-0458-42	1:200 (1 µl per test)
Antibody	CD64-AF488; clone 10.1; mouse monoclonal	Biolegend	Cat# 305010	1:200 (1 µl per test)
Antibody	CD163-PECy7; clone GHI/61; mouse monoclonal	Biolegend	Cat# 333613	1:100 (2 µl per test)
Antibody	ICAM1(CD54); unconjugated; clone HCD54; mouse monoclonal	Biolegend	Cat# 322704	1:40 (5 µl per test)
Antibody	LFA1(CD11a); unconjugated; clone TS2/4; mouse monoclonal	Biolegend	Cat# 350602	1:40 (5 µl per test)
Antibody	LFA1(CD11a/CD18)-AF647; clone m24; mouse monoclonal	Biolegend	Cat# 363412	1:40 (5 µl per test)

Subjects and samples

Samples from TB uninfected individuals were obtained from the University of California, San Diego Antiviral Research Center clinic (AVRC at UCSD, San Diego) and National Blood Center (NBC), Ministry of Health, Colombo, Sri Lanka, in an anonymous fashion as previously described (***Burel et al., 2017***). Samples from individuals with LTBI were obtained from AVRC at UCSD, San Diego, and the Universidad Peruana Cayetano Heredia (UPCH, Peru). Longitudinal active TB samples were obtained from National Hospital for Respiratory Diseases (NHRD), Welisara, Sri Lanka. Dengue previously infected samples were obtained from healthy adult blood donors from the National Blood Center (NBC), Ministry of Health, Colombo, Sri Lanka, in an anonymous fashion as previously described (***Weiskopf et al., 2013***). Acute dengue fever samples were collected at National Institute of Infectious Diseases, Gothatuwa, Angoda, Sri Lanka and the North Colombo Teaching Hospital, Ragama,

in Colombo, Sri Lanka. Longitudinal Tdap booster vaccination samples and non-vaccinated healthy samples were obtained from healthy adults from San Diego, USA. LTBI status was confirmed in subjects by a positive IFN-γ release assay (IGRA) (QuantiFERON-TB Gold In-Tube, Cellestis or T-SPOT. TB, Oxford Immunotec) and the absence of clinical and radiographic signs of active TB. TB uninfected control subjects were confirmed as IGRA negative. Active Pulmonary TB was defined as those exhibiting symptoms of TB, and are positive by sputum and culture as confirmed by the National Tuberculosis Reference Laboratory (NTRL, Welisara, Sri Lanka). Sputum was further confirmed positive for TB by PCR at Genetech (Sri Lanka). Active TB patients in this study were confirmed negative for HIV, HBV and HCV. Upon enrollment within seven days of starting their anti-TB treatment, active TB patients provided their first blood sample, followed by a second blood sample two months after initial diagnosis. Acute dengue fever and previously infected samples were classified by detection of virus (PCR+) and/or dengue-specific IgM and IgG in the serum. Laboratory parameters such as platelet and leukocyte counts, hematocrit, hemoglobulin, AST, ALT and if applicable an ultrasound examination of the chest and abdomen or an X-ray were used to further diagnose patients with either dengue fever (DF) or dengue hemorrhagic fever (DHF), a more severe form of disease, according to WHO's guidelines. Longitudinal Tdap booster vaccination samples were obtained from individuals vaccinated in childhood, and boosted with the DTP vaccine Tdap (Adacel). Blood samples were collected prior, one day, three days, seven days and fourteen days post boost. Longitudinal healthy samples were obtained from two consecutive bleeds of healthy adults, at seven days apart. For some latent TB and TB negative subjects, leukapheresis was performed instead of a whole blood donation in order to increase the number of PBMC obtained. Samples from Peru were exclusively collected by leukapheresis, whereas 65% of samples from San Diego were collected by leukapheresis (33 out of 51 subjects). No leukapheresis samples were collected in Sri Lanka. We have found no difference in T cell:monocyte complexes frequencies in samples collected by leukapheresis versus whole blood (data not shown), and have thus defined our cohorts based on TB diagnosis status (TB negative, latent TB or active TB), regardless of the blood draw technique. All blood samples were drawn in Lithium or Sodium heparin, except for the analysis of the effect of anti-coagulant (***Figure 2—figure supplement 1B***) where some healthy samples were also collected in EDTA. Clinical sites of Peru and Sri Lanka have been personally trained at La Jolla Institute for Immunology and all three sites follow the same operating procedures and protocols for blood processing. All blood samples were stored at room temperature for up to 12 hr before blood processing with a maximum processing time of three hours. Time of the day for blood draw was aleatory variable ranging from morning to afternoon for each site. For all cohorts, PBMC were obtained by density gradient centrifugation (Ficoll-Hypaque, Amersham Biosciences) according to the manufacturer's instructions. Cells were resuspended to 10 to 50 million cells per mL in FBS (Gemini Bio-Products) containing 10% dimethyl sulfoxide (Sigma) and cryopreserved in liquid nitrogen.

Magnetic RBC depletion

Magnetic RBC depletion was performed using the EasySep RBC depletion kit (STEMCELL technologies), according to the manufacturer's instructions. Briefly, 500 µl of whole blood was supplemented with 6 mM EDTA (final concentration) and 500 µl PBS + 2% FCS, and transferred into a 5 mL polystyrene round-bottom tube. After adding 25 µl of depletion reagent, the sample was incubated for 5 min on a EasySep magnet, and cell suspension was collected by inverting the magnet in one continuous motion into a new tube. Depletion was repeated once more by adding the same volume of depletion reagent. Cell suspension obtained after the second depletion (depleted of RBC) was directly used for flow cytometry staining.

Flow cytometry

Surface staining of fresh or frozen PBMC was performed as previously described in ***Burel et al. (2017)***. All centrifugations were performed at 600 g for 5 min. For cryopreserved PBMC, cells were quickly thawed by incubating each cryovial at 37˚C for 2 min, and cells transferred into 9 ml of cold medium (RPMI 1640 with L-Glutamin and 25 mM Hepes (Omega Scientific), supplemented with 5% human AB serum (GemCell), 1% Penicillin Streptomycin (Gibco) and 1% Glutamax (Gibco)) and 20 U/mL Benzonase Nuclease (Millipore) in a 15 ml conical tube. Cells were centrifuged and resuspended in medium to determine cell concentration and viability using Trypan blue and a hematocytometer.

Cells (1–10 million) were transferred into a 15 ml conical tube, centrifuged, resuspended in 100 µl of PBS containing 10% FBS and incubated for 10 min at 4˚C. Cells were then stained with 100 µl of PBS containing fixable viability dye eFluor506 (eBiosciences) and various combinations of the antibodies listed in ***Supplementary file 1*** for 20 min at room temperature. Each antibody was individually titrated for optimum staining, and dilutions/panels used in the study are available in ***Supplementary file 1*** . To assess the effect of Fc-receptor blocking on the formation of T cell:monocyte complexes (***Figure 2—figure supplement 1D***), 2 µl of Trustain FcR blocking reagent (BioLegend) was added along with the antibodies. After two washes in staining buffer (PBS containing 0.5% FBS and 2 mM EDTA (pH 8.0), cells were resuspended into 100–500 µl of staining Buffer, transferred into a 5 ml polypropylene FACS tube (BD Biosciences) and stored at 4˚C protected from light for up to 4 hr until flow cytometry acquisition. Acquisition was performed on a BD LSR-II cell analyzer (BD Biosciences) or on a BD FACSAria III cell sorter (BD Biosciences). Compensation was realized with single-stained beads (UltraComp eBeads, eBiosciences) in PBS using the same antibody dilution as for the cell staining. Performance of both instruments were checked daily by the flow cytometry core at La Jolla Institute for Immunology with the use of CS and T beads (BD Biosciences), and PMT voltages were manually adjusted for optimum fluorescence detection on each time it was used.

Imaging flow cytometry

For the visualization of CD3+CD14+ cells, frozen PBMC were thawed and stained with CD3-AF488 and CD14-AF647 (see ***Supplementary file 1*** for antibody details) as described in the flow cytometry section above. After two washes in PBS, cells were resuspended to 10×10^6 cells/mL in staining buffer containing 5 µg/mL Hoechst (Invitrogen) and 1 µg/mL 7-AAD (Biolegend) and stored at 4˚C protected from light until acquisition. Acquisition was performed with ImageStreamX MkII (Amnis) and INSPIRE software version 200.1.620.0 at 40X magnification and the lowest speed setting. A minimum of 4,000 CD3+CD14+ events in focus were collected. Data analysis was performed using IDEAS version 6.2.183.0.

Sample preparation for microscopy

For the visualization of LFA1/ICAM1 polarization on T cell:monocyte complexes, frozen PBMC were thawed as described in the flow cytometry section above and resuspended in blocking buffer (2% BSA, 10 mM EGTA, 5 mM EDTA, 0.05% Sodium Azide in 1X PBS) supplemented with 2 µl of Trustain FcR blocking reagent (BioLegend) for 10 min on ice. Antibodies (anti-human CD3-AF488, CD14-BV421, ICAM1-AF568, LFA1-CF633 or LFA1-AF647, see ***Supplementary file 1*** for antibody details) were added and incubated for 20 min on ice, and then washed twice with staining buffer (PBS containing 0.5% FBS and 2 mM EDTA, pH 8). Cells were fixed with 4% Paraformaldehyde, 0.4% Glutaldehyde, 10 mM EGTA, 5 mM EDTA, 0.05 Sodium Azide, 2% sucrose in PBS for 1 hr on ice, and then washed twice with MACS buffer. Cells were resuspended in 0.5–1 mL of MACS buffer, and kept at 4˚C until sorting. Cell sorting was performed on a BD Aria III/Fusion cell sorter (BD Biosciences). CD3+CD14+, CD3+CD14 T cells and CD14 +CD3 monocytes were sorted (see gating strategy ***Figure 1—figure supplement 2***) and each separately plated on a well of a µ-Slide 8 Well Glass Bottom chamber (Ibidi) that was freshly coated with poly-L-lysine (0.01%) for 30 min RT before use. For in-house antibody labeling, an Alexa Fluor 568 antibody labeling kit and a Mix-n-Stain CF633 Dye antibody labeling kit (Sigma) were used according to manufacturer's protocols.

Microscopy

Airyscan images were taken with a Plan-Apochromat 63x/1.4 Oil DIC M27 objective with a 152 µm sized pinhole with master gain 800 using a Zeiss LSM 880 confocal microscopy equipped with an Airyscan detector (Carl Zeiss). four laser lines at 405, 488, 561, and 633 nm and a filter set for each line were used for taking 20–25 series of z-plane Airyscan confocal images with a step of 0.185 µm or 0.247 µm for each channel. Pixel dwelling time was 2.33 µs and x and y step sizes were 43 nm. 3D-Airyscan processing was performed with the Zen Black 2.3 SP1 program. For some images, Z-plane linear transitional alignment was done by using the Zen Blue 2.5 program. Contrast of images for each fluorophores channel was adjusted based on FMO (Fluorescence minus one) control samples that were prepared and taken on the same day of each experiments. To visualize cell fragments, sorted CD3+CD14mid cells were immobilized using CyGel Sustain (Abcam) according to

manufacturer recommendations. Three-dimensional rendering of cellular fragments (***Figure 2D***) was created in Imaris 9.1 software (Bitplane).

Disruption of T cell:monocyte complexes

For the RBC lysis condition, immediately after thawing, PBMC were incubated for 10 min at room temperature with 4 mL of 1x RBC lysis reagent (Biolegend) according to the manufacturer's recommendations. After two washes in staining buffer, PBMC were then stained with fixable viability dye, anti-human CD3-AF700 and CD14-PE as described in the flow cytometry section above. For all other conditions, PBMC were first stained and then submitted to one of the following treatments: i) final resuspension in staining buffer at 10 mM EDTA (EDTA 10 mM), ii) vigorous pipetting up and down for 30 s after final resuspension in staining buffer (Pipette Up/Down) or iii) Sonication for 2 min and 30 s at 42 kHz (JSP Ultrasonic Cleaner) after final resuspension in staining buffer (Sonication).

Bulk memory CD4+ T cell sorting

Frozen PBMC were thawed and stained with fixable viability dye eFluor506 (eBiosciences) and various combinations of the antibodies listed in ***Supplementary file 1*** as described in the flow cytometry section above. Memory CD4 T cell sorting (see gating strategy ***Figure 1—figure supplement 1A***) was performed on a BD Aria III/Fusion cell sorter (BD Biosciences). 100,000 memory CD4+ T cells were sorted into TRIzol LS reagent (Invitrogen) for RNA extraction.

RNA sequencing and analysis

RNA sequencing and analysis of memory CD4+ T cells from LTBI infected subjects was performed as described in ***Picelli et al. (2013)***; ***Seumois et al. (2016)*** and quantified by qPCR as described previously (***Seumois et al., 2012***). 5 ng of purified total RNA was used for poly(A) mRNA selection, full length reverse-transcription and amplified for 17 cycles, following the smart-seq2 protocol (***Picelli et al., 2013***; ***Seumois et al., 2016***). After purification with Ampure XP beads (Ratio 0.8:1, Beckmann Coulter) and quantification (Picogreen assay, Invitrogen), 1 ng of cDNA was used to prepare a Nextera XT sequencing library with the Nextera XT DNA library preparation and index kits (Illumina). Samples were pooled and sequenced using the HiSeq2500 (Illumina) to obtain at least 12 million 50 bp single-end reads per library. The single-end reads that passed Illumina filters were filtered for reads aligning to tRNA, rRNA, and Illumina adapter sequences. The reads were then aligned to UCSC hg19 reference genome using TopHat (v 1.4.1) (***Trapnell et al., 2009***), filtered for low complexity reads, and parsed with SAMtools (***Li et al., 2009***). Read counts to each genomic feature were obtained using HTSeq-count program (v 0.6.0) (***Anders et al., 2015***) using the 'union' option. Raw counts were then imported to R/Bioconductor package DESeq2 (***Love et al., 2014***) to identify differentially expressed genes among samples.

Data deposition

Sequencing data is accessible online through Gene Expression Omnibus (accession numbers GSE84445 and GSE99373, https://www.ncbi.nlm.nih.gov/geo) and Immport (Study number SDY820, http://www.immport.org). All other data is available in the main text or the supplementary materials.

Acknowledgements

We thank Dr Chery Kim and all present and past members at the flow cytometry core facility at the La Jolla Institute for Immunology for assistance in cell sorting and technical discussion. We thank Dr Zbigniew Mikulski from the microscopy core at the La Jolla Institute for Immunology for assistance and technical advice on microscopy imaging. We thank Yoav Altman at the Sanford Burnham Prebys flow cytometry core for technical assistance with imaging flow cytometry. We thank Dr Joe Trotter from the R and D Advanced Technology Group at BD Biosciences for useful technical discussion about cytometry instrument fluidic systems. Research reported in this manuscript was supported by the National Institute of Allergy and Infectious Diseases division of the National Institutes of Health under award number U19AI118626, R01AI137681, HHSN272201400045C, P01HL078784, S10OD021831 and S10OD016262. The content is solely the responsibility of the

authors and does not necessarily represent the official views of the National Institutes of Health. Imaging flow cytometry was supported by the James B Pendleton Charitable trust.

Additional information

Funding

Funder	Grant reference number	Author
National Institute of Allergy and Infectious Diseases	U19AI118626	Alessandro Sette Bjoern Peters Pandurangan Vijayanand
National Institute of Allergy and Infectious Diseases	R01AI137681	Bjoern Peters
National Institutes of Health	HHSN272201400045C	Alessandro Sette
National Heart, Lung, and Blood Institute	P01HL078784	Klaus Ley

The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

Author contributions

Julie G Burel, Conceptualization, Investigation, Writing—original draft, Writing—review and editing; Mikhail Pomaznoy, Conceptualization, Investigation, Writing—review and editing; Cecilia S Lindestam Arlehamn, Conceptualization, Resources, Investigation, Writing—review and editing; Daniela Weiskopf, Ricardo da Silva Antunes, Yunmin Jung, Veronique Schulten, Gregory Seumois, Investigation, Writing—review and editing; Mariana Babor, Conceptualization, Writing—review and editing; Jason A Greenbaum, Sunil Premawansa, Gayani Premawansa, Ananda Wijewickrama, Dhammika Vidanagama, Bandu Gunasena, Rashmi Tippalagama, Aruna D deSilva, Robert H Gilman, Mayuko Saito, Randy Taplitz, Klaus Ley, Pandurangan Vijayanand, Resources, Writing—review and editing; Alessandro Sette, Conceptualization, Funding acquisition, Writing—review and editing; Bjoern Peters, Conceptualization, Funding acquisition, Writing—original draft, Writing—review and editing

Author ORCIDs

Julie G Burel https://orcid.org/0000-0003-1692-2758
Bjoern Peters https://orcid.org/0000-0002-8457-6693

Ethics

Human subjects: Ethical approval to carry out this work is maintained through the La Jolla Institute for Allergy and Immunology Institutional Review Board (protocols VD-143, VD-085, VD-090), the Medical Faculty of the University of Colombo (which served as a National Institutes of Health-approved institutional review board for Genetech, protocols EC-15-094, EC-15-002, EC.15.095), and the Johns Hopkins School of Public Health Institutional Review Board (RHG holds dual appointment at UPCH and JHU; protocol 00003804). All clinical investigations have been conducted according to the principles expressed in the Declaration of Helsinki. All participants, except anonymously recruited blood bank donors in Sri Lanka, provided written informed consent prior to participation in the study.

Decision letter and Author response

Decision letter https://doi.org/10.7554/eLife.46045.024
Author response https://doi.org/10.7554/eLife.46045.025

Additional files

Supplementary files

- Supplementary file 1. Fluorochrome-conjugated antibodies used in the study.

DOI: https://doi.org/10.7554/eLife.46045.017

- Transparent reporting form

DOI: https://doi.org/10.7554/eLife.46045.018

Data availability

RNA sequencing data of memory CD4 T cells have been deposited in GEO under accession codes GSE84445 and GSE99373. All data generated or analysed during this study are included in the manuscript and supporting files. An additional source data file has been provided for Figure 1.

The following datasets were generated:

Author(s)	Year	Dataset title	Dataset URL	Database and Identifier
Burel J, Lindenstam C, Seumois G, Fu Z, Greenbaum J, Sette A, Peters B, Vijayanand P	2016	Transcriptomic profile of circulating memory T cells can differentiate between latent tuberculosis individuals and healthy controls	https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE84445	NCBI Gene Expression Omnibus, GSE84445
Burel J, Lindenstam C, Seumois G, Fu Z, Greenbaum J, Sette A, Peters B, Vijayanand P	2018	Transcriptomic profile of circulating memory CD4 T cells can differentiate between latent tuberculosis individuals and healthy controls	https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE99373	NCBI Gene Expression Omnibus, GSE99373

References

Anders S, Pyl PT, Huber W. 2015. HTSeq–a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**:166–169. DOI: https://doi.org/10.1093/bioinformatics/btu638, PMID: 25260700

Arlehamn CL, Seumois G, Gerasimova A, Huang C, Fu Z, Yue X, Sette A, Vijayanand P, Peters B. 2014. Transcriptional profile of tuberculosis antigen-specific T cells reveals novel multifunctional features. *The Journal of Immunology* **193**:2931–2940. DOI: https://doi.org/10.4049/jimmunol.1401151, PMID: 25092889

Bongen E, Vallania F, Utz PJ, Khatri P. 2018. KLRD1-expressing natural killer cells predict influenza susceptibility. *Genome Medicine* **10**:45. DOI: https://doi.org/10.1186/s13073-018-0554-1, PMID: 29898768

Burel JG, Qian Y, Lindestam Arlehamn C, Weiskopf D, Zapardiel-Gonzalo J, Taplitz R, Gilman RH, Saito M, de Silva AD, Vijayanand P, Scheuermann RH, Sette A, Peters B. 2017. An integrated workflow to assess technical and biological variability of cell population frequencies in human peripheral blood by flow cytometry. *The Journal of Immunology* **198**:1748–1758. DOI: https://doi.org/10.4049/jimmunol.1601750, PMID: 28069807

Burel JG, Lindestam Arlehamn CS, Khan N, Seumois G, Greenbaum JA, Taplitz R, Gilman RH, Saito M, Vijayanand P, Sette A, Peters B. 2018. Transcriptomic analysis of $CD4^+$ T cells reveals novel immune signatures of latent tuberculosis. *The Journal of Immunology* **200**:3283–3290. DOI: https://doi.org/10.4049/jimmunol.1800118, PMID: 29602771

da Silva Antunes R, Babor M, Carpenter C, Khalil N, Cortese M, Mentzer AJ, Seumois G, Petro CD, Purcell LA, Vijayanand P, Crotty S, Pulendran B, Peters B, Sette A. 2018. Th1/Th17 polarization persists following whole-cell pertussis vaccination despite repeated acellular boosters. *Journal of Clinical Investigation* **128**:3853–3865. DOI: https://doi.org/10.1172/JCI121309, PMID: 29920186

Daubeuf S, Lindorfer MA, Taylor RP, Joly E, Hudrisier D. 2010. The direction of plasma membrane exchange between lymphocytes and accessory cells by trogocytosis is influenced by the nature of the accessory cell. *The Journal of Immunology* **184**:1897–1908. DOI: https://doi.org/10.4049/jimmunol.0901570, PMID: 20089699

Durbin AP, Vargas MJ, Wanionek K, Hammond SN, Gordon A, Rocha C, Balmaseda A, Harris E. 2008. Phenotyping of peripheral blood mononuclear cells during acute dengue illness demonstrates infection and increased activation of monocytes in severe cases compared to classic dengue fever. *Virology* **376**:429–435. DOI: https://doi.org/10.1016/j.virol.2008.03.028, PMID: 18452966

Dustin ML. 2014. The immunological synapse. *Cancer Immunology Research* **2**:1023–1033. DOI: https://doi.org/10.1158/2326-6066.CIR-14-0161, PMID: 25367977

Eberl M, Roberts GW, Meuter S, Williams JD, Topley N, Moser B. 2009. A rapid crosstalk of human gammadelta T cells and monocytes drives the acute inflammation in bacterial infections. *PLOS Pathogens* **5**:e1000308. DOI: https://doi.org/10.1371/journal.ppat.1000308, PMID: 19229322

Friedl P, Storim J. 2004. Diversity in immune-cell interactions: states and functions of the immunological synapse. *Trends in Cell Biology* **14**:557–567. DOI: https://doi.org/10.1016/j.tcb.2004.09.005, PMID: 15450978

Grifoni A, Costa-Ramos P, Pham J, Tian Y, Rosales SL, Seumois G, Sidney J, de Silva AD, Premkumar L, Collins MH, Stone M, Norris PJ, Romero CME, Durbin A, Ricciardi MJ, Ledgerwood JE, de Silva AM, Busch M, Peters

B, Vijayanand P, et al. 2018. Cutting edge: transcriptional profiling reveals multifunctional and cytotoxic antiviral responses of zika Virus-Specific CD8$^+$ T Cells. *The Journal of Immunology* **201**:3487–3491. DOI: https://doi.org/10.4049/jimmunol.1801090, PMID: 30413672

Harrington AM, Olteanu H, Kroft SH. 2012. A dissection of the CD45/side scatter "blast gate". *American Journal of Clinical Pathology* **137**:800–804. DOI: https://doi.org/10.1309/AJCPN4G1IZPABRLH, PMID: 22523220

Henry C, Ramadan A, Montcuquet N, Pallandre JR, Mercier-Letondal P, Deschamps M, Tiberghien P, Ferrand C, Robinet E. 2010. CD3+CD20+ cells may be an artifact of flow cytometry: comment on the article by wilk et al. *Arthritis & Rheumatism* **62**:2561–2563. DOI: https://doi.org/10.1002/art.27527, PMID: 20506165

HoWangYin KY, Caumartin J, Favier B, Daouya M, Yaghi L, Carosella ED, LeMaoult J. 2011. Proper regrafting of Ig-like transcript 2 after trogocytosis allows a functional cell-cell transfer of sensitivity. *The Journal of Immunology* **186**:2210–2218. DOI: https://doi.org/10.4049/jimmunol.1000547, PMID: 21242521

Jakubzick CV, Randolph GJ, Henson PM. 2017. Monocyte differentiation and antigen-presenting functions. *Nature Reviews Immunology* **17**:349–362. DOI: https://doi.org/10.1038/nri.2017.28, PMID: 28436425

Kitchen SG, Jones NR, LaForge S, Whitmire JK, Vu BA, Galic Z, Brooks DG, Brown SJ, Kitchen CM, Zack JA. 2004. CD4 on CD8(+) T cells directly enhances effector function and is a target for HIV infection. *PNAS* **101**:8727–8732. DOI: https://doi.org/10.1073/pnas.0401500101, PMID: 15173593

Kou Z, Quinn M, Chen H, Rodrigo WW, Rose RC, Schlesinger JJ, Jin X. 2008. Monocytes, but not T or B cells, are the principal target cells for dengue virus (DV) infection among human peripheral blood mononuclear cells. *Journal of Medical Virology* **80**:134–146. DOI: https://doi.org/10.1002/jmv.21051, PMID: 18041019

Kudernatsch RF, Letsch A, Stachelscheid H, Volk H-D, Scheibenbogen C. 2013. Doublets pretending to be CD34 + T cells despite doublet exclusion. *Cytometry Part A* **83A**:173–176. DOI: https://doi.org/10.1002/cyto.a.22247

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, 1000 Genome Project Data Processing Subgroup. 2009. The sequence alignment/Map format and SAMtools. *Bioinformatics* **25**:2078–2079. DOI: https://doi.org/10.1093/bioinformatics/btp352, PMID: 19505943

Lindestam Arlehamn CS, Gerasimova A, Mele F, Henderson R, Swann J, Greenbaum JA, Kim Y, Sidney J, James EA, Taplitz R, McKinney DM, Kwok WW, Grey H, Sallusto F, Peters B, Sette A. 2013. Memory T cells in latent mycobacterium tuberculosis infection are directed against three antigenic islands and largely contained in a CXCR3+CCR6+ Th1 subset. *PLOS Pathogens* **9**:e1003130. DOI: https://doi.org/10.1371/journal.ppat.1003130, PMID: 23358848

Lindestam Arlehamn CS, McKinney DM, Carpenter C, Paul S, Rozot V, Makgotlho E, Gregg Y, van Rooyen M, Ernst JD, Hatherill M, Hanekom WA, Peters B, Scriba TJ, Sette A. 2016. A quantitative analysis of complexity of human Pathogen-Specific CD4 T cell responses in healthy M. tuberculosis infected south africans. *PLOS Pathogens* **12**:e1005760. DOI: https://doi.org/10.1371/journal.ppat.1005760, PMID: 27409590

Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology* **15**:550. DOI: https://doi.org/10.1186/s13059-014-0550-8, PMID: 25516281

Monks CR, Freiberg BA, Kupfer H, Sciaky N, Kupfer A. 1998. Three-dimensional segregation of supramolecular activation clusters in T cells. *Nature* **395**:82–86. DOI: https://doi.org/10.1038/25764, PMID: 9738502

Nascimbeni M, Shin EC, Chiriboga L, Kleiner DE, Rehermann B. 2004. Peripheral CD4(+)CD8(+) T cells are differentiated effector memory cells with antiviral functions. *Blood* **104**:478–486. DOI: https://doi.org/10.1182/blood-2003-12-4395, PMID: 15044252

Pai M, Behr MA, Dowdy D, Dheda K, Divangahi M, Boehme CC, Ginsberg A, Swaminathan S, Spigelman M, Getahun H, Menzies D, Raviglione M. 2016. Tuberculosis. *Nature Reviews Disease Primers* **2**:16076. DOI: https://doi.org/10.1038/nrdp.2016.76, PMID: 27784885

Picelli S, Björklund ÅK, Faridani OR, Sagasser S, Winberg G, Sandberg R. 2013. Smart-seq2 for sensitive full-length transcriptome profiling in single cells. *Nature Methods* **10**:1096–1098. DOI: https://doi.org/10.1038/nmeth.2639, PMID: 24056875

Randolph GJ, Jakubzick C, Qu C. 2008. Antigen presentation by monocytes and monocyte-derived cells. *Current Opinion in Immunology* **20**:52–60. DOI: https://doi.org/10.1016/j.coi.2007.10.010, PMID: 18160272

Roy Chowdhury R, Vallania F, Yang Q, Lopez Angel CJ, Darboe F, Penn-Nicholson A, Rozot V, Nemes E, Malherbe ST, Ronacher K, Walzl G, Hanekom W, Davis MM, Winter J, Chen X, Scriba TJ, Khatri P, Chien YH. 2018. A multi-cohort study of the immune factors associated with M. tuberculosis infection outcomes. *Nature* **560**:644–648. DOI: https://doi.org/10.1038/s41586-018-0439-x, PMID: 30135583

Schmiedel BJ, Singh D, Madrigal A, Valdovino-Gonzalez AG, White BM, Zapardiel-Gonzalo J, Ha B, Altay G, Greenbaum JA, McVicker G, Seumois G, Rao A, Kronenberg M, Peters B, Vijayanand P. 2018. Impact of genetic polymorphisms on human immune cell gene expression. *Cell* **175**:1701–1715. DOI: https://doi.org/10.1016/j.cell.2018.10.022, PMID: 30449622

Seumois G, Vijayanand P, Eisley CJ, Omran N, Kalinke L, North M, Ganesan AP, Simpson LJ, Hunkapiller N, Moltzahn F, Woodruff PG, Fahy JV, Erle DJ, Djukanovic R, Blelloch R, Ansel KM. 2012. An integrated nano-scale approach to profile miRNAs in limited clinical samples. *American Journal of Clinical and Experimental Immunology* **1**:70–89. PMID: 23304658

Seumois G, Zapardiel-Gonzalo J, White B, Singh D, Schulten V, Dillon M, Hinz D, Broide DH, Sette A, Peters B, Vijayanand P. 2016. Transcriptional profiling of Th2 cells identifies pathogenic features associated with asthma. *The Journal of Immunology* **197**:655–664. DOI: https://doi.org/10.4049/jimmunol.1600397, PMID: 27271570

Sprangers S, Vries TJde, Everts V. 2016. Monocyte heterogeneity: consequences for Monocyte-Derived immune cells. *Journal of Immunology Research* **2016**:1–10. DOI: https://doi.org/10.1155/2016/1475435

Srivastava S, Ernst JD, Desvignes L. 2014. Beyond macrophages: the diversity of mononuclear cells in tuberculosis. *Immunological Reviews* **262**:179–192. DOI: https://doi.org/10.1111/imr.12217, PMID: 25319335

Thauland TJ, Parker DC. 2010. Diversity in immunological synapse structure. *Immunology* **131**:466–472. DOI: https://doi.org/10.1111/j.1365-2567.2010.03366.x, PMID: 21039474

Trapnell C, Pachter L, Salzberg SL. 2009. TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* **25**: 1105–1111. DOI: https://doi.org/10.1093/bioinformatics/btp120, PMID: 19289445

Tsai CY, Liong KH, Gunalan MG, Li N, Lim DS, Fisher DA, MacAry PA, Leo YS, Wong SC, Puan KJ, Wong SB. 2015. Type I IFNs and IL-18 regulate the antiviral response of primary human γδ T cells against dendritic cells infected with dengue virus. *The Journal of Immunology* **194**:3890–3900. DOI: https://doi.org/10.4049/jimmunol.1303343, PMID: 25732728

Wabnitz GH, Samstag Y. 2016. Multiparametric characterization of human T-Cell immune synapses by InFlow microscopy. *Methods in Molecular Biology* **1389**:155–166. DOI: https://doi.org/10.1007/978-1-4939-3302-0_10, PMID: 27460243

Weiskopf D, Angelo MA, de Azeredo EL, Sidney J, Greenbaum JA, Fernando AN, Broadwater A, Kolla RV, De Silva AD, de Silva AM, Mattia KA, Doranz BJ, Grey HM, Shresta S, Peters B, Sette A. 2013. Comprehensive analysis of dengue virus-specific responses supports an HLA-linked protective role for CD8+ T cells. *PNAS* **110**: E2046–E2053. DOI: https://doi.org/10.1073/pnas.1305227110, PMID: 23580623

Zak DE, Penn-Nicholson A, Scriba TJ, Thompson E, Suliman S, Amon LM, Mahomed H, Erasmus M, Whatney W, Hussey GD, Abrahams D, Kafaar F, Hawkridge T, Verver S, Hughes EJ, Ota M, Sutherland J, Howe R, Dockrell HM, Boom WH, et al. 2016. A blood RNA signature for tuberculosis disease risk: a prospective cohort study. *The Lancet* **387**:2312–2322. DOI: https://doi.org/10.1016/S0140-6736(15)01316-1

ISAC CYTOMETRY Journal of Quantitative Cell Science PART A

MIFlowCyt

The Challenge of Distinguishing Cell–Cell Complexes from Singlet Cells in Non-Imaging Flow Cytometry and Single-Cell Sorting

Julie G. Burel,[1*] Mikhail Pomaznoy,[1] Cecilia S. Lindestam Arlehamn,[1] Gregory Seumois,[1] Pandurangan Vijayanand,[1,2] Alessandro Sette,[1,2] Bjoern Peters[1,2]

[1]Division of Vaccine Discovery, La Jolla Institute for Immunology, La Jolla, California

[2]Department of Medicine, University of California San Diego, La Jolla, California

Received 31 January 2020; Revised 25 March 2020; Accepted 18 April 2020

Grant sponsor: National Institute of Allergy and Infectious Diseases, Grant number: S10OD016262, Grant number: S10OD021831, Grant number: U19AI118626; Grant sponsor: The Tullie Rickey Families SPARK awards at La Jolla Institute for Immunology

Additional Supporting Information may be found in the online version of this article.

*Correspondence to: Dr. Julie Burel, La Jolla Institute for Immunology, 9420 Athena Circle, La Jolla, CA 92037, USA. E Email: jburel@lji.org

Published online in Wiley Online Library (wileyonlinelibrary.com)

DOI: 10.1002/cyto.a.24027

Abstract

Our recent work has highlighted that care needs to be taken when interpreting single cell data originating from flow cytometry acquisition or cell sorting: We found that doublets of T cells bound to other immune cells are often present in the live singlet gate of human peripheral blood samples acquired by flow cytometry. This hidden "contamination" generates atypical gene signatures of mixed cell lineage in what is assumed to be single cells, which can lead to data misinterpretation, such as the description of novel immune cell types. Here, based on the example of T cell–monocyte complexes, we identify experimental and data analysis strategies to help distinguishing between singlets and cell–cell complexes in non-imaging flow cytometry and single-cell sorting. We found robust molecular signatures in both T cell–monocyte and T cell–B cell complexes that can distinguish them from singlets at both protein and mRNA levels. Imaging flow cytometry with appropriate gating strategy (matching the one used in cell sorting) and direct microscopy imaging after cell sorting were the two methods of choice to detect the presence of cell–cell complexes in suspicious dual-expressing cells. We finally applied this knowledge to highlight the likely presence of T cell–B cell complexes in a recently published dataset describing a novel cell population with mixed T cell and B cell lineage properties. © 2020 The Authors. *Cytometry Part A* published by Wiley Periodicals, Inc. on behalf of International Society for Advancement of Cytometry.

Key terms

flow cytometry; cell–cell complexes; doublets; single-cell immune profiling; single-cell RNA sequencing

MULTIPARAMETRIC flow cytometry is a powerful tool to unravel the phenotypic heterogeneity of immune cells in humans. When combined with cell sorting and sequencing, it can unravel both protein and RNA expression programs within cell populations, which has led to the discovery of many novel immune cell subsets and associated functions, in both healthy and disease settings (1). However, our recent work highlights an occasionally underappreciated challenge, namely that care needs to be taken when interpreting single-cell data originating from flow cytometry acquisition or cell sorting: We found that when analyzing human peripheral blood mononocular cells (PBMC), a small but reproducible proportion of presumed singlets by flow cytometry are tightly bound cell–cell complexes. These "contaminating" dual cell complexes can mislead subsequent interpretation of what is presumed to be single-cell data.

We first identified by flow cytometry a cell population in the live singlet gate of human PBMC from patients with dual expression for CD3 and CD14 (2). We found that the frequency of these CD3+CD14+ cells was modulated as a result of immune perturbations such as vaccination, disease treatment and disease severity. This cell

population expressed pan-markers of monocytes and T cells, both at the protein and the mRNA level, initially suggesting the discovery of a novel cell type with both T cell and monocyte lineage properties. However, subsequent analyses revealed that the CD3+CD14+ cell population did not consist of single cells bearing both T cell and monocyte lineage markers, but were either dual cell–cell complexes of T cells and monocytes, or T cells bound to smaller particles containing monocyte markers. Neither of these types of complexes was removed by conventional forward and side scatter gating approaches to avoid cell aggregates in flow cytometry. Importantly, the T cell–monocyte complexes we detected showed LFA-1/ICAM-1 polarization at their point of contact, could be isolated from fresh PBMC and whole blood, and were stable over time within a given individual, suggesting their presence in vivo, and not resulting from ex vivo sample manipulation. This makes studying the presence and composition of these complexes biologically important, and clearly refute our first interpretation that CD3+CD14+ cells could represent a novel cell type with mixed lineage properties.

Since T cell interactions are not restricted to monocytes, we expect to find them in complexes with other types of antigen-presenting cells (APCs). Indeed, others have previously reported on CD3+CD20+ cells detected by flow cytometry in human peripheral blood as doublets of T cells and B cells (3). Moreover, CD4+CD19+ cells have also been detected in draining lymph nodes of mice following infection, and found to be complexes of T follicular helper (Tfh) cells and B cells (4). Strikingly, the polarization of the Tfh cell and the immunoglobulin isotype class switch in the B cell were matching in each conjugate, and B cells in the conjugates were associated with a greater number of somatic hypermutations compared to singlets. Taken together, these results support our hypothesis that functional complexes of T cells and B cells are present in vivo and can be detected ex vivo by flow cytometry.

Here, we investigate how the presence of T cell–APC complexes in presumed single cell populations in flow cytometry can impact both data analysis and interpretation. We also highlight several experimental and data analysis strategies that can help identify complexes and thus avoid misinterpretations. We finally apply this strategy to follow up on a recently published study describing a novel immune cell population with T cell–B cell mixed lineage properties.

Results

First, we assessed the impact of the presence of complexes on the resulting protein and gene expression profiling data generated with single cell techniques such as flow cytometry and single-cell sort RNA sequencing. We took as an example CD3+CD14+ cells, since we have already validated in our previous study that this cell population mainly consist of T cell–monocyte complexes and not CD3/CD14 dual expressers (DEs). We elected the same gating strategy to identify CD3+CD14+ cells as in our previous study (Fig. 1A). A vs W was arbitrarily chosen to exclude doublets (since we have previously shown that T cell–monocyte complexes are consistently contained within the singlet gate regardless of the use of A vs H, A vs W or H vs W (2)). Expression of T cell and monocyte markers in CD3+CD14+ cells was compared to singlets monocytes and T cells at both protein and mRNA level. Since T cell and monocyte markers are exclusively expressed by each cell type, when analyzed by flow cytometry, the CD3+CD14+ population shows similar expression levels of canonical markers for T cells and monocytes when compared to their singlet counterpart (Fig. 1B). Similarly, at the mRNA level, when we performed single-cell sorts of the CD3+CD14+ population and subsequent RNA sequencing analysis of these cells, they showed positive expression for both T cell and monocyte markers (Fig. 1C) and in PCA they represented a "halfway" population between T cells and monocytes (Fig. 1D). This reflects the simultaneous detection of the RNA content of both the T cell and the monocyte present in each sorted complex. Taken together, solely based on typical protein and gene expression profiles, complexes that are present within the singlet gate population in flow cytometry will appear like a composite cell population combining features of its two cell type components, similar to what a DE would look like.

Next, we looked for additional parameters in data generated from flow cytometry or single-cell RNAseq (scRNAseq) that could differentiate complexes from singlets. We have found it overall challenging to differentiate between T cell-monocyte complexes from singlets using non-imaging flow cytometry. Nevertheless, our initial study identified a couple of parameters that could partially distinguish doublets from singlets, even within the live singlet population (2). In flow cytometry, whereas FSC-A vs SSC-A gating cannot completely separate CD3+CD14+ cells from bigger singlet cells such as monocytes, median fluorescence intensity (MFI) for FSC-A or SSC-A is increased for T cell-monocyte complexes when compared to singlet T cells or monocytes (Fig. 2A). Thus, a stringent FSC-A vs SSC-A gating on lymphocytes such as T cells will remove most of cell–cell complexes, but will also exclude bigger lymphocytes such as activated T cells, and the entire monocyte population (Fig. 2A). Similarly, despite overlapping gates, MFI for CD45 is increased in T cell-monocyte complexes compared to singlets, as both cells express this marker (Fig. 2B). In RNAseq, TPM-transformed expression values identify the relative concentration of different RNA species rather than providing an absolute quantification (such as in flow cytometry). Thus, in scRNAseq, markers that are exclusively found in T cells or exclusively found in monocytes are about half that expression level when evaluating complexes compared to singlet T cells and monocytes (Fig. 2C). Whereas these parameters are not sufficient to completely separate cell–cell complexes from single cells, they together represent useful signatures at both protein and mRNA level that can be checked by investigators after the data has been generated, and prompt for further cell–cell complexes analyses.

The by far preferred method to identify the presence of cell–cell complexes in a previously uncharacterized cell

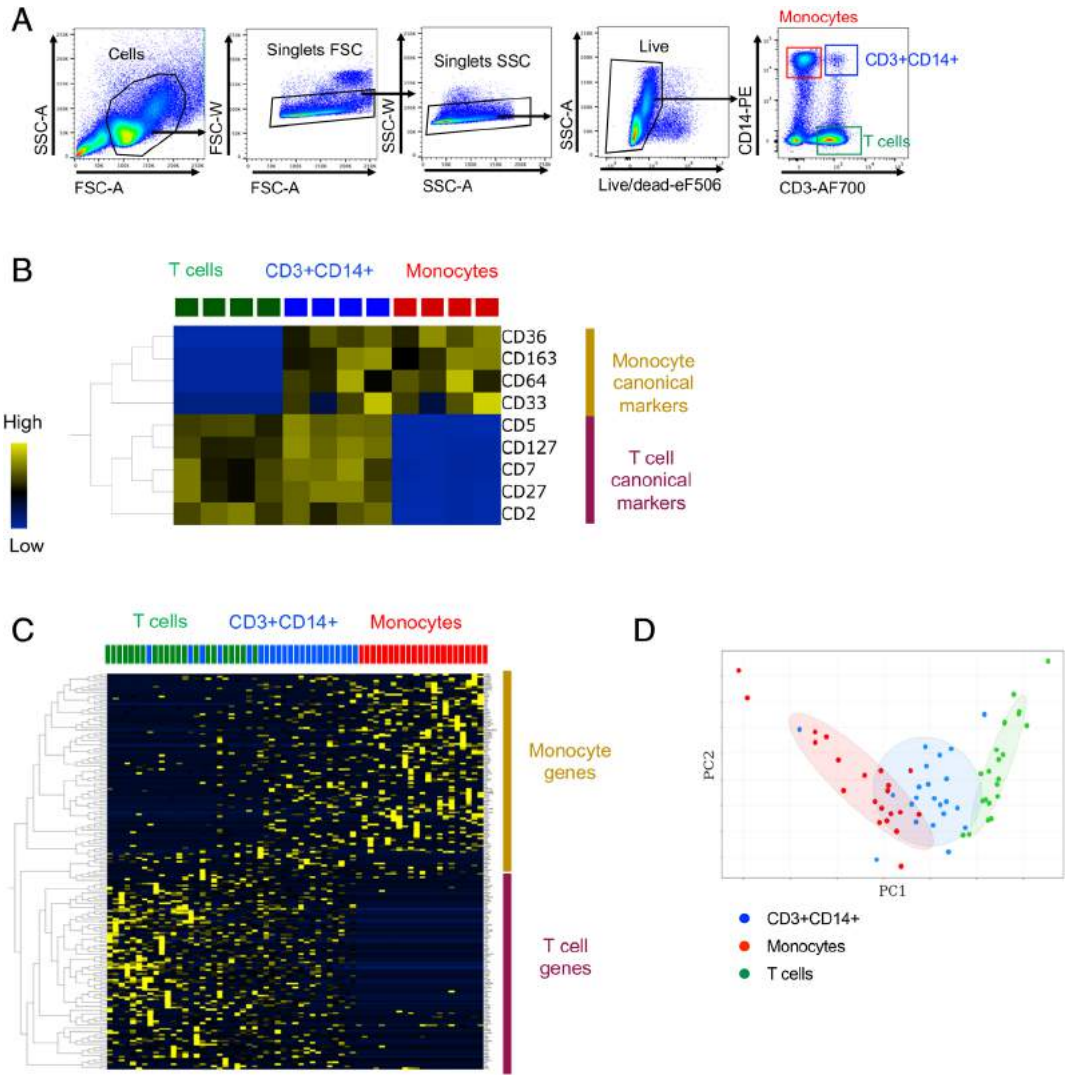


Fig 1. Cell–cell complexes appear like a "halfway" cell population with mixed lineage properties from its two cell type components. **A**, Gating strategy to identify T cells, B cells and CD3+CD14+ cells in human PBMC. **B**, Heatmap representing the protein expression of T cell and monocyte canonical markers in T cells, monocytes, and CD3+CD14+ T cells. Expression was measured by non-imaging flow cytometry and expressed as median fluorescent intensity. **C**, Heatmap representing single cell gene expression of the 100 genes most upregulated in monocytes vs T cells and vice versa for sorted monocytes, T cells, and CD3+CD14+ cells. **D**, PCA on all genes from the scRNAseq data of all three sorted populations. T cells, monocytes, and CD3 + CD14+ cells were identified as represented in Figure 1A. (A and B) Data were derived from PBMC of four healthy subjects. (C and D) Data were derived from 21 monocytes, 22 T cells, and 22 CD3 +CD14+ cells isolated from PBMC of one healthy subject, and are available under GEO accession number GSE117435.

population is to use imaging flow cytometry. The typical gating strategy is, similarly to non-imaging flow cytometry (Fig. 1A): First identify cells, followed by dead cells and doublets exclusion, to finally identify the double positive population (optimal [OPT] imaging gating, Figure 3A). However, imaging flow cytometry can identify doublets in a much more dispositive fashion than a standard flow cytometer by generating metrics derived from an actual picture of the event detected, which allows for the calculation of brightfield area and aspect ratio parameters. Additionally, the accuracy of the doublet exclusion gating can be cross checked by viewing an image gallery of all gated objects. When we applied the OPT imaging gating strategy to the detection of CD3+CD14+ cells, it detected a much lower frequency of double positive events compared to non-imaging flow cytometry (Fig. 3C). This confirms that the doublet exclusion in the imaging flow cytometer is far more sensitive than what we can expect from cell sorters relying on FSC and SSC based doublet exclusion, and results from the two cannot be directly compared. The first gating step for cell selection in imaging flow cytometry (SSC-A vs brightfield Area) is sufficient to eliminate bigger aggregates (but not tight doublets) in a similar fashion that A vs. W, W vs. H or A vs. H would perform in non-imaging flow cytometry (Fig. 3B). Using this cell sorting matching (CSM) gating strategy, we identified frequencies of CD3 +CD14+ cells similar to what cell sorters detected

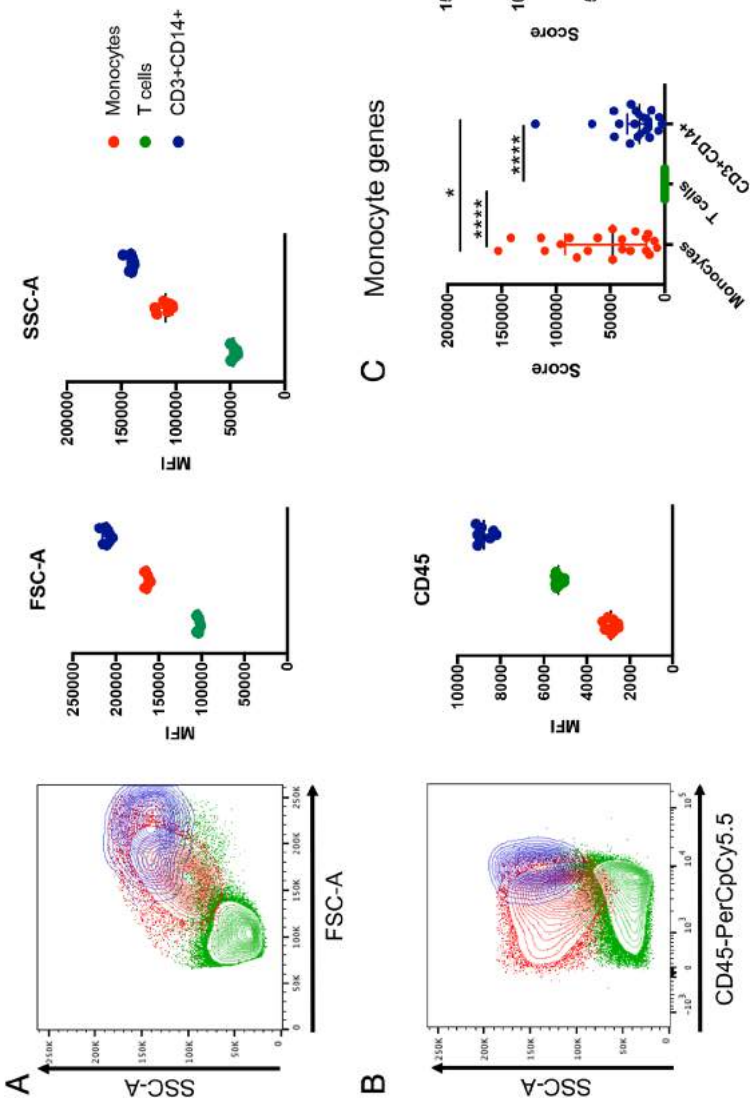


Fig 2. Molecular signatures of cell–cell complexes in flow cytometry and single-cell RNAseq derived from cell sorting. Two-dimensional representation plots and median fluorescence intensity (MFI) of **A**, FSC-A and SSC-A parameters and **B**, CD45 expression in monocytes, T cells, and CD3+CD14+ cells in non-imaging flow cytometry. Data were derived from PBMC of eight healthy subjects. **C**, Monocyte and T-cell genes scores derived from the scRNAseq data of sorted monocytes, T cells, and CD3+CD14+ cells as identified in Figure 1A. Data were derived from 21 monocytes, 22 T cells, and 22 CD3+CD14+ cells isolated from PBMC of one healthy subject, and are available under GEO accession number GSE117435.

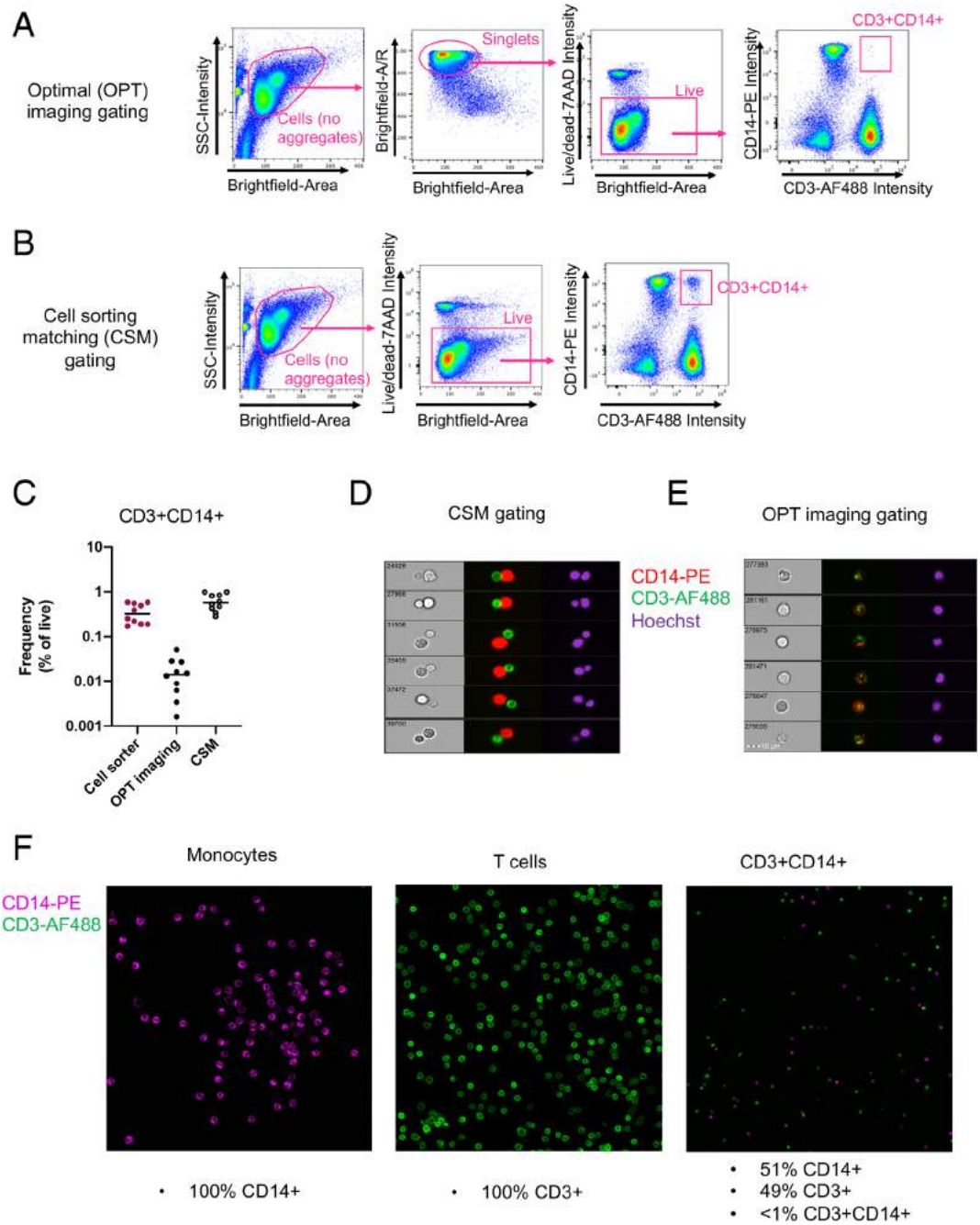


Fig 3. Imaging analyses to investigate the presence of cell–cell complexes in suspicious dual-expressing cell populations identified by non-imaging flow cytometry. **A**, Optimal (OPT) imaging and, **B**, cell sorting matching (CSM) gating strategies to identify CD3+CD14+ cells with imaging flow cytometry. **C**, Frequencies of CD3+CD14+ cells detected by imaging flow cytometry using the OPT imaging vs the CSM gating strategies, in comparison to cell sorting. Random galleries of CD3+CD14+ events identified by imaging flow cytometry with **D**, the CSM gating strategy and **E**, the OPT imaging gating strategy. **F**, Confocal microscopy analysis of sorted monocytes, T cells, and CD3+CD14+ cells as gated in Figure 1A. Data were derived from PBMC of 10 healthy subjects.

(Fig. 3C). Importantly, when exporting a gallery of events for both gating strategies, we identified that the CSM gating strategy identified mostly T cell–monocyte complexes (Fig. 3D), whereas the OPT imaging gating strategy identified predominantly DEs (Fig. 3E). Thus, the CSM gating strategy is more closely mimicking the population detected by cell sorters and non-imaging flow analyzers, and is a more reliable way to identify the possible presence of cell complexes and their relative proportion over true DEs within an unknown dual-positive cell population. In the case of CD3+CD14+ cells, this experiment demonstrates that whereas DEs do exist, the vast majority of dual positive events

identified by flow cytometry and cell sorting represent T cell–monocyte complexes.

To circumvent the hurdles associated with matching gating strategies between imaging flow cytometry and cell sorting aforementioned, an alternative experiment is to directly sort the dual-expressing cell population onto a slide and analyze them using microscopy. In the case of CD3+CD14+ cells, we observed that cells identified as DEs by FACS represented a perfect stoichiometric mix of individual T cells and monocytes, but not DEs under the microscope (Fig. 3F).

Finally, we applied our expertise in T cell-monocyte complexes detection to other types of cell–cell complexes that can be present in PBMC. A recent paper published in Cell (5) describes the discovery of a novel immune cell type expressing both T cell and B cell receptors in the peripheral blood of Type I diabetes patients. These findings reminded us of our initial assumption that we had identified a novel cell population with both T cell and monocyte lineage properties. Thus, we next wondered whether the DEs cells identified by Ahmed et al. indeed represented a completely novel population of a combined T cell and B cell lineage, or if they might be impacted by complexes of T cells and B cells tightly bound together that are circulating in vivo.

Following the same gating strategy as in the Ahmed et al. study, we were successfully able to detect DE cells (CD5+CD19+TCRab+ population within the live singlet gate) in human PBMC from healthy subjects (Fig. 4A). Like CD3+CD14+ cells, DE cells were on average associated with higher FSC/SSC values, as well as higher CD45 staining intensity compared to T cells, B cells and CD5+CD19+TCRab− cells (Fig. 4B,C). Imaging flow cytometry using the CSM gating strategy identified DE cells at a similar frequency traditional flow cytometry (Fig. 4D), and revealed that DE cells predominantly consist of T cell–B cell complexes (Fig. 4E). Contrastingly, the OPT imaging gating strategy identified a much lower frequency of double positive events (Fig. 4D), all of which were DEs (Fig. 4F). Thus, in our hands, DE cells in human PBMC display imaging and non-imaging flow cytometry characteristics of dual cell–cell complexes similar to what we found for T cell–monocyte complexes.

When examining the scRNAseq data generated by Ahmed et al. (GEO accession number GSE129112), we found that DE cells have a significantly lower RNA concentration for any marker unique to either T cells or B cells (Fig. 4G). Thus, the reported scRNAseq signature of sorted DE cells in Ahmed et al. is also consistent with cell–cell complexes.

Discussion

In this study, we demonstrated that cell–cell complexes pairing a T cell with another immune cell type such as a monocyte or a B cell appear at first glance to be a distinct "halfway" cell population with mixed lineage features at both protein and mRNA level. Further analyses identified key parameters from flow cytometry and scRNAseq data that are specific to cell complexes and thus indicate their presence within any cell population identified by flow cytometry. We also suggested experiments that can be used by researchers facing a case of suspicious dual-expressing cells, including imaging flow cytometry with appropriate gating strategy and direct microscopy analysis of sorted cells. Applying this knowledge to the DE cells identified in the study of Ahmed et al, we found that DE cells from healthy PBMC have a flow cytometry signature of doublets, which was also confirmed by imaging flow cytometry. Importantly, scRNAseq data from Ahmed et al also showed a DE signature of doublets. This observation emphasizes the importance of systematically and rigorously assessing the presence of cell–cell complexes in newly described cell populations, especially in cases where they display mixed lineage properties.

So why do we think that the Ahmed et al. paper might have overlooked the presence of complexes in their DE cell population expressing both T cell and B cell markers? The paper does show convincing evidence that single DE cells are present in imaging flow cytometry data. While the imaging data presented in the paper is limited to a single cell, the staining for both TCR and IgD is uniform on the cell surface, and it represents a convincing DE. Thus, our concern is not that the DE cell presented in the Ahmed study is not real, but rather that large parts of their subsequent experiments were solely based on non-imaging flow cytometry and the assumption that single-cell sorted DE cells are singlets. Given the limitations of non-imaging flow cytometry to distinguish cell–cell complexes from singlet cells aforementioned, this is potentially problematic. Their gating strategy for imaging flow cytometry was described as "focused cells were selected on the basis of gradient RMS and an aspect ratio that was consistent with single cell events and devoid of debris or multicellular events (doublets)" (5). Thus, it appears that the optimal imaging gating strategy was used here, which we have shown does not reflect adequately what is being detected by cell sorters. In this study, we have examined PBMCs from healthy donors and found that in non-perturbed settings, DE cells do have a flow cytometry signature of T cell-B cell complexes, with few dual expressing cells. The presence of DEs in the DE cell population could thus be unique to the Type I diabetes patients included in the Ahmed et al. study. Nevertheless, the fact that sorted DE cells in the Ahmed study have a scRNA-seq signature of complexes suggest that our conclusion might also hold for the Ahmed et al. study. Besides, the DEs identified by imaging flow cytometry might not necessarily represent a novel immune cell type, but could also be conventional T cells or B cells that have acquired the other lineage markers at the protein level following cell–cell interaction by trogocytosis, a common phenomenon described in immune cells (6, 7). The occurrence of such event could be estimated by collecting larger galleries of dual-expressing cells, and design imaging features to discriminate cells with uniform membrane staining vs areas of capping (the later suggesting trogocytosis).

Overall, given the challenges of distinguishing true DE single cells from cell–cell complexes in non-imaging flow cytometry and cell sorting, we think it is extremely important to include appropriate controls in future publications

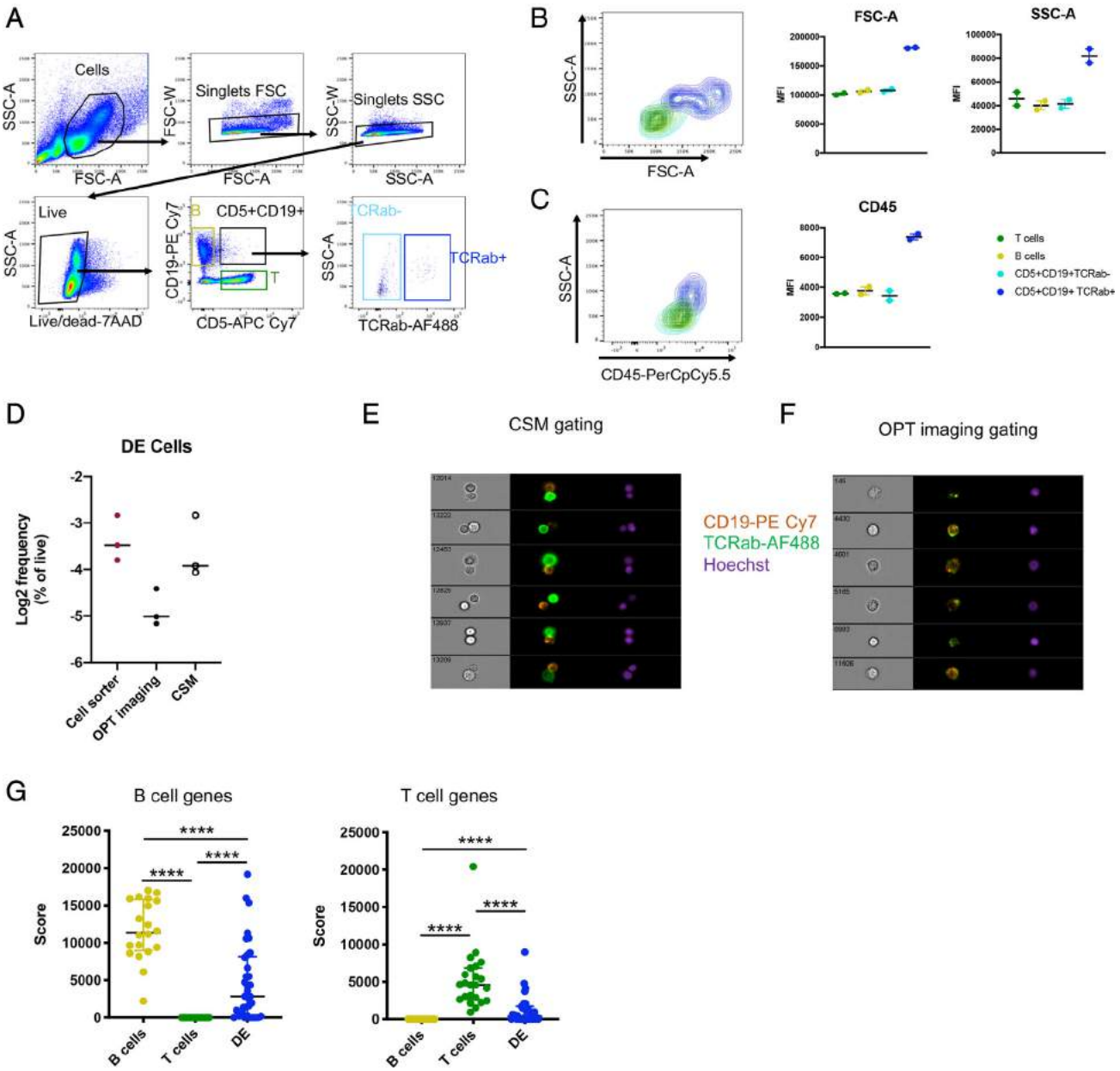


Fig 4. DE cells (CD5 + CD9 + TCRab+) in human PBMC have a flow cytometry and single- cell RNAseq signature of cell–cell complexes. **A**, Gating strategy to identify DE cells (CD5+CD19+TCRab+), CD5+CD19+TCRab- cells, T cells, and B cells in human PBMC. Representative contour plot and median fluorescent intensity (MFI) of **B**, FSC-A vs SSC-A and, **C**, CD45 in DE cells, CD5+CD19+TCRab– cells, T cells, and B cells measured by non-imaging flow cytometry. **D**, Frequencies of DE cells detected by imaging flow cytometry using the optimal (OPT) imaging and the cell sorting matching (CSM) gating strategies, in comparison to cell sorting. Random galleries of DE cells identified by imaging flow cytometry with **E**, the OPT imaging gating and, **F**, the CSM gating strategy. **G**, B cell and T cell gene scores derived from the scRNAseq data of sorted B cells, T cells, and DE cells as reported in (5), and available under GEO accession number GSE129112. Data were derived from (A–C) five healthy subjects and (D–F) three healthy subjects.

reporting on single cell immune profiling by flow cytometry. The presence of cell–cell complexes signatures in flow cytometry and scRNAseq data should be systematically checked and reported. Additional experiments such as imaging flow cytometry (with appropriate gating strategy) and direct microscopy analysis on sorted cells should be performed. The elegant study of Popescu et al. demonstrates the power of combining scRNAseq with imaging flow cytometry to confirm cellular identity and determine doublet prevalence across human fetal liver cell types (8). Imaging cell sorters are being developed but have not reached yet the speed and sensitivity required for the detection of rare events, including cell–cell complexes. Until such technology is made commercially available, these recommendations will be crucial to ensure reliable interpretation of

single cell datasets derived from flow cytometry, especially those reporting on the discovery of novel immune cell populations.

Material and Methods

Subjects and Samples

Blood samples were obtained from the University of California San Diego anti-viral research center clinic, and the normal blood donor program at La Jolla Institute for Immunology under ethical approval from the La Jolla Institute for Immunology Institutional Review Board. All participants provided written informed consent prior to participation in the study. PBMC were obtained from healthy subjects by density gradient centrifugation (Ficoll–Hypaque, Amersham Biosciences) from leukapheresis or whole blood samples, according to the manufacturer's instructions. Cells were resuspended to 10–50 million cells/ml in FBS (Gemini Bio-Products) containing 10% dimethyl sulfoxide (Sigma) and cryopreserved in liquid nitrogen. Cryopreserved PBMC were quickly thawed by incubating each cryovial at 37°C for 2 min, and cells transferred into 9 ml of cold medium (RPMI 1640 with L-Glutamin and 25 mM Hepes (Omega Scientific), supplemented with 5% human AB serum (GemCell), 1% Penicillin Streptomycin (Gibco) and 1% Glutamax (Gibco)) and 20 U/ml Benzonase Nuclease (Millipore) in a 15-ml conical tube. Cells were centrifuged and resuspended in medium to determine cell concentration and viability using Trypan blue and a hematocytometer and kept on ice until further analysis.

Non-imaging Flow Cytometry

PBMC staining and acquisition were performed as described in (2). Briefly, cryopreserved PBMC were quickly thawed as described above, then cells (1–10 million) were transferred into a 15-ml conical tube, centrifuged, resuspended in 100 µl of PBS containing 10% FBS and incubated for 10 min at 4°C. Cells were then stained with 100 µl of PBS containing antibody cocktails 1, 2, or 3 for 20 min at room temperature. Cocktail 1 was used to identify CD3+CD14+ cells and T cell markers and contained 3 µl of anti-humanCD3-AF700 antibody (clone UCHT1, Biolegend), 1 µl of anti-humanCD14-PE antibody (clone 61D3, eBiosciences), 3 µl of anti-humanCD45-PerCpCy5.5 (Clone HI30, TONBO Biosciences), 3 µl of anti-humanCD2-BV421 antibody (clone RPA-2.10, Biolegend), 3 µl of anti-humanCD7-APC antibody (clone CD7-6B7, Biolegend), 2 µl of anti-humanCD27-BV650 antibody (clone O323, Biolegend), 2 µl of anti-humanCD127-PE Cy7 antibody (clone eBioRDR5, eBisoceinces) and 0.2 µl of fixable viability dye eF506 (eBiosciences). Cocktail 2 was used to identify CD3 + CD14+ cells and monocyte markers and contained 3 µl of anti-humanCD3-AF700 antibody (clone UCHT1, Biolegend), 1 µl of anti-humanCD14-PE antibody (clone 61D3, eBiosciences), 3 µl of anti-humanCD45-PerCpCy5.5 (Clone HI30, TONBO Biosciences), 1 µl of anti-humanCD33-APC (clone WM53, Biolegend), 2 µl of anti-humanCD36-APC Cy7 (clone 5–271, Biolegend), 1 µl of anti-humanCD64-AF488 (clone 10.1, Biolegend), 2 µl of anti-humanCD163-PE Cy7 (clone GHI/61, Biolegend) and 0.2 µl of fixable viability dye eF506 (eBiosciences). Cocktail 3 was used to identify DE cells and contained 3 µl of anti-human-CD5-APCeF780 antibody (clone L17F12, Biolegend), 2 µl of anti-humanCD19-PECy7 antibody (clone HIB19, TONBO biosciences), 3 µl of anti-humanTCRab-AF488 antibody (clone IP26, Biolegend) and 7-AAD at 1 µg/ml (Biolegend). After two washes in staining buffer (PBS containing 0.5% FBS and 2 mM EDTA (pH 8.0), cells were resuspended into 100–500 µl of staining buffer, transferred into a 5 ml polypropylene FACS tube (BD Biosciences) and stored at 4°C protected from light for up to 4 h until flow cytometry acquisition. Acquisition was performed on a BD FACSAria III cell sorter (BD Biosciences). Performance of the cell sorter was checked daily by the flow cytometry core at La Jolla Institute for Immunology with the use of CS&T beads (BD Biosciences), and PMT voltages were manually adjusted for optimum fluorescence detection on each experiment day. Compensation was realized with single-stained beads (UltraComp eBeads, eBiosciences) in PBS using the same antibody volume as for the cell staining. Gating strategies are available on Figure 1A for T cells, monocytes and CD3+CD14 + cells and on Figure 4A for T cells, B cells, and DE cells.

Imaging Flow Cytometry

For the visualization of CD3+CD14+ cells, frozen PBMC were thawed and stained with 2 µl of anti-humanCD3-AF488 (clone UCHT1, Biolegend) and 1 µl of anti-humanCD14-PE (clone 61D3, eBiosciences) as described in the flow cytometry section above. For the visualization of DE cells, frozen PBMC were thawed and stained with 3 µl of anti-human-CD5-APCeF780 antibody (clone L17F12, Biolegend), 2 µl of anti-humanCD19-PECy7 antibody (clone HIB19, TONBO biosciences), and 3 µl of anti-humanTCRab-AF488 antibody (clone IP26, Biolegend), as described in the flow cytometry section above. After two washes in PBS, cells were resuspended to 10×10^6 cells/ml in staining buffer containing 5 µg/ml Hoechst (Invitrogen) and 1 µg/ml 7-AAD(Biolegend) and stored at 4°C protected from light until acquisition. Acquisition was performed with ImageStreamX MkII (Amnis) and INSPIRE software version 200.1.620.0 at 40× magnification and the lowest speed setting. A minimum of 4,000 CD3+CD14+ or CD5+CD19+ events in focus were collected. Compensation was performed using single stained cells. Data analysis was performed using IDEAS version 6.2.183.0.

Cell Sorting

PBMC were stained with fixable viability dye eFluor506 (eBiosciences), 3 µl of anti-humanCD3-AF700 antibody (clone UCHT1, Biolegend) and 1 µl of anti-humanCD14-PE antibody (clone 61D3, eBiosciences), as described in the flow cytometry section above. T cells, monocytes and CD3+CD14 + cells were identified based on the gating strategy presented in Figure 1A. Sorts were performed on a BD Aria III cell sorter directly into a 96-well PCR plate containing 4 µl cell lysis buffer for single cell sorts, or in Eppendorf tubes containing staining buffer for bulk sorts.

Single Cell RNA Sequencing

Single cell RNA sequencing was performed using Smart-seq2 as previously described (9). RNA loss was minimized by performing on-plate RNA capture, reverse-transcription and whole transcriptome pre-amplification (24 cycles) that results in ~1–30 ng of cDNA as previously described (9). Of note, 0.3–0.5 ng of pre-amplified cDNA was used to generate barcoded Illumina sequencing libraries (Nextera XT library preparation kit, Illumina) in an 8 μl reaction volume. Multiple quality-control steps were included to ensure consistency during the procedure for all samples. Samples that failed at quality-control steps as described in (9, 10), were eliminated from downstream procedures and analysis. Libraries were pooled and sequenced on the HiSeq2500 Illumina platform to obtain more than 2 million 50-bp single end reads per cell. Single cell RNA sequencing data were mapped against the human hg19 reference genome using TopHat (v1.4.1.,—library-type fr-secondstrand -C) and Gencode version 19 (GRCh37.p13) as gene model reference for alignment. Sequencing read coverage per gene was counted using HTSeq-count (-m union -s yes -t exon -i gene_id, http://www.huber.embl.de/users/anders/HTSeq/). Counts per gene are obtained by counting all the transcripts mapping to a gene and are together referred to as transcript. To normalize for sequencing depth and varying transcript length counts were TPM (transcripts per million reads) transformed. For PCA analysis, only genes with at least 20 TPM were considered. PCA was performed in Python using SciPy SVD algorithm and plotted with Matplotlib (11). Cell-type specific scores were calculated by summing the TPM counts of 100 selected genes for each cell type. Monocytes genes were selected as the 100 first genes with the highest TPM counts in monocytes, and null expression in T cells. Similarly, T cells genes were selected as the 100 first genes with the highest TPM counts in monocytes, and null expression in monocytes. Raw and normalized data are available under accession number GSE117435. T cell and B cell scores were calculated in a similar fashion, using the data available under GEO accession number GSE129112.

Microscopy

CD3−CD14+ monocytes, CD3+CD14− T cells and CD3+CD14+ cells were bulk sorted as described in the cell sorting section, and then transferred into individual chambers of a microscope slide. Fluorescence signals corresponding to CD3 and CD14 were detected with spectrally tuned PMTs using corresponding excitation lines of a Zeiss LSM 880 confocal microscope.

Acknowledgments

We thank Dr. Cheryl Kim and all present and past members at the flow cytometry core facility at the La Jolla Institute for Immunology for assistance in cell sorting and technical discussion. We thank Dr. Zbigniew Mikulski from the microscopy core at the La Jolla Institute for Immunology for assistance and technical advice on microscopy imaging. We thank Yoav Altman at the Sanford Burnham Prebys flow cytometry core for technical assistance with imaging flow cytometry. Research reported in this manuscript was supported by the National Institute of Allergy and Infectious Diseases division of the National Institutes of Health under award numbers U19AI118626, S10OD021831 and S10OD016262, and by The Tullie and Rickey Families SPARK Awards at La Jolla Institute for Immunology. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. Imaging flow cytometry was supported by the James B. Pendleton Charitable trust.

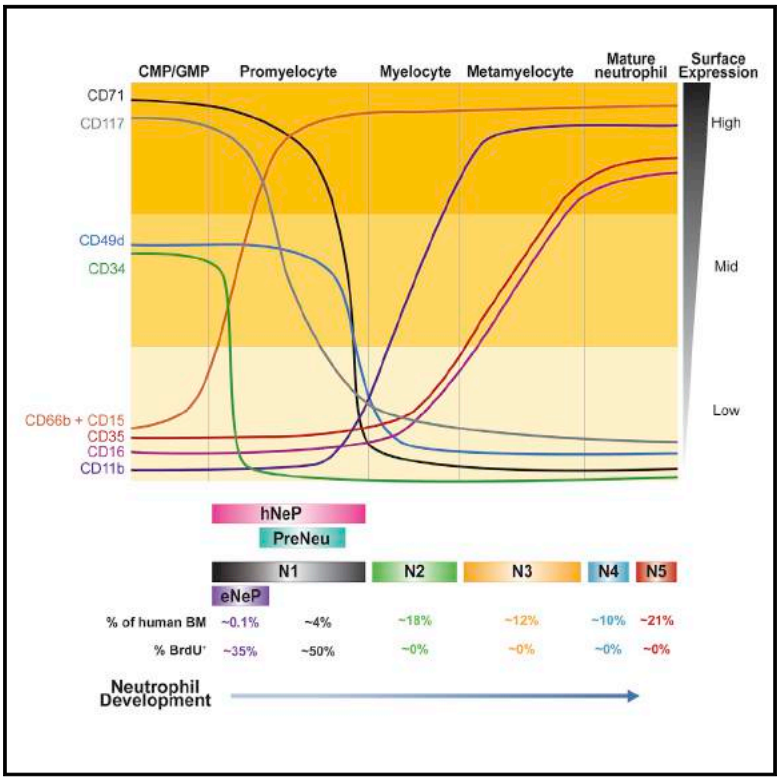
LITERATURE CITED

1. Stubbington MJT, Rozenblatt-Rosen O, Regev A, Teichmann SA. Single-cell transcriptomics to explore the immune system in health and disease. Science 2017;358(6359):58–63.
2. Burel JG, Pomaznoy M, Lindestam Arlehamn CS, Weiskopf D, da Silva AR, Jung Y, et al. Circulating T cell-monocyte complexes are markers of immune perturbations. Elife 2019;8:e46045.
3. Henry C, Ramadan A, Montcuquet N, Pallandre JR, Mercier-Letondal P, Deschamps M, Tiberghien P, Ferrand C, Robinet E. CD3+CD20+ cells may be an artifact of flow cytometry: Comment on the article by Wilk et al. Arthritis Rheum 2010;62(8):2561–2563. author reply 3-5.
4. Reinhardt RL, Liang HE, Locksley RM. Cytokine-secreting follicular T cells shape the antibody repertoire. Nat Immunol 2009;10(4):385–393.
5. Ahmed R, Omidian Z, Giwa A, Cornwell B, Majety N, Bell DR, et al. A public BCR present in a unique dual-receptor-expressing lymphocyte from type 1 diabetes patients encodes a potent T cell autoantigen. Cell 2019;177(6):1583–99.e16.
6. Daubeuf S, Lindorfer MA, Taylor RP, Joly E, Hudrisier D. The direction of plasma membrane exchange between lymphocytes and accessory cells by trogocytosis is influenced by the nature of the accessory cell. J Immunol (Baltimore, Md 1950) 2010;184(4):1897–1908.
7. HoWangYin KY, Caumartin J, Favier B, Daouya M, Yaghi L, Carosella ED, et al. Proper regrafting of Ig-like transcript 2 after trogocytosis allows a functional cell-cell transfer of sensitivity. J Immunol (Baltimore, Md 1950) 2011;186(4):2210–2218.
8. Popescu DM, Botting RA, Stephenson E, Green K, Webb S, Jardine L, Calderbank EF, Polanski K, Goh I, Efremova M, et al. Decoding human fetal liver haematopoiesis. Nature 2019;574(7778):365–371.
9. Picelli S, Bjorklund AK, Faridani OR, Sagasser S, Winberg G, Sandberg R. Smart-seq2 for sensitive full-length transcriptome profiling in single cells. Nat Methods 2013;10(11):1096–1098.
10. Patil VS, Madrigal A, Schmiedel BJ, Clarke J, O'Rourke P, de Silva AD, et al. Precursors of human CD4(+) cytotoxic T lymphocytes identified by single-cell transcriptome analysis. Sci Immunol 2018;3(19):eaan8664.
11. Hunter JD. Matplotlib: A 2D graphics environment. Comput Sci Eng 2007;9(3):90–95.

Immunity

Coexpression of CD71 and CD117 Identifies an Early Unipotent Neutrophil Progenitor Population in Human Bone Marrow

Graphical Abstract



Highlights

- Mass cytometry reveals a very early unipotent human bone marrow neutrophil progenitor
- The eNeP discovery provides a complete, unified model of human neutrophil development
- eNePs are defined by the distinct surface protein markers CD71 and CD117
- CD71^{+} neutrophils are expanded in blood and tumor of cancer patients

Authors

Huy Q. Dinh, Tobias Eggert, Melissa A. Meyer, ..., Ryan Llewellyn, Runpei Wu, Catherine C. Hedrick

Correspondence

hedrick@lji.org

In Brief

Human neutrophils are produced by bone marrow progenitors that were recently identified, but their developmental lineage remains unclear. Dinh et al. discover an early neutrophil progenitor population (termed "eNeP") marked by CD117^{+}CD71^{+} in human bone marrow and provide evidence for neutrophil progenitor expansion in human cancers.

Dinh et al., 2020, Immunity *53*, 319–334
August 18, 2020
https://doi.org/10.1016/j.immuni.2020.07.017

Immunity

CellPress

Article

Coexpression of CD71 and CD117 Identifies an Early Unipotent Neutrophil Progenitor Population in Human Bone Marrow

Huy Q. Dinh,[1,2] Tobias Eggert,[1,2] Melissa A. Meyer,[1,2] Yanfang Peipei Zhu,[1] Claire E. Olingy,[1] Ryan Llewellyn,[1] Runpei Wu,[1] and Catherine C. Hedrick[1,3,*]

[1]Center for Cancer Immunotherapy & Center for Autoimmunity and Inflammation, La Jolla Institute for Immunology, La Jolla, CA 92037, USA
[2]These authors contributed equally
[3]Lead Contact
*Correspondence: hedrick@lji.org
https://doi.org/10.1016/j.immuni.2020.07.017

SUMMARY

Neutrophils are the most abundant peripheral immune cells and thus, are continually replenished by bone marrow-derived progenitors. Still, how newly identified neutrophil subsets fit into the bone marrow neutrophil lineage remains unclear. Here, we use mass cytometry to show that two recently defined human neutrophil progenitor populations contain a homogeneous progenitor subset we term "early neutrophil progenitors" (eNePs) ($Lin^{-}CD66b^{+}CD117^{+}CD71^{+}$). Surface marker- and RNA-expression analyses, together with *in vitro* colony formation and *in vivo* adoptive humanized mouse transfers, indicate that eNePs are the earliest human neutrophil progenitors. Furthermore, we identified CD71 as a marker associated with the earliest neutrophil developmental stages. Expression of CD71 marks proliferating neutrophils, which were expanded in the blood of melanoma patients and detectable in blood and tumors from lung cancer patients. In summary, we establish $CD117^{+}CD71^{+}$ eNeP as the inceptive human neutrophil progenitor and propose a refined model of the neutrophil developmental lineage in bone marrow.

INTRODUCTION

Neutrophils are short-lived immune cells that play crucial roles in host defense and inflammation (Summers et al., 2010; Clark et al., 2007; Sierro et al., 2017; Ley et al., 2018). Because of their abundance and short lifespan in blood (Lahoz-Beneytez et al., 2016), these cells are constantly replenished from progenitors in the bone marrow and have long been considered a homogeneous population. Our current paradigm for understanding neutrophil development in human bone marrow has existed for decades (Bainton et al., 1971) and has lacked critical reassessment (Hidalgo et al., 2019). The introduction of technologies for high-dimensional characterization of cell populations such as mass cytometry (cytometry by time-of-flight; CyTOF) and single-cell RNA sequencing (scRNA-seq) makes evaluation of neutrophil heterogeneity possible. Neutrophils have recently received considerable attention because of their immunosuppressive functions as granulocytic myeloid-derived suppressor cells in cancer (Talmadge and Gabrilovich, 2013). Indeed, high densities of neutrophils within the tumor microenvironment are associated with reduced patient survival (Gentles et al., 2015), making these cells attractive targets for cancer immunotherapy.

The current model of neutrophil differentiation in human bone marrow describes several developmental stages, each with distinct gene-expression signatures (Ng et al., 2019). First, multipotent granulocyte-monocyte progenitors (GMPs), which develop from the common-myeloid progenitors (CMPs), give rise to neutrophil and monocyte precursors. Promyelocytes then arise immediately downstream of GMPs and express the neutrophil lineage marker CD66b (Hidalgo et al., 2019). Subsequent neutrophil development occurs through stepwise lineage and morphological maturation stages (Cowland and Borregaard 2016). Following upregulation of CD11b and CD16 expression, promyelocytes differentiate through myelocytes and metamyelocytes into banded and segmented neutrophils. Over the past decade, work has redefined CMPs and GMPs and their ability to generate downstream subsets (Mori et al., 2009; Yáñez et al., 2017; Drissen et al., 2019, 2016; Mori et al., 2015). Moreover, the dogma of hierarchical developmental stages in hematopoiesis has recently been challenged with a redefined model, in which adult human bone marrow contains only multipotent and unipotent progenitor cells (Notta et al., 2016).

We previously identified a neutrophil progenitor population in human bone marrow, termed "human neutrophil progenitors" (hNePs) ($Lin^{-}CD66b^{+}CD117^{+}$) (Zhu et al., 2018). These cells constitute 1%–3% of bone marrow neutrophils and contain a $CD34^{-}$ and $CD34^{+}$ population. When transferred into humanized NSG-SGM3 mice that support re-engraftment of human hematopoietic cells, both $CD34^{-}$ and $CD34^{+}$ hNePs produce only neutrophils, confirming their status as neutrophil progenitors.

Ng and colleagues have described a proliferative bone marrow neutrophil (neutrophil precursor; preNeu) population in both mice and humans (Evrard et al., 2018). These cells are $Lin^{-}CD66bCD34^{-}CD49d^{+}$ and constitute approximately 5% of total bone marrow neutrophils. PreNeus differentiate into non-proliferating immature and mature neutrophils.

However, the relationship between hNePs and preNeus, and their respective placement in the bone marrow neutrophil lineage, remains unclear. Here, we aimed to clarify the relationship between these recently defined neutrophil progenitors using mass cytometry. Specifically, we identified five neutrophil subsets, one of which was defined by expression of CD71, contained proliferating cells, and resembled both preNeus and promyelocytes in terms of morphology and surface marker expression. Furthermore, we identified within $CD71^{+}$ neutrophils a distinct $CD117^{+}$ subset as an inceptive progenitor of the neutrophil lineage. Our data support a developmental path for neutrophils from eNePs through downstream stages, thereby reconciling previously incompatible findings and introducing a unified model of neutrophil development in human bone marrow.

RESULTS

CyTOF Profiling of Human Bone Marrow Neutrophils Reveals Five Clusters with Distinct Surface Markers

To define neutrophil heterogeneity in human bone marrow, we used CyTOF mass cytometry to profile 35 protein surface markers (Figure 1A) across bone marrow from five healthy donors (Table S1). Applying a nonlinear dimensionality-reduction method, we projected 2,540,175 $CD45^{+}$ live cells (Table S1) onto a 2-dimensional map via the uniform manifold approximation and projection (UMAP) method (Becht et al., 2018). Clustering analysis using PhenoGraph (Levine et al., 2015; Chen et al., 2016) resulted in the identification of known immune cell types (Figures 1B–1D; Figure S1). Neutrophils made up nearly 90% of live cells in human bone marrow, consistent with previous reports (Evrard et al., 2018; Zhu et al., 2018).

Next, we selected neutrophils ($CD66b^{+}Siglec8^{-}$) for further clustering by PhenoGraph, producing 18 clusters (Figure S2A). Statistical evaluation based on relative changes (delta cutoff < 0.1; Figure S2B) in the area under cumulative distribution function (CDF) curve for each number of clusters from 2 to 15 (Wilkerson and Hayes 2010) identified 5 meta-clusters (N1–N5) from these PhenoGraph clusters (Figure 2A). Each meta-cluster exhibited a distinct combination of marker expression (Figure 2B). Subset N1 expressed a marker associated with neutrophil precursors CD49d (Evrard et al., 2018) and lacked maturation markers including CD11b, CD16, CD101, and CD10. The defining marker of subset N1 was CD71, a protein involved in iron metabolism (Aisen 2004) but not previously associated with neutrophils. Dimensionality reduction using a diffusion map (Coifman et al., 2005) suggested a potential neutrophil lineage trajectory from N1 toward N5 (Figure 2C) along the first diffusion component. As expected, expression of maturation markers (CD10, CD16) was consistent with the diffusion map-inferred lineage, whereas CD71 was exclusively expressed in the earliest neutrophil subset N1 (Figures 2D and 2E). Together, these data indicate that subset N1 contains neutrophils at very early stages of neutrophil development compared to the other four identified clusters. Expression of markers associated with distinct neutrophil developmental stages defined subsets N2 and N3 as immature neutrophils because of their lack of CD10 and CD16 expression. In contrast, subsets N4 and N5 expressed CD10 and CD16, identifying these subsets as mature neutrophils. Surface marker expression of subset N1 overlapped with the conventionally accepted definition of promyelocytes, based on the absence of CD11b and CD16 expression. Furthermore, subset N1 also overlapped with preNeus, based on their shared expression of CD49d. Thus, our subset N1 shared unmistakable similarities with promyelocytes and preNeus, according to their surface marker expression and frequency within bone marrow neutrophils. Moreover, CD35 (complement receptor type 1) is a marker of blood neutrophils in both healthy donors and in subjects with infectious diseases (Pliyev et al., 2012). Here, we show that CD35 was exclusively expressed on mature neutrophil subsets and was absent on the neutrophil precursor subset N1 as well as N2 (Figures 2D and 2E). Thus, CD35 can be used as a marker to characterize neutrophil maturation states in human bone marrow.

To analyze inter-donor variability, we determined the frequencies of the five neutrophil subsets in each individual donor. We found that all five clusters had similar frequencies in each donor, with N1 consistently representing the smallest population (approximately 5%) (Figure 2F). Thus, using high-dimensional analysis, we identified five neutrophil bone marrow subsets and their defining surface markers, including previously unappreciated patterns of CD71 and CD35 expression. Furthermore, subset N1 strongly resembled both promyelocytes and preNeu cells.

CyTOF-Identified Clusters Align with Bone Marrow Neutrophil Maturation Stages

In order to isolate the five bone marrow neutrophil populations for further investigation, we devised a flow cytometry gating strategy based on the top three most differentially expressed surface markers (CD71, CD16, and CD10) (Figures 2B and 3A). We defined subset N1 as $CD16^{-}CD10^{-}CD71^{+}$, subset N2 as $CD16^{-}CD10^{-}CD71^{-}$, subset N3 as $CD16^{+}CD10^{-}$, subset N4 as $CD16^{+}CD10^{int}$ and subset N5 as $CD16^{+}CD10^{hi}$. Biaxial gating of our CyTOF data based on these markers (Figures S2C and S2D) recapitulated our UMAP analysis (Figure 2A), validating our gating strategy. The frequency of these five subsets, as analyzed by flow cytometry in the bone marrow of five healthy donors, showed similar frequencies across individuals (Figure 3B). The surface marker expression pattern associated with each subset and the frequencies of subsets N1–N5 were very similar in a heterogeneous cohort of 10 different healthy donors (Table S1). Moreover, neutrophil subset frequencies were similar when two independent cohorts of five donors were assessed by either flow cytometry or CyTOF (Figures 2E and 3B), suggesting that this gating strategy was universally applicable. Furthermore, flow cytometry revealed similar expression of the surface markers utilized in the CyTOF panel on the five neutrophil bone marrow subsets (Figure 3C). For example, subset N1 and subset N2 not only differed in CD71 expression but also in CD101 expression, a surface marker for neutrophil maturation

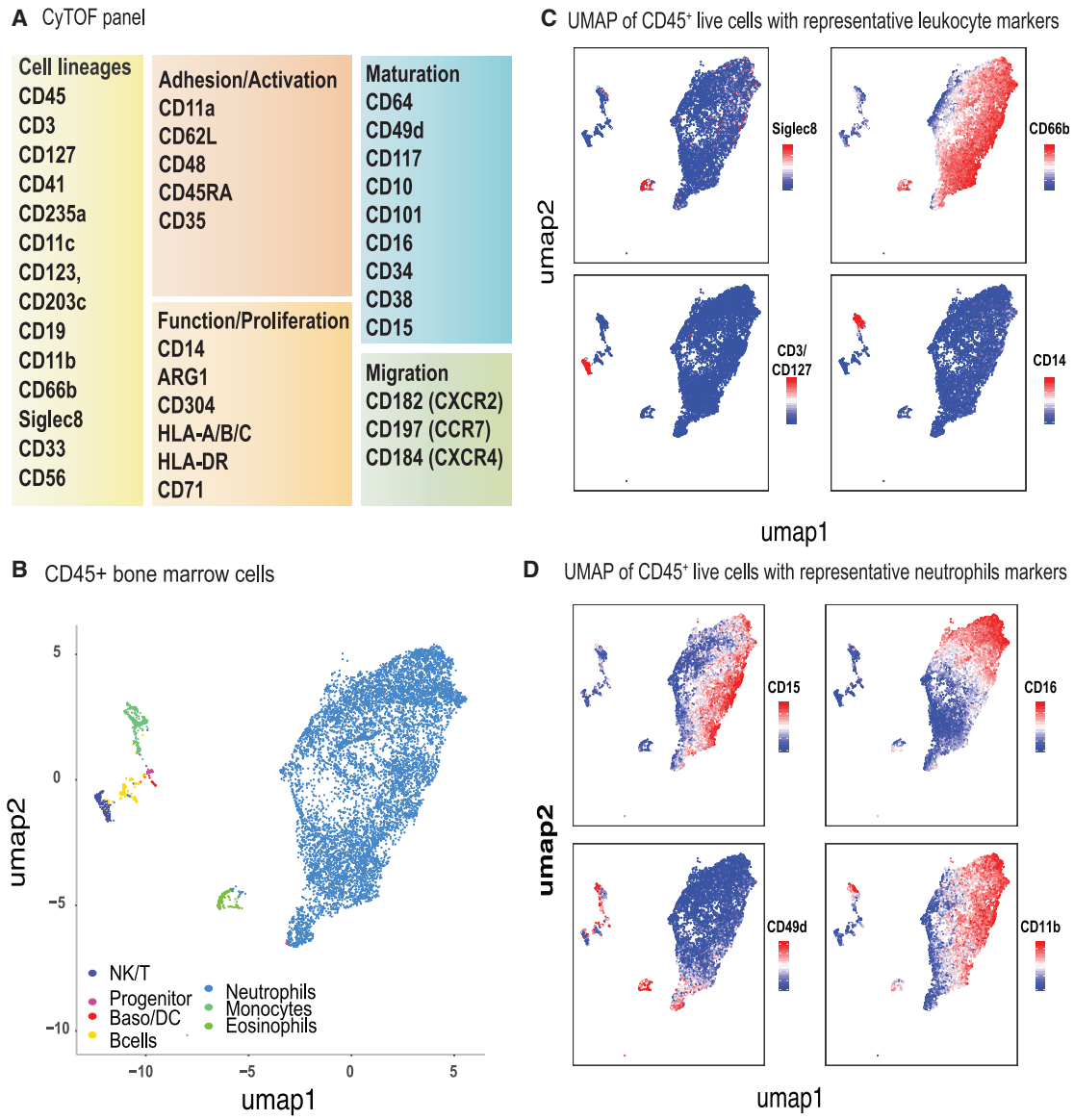


Figure 1. Study of Human Bone Marrow Neutrophil Heterogeneity by Mass Cytometry

(A) CyTOF antibody panel of human immune cell lineage markers and proteins associated with neutrophil function and maturation. CyTOF was performed on fresh human bone marrow from healthy donors (n = 5).

(B) Two-dimensional visualization of major leukocyte populations in human bone marrow (random sampling a total 50,000 cells in which 10,000 cells from each of five samples) using UMAP. Major populations were defined by PhenoGraph clustering of 2,540,175 CD45+ live cells and classified based on expression of lineage-specific surface markers. Please also see Figure S1.

(C) Single-cell expression of lineage surface markers Siglec-8 (eosinophils), CD66b (neutrophils), CD3/CD127 (NK/T cells), and CD14 (monocytes) used to define major leukocyte populations.

(D) Single-cell expression of neutrophil surface markers CD49d, CD14, CD11b, and CD16 highlights distinct expression within neutrophil clusters.

(Evrard et al., 2018). Importantly, flow cytometry also confirmed that CD71 expression was restricted to subset N1. Analogous to our findings with CyTOF, CD35 and CD16 were able to distinguish subsets N2 and N3, and subsets N4 and N5 did not show qualitatively but rather quantitatively differential surface marker expression. However, comparison of CD11b and CD101 analysis between flow cytometry and CyTOF showed a more pronounced distinction between positive and negative expression by the former.

Finally, by superimposing a published gating strategy for bone marrow neutrophil progenitors on the basis of CD16 and CD11b expression (Hidalgo et al., 2019) onto our five clusters, we showed that subset N1 represented promyelocytes, N2 represented myelocytes, N3 was similar to metamyelocytes, and N4 and N5 represented banded and mature neutrophils (Figure 3D). However, subsets N4 and N5 mostly overlapped and were only subtly distinguishable in the CD16 versus CD11b plot (Figure 3C). Applying these subsets onto the existing neutrophil

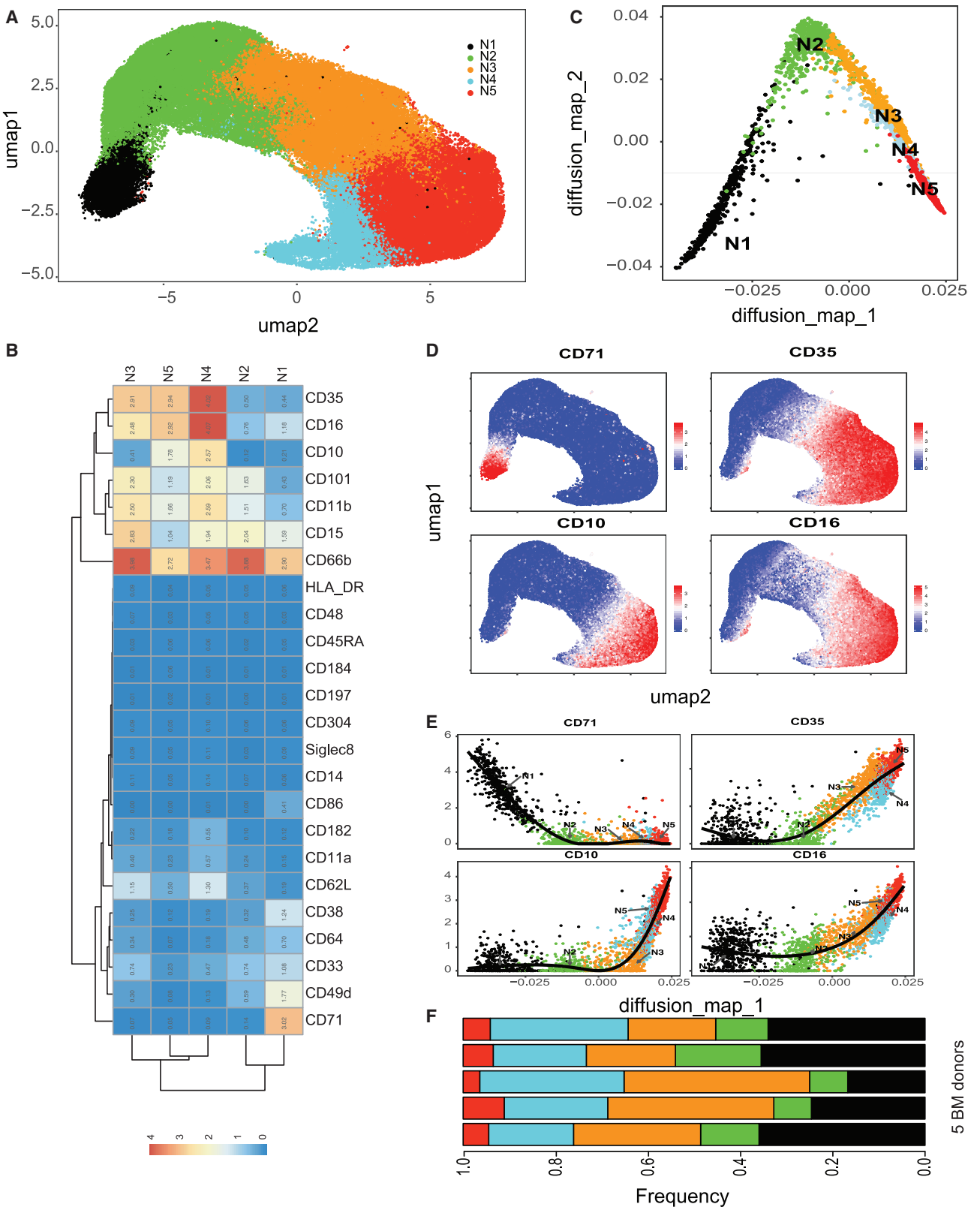


Figure 2. Human Bone Marrow Neutrophil Heterogeneity Identified by Unbiased Clustering

(A) Consensus clustering of PhenoGraph clusters identified five merged subsets of CD66b$^+$ neutrophils (labeled as N1–N5) visualized by UMAP projection of total CD66b$^+$ neutrophils.

(legend continued on next page)

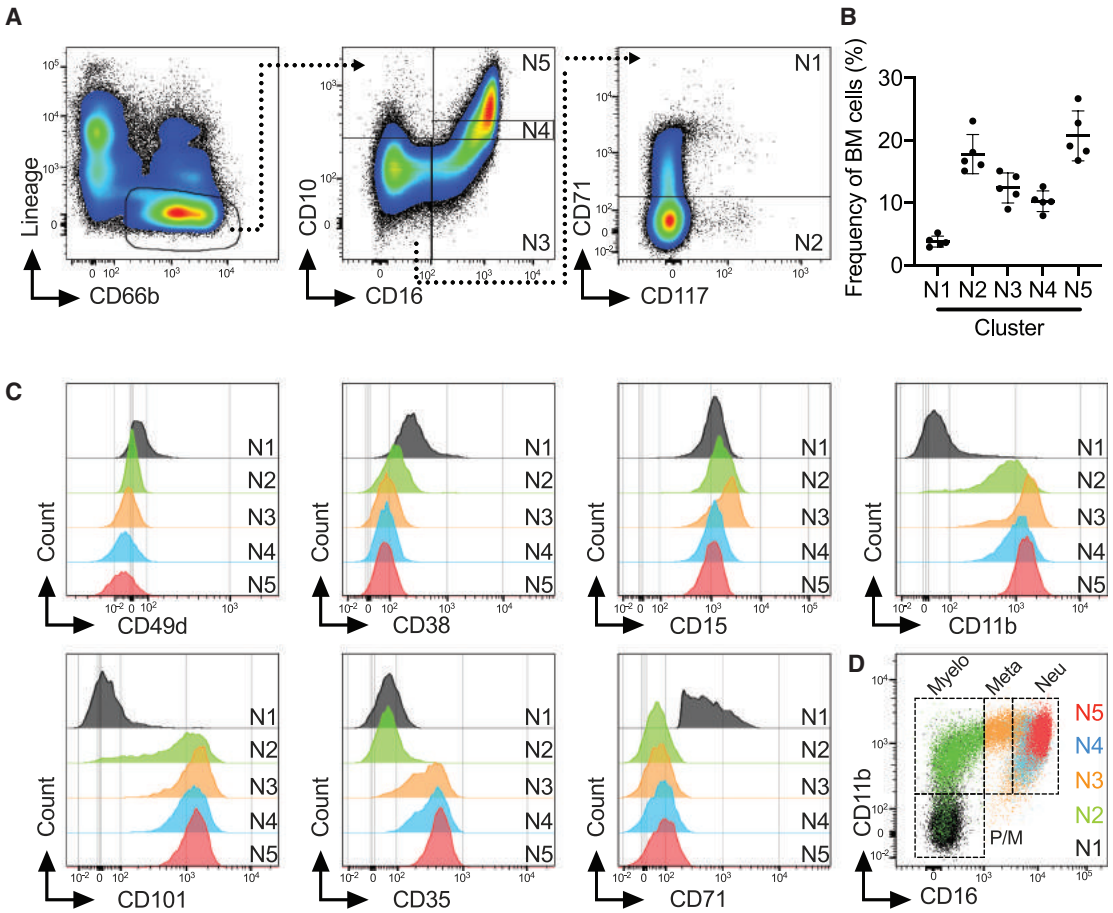


Figure 3. Flow Cytometry Confirms Five Human Bone Marrow Neutrophil Subsets

(A) Flow cytometry gating strategy based on marker expression from CyTOF data shown in Figure 2B to identify five bone marrow neutrophil clusters.

(B) Frequencies of five neutrophil subsets based on flow cytometry data for an independent cohort of five human bone marrow donors confirm frequencies observed in CyTOF data.

(C) Histograms of neutrophil markers across five neutrophil subsets based on flow cytometric analysis and gating strategy shown in (A).

(D) Overlay of five identified neutrophil subsets onto CD16 versus CD11b biaxial dot plot based on the conventional gating strategy for neutrophil progenitors (Hidalgo et al., 2019) identified subsets N1 to N5 as Promyelocytes (Pro), Myelocytes (Myelo), Metamyelocytes (Meta), and banded and segmented neutrophils (Neu).

See also Figure S2.

differentiation paradigm suggests a developmental trajectory from subset N1 toward subset N5, wherein N1 represents promyelocytes in both mass and flow cytometry analysis.

CD117 Identifies a Subpopulation of Subset N1, eNeP

We have previously identified a human CD66b^{+}CD117^{+} neutrophil progenitor population, hNePs (Zhu et al., 2018). Because our CyTOF clustering result did not discover these hNePs as a distinct population, we used our flow cytometry workflow, including incorporation of fluorescence minus one (FMO) controls (Figure S3), to determine the overlap between hNeP cells and our five called clusters. Stratification of hNeP by CD71, the marker defining subset N1, revealed CD71^{+} and CD71^{-} populations within hNePs (Figure 4A), suggesting that hNePs were more heterogeneous than previously appreciated and composed of at least two distinct populations. When our flow cytometry gating

(B) Heatmap depicting average surface marker expression of neutrophils, highlighting expression of subset-specific markers, such as CD71 on subset N1.

(C) Dimensionality reduction by diffusion map of CyTOF data revealed neutrophil lineage trajectory. Random samplings of 500 cells from each cluster N1–N5 were plotted on two diffusion components.

(D) Single-cell expression of early neutrophil marker CD71 and maturation markers CD35, CD10, and CD16 on human bone marrow neutrophils.

(E) Cell ordering inferred by diffusion map analysis represented by the first diffusion map component for early and mature markers (CD71, CD35, CD10, and CD16).

(F) Comparison of the frequencies of N1–N5 among healthy human bone marrow donors reveals minimal inter-donor heterogeneity. N1–N5 neutrophil subsets were present in all donors (n = 5).

See also Figure S2.

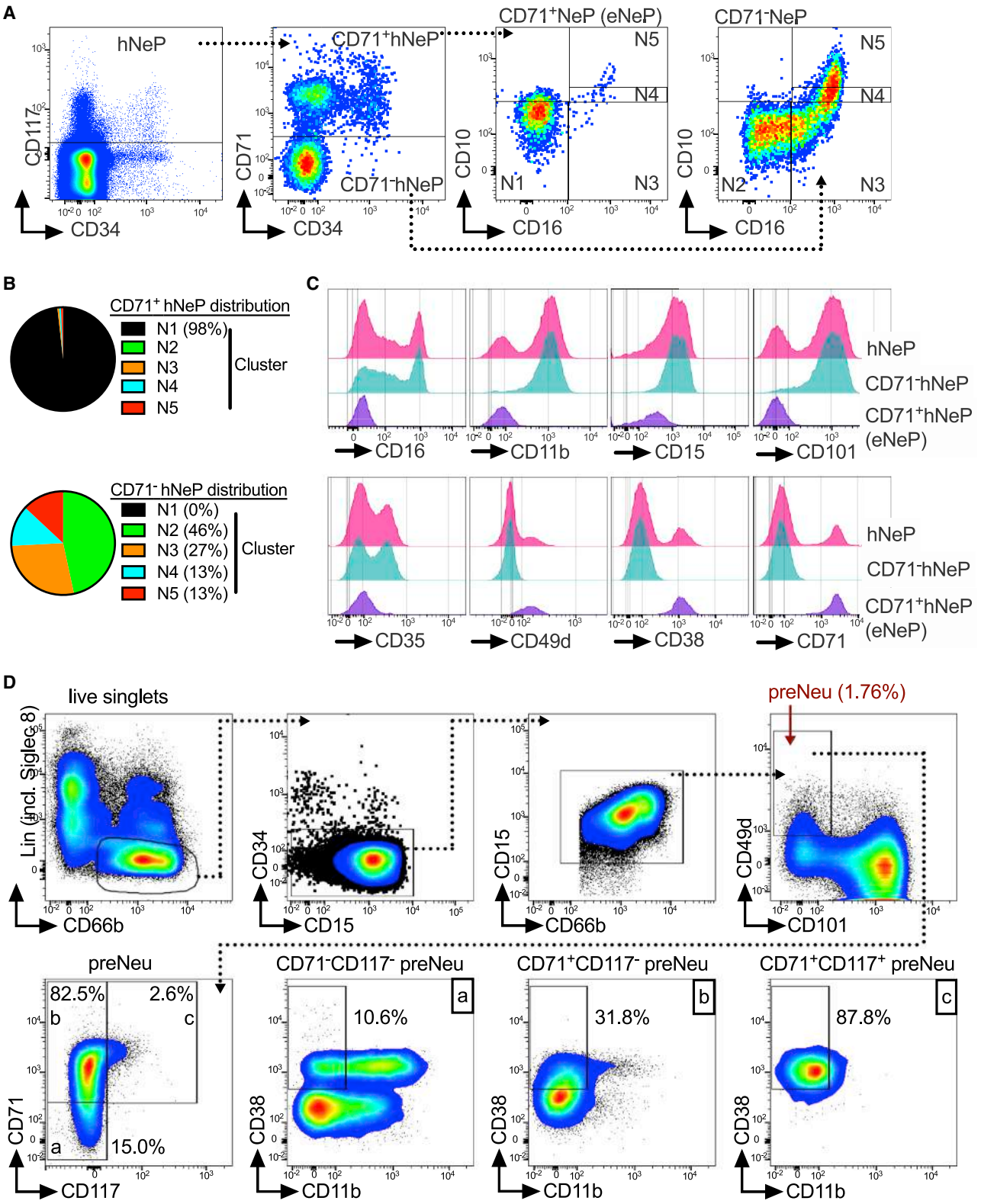


Figure 4. CD71⁺hNePs, Termed eNePs, Are a Distinct Subset of Progenitors

(A) Stratification of previously identified neutrophil progenitor population termed hNePs (Zhu et al., 2018) by promyelocyte (N1) marker CD71 identified CD71⁺hNePs and CD71⁻hNePs. Because CD71 was found to be a defining marker for neutrophil progenitors within promyelocytes (Figure 3), CD71⁺hNePs were

(legend continued on next page)

strategy was applied onto CD71$^+$ and CD71$^-$ hNePs, we found that the CD71$^-$ hNeP subset specifically contained cells belonging to all neutrophil subsets except N1 (Figure 4B). Accordingly, the CD71$^-$ hNeP consisted of more mature neutrophil subsets that express CD101 but lacked expression of CD49d (Figure 4C). We have previously shown that hNePs are CD117$^+$ (Zhu et al., 2018). However, CD117 expression in our CyTOF data was not entirely distributed in the neutrophil progenitors, as shown in our flow cytometry results (Figure S4A). We observed relatively high expression of CD117 in a fraction of subset N1, which failed to be recognized by PhenoGraph (Figure S2A) even after isolating subset N1 for further unbiased clustering (data not shown). Because hNePs (CD66b$^+$CD117$^+$) were shown to be unipotent neutrophil progenitors in NSG-SGM3 mouse transfer experiments (Zhu et al., 2018), we hypothesized that CD117$^+$CD71$^+$ cells within subset N1 were the pure population of neutrophil progenitors. Therefore, expression of surface markers not included in our gating strategy were analyzed on both CD71$^+$ (CD117$^+$CD71$^+$ cells) and CD71$^-$ hNeP cells (Figure 4C). Again, total hNePs showed heterogeneous expression for CD11b, CD101, CD35, CD49d, and CD38, whereas CD71$^+$ hNePs represented a pure population, with uniformly positive expression of progenitor markers CD49d and CD38 and a lack of maturation markers CD11b, CD101, and CD35. We therefore termed these CD71$^+$ hNePs "early neutrophil progenitors" (eNePs). The neutrophil maturation marker CD101 was positive on CD71$^-$ hNePs but negative on eNePs, indicating that eNePs might be, as their affiliation with subset N1 suggests, the earliest bone marrow neutrophil progenitors within heterogeneous hNePs. eNePs represent ~2.5% of subset N1 and only ~0.14% of all bone marrow neutrophils (Figure S4B), making these cells a small population within the bone marrow. Mapping eNePs on the diffusion map trajectory (Figure 2C) showed that eNePs were likely the earliest identified cells found within the neutrophil developmental lineage, and eNePs possessed these early progenitor surface markers as shown in the trajectory heatmap (Figures S4C and S4D). Thus, by using flow cytometry, we further characterized subset N1 and stratified it according to CD117 expression into the smaller eNePs and the bigger CD117$^-$CD71$^+$ fraction (N1 w/o eNePs).

Next, we applied the CD16 versus CD11b gating strategy for bone marrow neutrophil progenitors (Hidalgo et al., 2019) to eNePs and CD71$^-$ hNePs (Figures S4E and S4F). eNePs were located within the promyelocyte gate whereas CD71$^-$ hNePs were dispersed out over the more mature subsets from myelocytes to banded and mature neutrophils. eNePs, however, constituted less than 1% of the entire promyelocyte population, and subset N1 represented around 70% of promyelocytes, demonstrating that eNePs are a small precursor subpopulation.

We then asked whether the recently discovered preNeus (Evrard et al., 2018) were part of our subset N1. Using the published preNeu gating strategy (Evrard et al., 2018) in combination with our CyTOF and flow cytometry data, we identified CD66b$^+$ CD15$^+$CD49$^+$Siglec8$^-$CD34$^-$CD101$^-$ preNeus and analyzed their expression of CD71 and CD117 (Figure 4D; Figure S5A). PreNeus stratified by CD71 and CD117 revealed three subsets: a CD71$^-$CD117$^-$ subset (labeled 'a' in Figure 4D), a CD71$^+$ CD117$^-$ subset (labeled 'b' in Figure 4D), and a small CD71$^+$CD117$^+$ subset (labeled 'c' in Figure 4D). These three preNeu subsets exhibited distinct expression of CD38, a marker associated with myeloid progenitor populations, with the CD71$^+$CD117$^+$ subset being uniformly positive for this marker. Approximately 70% and ~85% of preNeus were positive for CD71 as determined by CyTOF and flow cytometric analysis, respectively, confirming that the majority of preNeus were a part of subset N1. Furthermore, eNePs were part of preNeus (CD71$^+$CD117$^+$ preNeu subset) and represented 4.4% and 2.6% of preNeus as determined by CyTOF and flow cytometry, respectively (Figure 4D; Figure S5B). In summary, analysis of CD117 and CD71 coexpression by flow cytometry-identified eNePs, a homogeneous neutrophil progenitor subset in humans, which potentially represent the hierarchically "earliest" unipotent neutrophil progenitor cells within promyelocytes and preNeus.

eNePs Are Proliferating Neutrophil Progenitors

To determine the progenitor characteristics of eNePs, we probed expression of the stem cell marker CD34 (Sidney et al., 2014) on eNePs and the rest of hNePs (CD71$^-$ hNeP) (Figures 5A and 5B), as hNePs were shown to be composed of both CD34$^+$ and CD34$^-$ subsets (Figure 4A) (Zhu et al., 2018). Although a sizable fraction (~8%) of eNePs stained positive for CD34, this marker was absent on CD71$^-$ NeP. None of the five CyTOF-identified clusters displayed more than 0.01% of CD34$^+$ cells (Figure 5B), indicating that only eNePs are enriched for CD34$^+$ cells. To confirm that these cells indeed had features of progenitor cells and possessed proliferative potential, we measured BrdU incorporation in bone marrow neutrophils (Figure 5C). Although only N1 and eNePs contained BrdU$^+$ cells, the surface expression of CD66b differed between these two proliferating subsets (Figure 5C). The majority of non-proliferating N1 and eNeP cells were CD66b^{lo}. However, unlike proliferating eNePs, which were also low in CD66b expression, the majority of BrdU$^+$ proliferating N1 cells were CD66b^{hi}, perhaps reflecting ongoing maturation during the proliferation process. We also found all proliferating BrdU$^+$ cells to be positive for CD71 (Figure 5D), in line with findings of CD71 and Ki67 expression on activated T cells (Motamedi et al., 2016).

In our previous hNeP study, we showed that both CD34 positive and negative hNePs exclusively generate neutrophils after

relabeled as early NePs (eNePs). Subsequently, eNePs (CD71$^+$hNePs) and CD71$^-$NePs were backgated using the gating strategy depicted in Figure 3A, revealing that only eNePs were a pure population and negative for maturation markers CD10 and CD16.

(B) Frequencies of five defined neutrophil subsets (N1 to N5) in eNePs (CD71$^+$hNePs) and CD71$^-$hNePs.

(C) Comparison of selected neutrophil markers based on histograms of flow cytometric analysis of hNePs, eNePs (CD71$^+$NePs) and CD71$^-$NePs.

(D) Top: flow cytometry gating strategy to identify previously described preNeus (Evrard et al., 2018). Bottom: preNeus were further stratified according to CD71 and CD117 expression, identifying eNePs as a small subpopulation of preNeus (2.6%; gate c). The resulting subsets were then examined for neutrophil maturation marker CD11b and progenitor marker CD38.

See also Figures S3, S4, and S5.

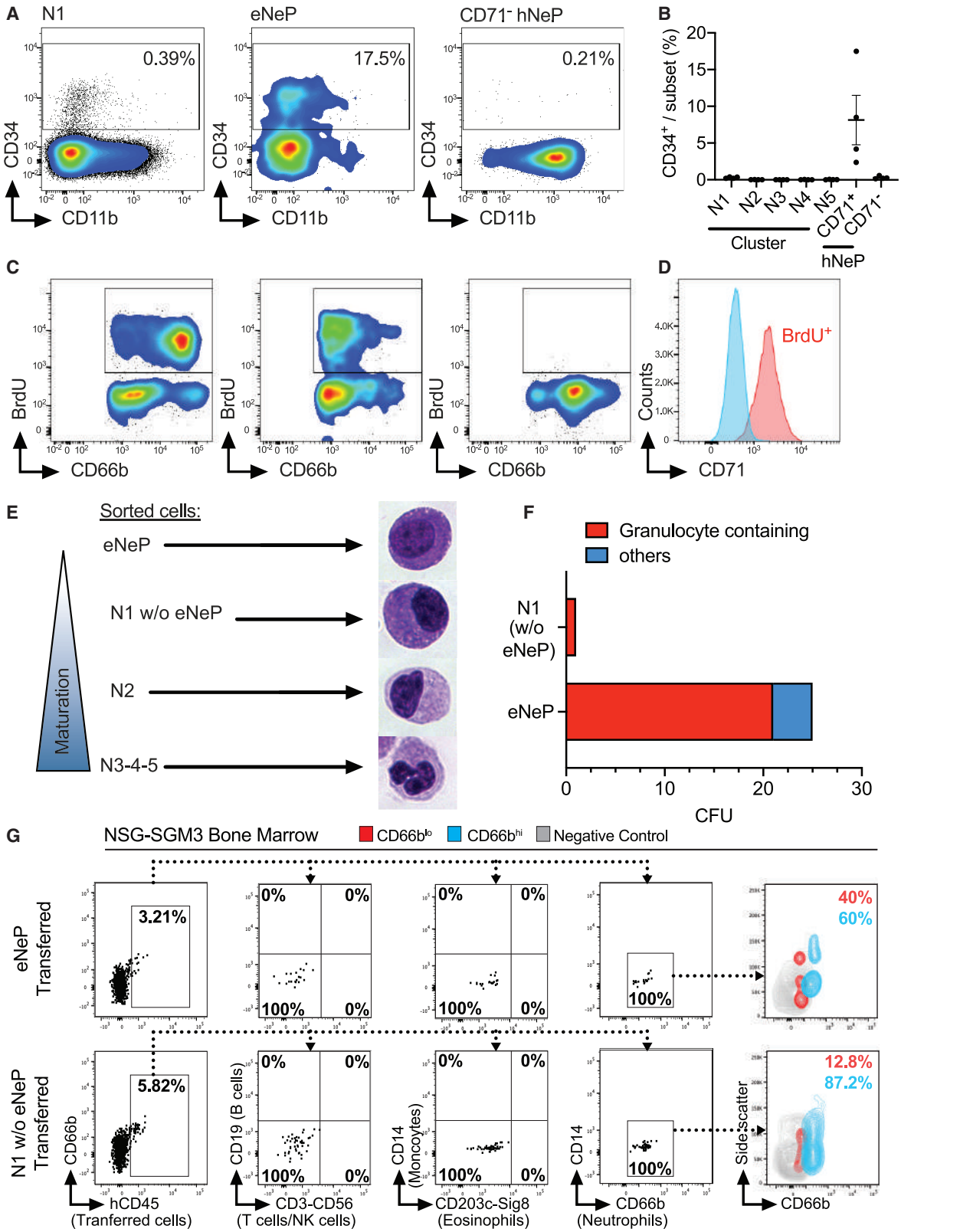


Figure 5. eNePs Are Early Unipotent Proliferating Neutrophil Progenitors

(A) Analysis of CD34 expression on early neutrophil subsets by flow cytometry.

(B) Frequency of CD34$^+$cells among five neutrophil clusters, eNePs, and CD71$^-$NeP analyzed by flow cytometry as shown in (A). Please also see Figure S6.

(legend continued on next page)

transfer into NSG-SGM3 mice (Zhu et al., 2018). Hence, we hypothesized that both $CD34^+$ and $CD34^-$ hNePs should be at least partially composed of neutrophil progenitors. We found that $CD34^+$ cells were restricted to eNePs ($CD34^+CD71^+$ hNePs or $CD34^+CD71^+CD117^+$ cells) and absent from all other clusters (Figure 5B; Figure S6A), indicating that all previously reported $CD34^+$ hNePs fell within eNePs. In short, $CD34^+$ hNePs were identical to the herein identified $CD34^+$ eNePs. Approximately 1/3 of these $CD34^+$ eNePs are $BrdU^+$ and proliferating (Figure S5C). The fact that $CD34^+$ eNePs were both proliferative and non-proliferative could perhaps be explained by the existence of a quiescent state of these neutrophil progenitors. Taken together, these results suggest that $CD34^+$ eNePs possess unipotent differentiation potential *in vivo*.

We therefore aimed to further characterize $CD34^-$ hNePs, which also produced neutrophils in NSG-SGM3 mice (Zhu et al., 2018). $CD34^-$ hNePs consisted of $CD71^+CD101^-$ and $CD71^-CD101^+$ populations (Figure S6A), of which only the former localized in the promyelocyte gate. The $CD71^-CD101^+$ population aligned with more mature subsets (Figure S6A). Based on these findings, we believe that only the $CD71^+CD101^-$ fraction was responsible for the neutrophil lineage reconstitution observed after transfer of $CD34^-$ hNePs into NSG-SGM3 mice (Zhu et al., 2018).

Thus, we next sought to determine whether neutrophils derive from all $CD71^+$ bone marrow neutrophils (N1) or solely the $CD117^+$ eNeP within N1. To address this question, we first examined the morphology of bone marrow neutrophil subsets via cytospin (Figures 5E and 6B; Figure S7B). eNePs displayed morphological features of progenitor cells, including large, round nuclei and a high nucleus-to-cytoplasm ratio (Figure 5E). Nuclei of the rest of the cells within N1 appeared flatter and the nucleus-to-cytoplasm ratio decreased, a morphology that continued through subset N2. Cells of subsets N3, N4, and N5 showed morphology more similar to mature neutrophils, in line with their surface marker expression. We then sorted eNePs and N1 w/o eNePs into cytokine-supplemented methylcellulose and counted progeny colonies. As expected, eNePs produced granulocyte-containing colonies (Figure 5F). In contrast, the N1 w/o eNeP subset generated almost no colonies, demonstrating that only eNePs are potent neutrophil progenitors. As subset N1 was the only cluster with surface marker expression consistent with progenitor cells (CD38, CD49d, and a lack of CD11b and CD101), these data also suggested that eNePs are the only progenitors within $CD66b^+$ bone marrow neutrophils. In order to provide more conclusive evidence *in vivo* that eNePs indeed differentiate into their downstream subset(s), we adoptively transferred fluorescence-activated cell sorting (FACS)-sorted eNePs ($CD71^+CD117^+$) and N1 w/o eNePs ($CD71^{+-}CD117^-$; i.e., lacking eNePs) from healthy human bone marrow into recipient NSG-SGM3 mice. This strategy permitted engraftment of human hematopoietic stem cells, particularly those of the myeloid lineage. In order to track the lineage, we examined the bone marrow and blood early, 5 days after adoptive transfer (Zhu et al., 2018). First, we observed that eNePs produced only neutrophil-lineage cells in bone marrow 5 days following adoptive transfer (see Figure 5G, S5D, and S5E). We found no production of T cells, natural killer (NK) cells, eosinophils, or monocytes by eNePs. Second, we observed that a higher frequency of eNeP-derived progeny in the bone marrow expressed lower frequency of CD66b and slightly lower side-scatter compared to progeny derived from N1 w/o eNePs (40% $CD66b^{lo}$ in eNePs versus 12.8% in N1 w/o eNePs) (Figure 5G). These data suggest that the eNePs were positioned earlier in the neutrophil developmental lineage than the N1 without eNePs, as the eNeP progeny appeared to be less mature (more cells with $CD66b^{lo}$ and fewer with $CD66b^{hi}$), averaging only 49% $CD66b^{hi}$ for eNePs versus 71% $CD66b^{hi}$ for N1 w/o eNePs, over the given time period (Figure S5E). We also found the presence of only human neutrophil progeny from transferred eNePs in the blood of these mice, albeit in very low numbers at this early time point. Although we tried to apply the eNePs and N1–N5 gating strategy to these neutrophil progeny, we were unable to see progeny that contributed to the eNeP or N1 populations. We did observe N2–N4 progeny, again with the eNeP-derived progeny appearing slightly less mature than the N1 w/o eNeP-derived progeny (Figures S5D and S5F). Figures S5E and S5F show replication with consistent results of three experiments of the adoptive transfer of human progenitors into recipient NSG-SGM3 mice. In all experiments, eNePs (red dots) produced only neutrophils (Figures S5E and S5F), and the eNeP progeny appeared to be less mature than N1-eNeP (blue dots) progeny (Figure S5E). These data are in line with previous observations with hNeP transfers (Zhu et al., 2018) showing unipotency, and these data on eNePs show functional and unipotent production of neutrophil progeny by eNePs. Together with our *in vitro* differentiation assay (Figures 5E and 5F), these data indicate that eNePs are unipotent neutrophil progenitors and are likely more immature in their lineage development than N1 w/o eNePs.

These findings, however, raised the question as to how eNePs are distinct from more upstream progenitors. To visualize defining characteristics between eNePs and $CD34^+CD38^+$ myeloid progenitor cells, such as CMPs and GMPs, we compared the expression of neutrophil-defining surface markers

(C) Quantification of proliferation by measuring BrdU incorporation with flow cytometry. Neutrophil subsets not depicted here contained <0.2% proliferating cells (not shown).

(D) Histogram of BrdU staining in bone marrow neutrophils stratified by CD71. Gated on $BrdU^+$ or $BrdU^-$ live neutrophils after exclusion of doublets (not shown).

(E) Microscopic analysis of morphology of fluorescence-activated cell sorted neutrophil subsets following Cytospin and Hema 3 staining identified the typical morphology of neutrophil maturation gradually developing along identified neutrophil subsets. Please also see Figure S6.

(F) Quantification of *in vitro* progenitor differentiation assay of fluorescence-activated cell sorted eNePs and cluster N1 (w/o eNePs). Colonies were identified by side-by-side comparison with representative images of pre-defined colonies from three independent plates.

(G) eNePs or N1 w/o eNePs human bone marrow cells were sorted and adoptively transferred into irradiated NSG-SGM3 mice. Recipient bone marrow was analyzed by flow cytometry after 5 days. Plots show the absence of T cells (CD3), NK cells (CD56), B cells (CD19), monocytes (CD14), and eosinophils (Siglec8 and CD203c) and sole presence of neutrophils (CD66b) in human CD45+ (hCD45) progeny. Contour plot shows $CD66b^{lo}$ (red) and $CD66b^{hi}$ (blue) $hCD45^+$ progeny. Grey contour indicates negative control (mouse $CD45^+$ cells). Flow cytometry is representative of three independent experiments (n = 3 mice per group). Please also see Figure S5.

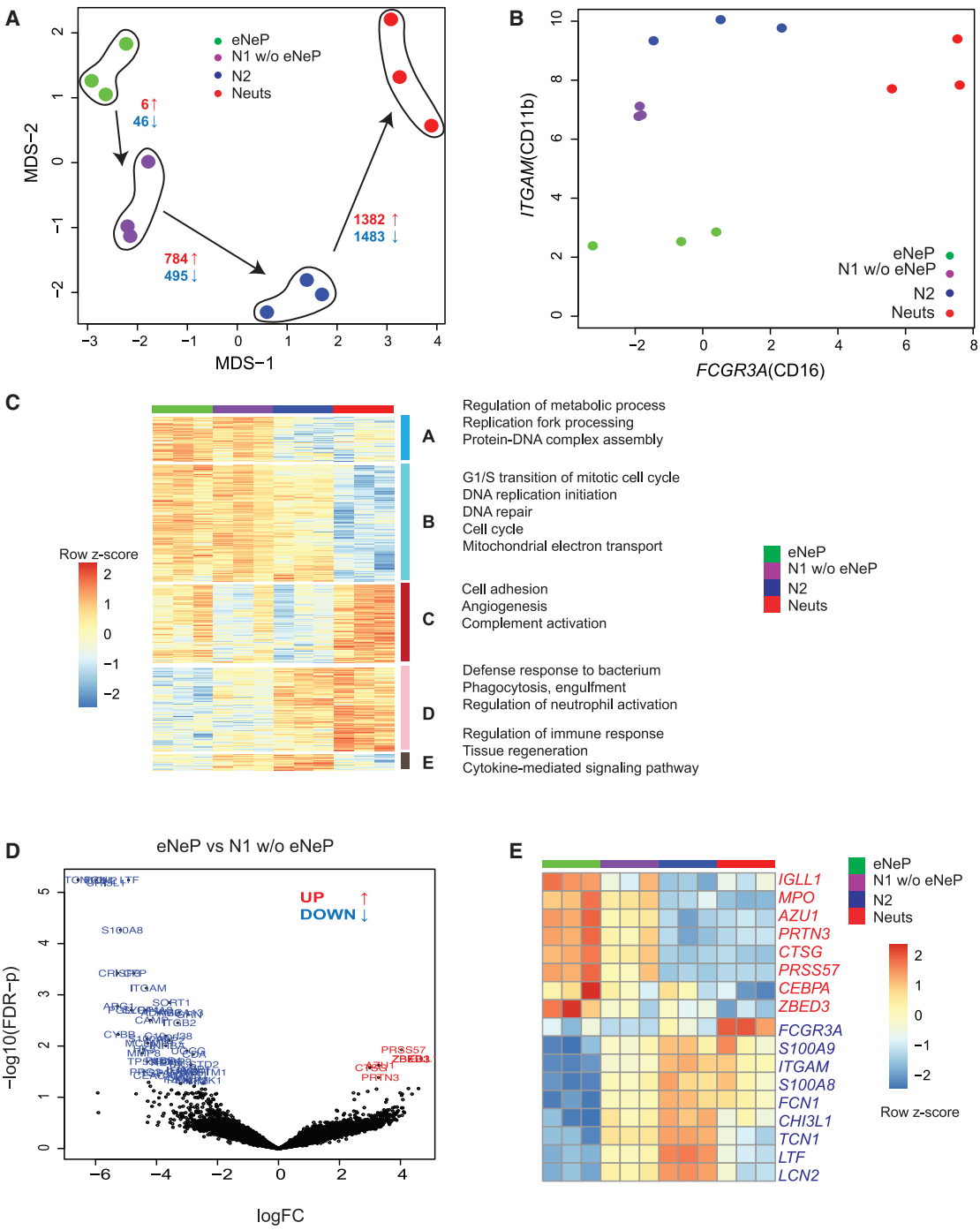


Figure 6. Transcriptome Analysis Reveals Distinct Gene Expression of eNePs

(A) eNeps, N1 w/o eNePs, and N2 and N3-4-5 (Neuts) were fluorescence-activated cell sorted, followed by bulk RNA-seq. Please also see Figure S6. MDS plot of eNePs and preNeu-like immature and mature neutrophils (three replicates each) suggested the developmental stages starting with eNePs toward more mature neutrophils. Number of statistically significant differentially expressed (logFC 2 and false discovery rate [FDR]-corrected p values cutoff 0.05) genes (red: up-regulated, blue: downregulated) for the comparison between eNep-N1(w/o eNePs), N1(w/o eNePs) - N2, N2 - mature neutrophils (N3-4-5).

(B) Log2 expression of CD11b (*ITGAM*) and CD16 (*FCGR3A*) of the four sorted neutrophil subsets shown in (A) aligned with the conventional neutrophil development gating strategy (Hidalgo et al., 2019) of neutrophil maturation from promyelocytes to mature neutrophils (compare to Figure 3D).

(C) Five gene clusters from pairwise differentially expression analysis between two neutrophil subsets next to each other in the developmental lineages (eNeP-N1 [w/o eNePs], N1 [w/o eNeP] - N2, N2 - mature neutrophils [N3-4-5]). Biological process GO terms that were enriched in each of five subsets.

(legend continued on next page)

CD66b and CD15 on these subsets (Figures S7A and S7B). Although $CD34^+CD38^+$ myeloid progenitor cells were negative for CD66b and CD15 in both flow cytometry and CyTOF data (Cluster 8 in Figures S1, S7A, and S7B), eNePs showed marker expression between that of myeloid progenitors and more mature neutrophils (subset N2). These data confirm eNePs as the earliest progenitors in human bone marrow expressing neutrophil lineage markers CD66b and CD15.

Transcriptomic Analysis Reveals Gene Signatures of Human Neutrophil Development

To understand the molecular characteristics of different neutrophil subsets, we used RNA-seq to profile sorted subsets (Figure S6B): eNePs, N1 w/o eNePs, N2, and mature neutrophil subsets (N3+4+5 combined). Multidimensional scaling (MDS) analysis based on differentially expressed genes (DEGs; see STAR Methods) between all pairs of 12 samples showed that sorted neutrophil subsets grouped together according to the clusters we defined based on CyTOF and flow cytometry (Figure 6A). According to their gene expression, eNePs shared more similarities with subset N1 ($CD71^+$ neutrophils) than with immature and mature neutrophils ($CD16^+$). The number of DEGs (both up- and downregulated) between two neighboring neutrophil subsets increased from 52 genes between eNePs and subset N1 to 2,865 genes between immature ($CD16^-$) and mature ($CD16^+$) neutrophils with the same statistical cutoffs (Figure 6A). These findings reflected the heterogeneity of transcriptional output from neutrophils at different developmental stages, which was similar to the hierarchical trajectory inferred from our CyTOF data (Figure 2C). The expression of transcripts for surface markers CD11b (*ITGAM*) and CD16 (*FCGR3A*) by sorted neutrophil subsets confirmed flow cytometry data (Figure 3D) and suggested that eNePs were distinct from the rest of subset N1 and possessed the lowest CD11b expression (Figure 6B). Thus, our transcriptomic data support the distinct early development characteristics of eNePs.

Next, we combined all 5,446 DEGs from each of the five pairwise comparisons (eNePs versus N1 w/o eNePs, N1 w/o eNePs versus N2, N2 versus mature neutrophils, eNePs versus mature neutrophils) for hierarchical clustering analysis. We identified five different groups of genes with different neutrophil differentiation expression patterns (Figure 6C). We found 2,586 genes highly expressed in early neutrophils (N1 including eNePs) compared to later developmental neutrophils (gene clusters A and B) with enriched Gene Ontology (GO) terms relating to early cellular features such as cell cycle, DNA replication and repair, and mitochondria (Table S2). Genes enriched in mature neutrophils were subsequently grouped into two clusters (gene clusters C and D) and classified by GO terms complement activation, cell-matrix adhesion, and regulation of neutrophil activation. 266 genes (cluster E) were specific to immature neutrophils and associated with immune response, tissue regeneration, and cytokines (Table S2). Collectively, our transcriptomic data are consistent with the notion that eNePs and subset N1 form the early stages of neutrophil development in bone marrow. We next interrogated the genes between eNePs and subset N1 in order to determine which subset precedes the other in neutrophil maturation. Comparing eNePs and subset N1 revealed only six upregulated genes in eNePs (Figures 6D and 6E), including a few involved in early neutrophil development (*AZU1*, *PRTN3*, *CTSG*, and *MPO*) according to recent published RNA-seq data (Grassi et al., 2018). These results indicate that eNePs are a more homogeneous early neutrophil progenitor subset than other established subsets, such as promyelocytes or preNeus. In addition, 46 genes including genes typically enriched in mature neutrophils (*S100A8*, *S100A9*, *ITGAM*, *FCN1*, and *MMP8*) were statistically downregulated in eNePs compared to subset N1, providing additional evidence that eNePs are upstream progenitors of subset N1 or preNeus. In addition, zinc finger BED domain-containing protein 3 (*ZBED3*) was identified as an upregulated gene in comparison to all other neutrophil subsets (Figure 6E). In addition, *ZBED3* was lowly detected in the recent RNA-seq data of four conventional neutrophil subsets (Grassi et al., 2018). In summary, transcriptomic analysis revealed eNeP as an early neutrophil progenitor with a gene expression signature distinct from that of preNeus and promyelocytes.

Expansion of the $CD71^+$ Subset in Cancer Patients

Cancer is associated with an increase of neutrophils and neutrophil precursors in the blood, due to aberrant expression of chemokines and cytokines. Likewise, neutrophil subsets egress from the bone marrow at earlier stages than under physiological steady-state conditions during cancer (Talmadge and Gabrilovich 2013). We thus asked whether we could detect an expansion of eNePs and subset N1 in cancer patients. We analyzed blood from healthy donors and melanoma patients (Table S3) with our CyTOF-antibody panel and identified subset N1 by gating on $Lin^-CD66b^+CD16^-CD10^-CD71^+$ cells (Figures 7A and 7B). We were unable to confidently identify eNePs in our CyTOF data from patient blood, likely because of their scarcity even in the bone marrow (only ~3% of subset N1). Nevertheless, we found a statistically significant increase of $CD71^+$ neutrophils (N1) in blood of melanoma patients compared to healthy donors (Figure 7B). Furthermore, our analysis of differentially expressed protein markers from mass cytometry data between $CD71^+$ versus $CD71^-$ neutrophils across 24 melanoma patients (Wilcoxon rank sum test) revealed that $CD71^+$ neutrophils were distinct from $CD71^-$ neutrophils (Figure 7C; Figure S7C). Specifically, $CD71^+$ neutrophils highly expressed immature and progenitor markers (CD38, CD49d, and CD48) and poorly expressed markers of maturation (CD16, CD10, and CD101). In addition, $CD71^+$ neutrophils had higher expression of CD304, a VEGFR2 co-receptor that is associated with hypoxia-induced polarization of alternatively activated macrophages (Gelfand et al., 2014; Chen et al., 2019), suggesting a pro-tumoral function. We also observed that $CD71^+$ neutrophils had higher expression of antigen-presenting proteins (HLA-DR,

(D) Volcano plot of differentially expressed genes between the earliest neutrophil progenitor subsets eNePs and N1 w/o eNePs (−log10 adjusted p values: y axis and log2FC: x axis) revealed six upregulated genes in eNePs.

(E) Top up- and downregulated genes (logFC 2, FDR-corrected p values 0.05) in eNePs from (D) and neutrophil genes (transcriptional factors *CEBPA/E,* neutrophil marker *MPO*, *AZU1*, *FCGR3A*, and *S100A8/9*) confirmed early developmental stage of eNePs.

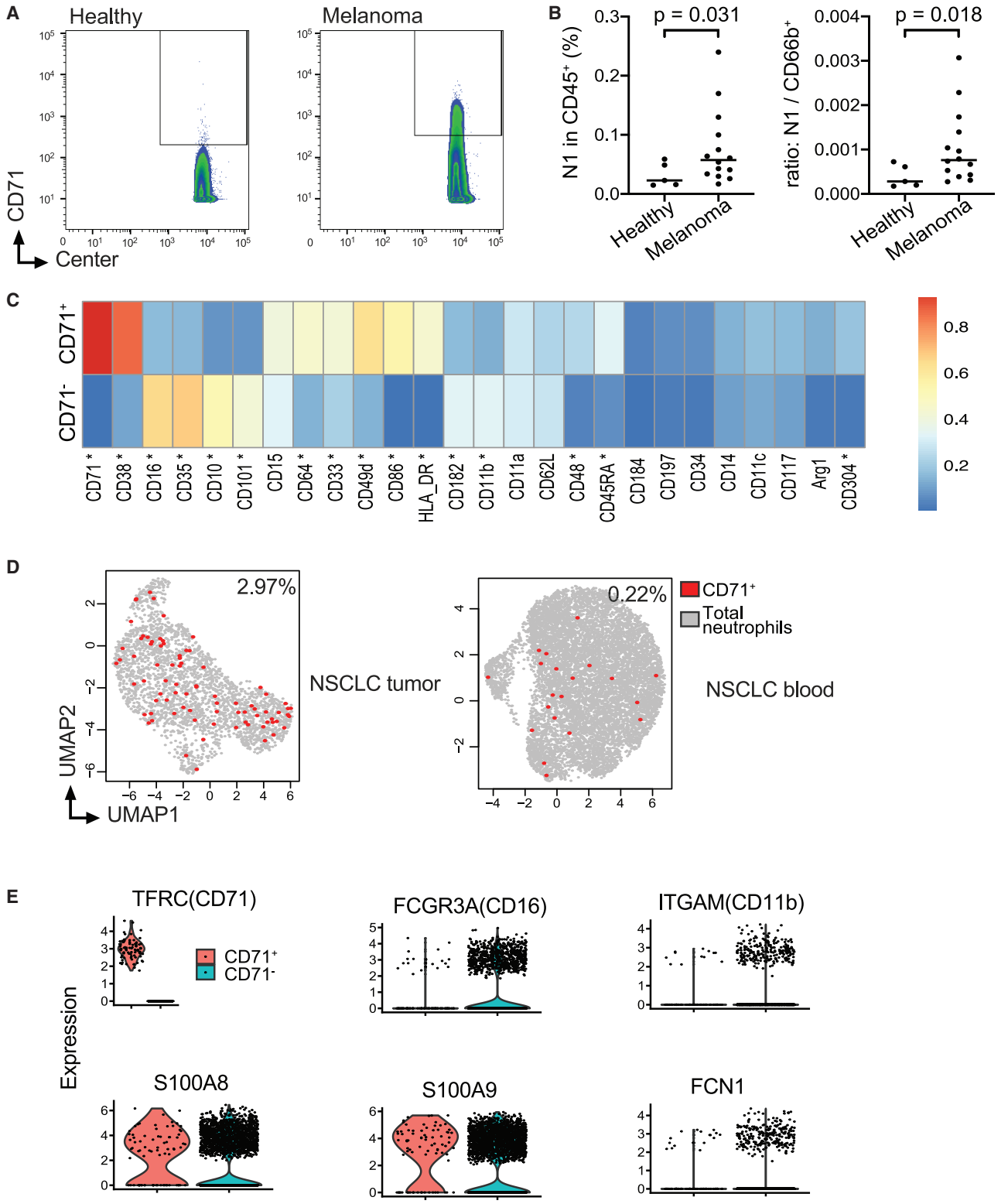


Figure 7. Expansion of CD71+ Neutrophils in Cancer Patients

(A) Representative contour plot of CyTOF analysis of 14 melanoma samples and 5 healthy controls to identify CD71 positive cells among CD66b^{+}CD15^{+} CD16^{-}CD10^{-} neutrophils in blood of healthy donors and melanoma patients suggested expansion of neutrophil progenitor subset N1 in blood of cancer patients. Gating also included exclusion of dead cells and cells expressing CD3, CD19, CD56, and CD14 (not shown).

(legend continued on next page)

CD86, and CD64) compared to CD71$^-$ neutrophils in melanoma patients. Expression of HLA-DR, CD86, and CD64 was not detected in CD71$^+$ or CD71$^-$ neutrophils from healthy human bone marrow, suggesting a capacity to modulate T cell-mediated adaptive immunity may be acquired in melanoma. We also took advantage of published neutrophil scRNA-seq data from patients with non-small cell lung carcinoma (NSCLC) (Zilionis et al., 2019). Despite the potential dropout of scRNA-seq, we were able to identify a similar percentage of CD71$^+$ neutrophils in the blood of NSCLC patients (Figure 7D). Noticeably, we found even higher numbers of these CD71$^+$ neutrophils in NSCLC tumors (10-fold change; Figure 7D), which indicated a possible accumulation of proliferative neutrophils in the tumor microenvironment compared to that in the circulation. Moreover, CD71$^+$ neutrophils in tumors lowly expressed maturation genes (*FCGR3A*, *ITGAM*, and *FCN1*) while maintaining high expression of neutrophil markers *S100A8/9* (Figure 7E), confirming the CyTOF data from melanoma blood. Together, these results reveal an expansion of early bone marrow neutrophil progenitors in the blood and tumors of cancer patients.

DISCUSSION

By using mass cytometry and single-cell clustering, we identified five CD66b$^+$Siglec8$^-$ neutrophil subsets in human bone marrow. We have updated the current paradigm of neutrophil bone marrow development while also revealing defining surface markers of specific stages in this process. The preNeu-like subset contained a rare CD71$^+$CD117$^+$ population: eNePs. eNePs were identified here as the earliest neutrophil progenitors in human bone marrow, as evidenced by their morphology, expression of neutrophil surface markers, colony-forming unit development, their ability to produce solely neutrophils *in vivo*, and gene-expression signature.

CD71, a transferrin receptor facilitating cellular iron uptake, is commonly used to identify erythroid progenitor cells (Mori et al., 2015; Marsee et al., 2010) and is a useful diagnostic marker for certain leukemias and lymphomas (Feremans et al., 1991; Beguin et al., 1993; Wu et al., 2006). More recently, CD71 expression in non-hematopoietic cancers has been reported to be associated with aggressive tumor behavior and poor patient outcome (Leung et al., 2019; Chan et al., 2014; Habashy et al., 2010). Here, CD71$^+$ neutrophil progenitors were expanded in blood from melanoma patients and found in both blood and tumor of lung cancer patients, in line with the concept of cancer-associated early egress of myeloid progenitor cells from bone marrow (Talmadge and Gabrilovich, 2013).

One of the main objectives of this study was to resolve conflicts concerning bone marrow neutrophil progenitors. Clustering of CyTOF data revealed that subset N1 showed similar surface marker expression to a recently identified neutrophil precursor: preNeu (Evrard et al., 2018). PreNeus were originally defined as proliferating bone marrow neutrophils in mice. Their ability to differentiate into immature and mature neutrophils, their expression of transcription factors associated with neutrophil development, and their close proximity to CXCL12-producing reticular cells in the bone marrow confirmed their status as murine neutrophil progenitors. In human bone marrow, Evrard and colleagues stratified CD15$^+$CD66b$^+$ neutrophils according to their ability to proliferate, as measured by IdU incorporation. Surface markers detected on proliferating neutrophils were then used to create a gating strategy to identify preNeus. PreNeus represented around 5% of bone marrow neutrophils, similar to the frequency of our N1 (preNeu-like) subset. Nevertheless, stratifying bone marrow neutrophils according to proliferation excludes quiescent pool progenitor cells, which are a critical component of the hematopoietic progenitor compartment (Chen et al., 2017). Our eNeP progenitor population, on the other hand, contained both dividing and non-dividing fractions, demonstrating the existence of an actively proliferating and resting or less-proliferative pool. Furthermore, eNePs were more potent in producing granulocyte colonies *in vitro* compared with preNeu-like cells, collectively suggesting that eNePs preceded preNeus in neutrophil development. Upstream of eNePs are GMPs, which are also CD34$^+$ and CD38$^+$ similar to CD34$^+$ eNePs. However, GMPs differ from eNePs by the absence of neutrophil lineage markers CD66b and CD15 (Olweus et al., 1995; Scheinecker et al., 1995).

We previously identified a neutrophil progenitor subset termed hNePs. Like preNeus, hNePs were first discovered in mice and subsequently confirmed in human bone marrow. Human hNePs were CD66b$^+$CD117$^+$ and consisted of a CD34 positive and negative population, both of which produce only neutrophils in transfer experiments into NSG-SGM3 mice (Zhu et al., 2018). These findings, combined with an hNeP frequency less than 3% of neutrophils in human bone marrow, led us to conclude that hNePs are a pure progenitor population. CyTOF analysis of surface markers, however, revealed that hNePs are a heterogeneous population of neutrophils, owing to the expression of CD117 within each of our five bone marrow subsets. Flow cytometry confirmed CD117 coexpression on neutrophils that simultaneously expressed maturation markers such as CD11b, CD101, and CD16. However, the combination of CD117 with CD71 identified eNePs as pure neutrophil progenitors, which represent approximately 0.15% of bone marrow neutrophils. Surface marker expression and morphology of eNePs are

(B) Frequency of neutrophil progenitor subset N1 within all leukocytes or ratio of subset N-to-CD66b$^+$ neutrophils in blood of healthy donors and melanoma patients. Data analyzed according to results shown in (A). n = 5 healthy donors, n = 14 melanoma patients. Welch's t test was performed, and p value shown.
(C) Differential expression test of protein marker expression between CD71− and CD71+ neutrophils (Wilcoxon rank sum test; Figure S7C) revealed a number of important markers including progenitor markers CD38, CD48, and CD49d; maturation markers CD16, CD10, CD35, and CD101; antigen-presenting markers CD86, HLA-DR, and CD64; and angiogenesis-associated marker CD304 (star indicating FDR-corrected p values < 0.01, 2-fold change difference in the expression median; Figure S7C). Expression value was normalized from 0–1 for each marker.
(D) CD71$^+$ eNePs are found in the blood of NSCLC patients and in lung tumor samples (from our reanalysis of scRNA-seq datasets; Zilionis et al., 2019). Each dot is one cell, red dots representing positive CD71 expression.
(E) Violin plot from scRNA-seq of CD71$^+$ neutrophils and other neutrophils (CD71$^-$, dataset also depicted in D).
See also Figure S7.

consistent with neutrophil progenitors and, importantly, they are endowed with the ability to form granulocyte colonies. Furthermore, reanalysis of CD34 positive and negative hNePs that produced only neutrophils in NSG-SGM3 mice confirmed the presence of $CD71^+$ cells in $CD34^+$ hNePs and $CD34^-$ hNePs and the progenitor potency of eNePs. Therefore, CD117 expression is not specific for progenitor populations but, in combination with CD66b and CD71, reliably identifies human neutrophil progenitors. The significance of CD117 in the expansion of human bone marrow progenitor populations has been known for decades (Brandt et al., 1992) Recently, depletion of $CD117^+$ leukocytes in non-human primates or humanized NSG mice successfully eliminated hematopoietic stem cells and allowed for the engraftment of donor hematopoietic stem cells (Kwon et al., 2019).

Our RNA-seq-based profiling of eNePs and different neutrophil subsets supports that eNePs are developmentally early neutrophils, compared to other subsets. eNePs have the highest expression of neutrophil progenitor genes *MPO*, *AZU1*, *PRTN3*, *CTSG*, *PRSS57*, and *CEBPA* and the lowest expression of *FCGR3A (CD16)*, *ITGAM (CD11b)*, and *S100A8/9.* These genes were recently defined in neutrophil RNA-seq data (Grassi et al., 2018). Although the transcriptome of eNePs was similar to the preNeu-like subset, we showed that eNePs had higher expression of early neutrophil genes and lower expression of genes specific to mature neutrophils compared to the preNeu-like subset. Transcriptomic data revealed gene set enrichment for cell adhesion and angiogenesis GO categories that are expressed in both eNePs and mature neutrophils, but not in N1 w/o eNePs. Although these genes may help mature neutrophils traffic and function in the periphery, these gene sets may also provide insights into how neutrophil progenitors are retained within the bone marrow during homeostasis, given that preNeus are located in close proximity to niche reticular cells and endothelial cells (Evrard et al., 2018). These findings show that the eNeP is likely an upstream progenitor of the preNeu subset.

Single-cell transcriptomic approaches have recently been used to identify six neutrophil subsets from peripheral blood and five neutrophil subsets in tumors from lung cancer patients (Zilionis et al., 2019). We observed the presence of the $CD71^+$ neutrophil progenitor subset in blood and tumor samples. However, most early neutrophil markers and transcription factors were not prominently expressed, likely because of the challenges of profiling neutrophil transcripts at single-cell resolution. Whether each of these subsets is endowed with specific functions and contributes to immunosuppression in cancer remains to be studied.

In summary, we provide a framework for neutrophil development in human bone marrow, beginning with eNeP as the earliest unipotent neutrophil progenitor. This framework will help for future studies of neutrophil heterogeneity in human diseases.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.immuni.2020.07.017.

ACKNOWLEDGMENTS

We thank Daniel Araujo for critical editing and helpful discussions. Sequencing data were generated from a grant supported by the Tullie and Rickey Families' SPARK Awards program at La Jolla Institute (LJI) for Immunology to H.Q.D. M.A.M acknowledges support from the National Institutes of Health (2T32AR064194). This work was supported by NIH grants U01 CA224766, R01 CA202987, R01 HL134236, and P01 HL136275 (all to C.C.H.). We would like to thank the LJI Flow Cytometry Core Facility for assistance with cell sorting. The FACSAria-3 cell sorter and the Fluidigm CyTOF Helios were acquired through the NIH Shared Instrumentation Grant (SIG) Program S10 (FACSAria II Cell Sorter S10 RR027366 and CyTOF Mass Cytometer S10 OD018499 to LJI).

AUTHOR CONTRIBUTIONS

H.Q.D, T.E., and C.C.H. conceptualized the project. H.Q.D., T.E., and C.E.O. analyzed CyTOF and RNA-seq data. Y.P.Z. designed the CyTOF panel. T.E., Y.P.Z., R.L., M.A.M., R.W. performed CyTOF and flow cytometry experiments. T.E. prepared samples for RNA-seq, analyzed flow cytometry data, and performed and analyzed Methocult experiments. M.A.M. performed the NSG-SGM3 adoptive transfer experiments and analyzed the data. H.Q.D. and T.E. drafted the original figures and manuscript with significant contributions from M.A.M. and C.E.O. All authors reviewed and edited the manuscripts. C.C.H. and H.Q.D. acquired funding and supervised the project.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: January 4, 2020
Revised: April 18, 2020
Accepted: July 23, 2020
Published: August 18, 2020

REFERENCES

Aisen, P. (2004). Transferrin receptor 1. Int. J. Biochem. Cell Biol. *36*, 2137–2143.

Alexa, A., Rahnenführer, J., and Lengauer, T. (2006). Improved scoring of functional groups from gene expression data by decorrelating GO graph structure. Bioinformatics *22*, 1600–1607.

Bainton, D.F., Ullyot, J.L., and Farquhar, M.G. (1971). The development of neutrophilic polymorphonuclear leukocytes in human bone marrow. J. Exp. Med. *134*, 907–934.

Becht, E., McInnes, L., Healy, J., Dutertre, C.-A., Kwok, I.W.H., Ng, L.G., Ginhoux, F., and Newell, E.W. (2018). Dimensionality reduction for visualizing single-cell data using UMAP. Nat. Biotechnol. https://doi.org/10.1038/nbt.4314.

Beguin, Y., Lampertz, S., De Groote, D., Igot, D., Malaise, M., and Fillet, G. (1993). Soluble CD23 and other receptors (CD4, CD8, CD25, CD71) in serum of patients with chronic lymphocytic leukemia. Leukemia *7*, 2019–2025.

Benjamini, Y., and Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. J. R. Stat. Soc. B *57*, 289–300.

Brandt, J., Briddell, R.A., Srour, E.F., Leemhuis, T.B., and Hoffman, R. (1992). Role of c-kit ligand in the expansion of human hematopoietic progenitor cells. Blood *79*, 634–641.

Chan, K.T., Choi, M.Y., Lai, K.K.Y., Tan, W., Tung, L.N., Lam, H.Y., Tong, D.K.H., Lee, N.P., and Law, S. (2014). Overexpression of transferrin receptor CD71 and its tumorigenic properties in esophageal squamous cell carcinoma. Oncol. Rep. *31*, 1296–1304.

Chen, H., Lau, M.C., Wong, M.T., Newell, E.W., Poidinger, M., and Chen, J. (2016). Cytofkit: A Bioconductor Package for an Integrated Mass Cytometry Data Analysis Pipeline. PLoS Comput. Biol. *12*, e1005112.

Chen, X., Deng, H., Churchill, M.J., Luchsinger, L.L., Du, X., Chu, T.H., Friedman, R.A., Middelhoff, M., Ding, H., Tailor, Y.H., et al. (2017). Bone Marrow Myeloid Cells Regulate Myeloid-Biased Hematopoietic Stem Cells via a Histamine-Dependent Feedback Loop. Cell Stem Cell *21*, 747–760.e7.

Chen, X.J., Wu, S., Yan, R.M., Fan, L.S., Yu, L., Zhang, Y.M., Wei, W.F., Zhou, C.F., Wu, X.G., Zhong, M., et al. (2019). The role of the hypoxia-Nrp-1 axis in the activation of M2-like tumor-associated macrophages in the tumor microenvironment of cervical cancer. Mol. Carcinog. *58*, 388–397.

Clark, S.R., Ma, A.C., Tavener, S.A., McDonald, B., Goodarzi, Z., Kelly, M.M., Patel, K.D., Chakrabarti, S., McAvoy, E., Sinclair, G.D., et al. (2007). Platelet TLR4 activates neutrophil extracellular traps to ensnare bacteria in septic blood. Nat. Med. *13*, 463–469.

Coifman, R.R., Lafon, S., Lee, A.B., Maggioni, M., Nadler, B., Warner, F., and Zucker, S.W. (2005). Geometric diffusions as a tool for harmonic analysis and structure definition of data: diffusion maps. Proc. Natl. Acad. Sci. USA *102*, 7426–7431.

Cowland, J.B., and Borregaard, N. (2016). Granulopoiesis and granules of human neutrophils. Immunol. Rev. *273*, 11–28.

Drissen, R., Buza-Vidas, N., Woll, P., Thongjuea, S., Gambardella, A., Giustacchini, A., Mancini, E., Zriwil, A., Lutteropp, M., Grover, A., et al. (2016). Distinct myeloid progenitor-differentiation pathways identified through single-cell RNA sequencing. Nat. Immunol. *17*, 666–676.

Drissen, R., Thongjuea, S., Theilgaard-Mönch, K., and Nerlov, C. (2019). Identification of two distinct pathways of human myelopoiesis. Sci. Immunol. *4*, eaau7148.

Evrard, M., Kwok, I.W.H., Chong, S.Z., Teng, K.W.W., Becht, E., Chen, J., Sieow, J.L., Penny, H.L., Ching, G.C., Devi, S., et al. (2018). Developmental Analysis of Bone Marrow Neutrophils Reveals Populations Specialized in Expansion, Trafficking, and Effector Functions. Immunity *48*, 364–379.e8.

Feremans, W., Bujan, W., Neve, P., Delville, J.P., and Schandene, L. (1991). CD71 phenotype and the value of gallium imaging in lymphomas. Am. J. Hematol. *36*, 215–216.

Finck, R., Simonds, E.F., Jager, A., Krishnaswamy, S., Sachs, K., Fantl, W., Pe'er, D., Nolan, G.P., and Bendall, S.C. (2013). Normalization of mass cytometry data with bead standards. Cytometry *83A*, 483–494.

Gelfand, M.V., Hagan, N., Tata, A., Oh, W.J., Lacoste, B., Kang, K.T., Kopycinska, J., Bischoff, J., Wang, J.H., and Ga, C. (2014). Neuropilin-1 functions as a VEGFR2 co-receptor to guide developmental angiogenesis independent of ligand binding. Elife *3*, e03720.

Gentles, A.J., Newman, A.M., Liu, C.L., Bratman, S.V., Feng, W., Kim, D., Nair, V.S., Xu, Y., Khuong, A., Hoang, C.D., et al. (2015). The prognostic landscape of genes and infiltrating immune cells across human cancers. Nat. Med. *21*, 938–945.

Grassi, L., Pourfarzad, F., Ullrich, S., Merkel, A., Were, F., Carrillo-de-Santa-Pau, E., Yi, G., Hiemstra, I.H., Tool, A.T.J., Mul, E., et al. (2018). Dynamics of Transcription Regulation in Human Bone Marrow Myeloid Differentiation to Mature Blood Neutrophils. Cell Rep. *24*, 2784–2794.

Habashy, H.O., Powe, D.G., Staka, C.M., Rakha, E.A., Ball, G., Green, A.R., Aleskandarany, M., Paish, E.C., Douglas Macmillan, R., Nicholson, R.I., et al. (2010). Transferrin receptor (CD71) is a marker of poor prognosis in breast cancer and can predict response to tamoxifen. Breast Cancer Res. Treat. *119*, 283–293.

Hidalgo, A., Chilvers, E.R., Summers, C., and Koenderman, L. (2019). The Neutrophil Life Cycle. Trends Immunol. *40*, 584–597.

Kwon, H.-S., Logan, A.C., Chhabra, A., Pang, W.W., Czechowicz, A., Tate, K., Le, A., Poyser, J., Hollis, R., Kelly, B.V., et al. (2019). Anti-human CD117 antibody-mediated bone marrow niche clearance in nonhuman primates and humanized NSG mice. Blood *133*, 2104–2108.

Lahoz-Beneytez, J., Elemans, M., Zhang, Y., Ahmed, R., Salam, A., Block, M., Niederalt, C., Asquith, B., and Macallan, D. (2016). Human neutrophil kinetics: modeling of stable isotope labeling data supports short blood neutrophil half-lives. Blood *127*, 3431–3438.

Law, C.W., Chen, Y., Shi, W., and Smyth, G.K. (2014). voom: Precision weights unlock linear model analysis tools for RNA-seq read counts. Genome Biol. *15*, R29.

Leung, T.H.-Y., Tang, H.W.-M., Siu, M.K.-Y., Chan, D.W., Chan, K.K.-L., Cheung, A.N.-Y., and Ngan, H.Y.-S. (2019). $CD71^{+}$ Population Enriched by HPV-E6 Protein Promotes Cancer Aggressiveness and Radioresistance in Cervical Cancer Cells. Mol. Cancer Res. *17*, 1867–1880.

Levine, J.H., Simonds, E.F., Bendall, S.C., Davis, K.L., Amir, A.D., Tadmor, M.D., Litvin, O., Fienberg, H.G., Jager, A., Zunder, E.R., et al. (2015). Data-Driven Phenotypic Dissection of AML Reveals Progenitor-like Cells that Correlate with Prognosis. Cell *162*, 184–197.

Ley, K., Hoffman, H.M., Kubes, P., Cassatella, M.A., Zychlinsky, A., Hedrick, C.C., and Catz, S.D. (2018). Neutrophils: New insights and open questions. Sci. Immunol. *3*, eaat4579.

Liao, Y., Smyth, G.K., and Shi, W. (2013). The Subread aligner: fast, accurate and scalable read mapping by seed-and-vote. Nucleic Acids Res. *41*, e108.

Liao, Y., Smyth, G.K., and Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics *30*, 923–930.

Marsee, D.K., Pinkus, G.S., and Yu, H. (2010). CD71 (transferrin receptor): an effective marker for erythroid precursors in bone marrow biopsy specimens. Am. J. Clin. Pathol. *134*, 429–435.

Mori, Y., Iwasaki, H., Kohno, K., Yoshimoto, G., Kikushige, Y., Okeda, A., Uike, N., Niiro, H., Takenaka, K., Nagafuji, K., et al. (2009). Identification of the human eosinophil lineage-committed progenitor: revision of phenotypic definition of the human common myeloid progenitor. J. Exp. Med. *206*, 183–193.

Mori, Y., Chen, J.Y., Pluvinage, J.V., Seita, J., and Weissman, I.L. (2015). Prospective isolation of human erythroid lineage-committed progenitors. Proc. Natl. Acad. Sci. USA *112*, 9638–9643.

Motamedi, M., Xu, L., and Elahi, S. (2016). Correlation of transferrin receptor (CD71) with Ki67 expression on stimulated human and mouse T cells: The kinetics of expression of T cell activation markers. J. Immunol. Methods *437*, 43–52.

Ng, L.G., Ostuni, R., and Hidalgo, A. (2019). Heterogeneity of neutrophils. Nat. Rev. Immunol. *19*, 255–265.

Notta, F., Zandi, S., Takayama, N., Dobson, S., Gan, O.I., Wilson, G., Kaufmann, K.B., McLeod, J., Laurenti, E., Dunant, C.F., et al. (2016). Distinct routes of lineage development reshape the human blood hierarchy across ontogeny. Science *351*, aab2116.

Olweus, J., Lund-Johansen, F., and Terstappen, L.W. (1995). CD64/Fc gamma RI is a granulo-monocytic lineage marker on CD34+ hematopoietic progenitor cells. Blood *85*, 2402–2413.

Pliyev, B.K., Shmidt, E.I., Ivanova, A.V., and Menshikov, M. (2012). Circulating CD35(-)/CD49d(+) neutrophils in influenza virus infection patients. Hum. Immunol. *73*, 1087–1090.

Robinson, M.D., McCarthy, D.J., and Smyth, G.K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics *26*, 139–140.

Scheinecker, C., Strobl, H., Fritsch, G., Csmarits, B., Krieger, O., Majdic, O., and Knapp, W. (1995). Granulomonocyte-associated lysosomal protein expression during in vitro expansion and differentiation of CD34+ hematopoietic progenitor cells. Blood *86*, 4115–4123.

Sidney, L.E., Branch, M.J., Dunphy, S.E., Dua, H.S., and Hopkinson, A. (2014). Concise review: evidence for CD34 as a common marker for diverse progenitors. Stem Cells *32*, 1380–1389.

Sierro, F., Evrard, M., Rizzetto, S., Melino, M., Mitchell, A.J., Florido, M., Beattie, L., Walters, S.B., Tay, S.S., Lu, B., et al. (2017). A Liver Capsular Network of Monocyte-Derived Macrophages Restricts Hepatic Dissemination of Intraperitoneal Bacteria by Neutrophil Recruitment. Immunity *47*, 374–388.e6.

Summers, C., Rankin, S.M., Condliffe, A.M., Singh, N., Peters, A.M., and Chilvers, E.R. (2010). Neutrophil kinetics in health and disease. Trends Immunol. *31*, 318–324.

Talmadge, J.E., and Gabrilovich, D.I. (2013). History of myeloid-derived suppressor cells. Nat. Rev. Cancer *13*, 739–752.

Wilkerson, M.D., and Hayes, D.N. (2010). ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking. Bioinformatics *26*, 1572–1573.

Wu, J.M., Borowitz, M.J., and Weir, E.G. (2006). The usefulness of CD71 expression by flow cytometry for differentiating indolent from aggressive CD10+ B-cell lymphomas. Am. J. Clin. Pathol. *126*, 39–46.

Yáñez, A., Coetzee, S.G., Olsson, A., Muench, D.E., Berman, B.P., Hazelett, D.J., Salomonis, N., Grimes, H.L., and Goodridge, H.S. (2017). Granulocyte-Monocyte Progenitors and Monocyte-Dendritic Cell Progenitors Independently Produce Functionally Distinct Monocytes. Immunity *47*, 890–902.e4.

Zhu, Y.P., Padgett, L., Dinh, H.Q., Marcovecchio, P., Blatchley, A., Wu, R., Ehinger, E., Kim, C., Mikulski, Z., Seumois, G., et al. (2018). Identification of an Early Unipotent Neutrophil Progenitor with Pro-tumoral Activity in Mouse and Human Bone Marrow. Cell Rep. *24*, 2329–2341.e8.

Zhu, Y.P., Padgett, L., Dinh, H.Q., Marcovecchio, P., Wu, R., Hinz, D., Kim, C., and Hedrick, C.C. (2019). Preparation of Whole Bone Marrow for Mass Cytometry Analysis of Neutrophil-lineage Cells. J. Vis. Exp. *148*, e59617.

Zilionis, R., Engblom, C., Pfirschke, C., Savova, V., Zemmour, D., Saatcioglu, H.D., Krishnan, I., Maroni, G., Meyerovitz, C.V., Kerwin, C.M., et al. (2019). Single-Cell Transcriptomics of Human and Mouse Lung Cancers Reveals Conserved Myeloid Populations across Individuals and Species. Immunity *50*, 1317–1334.e10.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
CyTOF Antibodies		
Anti-human CD45 (Clone: HI30) - 89Y	Fluidigm	Cat#: 3089003B, RRID: AB_2661851
Anti-human CD3 (Clone: UCHT1 - Purified	Biolegend	Cat#: 300443, RRID: AB_2562808
Anti-human CD127 (Clone: A019D5) - Purified	Biolegend	Cat#: 351302, RRID: AB_10718513
Anti-human CD41 (Clone: HIP8) - Purified	Biolegend	Cat# 303702, RRID: AB_314372
Anti-human CD235a (Clone: HIR2) – Purified	Biolegend	Cat# 306602, RRID: AB_314620
Anti-human CD11c (Clone: 3.9) – Purified	Biolegend	Cat#: 301602, RRID: AB_314172
Anti-human CD11a (Clone: HI111) – 142Nd	Fluidigm	Cat#: 3142006B
Anti-human CD123 (Clone: 6H6) – 143Nd	Fluidigm	Cat#: 3143014B, RRID: AB_2811081
Anti-human CD203c (Clone: NP4D6) – Purified	Biolegend	Cat# 324602, RRID: AB_756040
Anti-human CD19 (Clone: HIB19) – Purified	Biolegend	Cat#: 302202, RRID: AB_314232
Anti-human CD62L (Clone: DREG-56) – Purified	Biolegend	Cat#: 304802, RRID: AB_314462
Anti-human CD64 (Clone: 10.1) – 146Nd	Fluidigm	Cat#: 3146006B, RRID: AB_2661790
Anti-human CD182 (Clone: 5E8/CXCR2) – 147Sm	Fluidigm	Cat#: 3147010B
Anti-human/anti-mouse CD11b (Clone: M1/70) – Purified	Biolegend	Cat#: 101214 (Discontinued), RRID: AB_312797
Anti-human CD48 (Clone: BJ40) – Purified	Biolegend	Cat#: 336702, RRID: AB_1227561
Anti-human CD49d (Clone: 9F10) – Purified	Biolegend	Cat#: 304319, RRID: AB_2563755
Anti-human CD66b (Clone: 80H3) – 152Sm	Fluidigm	Cat#: 3152011B, RRID: AB_2661795
Anti-human CD14 (Clone: M5E2) – Purified	Biolegend	Cat#: 301802, RRID: AB_314184
Anti-human CD117 (Clone: 104D2) – Purified	Biolegend	Cat#: 313202, RRID: AB_314981
Anti-human CD45RA (Clone: HI100) – 155Gd	Fluidigm	Cat#: 3155011B, RRID: AB_2810246
Anti-human CD10 (Clone: HI10a) – 156Gd	Fluidigm	Cat#: 3156001B, RRID: AB_2802107
Anti-human CD101 (Clone: BB27) – 158Gd	Fluidigm	Cat#: 3158020B
Anti-human CD197 (Clone G043H7) – 159Tb	Fluidigm	Cat#: 3159003A, RRID: AB_2714155
Anti-human CD79b (Clone: CB3-1) – Purified	BD	Cat#: 555678, RRID: AB_396031
Anti-human Arginase I (Clone: 14D2C43) -Purified	Biolegend	Cat#: 369702, RRID: AB_2571898
Anti-human CD71 (Clone: CY1G4) – Purified	Biolegend	Cat#: 334102, RRID: AB_1134247
Anti-APC (Clone: APC003) – 163Dy	Fluidigm	Cat#: 3163001B, RRID: AB_2687636
Anti-human CD86 (Clone: BU63) – APC	Biolegend	Cat#: 374208, RRID: AB_2721449
Anti-human Siglec 8 (Clone: 7C9) – 164Dy	Fluidigm	Cat#: 3164017B
Anti-human CD16 (Clone: 3G8) – 165Ho	Fluidigm	Cat#: 3165001B, RRID: AB_2802109
Anti-human CD34 (Clone: 581) – 166Er	Fluidigm	Cat#: 3166012B, RRID: AB_2756424
Anti-human CD38 (Clone: HIT2) – 167Er	Fluidigm	Cat#: 3167001B, RRID: AB_2802110
Anti-human CD304 (Clone: BDCA-4) – Purified	Biolegend	Cat#: 354502, RRID: AB_2564475
Anti-human CD33 (Clone: WM53) – 169Tm	Fluidigm	Cat#: 3169010B, RRID: AB_2802111
Anti-human CD35 (Clone: E11) – Purified	Biolegend	Cat#: 333402, RRID: AB_1089032
Anti-human HLA-ABC (Clone: W6/32) – Purified	Biolegend	Cat#: 311402, RRID: AB_314871
Anti-human CD15 (Clone: W6D3) – 172Yb	Fluidigm	Cat#: 3172021B, RRID: AB_2802113
Anti-human HLA-DR (Clone: L243) – Purified	Biolegend	Cat#: 307602, RRID: AB_314680
Anti-human CD184 (Clone: 12G5) – 175Lu	Fluidigm	Cat#: 3175001B
Anti-human CD56 (Clone: R19760) – 176Yb	Fluidigm	Cat#: 3176013B
Flow Cytometry Antibodies		
Anti-human CD34 (Clone: 561) – Brilliant Violet 421	Biolegend	Cat#: 343610,RRID: AB_2561358
Anti-mouse CD45 (Clone: 30-F11) – Brilliant Violet 421	Biolegend	Cat#: 103134, RRID: AB_2562559

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Anti-human CD10 (Clone: HI10a) – Brilliant Violet 510	Biolegend	Cat#: 312220, RRID: AB_2563835
Anti-human CD45 (Clone: HI30) – Brilliant Violet 570	Biolegend	Cat#: 304034, RRID: AB_2563426
Anti-human CD35 (Clone: E11) – Brilliant Violet 605	BD	Cat#: 744276, RRID: AB_2742114
Anti-human CD71 (Clone: CY1G4) – Brilliant Violet 650	Biolegend	Cat#: 334116, RRID: AB_2687103
Anti-human CD49d (Clone: 9F10) – Brilliant Violet 711	Biolegend	Cat#: 304332, RRID: AB_2687198
Anti-mouse/anti-human CD11b (Clone: M1/70) – Brilliant Violet 785	Biolegend	Cat#: 101243, RRID: AB_2561373
Anti-human CD117 (Clone: 104D2) – Brilliant Violet 785	Biolegend	Cat#: 313238, RRID: AB_2629837
Anti-human CD7 (Clone: M-T701) – BB515	BD	Cat#: 565211, RRID: AB_2739113
Anti-human CD127 (Clone: A019D5) – FITC	Biolegend	Cat#: 351312, RRID: AB_10933247
Anti-human CD90 (Clone: 5E10) – FITC	Biolegend	Cat# 328108, RRID: AB_893438
Anti-human CD3 (Clone: HIT3a) – FITC	Biolegend	Cat#: 300306, RRID: AB_314042
Anti-human CD19 (Clone: HIB19) – FITC	Biolegend	Cat#: 302206, RRID: AB_314236
Anti-human CD56 (Clone: B159) – BB515	Biolegend	Cat#: 564488
Anti-human CD161 (Clone: HP-3G10) – Alexa Fluor 488	Biolegend	Cat#: 339924, RRID: AB_2563939
Anti-human CD41 (Clone: HIP8) – FITC	Biolegend	Cat#: 303704, RRID: AB_314374
Anti-human CD235a (Clone: HI264) – FITC	Biolegend	Cat#: 349104, RRID: AB_10613463
Anti-human CD123 (Clone: 6H6) – FITC	Biolegend	Cat#: 306014, RRID: AB_2124259
Anti-human CD11c (Clone: Bu15) – FITC	Biolegend	Cat#: 337214, RRID: AB_2129792
Anti-human Siglec-1 (Clone: 7-239) – BB515	BD	Cat#: 565353, RRID: AB_2739204
Anti-human Siglec-8 (Clone: 7C9) – FITC	Miltenyi Biotec	Cat#: 130-098-716, RRID: AB_2653426
Anti-human FceRIa (Clone: AER-37) – FITC	Biolegend	Cat#: 334608, RRID: AB_1227653
Anti-human CD86 (Clone: 2331 (FUN-1)) – BB515	BD	Cat#: 564544, RRID: AB_2744453
Anti-human CD14 (Clone: M5E2) - FITC	Biolegend	Cat#: 301804, RRID: AB_314186
Anti-human CD125 (Clone: 26815) – Alexa Fluor 488	R&D Systems	Cat#: FAB253G-100UG
Anti-human CD66b (Clone: G10F5) – PE	Biolegend	Cat#: 305106, RRID: AB_2077857
Anti-human CD38 (Clone; HB-7) – PE/Dazzle 594	Biolegend	Cat#: 356630, RRID: AB_2650757
Anti-human CD203c (Clone: NP4D6) – PE/Dazzle 594	Biolegend	Cat#: 324624, RRID: AB_2566235
Anti-human Siglec-8 (Clone: 7C9) – PE/Dazzle 594	Biolegend	Cat#: 347110, RRID: AB_2629718
Anti-human CD16 (Clone: 3G8) – PerCP	Biolegend	Cat#: 302030, RRID: AB_940380
Anti-human CD3 (Clone: OKT3) – PerCP-Cy5.5	Tonbo Bioscience	Cat#: 65-0037-T100
Anti-human CD56 (Clone: MEM-188) – PerCP-Cy5.5	Biolegend	Cat#: 304626, RRID: AB_10641700
Anti-human CD101 (Clone: BB27) – PE-Cy7	Biolegend	Cat#: 331014, RRID: AB_2716109
Anti-human CD19 (Clone: HIB19) – PE-Cy7	Biolegend	Cat#: 302216, RRID: AB_314246
Anti-human CD15 (Clone: HI98) – APC	Biolegend	Cat#: 301908, RRID: AB_314200
Anti-mouse CD45 (Clone: 30-F11) – APC-Cy7	Biolegend	Cat#: 103116, RRID: AB_312981
Anti-human CD117 (Clone: YB5.B8) – APC-R700	BD	Cat#: 565195, RRID: AB_2716871
Anti-human CD45 (Clone: HI30) – Alexa Fluor 700	Biolegend	Cat#: 304024, RRID: AB_493761
Biological Samples		
Healthy human bone marrow	AllCells, Inc.	BM, FR, 3mL
Healthy human peripheral blood	La Jolla Institute for Allergy and Immunology, Clinical studies core, Normal Blood Donor Program (NBDP)	https://www.lji.org/faculty-research/scientific-cores/clinical-studies
Melanoma patient peripheral blood	University of Kansas Cancer Center, Biospecimen Repository Core Facility (BRCF)	http://www.kumc.edu/school-of-medicine/biospecimen.html
Chemicals, Peptides, and Recombinant Proteins		
MAXPAR® Antibody Labeling Kits	Fluidigm	https://www.dvssciences.com/product-catalog-maxpar.php

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Cisplatin-194Pt	Fluidigm	Cat# 201194
Intercalator-Ir	Fluidigm	Cat# 201192B
Hyclone Phosphate Buffered Saline	Fisher Scientific	Cat#: SH30256FS
Ethylenediaminetetraacetic acid (EDTA)	ThermoFisher	Cat# AM9260G
RBC Lysis Buffer, 10x	Biolegend	Cat#: 420302
Antibody Stabilizer	CANDOR Bioscience	Cat# 130050
Sodium azide	Sigma-Aldrich	Cat# S2002
Paraformaldehyde	Sigma-Aldrich	Cat# 158127
Bovine Serum Albumin Solution	Millipore Sigma	Cat#: A9576
Maxpar Fix and Perm Buffer	Fluidigm	Cat#: 201067
Maxpar Cell Acquisition Solution	Fluidigm	Cat#: 201240
EQ Four Element Calibration Beads	Fluidigm	Cat# 201078
Human AB Serum, Heat Inactivated	Omega Scientific	Cat# HS-25
Flavopiridol hydrocholoride hydrate	Milipore Sigma	Cat# F3055-25MG
Recombinant ribonuclease inhibitor	Takara	Cat#: 2313B
Fetal Bovine Serum	Omega Scientific	Cat# FB-02
TRIzol LS Reagent	Thermo Fisher	Cat#: 10296028
Methocult SF H4436	StemCell	Cat#: 04436
Critical Commercial Assays		
Zombie NIR Fixable Viability Kit	Biolegend	Cat#: 423105
LIVE/DEAD Fixable Blue Dead Cell Stain Kit, for UV excitation	Thermo Fisher	Cat#: L23105
Cell ID IdU	Fluidigm	Cat#: 201127
APC BrdU Flow Kit	BD	Cat#: 552598
EasySep Mouse CD11b Positive Selection Kit II	StemCell	Cat#: 18970
Hema 3 Manual Staining System	Fisher Scientific	Cat#: 23-123869
Direct-zol RNA Kit	Zymo Research	Cat#: R2051
Deposited Data		
Sorted human bone marrow RNA-Seq	This study	GSE153263
Lung cancer scRNA-Seq	Zilionis et al., 2019	GSE127465
Experimental Models		
Mouse: NOD.Cg-*Prkdc*^{scid} *Il2rg*^{tm1Wjl}Tg (CMV-IL3,CSF2,KITLG)1Eav/MloySzJ	The Jackson Laboratory	Stock No: 013062
Software and Algorithms		
Bead-based Normalizer	Finck et al., 2013	https://med.virginia.edu/flow-cytometry-facility/wp-content/uploads/sites/170/2015/10/3_Finck-Rachel_CUGM_May2013.pdf
RSubread R package v1.30.5	Liao et al., 2013	http://bioconductor.org/packages/release/bioc/html/Rsubread.html
edgeR v3.22.3	Robinson et al., 2010	https://bioconductor.org/packages/release/bioc/html/edgeR.html
FeatureCount R package v1.22.2	Liao et al., 2013	https://rdrr.io/bioc/Rsubread/man/featureCounts.html
Pheatmap	CRAN	https://cran.r-project.org/web/packages/pheatmap/index.html
Umap	CRAN	https://cran.r-project.org/web/packages/umap/index.html
Limma	Law et al., 2014	https://www.bioconductor.org/packages/release/bioc/html/limma.html

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Cytofkit	Chen et al., 2016	https://bioconductor.riken.jp/packages/3.7/bioc/html/cytofkit.html
topGO	Alexa et al., 2006	https://bioconductor.org/packages/release/bioc/html/topGO.html
ConsensusClusterPlus	Wilkerson and Hayes, 2010	https://bioconductor.org/packages/release/bioc/html/ConsensusClusterPlus.html

RESOURCE AVAILABILITY

Lead Contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Catherine C. Hedrick (hedrick@lji.org).

Materials Available

This study did not generate new unique reagents.

Data and Code Availability

The accession number for the bulk RNA-Seq dataset reported in this paper is Gene Expression Omnibus (GEO) GSE153263.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Mice

NSG-SGM3 mice were purchased from The Jackson Laboratory. Mice were housed in pathogen-free conditions in microisolator cages and fed standard rodent chow diet. Mice were euthanized by CO2 inhalation followed by cervical dislocation. All experiments were executed in accordance with La Jolla Institute for Immunology Animal Care and Use Committee guidelines. Approval for use of rodents was obtained from the La Jolla Institute for Immunology according to criteria outlined in the Guide for the Care and Use of Laboratory Animals from the National Institutes of Health. Male animals 7-8 weeks of age were used for experiments described in the NSG-SGM3 Adoptive Transfer section.

Human bone marrow cells

Fresh bone marrow samples from anonymous healthy adult donors were obtained from AllCells, Inc. (Alameda, CA). For information on donor characteristics, refer to Table S1. Upon arrival, the fresh cells were immediately stained for mass cytometry, flow cytometry or cell sorting. Sorted cells were used for *in vitro* progenitor differentiation assay, Cytospin with subsequent staining and imaging, RNA isolation, and NSG-SGM3 adoptive transfer.

Human Peripheral Blood Collection

Blood from untreated melanoma patients (no prior melanoma-related treatment) was collected in EDTA-tubes by the Biospecimen Repository Core Facility (BRCF) at the University of Kansas Cancer Center, shipped overnight and processed for mass cytometry staining immediately upon arrival. EDTA-coated blood from healthy volunteers was obtained after written informed consent under the guidelines of the Institutional Review Board of the La Jolla Institute for Allergy and Immunology and in accordance with US Department of Health and Human Services Policy for the protection of Human Research Subjects (VD-057-0217). Both healthy donors and melanoma patient blood samples were processed at the same time and were stained for mass cytometry followed by the protocol described in the Mass Cytometry (CyTOF) section. For information on donor characteristics, refer to Table S3.

METHOD DETAILS

Cell suspension for mass cytometry, flow cytometry and cell sorting

All samples (both bone marrow and peripheral blood) were collected in ice cold D-PBS (GIBCO) with 2 mM EDTA to prevent cation-dependent cell-cell adhesion. Cells were subsequently subjected to a red blood cell lysis (RBC lysis buffer, eBiosciences) at room temperature (5 min x 2 for BM cells, 10 min x 2 for blood cells), washed and filtered through a 70 um strainer to yield a single cell suspension.

Mass Cytometry Antibodies

Metal-conjugated antibodies were purchased directly from Fluidigm for available targets. For all other targets, purified antibodies were purchased from the companies as described before (Zhu et al., 2018). Antibody conjugations were prepared using the Maxpar Antibody Labeling Kit according to the recommended protocol provided by Fluidigm. Maxpar-conjugated antibodies

were stored in PBS-based antibody stabilization solution (Candor Biosciences) supplemented with 0.05% NaN3 at 4°C. All antibodies were titrated before use.

Mass Cytometry (CyTOF)

CyTOF was performed following previously described protocols (Zhu et al., 2019). For viability staining, cells were washed in PBS and stained with Cisplatin (Fluidigm) to a final concentration of 5 μM. Prior to surface staining, RBC-lysed WB cells were resuspended in staining buffer to block the Fc receptors for 15 min at RT. The surface antibody cocktail, containing CyTOF antibodies listed in Key Resource Table, was added into the cell suspension for 1 h at 4°C. The cells were then washed and fixed with 1.6% paraformaldehyde (Thermo Fisher) for 15 min at RT. Prepare 1 mL of intercalation solution for each sample by adding Cell-ID Intercalator-Ir (Fluidigm) into Maxpar Fix and Perm Buffer (Fluidigm) to a final concentration of 125 nM (a 1,000X dilution of the 125 μM stock solution) and vortex to mix. After fixation, resuspend the cells with the 1mL intercalation solution and incubate overnight at 4°C. Cells were then washed in staining buffer, with subsequent washes in Cell Acquisition Solution (CAS) (Fluidigm) to remove buffer salts. Cells were resuspended in CAS with a 1:10 dilution of EQ Four Element Calibration beads (Fluidigm) and filtered through a 35 μm nylon mesh filter cap (Corning, Falcon). Samples were analyzed on a Helios 2 CyTOF Mass Cytometer (Fluidigm) equipped with a Super Sampler (Victorian Airship & Scientific Apparatus) at an event rate of 500 events/second or less. Mass cytometry data files were normalized using the bead-based Normalizer (Finck et al., Cytometry A 83:48).

Computational analysis of mass cytometry

CyTOF data was analyzed using R and Bioconductor packages. Protein expression was normalized using arcsinh transformation (cofactor = 5). We used Phenograph clustering (Levine et al., 2015) implemented in cytofkit package (Chen et al., 2016) to identify neutrophils (CD66b+Siglec8-) and other leukocyte cell-types in bone marrow. Then we selected neutrophils out for further Phenograph clustering. Consensus clustering was used to justify the number of clusters from k = 2 to k = 15 and we identify k = 5 as the most optimal number of clusters based on relatively decrease in area under the CDF (Cumulative Distribution Function) curve (Wilkerson and Hayes 2010). Phenograph clusters were merged into 5 major clusters based on consensus clustering. Diffusion map was done using cytofkit package (Chen et al., 2016). UMAP deduction was done using umap R package (v0.2.3.1), a wrapper for Python package 'umap-learn'. Heatmap was produced using pheatmap R package (v0.2).

Flow Cytometry and Cell Sorting

Staining for flow cytometry and cell sorting was performed in FACS buffer (D-PBS + 1% human serum + 0.1% sodium azide + 2 mM EDTA) on ice. Staining for cell sorting and subsequent RNA isolation was performed in FACS buffer, supplemented with Flavopiridol (100 nM; Sigma-Aldrich) and recombinant RNase inhibitor (40 U/mL; Takara Bio USA, Inc.). Cells were stained with 100 ul of titrated antibody master mix including viability dye per 5 million cells, followed by a 30-min incubation on ice in the dark. Sorted cells were collected in receiving buffer (D-PBS + 10% FBS). For RNA isolation, cells were sorted into TRIzol™LS Reagent (ThermoFisher Scientific), mixed well and immediately frozen at −80°C until RNA isolation.

Flow cytometric analysis was carried out using a LSR-II (BD biosciences) or LSR Fortessa (BD bioscience) and cell sorting was performed on a FACS Aria-II (BD Biosciences). Both flow cytometric analyses and cell sorting were performed on live cells after exclusion of doublets as shown in Figure S4B. Furthermore, cells expressing the following lineage markers were also excluded: CD7, CD127, CD90, CD3, CD19, CD56, CD161, CD41, CD235a, CD123, CD125, CD11c, CD169, Siglec8, FceRIa, CD86.

Following sorting, purity of sorted fractions was checked visually and by FACS reanalysis of the surface markers. Data were analyzed using FlowJo software (version 10.5.3, BD).

NSG-SGM3 Mouse Adoptive Transfer

NSG-SGM3 recipient mice were kept under sterile conditions prior to and after adoptive transfer. Recipient mice were sub-lethally irradiated (240 rads). CD11b depletions was performed on fresh human bone marrow using CD11b Positive Selection Kit II (StemCell). eNeP and N1-eNeP were sorted from fresh human bone marrow using FACS as described above and collected into receiving buffer (D-PBS + 10% FBS). Sorted eNeP or N1-eNeP (32,000-154,000 progenitor cells isolated from 10 mL human bone marrow) were injected in 100ul D-PBS retro-orbitally. In each experiment, mice received equal numbers of either eNeP or N1-eNeP. We attempted 6 experiments. Two FACS sort of human bone marrow resulted in too few cells to inject. Of the other 4 experiments, each with adoptive transfer of either 32,000 eNeP and N1-eNeP, 50,000 eNeP and N1-eNeP, 48,000 eNeP and N1-eNeP, or 154,000 eNeP and N1-eNeP into recipient mice, three of the experiments engrafted. After 5 days, blood and bone marrow were analyzed by flow cytometry as described above.

Proliferation analysis measuring BrdU incorporation

To assess cellular proliferation, we performed a BrdU assay using the APC BrdU Flow Kit (BD Bioscience) according to manufacturer's instructions. Briefly, fresh bone marrow cells were incubated with culture medium containing BrdU (50 uM) and incubated for 1 h at 37°C at 5% CO_2. Following surface marker staining, the cells were fixed and permeabilized, treated with DNase for 1 h at 37°C to expose incorporated BrdU and stained with anti-BrdU antibody. The stained cells were then analyzed using flow cytometry. BrdU positivity indicated proliferation.

In vitro progenitor differentiation assay

Sorted progenitor cells were seeded into 6-well plates and cultured for 14 days in complete Methocult™SF H4436 media (Stem Cell Technologies) according to the manufacturer's protocol. The cytokine cocktail of this methylcellulose assay consists of rhSCF, rhGM-CSF, rhIL-3, rhIL-6, rhG-CSF, rhEPO and serum substitute. Following incubation, microscopic images with 50x and 100x magnification were taken of all established colonies. Colonies were counted and identified by side-by-side comparison with representative images from 3 independent plates provided in the manufacturer's technical manual.

Cytospin

Sorted bone marrow cells were cytocentrifuged onto microscope slides using a Cytospin 4 centrifuge (Thermo Shandon), stained with Hema 3 Manual Staining System (Fisher Scientific) and analyzed by light microscopy.

RNA isolation and RNA sequencing

Sorted bone marrow cells in TRIzol™LS Reagent were thawed on ice. RNA isolation was performed using the Direct-zol™RNA MiniPrep (Zymo Research) according to manufacturer's instructions. RNA concentration was measured on a Agilent 2200 TapeStation using the Agilent High Sensitivity RNA ScreenTape System (Agilent Technologies). Cells will be sorted directly in TriZOL LS and processed immediately into RNA using DirectZOL RNA Miniprep (Zymo Research). Library preparation and sequencing will be performed according to the Smart-Seq2 protocol in the La Jolla Institute for Immunology sequencing core.

RNA-Seq analysis

The RNA-Seq short reads were mapped to the human genome (hg38) using subread-align from the Subread R package (Liao, Smyth, and Shi 2013). Uniquely mapped reads were annotated with NCBI RefSeq annotation using featureCounts R package (v1.5.3) (Liao, Smyth, and Shi 2013, 2014) were used for downstream differentially analysis. Multidimensional scaling (MDS) analysis and differential expression testing was performed using linear model analysis (function voom from limma R package; v3.33.7) (Law et al., 2014) with scaling normalization factors estimated using edgeR package (Robinson, McCarthy, and Smyth 2010). Gene Ontology enrichment was done using Fisher exact test implemented in topGO Bioconductor package (Alexa, Rahnenführer, and Lengauer 2006). Multi-test correction using Benjamini and Hochberg method (Benjamini and Hochberg 1995) implemented in p.adjust function in R.

QUANTIFICATION AND STATISTICAL ANALYSIS

Data for all experiments were analyzed with Prism software (GraphPad). Paired and unpaired t- tests and one-way or two-way ANOVA analysis were used for comparison of experimental groups as indicated in figure legends. P values of less than 0.05 were considered significant. All data replicates can be found indicated by individual points in the figures or listed in the figure legends. Mean value is indicated as center on graphs with corresponding standard error, unless otherwise noted. More information about statistical tests and results for each experiment can be found in the figure legends.

The Journal of Immunology

Functional Analysis of Immune Signature Genes in Th1* Memory Cells Links ISOC1 and Pyrimidine Metabolism to IFN-γ and IL-17 Production

Yulia Kushnareva, Ian T. Mathews, Alexander Y. Andreyev, Gokmen Altay, Cecilia S. Lindestam Arlehamn, Vijayanand Pandurangan, Roland Nilsson, Mohit Jain, Alessandro Sette, Bjoern Peters and Sonia Sharma

J Immunol February 5, 2021, ji2000672; DOI: https://doi.org/10.4049/jimmunol.2000672

KEY POINTS

- Functional analysis of Th1* immune signature genes uncovers novel regulators of IFN-γ and IL-17.
- Loss of ISOC1 impairs $CD4^+$ T cell pyrimidine metabolism, bioenergetics, and cytokine production.

Abstract

$CCR6^+CXCR3^+CCR4^-CD4^+$ memory T cells, termed Th1*, are important for long-term immunity to *Mycobacterium tuberculosis* and the pathogenesis of autoimmune diseases. Th1* cells express a unique set of lineage-specific transcription factors characteristic of both Th1 and Th17 cells and display distinct gene expression profiles compared with other $CD4^+$ T cell subsets. To examine molecules and signaling pathways important for the effector function of Th1* cells, we performed loss-of-function screening of genes selectively enriched in the Th1* subset. The genetic screen yielded candidates whose depletion significantly impaired TCR-induced IFN-γ production. These included genes previously linked to IFN-γ or *M. tuberculosis* susceptibility and novel candidates, such as *ISOC1*, encoding a metabolic enzyme of unknown function in mammalian cells. ISOC1-depleted T cells, which produced less IFN-γ and IL-17, displayed defects in oxidative phosphorylation

and glycolysis and impairment of pyrimidine metabolic pathway. Supplementation with extracellular pyrimidines rescued both bioenergetics and IFN-γ production in ISOC1-deficient T cells, indicating that pyrimidine metabolism is a key driver of effector functions in $CD4^+$ T cells and Th1* cells. Results provide new insights into the immune-stimulatory function of ISOC1 as well as the particular metabolic requirements of human memory T cells, providing a novel resource for understanding long-term T cell–driven responses.

Footnotes

- This work was supported by National Cancer Institute Grants R01CA199376 (to S.S.) and F31CA236405 (to I.T.M.), National Institute of Dental and Craniofacial Research Grant U01DE028227 (to S.S.), an Infrastructure Operating Fund grant (to S.S.) under the Human Immune Profiling Center of the National Institute of Allergy and Infectious Diseases Grant U19AI118626 (to A.S.), National Institutes of Health Office of the Director Grant S10OD020025 (to M.J.), National Institute of Environmental Health Sciences Grant R01ES027595 to (M.J.), National Institute of General Medical Sciences T32GM007752 (to I.T.M.), and a SPARK award from the La Jolla Institute for Immunology Board of Directors (to I.T.M.).

- The online version of this article contains supplemental material.

Received June 8, 2020.

Accepted January 4, 2021.

Shuyu Yao et al., J Immunol, 2014

Distinct, Specific IL-17- and IL-22-Producing CD4+ T Cell Subsets Contribute to the Human Anti-Mycobacterial Immune Response
Thomas J. Scriba et al., J Immunol, 2008

The central memory CD4+ T cell population generated during Leishmania major infection requires IL-12 to produce IFN-gamma.
Nazzy Pakpour et al., J Immunol, 2008

Systemic administration of IL-23 induces potent antitumor immunity primarily mediated through Th1-type response in association with the endogenously expressed IL-12.
Teruo Kaiga et al., J Immunol, 2007

A key role for Itk in both IFN gamma and IL-4 production by NKT cells.
Byron B Au-Yeung et al., J Immunol, 2007

innate TH1-like effector function
Takeshi Kawabe et al., Sci Immunol, 2017

Regulatory T cells inhibit acute IFN-γ synthesis without blocking T-helper cell type 1 (Th1) differentiation via a compartmentalized requirement for IL-10.
Dorothy K Sojka et al., Proc Natl Acad Sci U S A, 2011

Agricultural lime disturbs natural strontium isotope variations: Implications for provenance and migration studies
Erik Thomsen et al., Sci Adv, 2019

CYTOKINES AS REGULATORS OF CD4+ T CELL RESPONSES
World Scientific

Previous Next

Back to top

In this issue

The Journal of Immunology
Vol. 206, Issue 11
1 Jun 2021

Top Reads

PubMed Google Scholar

▸ Cited By...

▸ More in this TOC Section

▸ Similar Articles

ImmunoCasts

Twitter

Print ISSN 0022-1767 Online ISSN 1550-6606

IMMUNOLOGY

Cellular sensing of extracellular purine nucleosides triggers an innate IFN-β response

Rekha Dhanwani[1]*, Mariko Takahashi[1]*, Ian T. Mathews[1,2,3]*, Camille Lenzi[1], Artem Romanov[1], Jeramie D. Watrous[2,3], Bartijn Pieters[1], Catherine C. Hedrick[1], Chris A. Benedict[1], Joel Linden[1,3], Roland Nilsson[4,5], Mohit Jain[2,3], Sonia Sharma[1†]

Mechanisms linking immune sensing of DNA danger signals in the extracellular environment to innate pathways in the cytosol are poorly understood. Here, we identify a previously unidentified immune-metabolic axis by which cells respond to purine nucleosides and trigger a type I interferon-β (IFN-β) response. We find that depletion of ADA2, an ectoenzyme that catabolizes extracellular dAdo to dIno, or supplementation of dAdo or dIno stimulates IFN-β. Under conditions of reduced ADA2 enzyme activity, dAdo is transported into cells and undergoes catabolysis by the cytosolic isoenzyme ADA1, driving intracellular accumulation of dIno. dIno is a functional immunometabolite that interferes with the cellular methionine cycle by inhibiting SAM synthetase activity. Inhibition of SAM-dependent transmethylation drives epigenomic hypomethylation and overexpression of immune-stimulatory endogenous retroviral elements that engage cytosolic dsRNA sensors and induce IFN-β. We uncovered a previously unknown cellular signaling pathway that responds to extracellular DNA–derived metabolites, coupling nucleoside catabolism by adenosine deaminases to cellular IFN-β production.

INTRODUCTION

Innate immunity is a universal cellular response to pathogenic threats. Upon sensing infection, damage, or genotoxic stress, susceptible cells activate innate immune signaling cascades such as the interferon-β (IFN-β) axis. IFN-β is a pleiotropic cytokine that signals via the type I IFN receptor (IFNAR) to exert autocrine and paracrine effects on cellular growth, apoptosis, and immune cell activation, ultimately playing an essential role in propagation and resolution of the inflammatory response (*1*). IFN-β is produced during the course of infection, neoplastic transformation, cancer immunotherapy, and autoimmune disease due to cellular recognition of atypical or mislocalized nucleic acids, which involves both intrinsic and extrinsic sensing mechanisms (*2*, *3*). Cell-intrinsic RNA or DNA ligands derive from internalized viruses and bacteria or from host-cellular sources such as damaged mitochondria, stalled replication forks, or endogenous retroviral elements encoded within the genome. Inside the cell, these aberrant nucleic acids are recognized by cytoplasmic sensors such as retinoic acid–inducible gene-I (RIG-I), melanoma differentiation–associated protein 5 (MDA5), or cyclic guanosine monophosphate–adenosine monophosphate synthetase (cGAS). Engagement of innate cytosolic sensors in turn activates key signaling adaptor proteins including mitochondrial antiviral signaling (MAVS) or stimulator of IFN genes (STING), which drive IFN-β production.

How extracellular nucleic acids engage cytosolic sensors is incompletely understood. Extracellularly, Mammalian DNA is a damage-associated molecular pattern (DAMP) with immune-stimulatory activity that largely depends recognition by internal sensors (*4*). Specialized phagocytic cells, including monocytes, macrophages, and dendritic cells, may engulf extracellular DNA fragments to access the cytosolic machinery. However, it is unclear how nonphagocytic or stromal cells respond to extracellular DNA danger signals or whether extracellular DNA–derived metabolites generated by the activity of endonucleases and other nucleic acid modifying enzymes can act in an extrinsic or paracrine manner to modulate the innate immune response.

Here, we describe a previously unidentified cellular response to extracellular deoxy-adenosine (dAdo), which stimulates IFN-β and an IFN-β–driven innate immune response in a paracrine manner. We show that loss of extracellular adenosine deaminase 2 (ADA2) enzyme activity in stromal endothelial cells drives increased cellular uptake of extracellular dAdo into stromal endothelial cells. Upon transport into the cell, dAdo is catabolized by the cytosolic isoenzyme ADA1 to yield dIno, which specifically accumulates inside the ADA2-depleted cells. dIno is a functional immunometabolite that perturbs the cellular metabolic reactions suppressing expression of endogenous retroviral elements, immune-stimulatory molecules that directly engage innate cytosolic double stranded RNA (dsRNA) sensors. Our results suggest that metabolism of extracellular DNA or nucleoside danger signals, which are released due to ischemia, infection, and tumor growth, is a key regulatory checkpoint during physiological inflammation. These regulatory mechanisms are likely abolished in human metabolic syndromes such as ADA2 deficiency or purine nucleoside phosphorylase (PNP) deficiency diseases, which manifest as profound immunological dysregulation characterized by systemic inflammation and inappropriate induction of IFN-β and other innate cytokines.

RESULTS

Loss of ADA2 triggers spontaneous IFNβ production

Gene mutations in molecules of the innate cytosolic sensing machinery drive systemic inflammatory diseases in humans due to excessive IFN-β production (*5*). We reasoned that new mechanistic insights into innate immune sensing and molecular triggers of inflammation could be gained by examining functionally uncharacterized disease-associated genes. Search of the Online Mendelian Inheritance in

[1]La Jolla Institute for Immunology, La Jolla, CA 92037, USA. [2]Department of Medicine, University of California, San Diego, La Jolla, CA 92093, USA. [3]Department of Pharmacology, University of California, San Diego, La Jolla, CA 92093, USA. [4]Cardiovascular Medicine Unit, Department of Medicine, Karolinska Institutet, Karolinska University Hospital, SE-17176 Stockholm, Sweden. [5]Center for Molecular Medicine, Karolinska Institutet, Karolinska University Hospital, SE-17176 Stockholm, Sweden.
*These authors contributed equally to this work.
†Corresponding author. Email: soniasharma@lji.org

Man database identified 38 syndromes of systemic inflammation each resulting from a single, sequence-verified gene mutation (table S1). To examine whether any of the candidate disease genes regulate IFN-β pathway activation, in situ immunofluorescence was used to assess inducible nuclear translocation of the transcription factor IFN regulatory factor 3 (IRF3) (*6*), which directly transactivates the IFN-β promoter locus. Upon transfection with immune-stimulatory poly (dA:dT) DNA, IRF3 translocation and IFN-β mRNA levels were quantified in primary human umbilical vein endothelial cells (HUVECs) and other endothelial lineage cells (fig. S1, A and B). Forward genetic screening in HUVEC confirmed STING and three prime repair exonuclease I (TREX1), an inhibitor of DNA sensing through STING, as positive and negative regulators of IRF3/IFN-β, respectively. Results identify ADA2 as a new negative regulator of IRF3 activation (Fig. 1A and table S1).

To ascertain whether loss of ADA2 expression drives spontaneous IRF3 activation in the absence of exogenous stimulation, phosphorylation of IRF3 and autophosphorylation of the IRF3 kinase Tank-binding kinase 1 (TBK1) were examined in ADA2 knockdown cells. Western blotting with phospho-specific antibodies demonstrated spontaneous phosphorylation of IRF3 and TBK1 in small interfering RNA targeting ADA2 (siADA2)–treated or small interfering RNA targeting TREX1 (siTREX1)–treated cells compared to a control, nontargeting small interfering RNA (siControl) (Fig. 1B). *IFN*-β mRNA levels were also spontaneously elevated in ADA2-depleted cells (Fig. 1C) and were further enhanced upon infection with human cytomegalovirus (hCMV), a β-herpesvirus that induces a robust cellular IFN-β response (*6*). Induction of an IFN-β–stimulated gene signature in unstimulated siADA2-treated cells was observed using whole-genome transcriptome analysis (Fig. 1D). Differentially expressed transcripts in this dataset included genes encoding proinflammatory cytokines, chemokines, and innate immune response proteins (Fig. 1E). These findings are concordant with previous observations that IFN-associated gene expression signatures are detectable in peripheral blood cells from ADA2-deficient patients (*7*). Application of a neutralizing antibody for IFN-β confirmed autocrine/paracrine IFN-β–driven signaling in ADA2-depleted cells (Fig. 1F). Together, results establish ADA2 as an inhibitor of spontaneous IFN-β production.

Human endothelial cells express functional ADA2

We next sought to determine the mechanism by which ADA2 regulates innate immunity. Human ADA1 and ADA2 are isoenzymes that catalyze the irreversible deamination of adenosine (Ado) and deoxyadenosine (dAdo) to inosine (Ino) and deoxyinosine (dIno), respectively (*8*, *9*). *ADA2* mRNA transcripts are broadly expressed in human immune cells (fig. S2A) (*10*) and were detected in the monocytic cell line U937 and multiple primary human endothelial lineage cells by quantitative reverse transcription polymerase chain reaction (qRT-PCR) (Fig. 2A) and RNA sequencing (RNA-seq) in HUVEC (fig. S2B). A prior report failed to detect intracellular ADA2 protein or enzyme activity in HUVEC (*11*). However, ADA2 is an ectoenzyme having a classical signal peptide (*12*), and extracellular ADA2 protein was detectable in cultured supernatants of U937 and human endothelial cells by enzyme-linked immunosorbent assay (ELISA) (Fig. 2B) and Western blotting (Fig. 2C). In HUVEC, siADA2 reduced *ADA2* mRNA levels by ~75%, with no cross-reactivity against *ADA1* (fig. S2C). Similarly, siADA1 reduced *ADA1* mRNA levels by ~90%, with no cross-reactivity toward *ADA2* (fig. S2D). Analysis of deaminase activity, measured through de novo conversion of ^{15}N-labeled dAdo to dIno by mass spectrometry (MS), demonstrated that >90% of intracellular ADA activity measured in HUVEC whole-cell lysate was specifically reduced by ADA1 depletion and unaffected by ADA2 depletion (Fig. 2D). In contrast, extracellular ADA activity measured in HUVEC supernatant was reduced >75% by specific depletion of ADA2, with ADA1 accounting for a minority of extracellular activity (Fig. 2D), recapitulating results obtained in human plasma (*13*). To confirm ADA2-specific activity in HUVEC supernatants, dAdo to dIno conversion was measured in the presence of ADA2 siRNA and 50 μM erythro-9-(2-hydroxy-3-nonyl) adenine (EHNA) to selectively inhibit ADA1 activity (fig. S2E). Together, experiments demonstrate expression of functional ADA2 by HUVEC.

ADA2 and dAdo are paracrine modulators of IFNβ

To examine whether extracellular ADA2 activity is involved in IFN-β regulation, enzyme add-back experiments were performed in ADA2-depleted cells using recombinant ADA (rADA) proteins. Addition of rADA2 inhibited spontaneous induction of *IFN*-β, *CCL5*, and *CXCL10* mRNA in siADA2-treated HUVEC (Fig. 2E). Notably, rescue effects accompanied reconstitution of extracellular ADA activity to wild-type levels and were inhibited by pentostatin, a general inhibitor of ADA activity. Prior reports suggest that extracellular ADA2 may function independently of its nucleoside deaminase activity by exerting growth factor–like effects on cells (*8*, *12*). To confirm involvement of ADA activity in immune activation, rADA1 was added to ADA2-depleted cells, resulting in rescue of cytokine expression and extracellular ADA activity (Fig. 2E). Given the essential role of extracellular purine nucleoside deaminase activity in suppressing expression of innate cytokines, U937 cells were supplemented with exogenous nucleosides, and innate immune responses were examined. Supplementation of dAdo specifically induced IFN-β (Fig. 2F), suggesting that both ADA2 and dAdo act in a paracrine manner to modulate cellular innate responses. Accordingly, coculture of siADA2-treated HUVEC with untreated U937 cells provoked spontaneous IFN-β production in bystander U937 cells (Fig. 2G), confirming paracrine or non–cell autonomous effects of ADA2 in HUVEC.

Loss of ADA2 drives intracellular dAdo catabolism and accumulation of dIno

Extracellular Ado and dAdo are bioactive molecules that signal to immune cells during pathological inflammation (*14*, *15*). Ado and dAdo are released into the microenvironment in an autocrine/paracrine fashion or produced upon degradation of extracellular nucleic acids. Ado, but not dAdo, primarily signals via heterotrimeric GTP-binding protein (G protein)–coupled purinergic adenosine receptors (AdoR) (Fig. 3A), which elicit pleiotropic immune-modulatory effects. Alternatively, dAdo may be transported into cells via the action of nucleoside transporters. Vascular endothelial cells express relatively high levels of equilibrative nucleoside transporters (ENT) (Fig. 3B) that enable rapid and efficient clearance of dAdo from the blood (*16*) and may render these cells particularly sensitive to the effects of ADA2 deficiency and dysregulation of extracellular dAdo metabolism. Accordingly, incubation of HUVEC with ^{15}N-labeled dAdo resulted in rapid cellular uptake and equilibration of intracellular and extracellular dAdo pools within 5 min (Fig. 3C). Pretreatment with dipyridamole (DPM), a pharmacological inhibitor of ENT-mediated dAdo transport, reduced expression of the IFN-β and the IFN-β–driven gene signature in ADA2-depleted cells (Fig. 3D). Results indicate

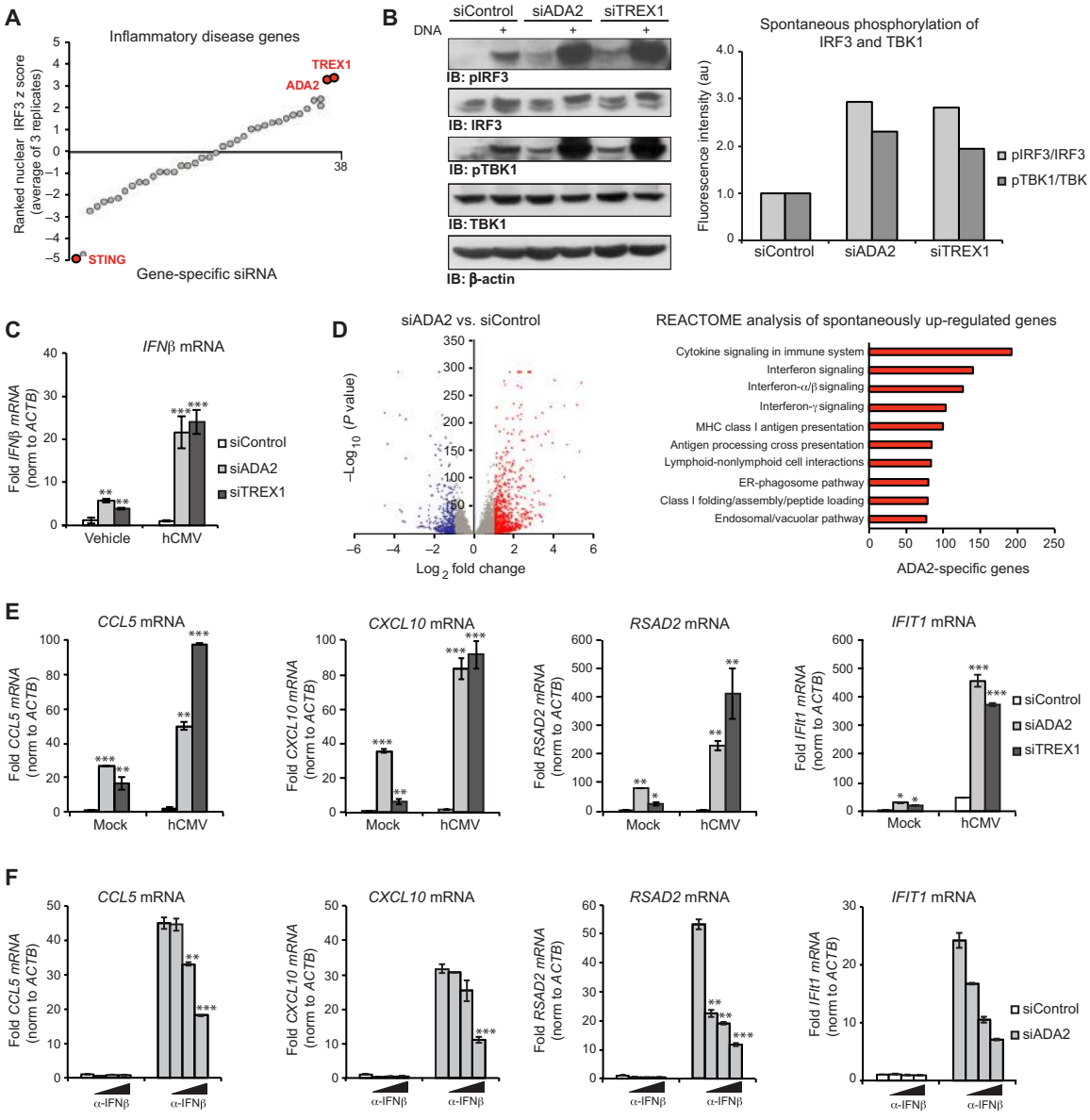


Fig. 1. Loss of ADA2 triggers spontaneous IFNβ production. (**A**) siRNA screen of 38 human disease genes in HUVEC (table S1), quantifying IRF3 nuclear translocation upon poly (dA:dT) DNA treatment (200 ng/ml for 3 hours). Averaged ranked *z* score for each gene-specific siRNA oligonucleotide pool is represented. STING, TREX1, and ADA2 scores are indicated in red ($n = 3$ technical replicates for $n = 3$ biological replicates, >200 single cells per technical replicate). (**B**) Western blot analysis of siControl, siADA2-, or siTREX-transfected HUVEC lysates (35 μg per lane). ImageJ quantification of the intensity ratio between phospho-IRF3/pan IRF3 and phospho-TBK1/TBK in unstimulated cells is shown in bar graphs. IB, immunoblot.; au, arbitrary units. (**C**) *IFN*-β mRNA levels, measured by qRT-PCR, in siControl, siADA2-, or siTREX1-transfected HUVEC upon treatment with vehicle or infection with hCMV [multiplicity of infection (MOI) = 1 for 3 hours] ($n = 3$ technical replicates). (**D**) Differential gene expression between siControl and siADA2-transfected HUVEC measured by polyadenylated poly (A^+)–enriched RNA-seq ($n = 3$ biological replicates) and REACTOME pathway analysis of the ADA2-specific genes. Red and blue dots denote genes significantly up- or down-regulated >2-fold. (**E**) Expression levels of IRF3-driven or IFN-β–driven genes, measured by qRT-PCR, in siControl, siADA2-, or siTREX1-transfected HUVEC upon mock or hCMV infection (MOI = 1 for 3 hours). (**F**) Expression levels of IRF3-driven or IFN-β–driven genes, measured by qRT-PCR, in siControl or siADA2-transfected HUVEC treated with IFN-β–neutralizing antibody (10, 20, and 40 U/ml) ($n = 3$ technical replicates). All results were replicated in three independent experiments. Values are presented as means ± SD. $^*P < 0.05$, $^{**}P < 0.01$, and $^{***}P < 0.001$. MHC, major histocompatibility complex; ER, endoplasmic reticulum; IB, immunoblot.

that cellular nucleoside transport of dAdo is essential for IFN-β production driven by reduced ADA2 enzyme activity.

Once taken up by cells, dAdo is metabolized through intracellular purine salvage and/or degradation pathways (Fig. 4A) (*15*), which may yield downstream effector molecules that influence innate signaling. To assess the potential intracellular fates of consumed extracellular dAdo, global MS-based metabolomics was performed (*17*). Comprehensive analysis of metabolites revealed intracellular accumulation of dIno and its downstream catabolite hypoxanthine in ADA2-depleted cells relative to control cells (Fig. 4, B and C), while

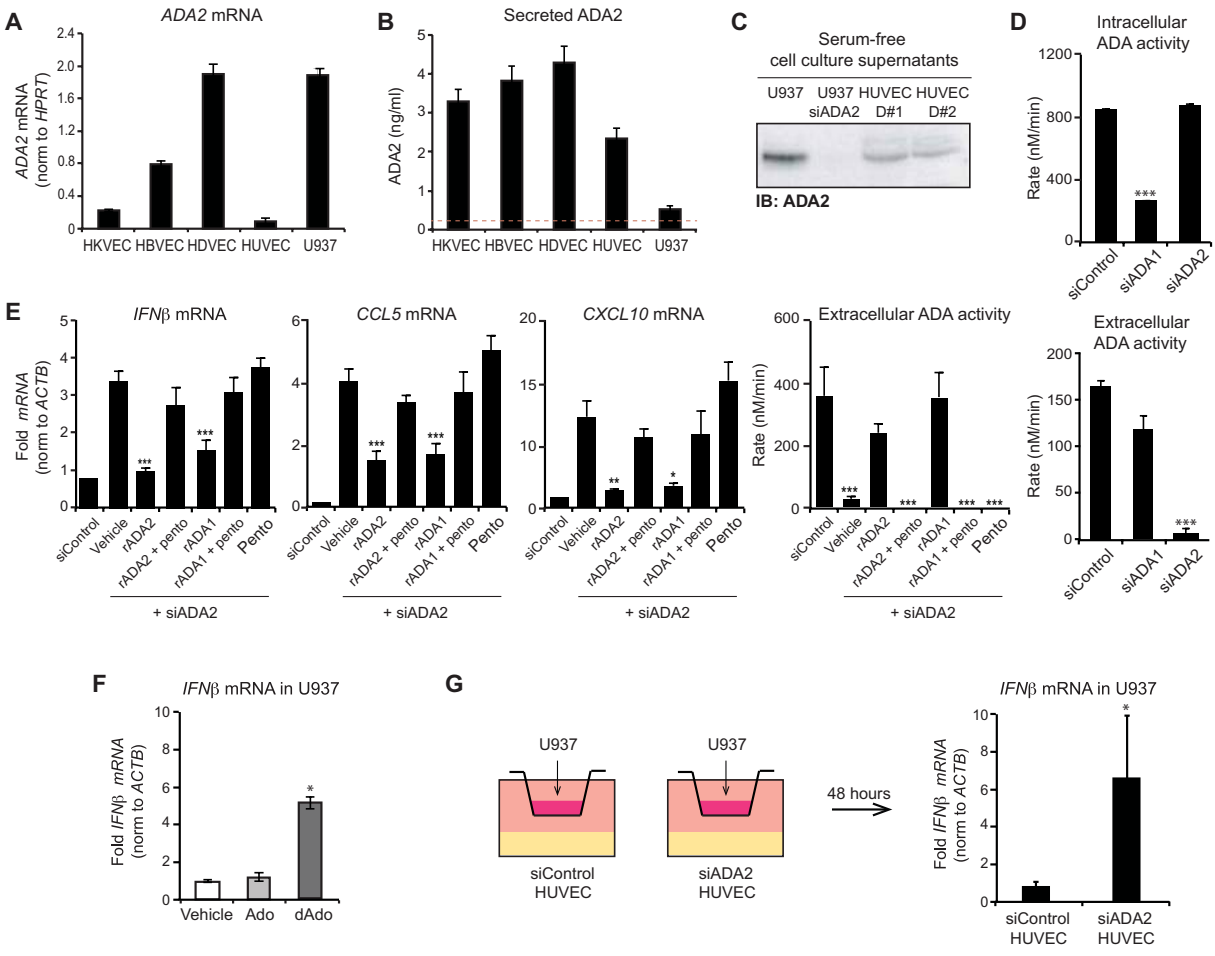


Fig. 2. Extracellular ADA2 and dAdo are paracrine modulators of IFNβ. (**A**) *ADA2* mRNA levels, measured by qRT-PCR, in primary endothelial cells and U937 monocytic cells ($n = 3$ technical replicates). (**B**) Secreted ADA2 protein, measured by ELISA, in primary endothelial cells and U937 monocytic cells ($n = 3$ technical replicates). (**C**) Secreted ADA2 protein, assessed by Western blotting of serum-free cell-conditioned supernatants, from HUVEC and U937 monocytic cells. (**D**) ADA activity, measured by de novo conversion of 1 mM isotopically labeled dAdo to dIno in whole-cell lysates (intracellular) or cell-conditioned supernatants (extracellular) from siControl, siADA1-, or siADA2-transfected HUVEC. Values are normalized to cell number and protein concentration ($n = 5$ technical replicates). (**E**) Expression levels of IRF3-driven or IFN-β–driven genes, measured by qRT-PCR, and extracellular ADA activity, measured by de novo conversion of 1 mM isotopically labeled dAdo to dIno in cell-conditioned supernatants, from siControl or siADA2-transfected HUVEC supplemented with rADA1 or rADA2 and pretreated with vehicle or pentostatin (10 μM for 30 min). Activity values are normalized to cell number and protein concentration ($n = 5$ technical replicates). (**F**) *IFN*-β mRNA, measured by qRT-PCR, in U937 monocytic cells supplemented with adenosine (Ado) or deoxyadenosine (dAdo) (100 μM for 48 hours) ($n = 3$ technical replicates). (**G**) *IFN*-β mRNA levels, measured by qRT-PCR, in U937 cocultured in Transwell inserts with siControl or siADA2-treated HUVEC for 48 hours ($n = 3$ technical replicates). All results were replicated in three independent experiments. Values are presented as means ± SD. $^{*}P \leq 0.05$, $^{**}P \leq 0.01$, and $^{***}P \leq 0.001$. HDVEC, dermal microvascular endothelial cells; HBVEC, brain microvascular endothelial cells; HKVEC, kidney microvascular endothelial cells.

other purine metabolites including dAdo and Ado were not changed (fig. S3A). These results show that loss of ADA2 activity dysregulates cellular purine catabolism and imply a compensatory mechanism by which extracellular dAdo is catabolized inside the cell. Consistent with this, addition of extracellular ^{15}N-labeled dAdo drove rapid intracellular accumulation of labeled dIno product in ADA2-depleted cells (Fig. 4C). Intracellularly, dIno is produced from consumed dAdo by ADA1 (Fig. 4A). Depletion of ADA1 in ADA2 knockdown cells suppressed both IFN-β expression and dIno production, demonstrating a functional role for ADA1 in producing intracellular dIno in the setting of ADA2 deficiency (Fig. 4D and fig. S3, B and C). In contrast, increasing intracellular dIno pools by silencing its catabolic enzyme PNP or silencing the cellular enzyme deoxycytidine kinase (dCK), an alternative enzyme for intracellular metabolism of dAdo (Fig. 4A), increased IFN-β levels (Fig. 4D and fig. S3C). Consistent with these observations, supplementation of exogenous dIno, but not Ino or hypoxanthine, was sufficient to trigger *IFN*-β mRNA production (Fig. 4E and fig. S3D), supporting a novel functional role for dIno in the cellular innate immune response. Notably, the immune-stimulatory effect of exogenous dIno was sensitive to DPM (fig. S3D) and was enhanced by pretreatment with 9-deazaguanine (Fig. 4E), which inhibits intracellular and extracellular catabolism of dIno by PNP. Data show that dIno accumulates intracellularly in ADA2-depleted cells due to the activity of ADA1 and functions as an immune-metabolite that is both necessary and sufficient for IFN-β induction.

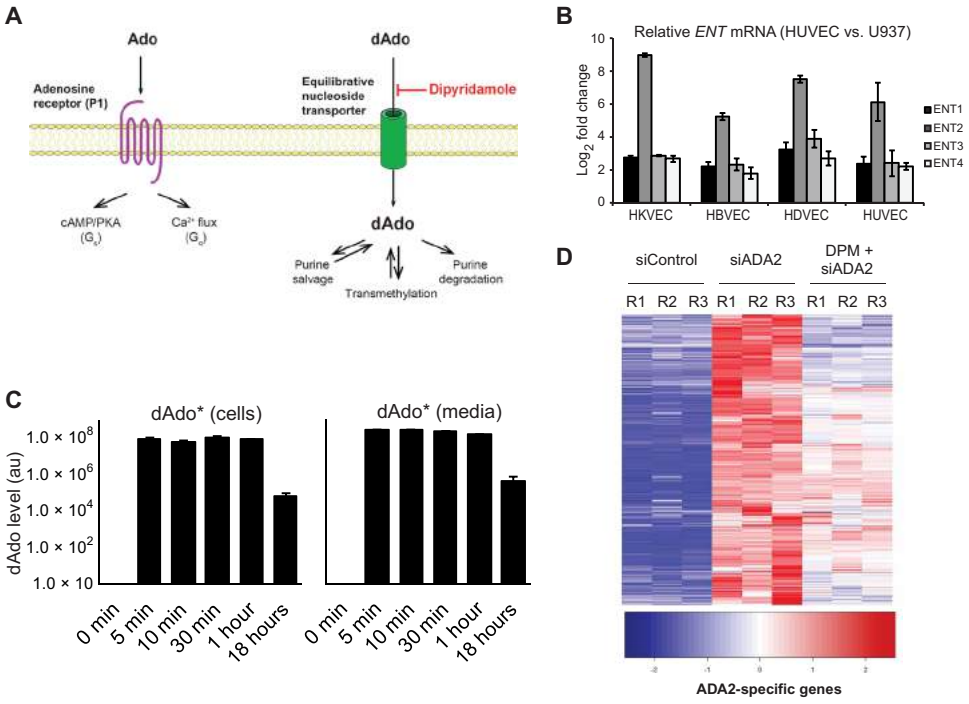


Fig. 3. Cellular nucleoside transport is required for induction of IFNβ upon depletion of ADA2. (**A**) Schematic representation of extracellular Ado receptor signaling and dAdo uptake/intracellular metabolism. cAMP, cyclic adenosine 3′,5′-monophosphate; PKA, cAMP-dependent protein kinase. (**B**) Equilibrative nucleoside transporter (*ENT*) mRNA levels, measured by qRT-PCR, in HUVEC and U937 ($n = 3$ technical replicates). (**C**) Extracellular and intracellular levels of isotope-labeled dAdo, measured by LC-MS/MS ($n = 5$ biological replicates). (**D**) Differential gene expression in siControl, siADA2-, or siADA2/DPM-treated (40 μM for 48 hours) HUVEC, measured by polyA$^+$ RNA-seq ($n = 3$ biological replicates). All results were replicated in three independent experiments. Values are presented as means ± SD. *$P \leq 0.05$.

dIno inhibits the cellular methionine cycle via SAM synthetase

We next sought to determine the mechanism by which dIno acts as an immunometabolite and functional regulator of innate immune signaling. Intracellularly, nucleosides and their analogs or metabolites interface with multiple cellular processes including transmethylation (*18*). Loss of ADA2 was associated with steady-state reduction in relative levels of L-cystathionine, a downstream metabolite in the methionine cycle, in siADA2-treated cells compared to controls (fig. S4A). To further examine methionine metabolism, time-course stable isotope tracing was performed with U-^{13}C methionine (*19*, *20*). Intracellular methionine metabolite pools were fully labeled within minutes of exposure to extracellular label, while *S*-adenosyl-methionine (SAM) and its downstream metabolite *S*-adenosylhomocysteine (SAH) acquired labeling over time (fig. S4B). The methionine M+4 mass isotopomer was consistently measured at ~2% (fig. S4C), suggesting a truncated cycle with little remethylation of homocysteine into methionine. Upon ADA2 depletion, SAM and SAH pools exhibited slower labeling over time relative to controls, and model-based metabolic flux analysis (*19*) revealed impaired cycle flux through the SAM synthetase [methionine adenosyltransferase (MAT)] and methyl-transferase (MT) reactions (Fig. 4F). In vitro enzyme activity assays demonstrated that dIno and, to a lesser extent, dAdo, but not the nucleoside cytidine, directly inhibited MAT activity (Fig. 4G). Accordingly, pharmacological inhibition of MAT using cycloleucine induced IFN-β (Fig. 4H), phenocopying the effects of ADA2 deficiency or dIno accumulation and suggesting a role for MAT inhibition in the immune-stimulatory effects of dIno. The MAT product SAM serves as a methyl donor for multiple enzymatic reactions, including DNA MT (DNMT)–dependent methylation of genomic cytosine bases (Fig. 4F). Accordingly, inhibition of DNMT using 5-azacytidine induced IFN-β (Fig. 4H). To examine the status of genomic transmethylation in ADA2-depleted cells, labeled methionine was used to monitor incorporation of methyl groups onto cytosine residues in genomic DNA. Control cells steadily accumulated labeled methyl-cytosine, reaching ~20 to 25% of genomic methyl-cytosine over 72 hours. In contrast, cells depleted of ADA2 or DNMT (Fig. 4I) and dIno-supplemented cells (fig. S4D) all demonstrated reduced de novo DNA methylation as early as 24 hours after the addition of label. These data demonstrate that dIno directly inhibits the activity of MAT and the methionine/SAM cycle, with subsequent reduction of genomic DNA transmethylation upon loss of ADA2 or supplementation with dIno.

Induction of methylation-sensitive endogenous retrovirus elements triggers IFNβ

Endogenous retrovirus elements (*ERV*) encode a diverse family of germline immune-stimulatory dsRNA-like molecules whose dynamic expression is particularly sensitive to genomic DNA hypomethylation. Overexpression of *ERV* in cells is directly linked to IFN-β production due to engagement of innate immune dsRNA sensing pathways (*21*, *22*). Loss-of-function studies confirmed that the cytosolic dsRNA sensors RIG-I and MDA5 and the dsRNA signaling adaptor MAVS drive spontaneous IFN-β production upon depletion of ADA2

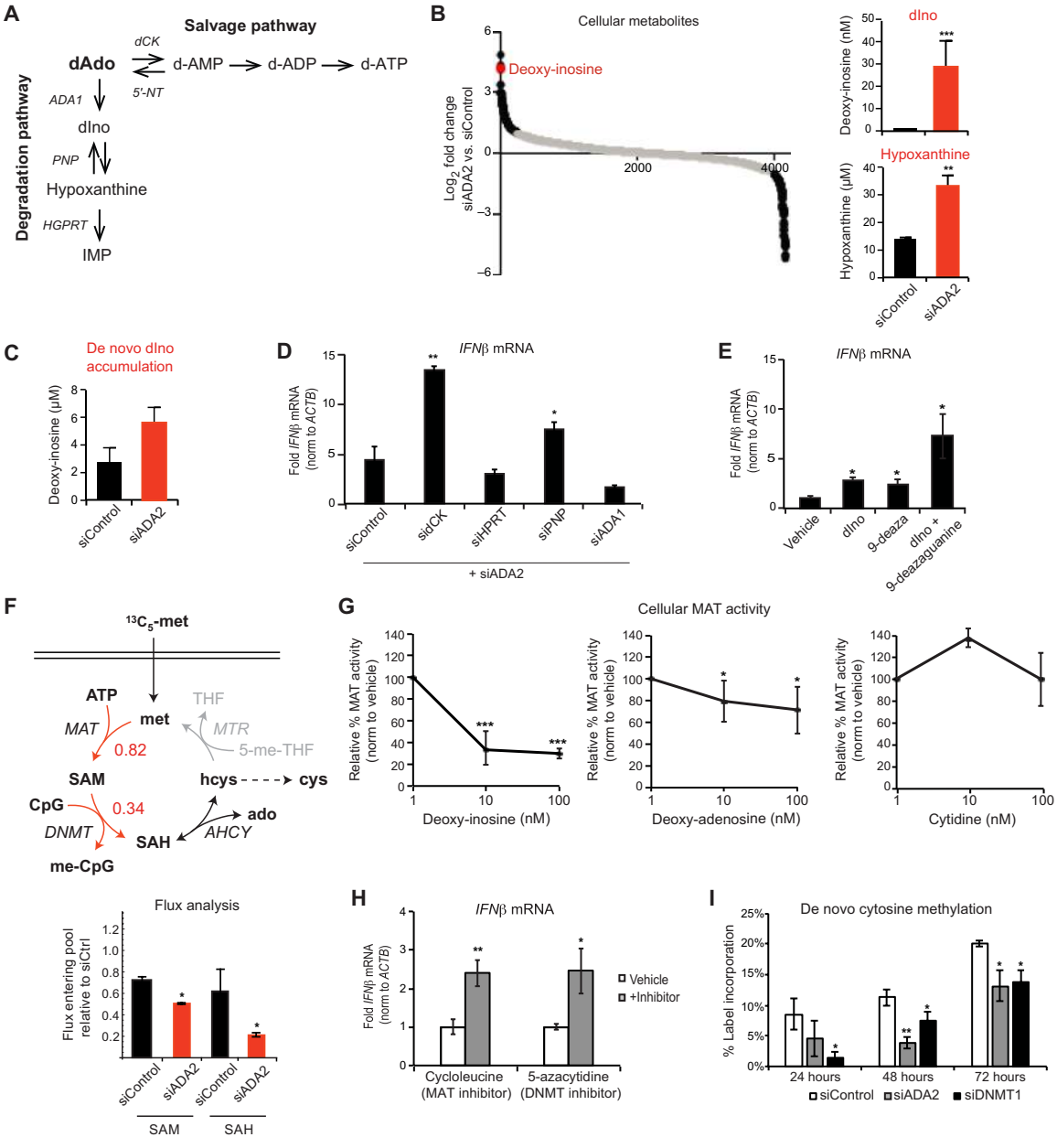


Fig. 4. Loss of ADA2 drives intracellular dAdo catabolism and accumulation of dIno. (**A**) Intracellular metabolic pathways of purine degradation [ADA1 (protein); PNP (protein); and HGPRT (protein), hypoxanthine phosphoribosyltransferase] and purine salvage [ADK (protein), adenosine kinase; dCK (protein), deoxycytidine kinase; and *5′-NT*, 5′-nucleotidase]. (**B**) Relative levels and concentrations of small-molecule polar metabolites measured by LC-MS/MS quantification of siControl or siADA2-transfected HUVEC ($n = 5$ biological replicates). (**C**) De novo accumulation of ^{15}N-dAdo–labeled dIno measured by LC-MS/MS quantification of siControl or siADA2-treated HUVEC ($n = 5$ biological replicates). (**D**) *IFN*-β mRNA levels, measured by qRT-PCR, in HUVEC transfected with siRNAs targeting ADA2 and ADA1, PNP, HGPRT (protein), or dCK. (**D**) *IFN*-β mRNA levels, measured by qRT-PCR, in HUVEC supplemented with dIno (500 μM for 24 hours) or (**E**) pretreated with the PNP inhibitor 9-deazaguanine (100 μM for 30 min) before dIno supplementation (500 μM for 24 hours) ($n = 3$ technical replicates). (**F**) Flux analysis of the methionine/SAM cycle, measured by LC-MS/MS quantification of U-^{13}C methionine–labeled methionine, SAM, and SAH, in siControl or siADA2-treated HUVEC ($n = 5$ biological replicates). (**G**) Methionine adenosyltransferase (MAT) enzyme activity, measured by LC-MS/MS quantification of SAM, in HUVEC lysates supplemented with nucleosides ($n = 3$ technical replicates). (**H**) *IFN*-β mRNA levels, measured by qRT-PCR, in HUVEC pretreated with cycloleucine (MAT inhibitor, 20 mM for 30 min) or 5-azacytidine (DNMT inhibitor, 500 nM for 30 min) ($n = 3$ technical replicates). (**I**) De novo DNA methylation, measured by LC-MS/MS quantification of U-^{13}C-methionine labeled 2′-deoxy-5-methylcytidine in siControl, siADA2-, or siDNMT1-treated HUVEC ($n = 3$ biological replicates). Values are represented as the percentage of total label incorporation. All results were replicated in three independent experiments. Values are presented as means ± SD. $^*P \le 0.05$, $^{**}P \le 0.01$, and $^{***}P \le 0.001$. ADP, adenosine diphosphate; IMP, inosine monophosphate; THF, tetrahydrofolate; MTR, 5-methyltetrahydrofolate-homocysteine methyltransferase or methionine synthase.

(Fig. 5A and fig. S5A). Analysis of transcriptomic data obtained in ADA2-depleted cells to specifically examine expression of long terminal repeat (LTR)–containing genomic sequences, which comprise a large portion *ERV* elements, revealed overexpression of >30 LTR-containing *ERV*-like molecules (Fig. 5B). Furthermore, *ERVFRD-1* and *ERVK28* genes were highly induced after supplementation with dIno or depletion of ADA2 (Fig. 5, C and D). Augmented expression of *ERV* in dIno-treated cells (24 hours; Fig. 5C) or ADA2-depleted cells (24 to 48 hours; Fig. 5D) was detectable prior to induction of *IFN*-β mRNA (72 hours; fig. S5B) and occurred independently of autocrine/paracrine IFN-β–driven signaling (Fig. 5E and fig. S5F). To confirm the immune-stimulatory effects of *ERV*, transcript variants of *ERVK28* were overexpressed in HUVEC, resulting in spontaneous induction of *IFN*-β, *CCL5*, and *CXCL10* (Fig. 5F). Results demonstrate that overexpression of immune-stimulatory *ERV* in ADA2-depleted cells drives the IFN-β response.

DISCUSSION

Here, we report a new cellular strategy for triggering IFN-β production by extracellular dAdo (Fig. 6). Loss of the ectoenzyme ADA2 or direct supplementation of exogenous dAdo drives cellular transport and catabolism by the cytosolic enzyme ADA1, driving de novo dIno production and dIno accumulation inside ADA2-depleted cells. dIno is a functional immune-metabolite that directly inhibits the cellular methionine cycle at the level of MAT, thereby suppressing cellular transmethylation reactions. Genomic hypomethylation provokes overexpression of immune-stimulatory *ERV* transcripts, including *ERVK28* and *ERVFRD-1*, which trigger IRF3 activation and IFN-β. Thus, purine nucleoside metabolism by ADA2 and other enzymes involved in extracellular DNA metabolism is likely an important determinant of IFN-β and IFN-β–driven gene expression in inflammatory microenvironments, where inappropriate release and accumulation of cell-free DNA and its nucleoside

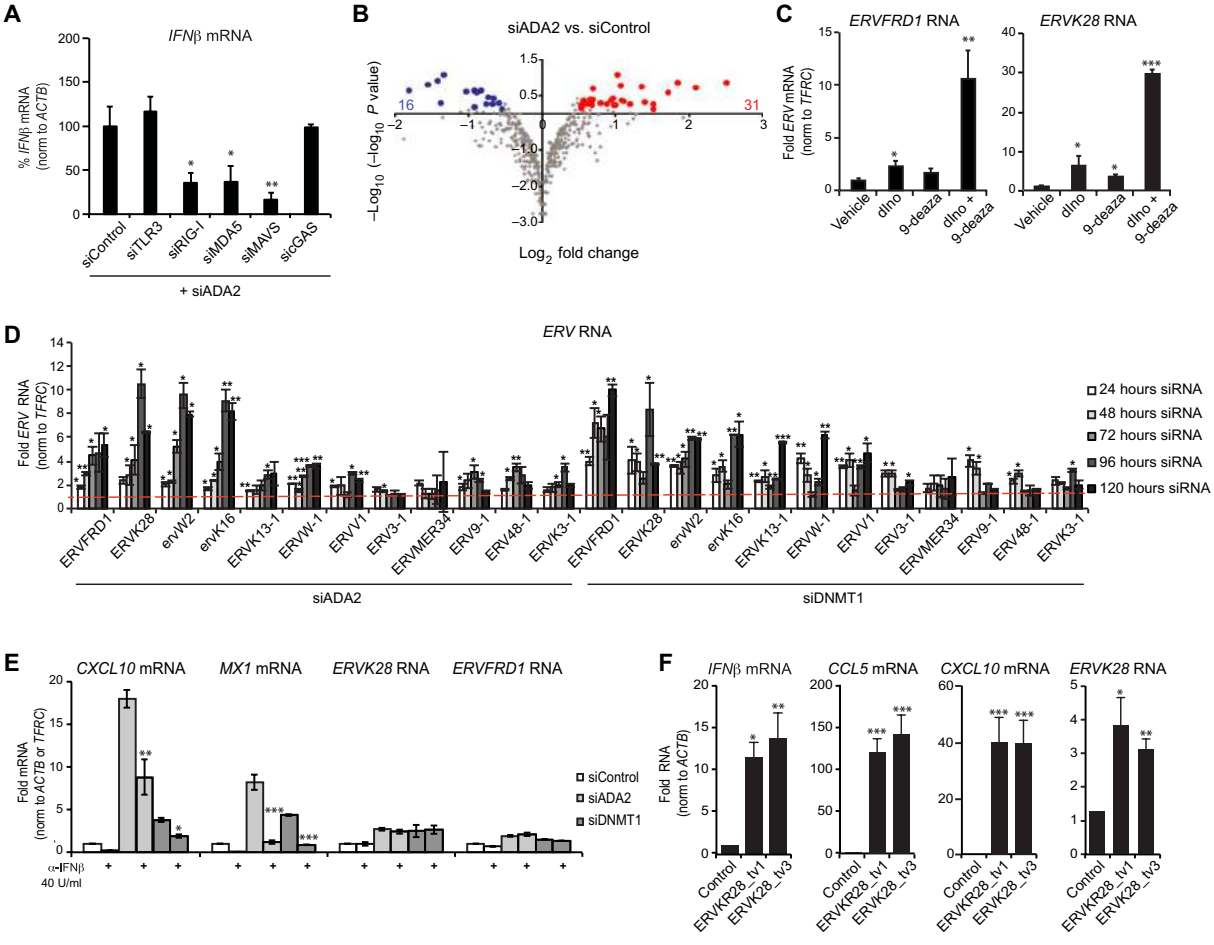


Fig. 5. Induction of methylation-sensitive ERV triggers cytosolic dsRNA signaling and IFNβ production. (**A**) *IFN*-β mRNA levels, measured by qRT-PCR, in HUVEC transfected with siRNAs targeting ADA2 and innate sensing molecules ($n = 3$ technical replicates). (**B**) Differential expression of LTR-containing *ERV* genes between siControl and siADA2-transfected HUVEC, measured by polydenylated poly(A^+)–enriched RNA-seq ($n = 3$ biological replicates). Red and blue dots denote *ERV* significantly up- or down-regulated ≥1.5-fold. (**C**) mRNA levels of *ERVFRD1* and *ERVK28*, measured by qRT-PCR, in HUVEC supplemented for 24 hours with dIno (500 μM) or pretreated for 30 min with the PNP inhibitor 9-deazaguanine (100 μM) before dIno supplementation for 24 hours ($n = 3$ technical replicates). (**D**) mRNA levels of *ERV* at 24 to 120 hours, measured by qRT-PCR, in siADA2- or siDNMT1-transfected HUVEC ($n = 3$ technical replicates). (**E**) mRNA levels of *ISG* and *ERV*, measured by qRT-PCR, in siADA2-transfected HUVEC treated with anti–IFN-β–neutralizing antibody (40 U/ml) ($n = 3$ technical replicates). (**F**) *ERV*, *IFN*-β, *CCL5*, and *CXCL10* mRNA levels, measured by qRT-PCR, in HUVEC transfected with control plasmid or plasmids encoding *ERVK28* transcript variants. Values are presented as means ± SD. $^*P \le 0.05$, $^{**}P \le 0.01$,a nd $^{***}P \le 0.001$.

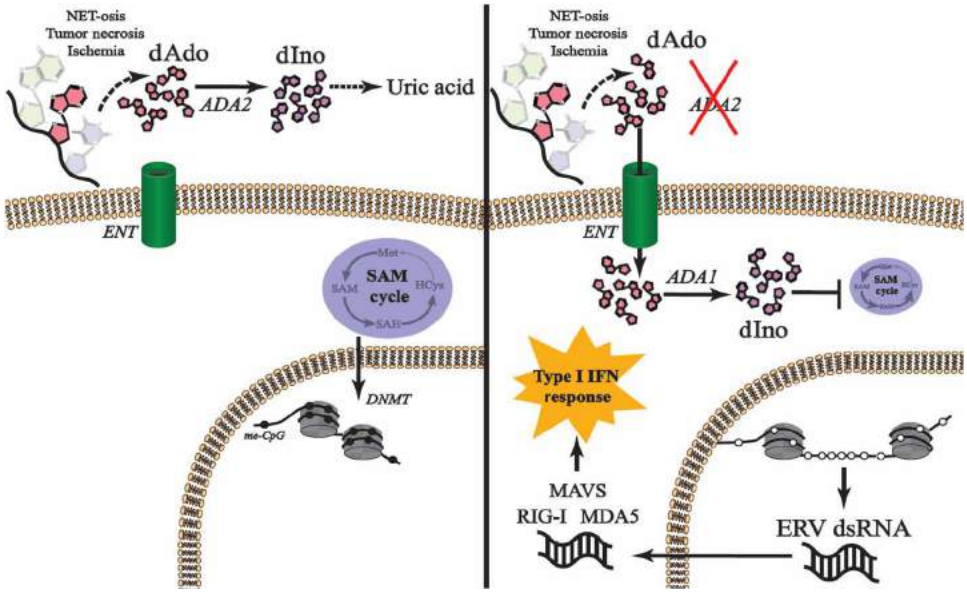


Fig. 6. Cellular sensing of extracellular purine nucleosides triggers an innate immune response. Loss of extracellular ADA2 activity or an excess of extracellular purine nucleosides drives uptake of dAdo and intracellular catabolism by ADA1, yielding dIno, an immune-metabolite that directly inhibits MAT activity, the methionine cycle, and DNA methylation. Genomic hypomethylation drives up-regulation of *ERV*, dsRNA molecules that engage cytosolic dsRNA sensors RIG-I and MDA5 via the signaling adaptor MAVS.

by-products accompanies infection, ischemic injury, and tumor growth.

The ADA2/ADA1 axis controls bioactivity of purine nucleoside metabolites

Results presented here show that cellular innate immunity is triggered by a unique immune-metabolic axis that balances competing, compartmentalized ADA activities in the extracellular and cytosolic spaces. Differential effects of ADA2 and ADA1 depletion on IFN-β expression in HUVEC underscore the complex nature of bioactive purines in modulating immune cells and inflammation. In the extracellular space, Ado exerts pleiotropic effects on innate and adaptive immune cells from signals generated via purinergic G protein–coupled receptors or ligand-gated ion channels (*14*). Because of the broad expression of the A2A and A2B adenosine receptor isoforms (AdoR) in the immune compartment, extracellular Ado exerts substantial anti-inflammatory effects on cells. Here, we show that extracellular dAdo exerts a proinflammatory effect due to triggering of cellular IRF3 activation, IFN-β production, and an IFN-β–driven innate immune response that up-regulates an array of proinflammatory cytokines, chemokines, and signaling proteins (Fig. 1D). Ultimately, the magnitude, duration, and outcome of competing immune-suppressive and immune-stimulatory purine signals at a given immunological niche will depend on metabolic activities of cellular purine salvage or degradation enzymes, as well as ectonucleases and ectonucleotidases such as deoxyribonuclease I (DNAse I), CD39, and CD73 that catalyze Ado and dAdo production from cell-free RNA or DNA in circulation. These enzymes are the key targets for modulating the purine pathway and thus developing novel strategies for immunotherapy.

In HUVEC, ADA activities are compartmentalized such that ADA2 accounts for >75% of extracellular activity, and ADA1 accounts for all intracellular activity (Fig. 2D). Extracellular activity in HUVEC-conditioned supernatants accurately reflects human plasma, where total circulating ADA activity measured in healthy individuals specifically correlates with the ADA2-specific activity (*13*, *23*). Median ADA1 activity measured in normal human plasma accounts for ~25% (*13*), recapitulating results obtained here using gene-specific siRNAs. Despite data from multiple studies demonstrating that ADA2 accounts for most of extracellular ADA activity measured in human plasma, it has nonetheless been postulated that ADA2 plays a minimal role regulating purine metabolism in vivo. This is variously attributed to the particular biochemical properties of ADA2, putative growth factor–like properties of ADA2, and lack of Ado/dAdo elevation in ADA2 deficiency disease (DADA2) patient plasma. Data obtained using rADA2 and rADA1 proteins (Fig. 2E) or dAdo supplementation (Fig. 2F) unequivocally demonstrate that ADA enzymatic activity regulates IFNβ expression. Regarding the biochemistry of ADA2, many early studies reporting enzyme reaction rate and substrate affinity were obtained using purified protein preparations that are highly unstable (*9*) and may have underestimated the actual enzyme activity in vivo. Furthermore, ADA2 activity assays in plasma and cell-conditioned supernatants typically use relatively high nucleoside concentrations (e.g., 1 mM); however, our results were fully recapitulated at a range of dAdo concentrations as low as 100 nM. Notably, ADA2 exhibits a broadly optimal pH range compared to ADA1 (*8*, *9*), suggesting that it may be specifically suited to pathological conditions when release of lactic acid or reduced blood flow/hypoxia drives pH down, which substantially impairs normal metabolic activities and potentially impairs ADA1. Circulating ADA2 activity is specifically increased upon HIV infection (*24*), chronic hepatitis (*25*), rheumatoid arthritis (*26*), and tuberculosis (*27*) and has repeatedly been proposed as a diagnostic tool or surrogate biomarker for multiple human inflammatory diseases. Our new mechanistic insights regarding cellular innate immunity provide a key missing component in our understanding of ADA2 as

a functional regulator of immunity and inflammation rather than a passive biomarker.

While ADA2 has a signal peptide that directs trafficking through classical protein secretion pathways, ADA2 intracellular protein stores appear to be selectively retained in the cytoplasm of myeloid cells and mobilized upon cellular stimulation. In contrast to endothelial cells in which intracellular ADA2 protein (*11*) and enzyme activity (Fig. 2D) is undetectable, monocytes and macrophages retain ADA2 inside cells (*11*). Upon cellular stimulation with phorbol esters and calcium ionophore, intracellular stores of ADA2 are acutely mobilized into the extracellular space (*28*), likely contributing to increased ADA2 activity observed in vivo in human plasma. Transcriptional profiles obtained from unbiased analysis of human hematopoietic cell subsets (*10*) revealed broad expression of ADA2 mRNA in many immune cell types (fig. S2A), where regulated retention and release of ADA2 might be at play during the course of physiological inflammation in vivo. Elucidating the specific signals that regulate inducible ADA2 activity during inflammation will require developing an appropriate, multicellular immunocompetent system, in which ADA1 and ADA2 expression and compartmentalization are conserved.

Extrinsic sensing of DNA danger signals

Like ADA2, extracellular accumulation of bioactive nucleosides also occurs during inflammation upon acute damage to cells and tissues (*14*). Driving local concentrations of extracellular nucleotides and nucleosides well above their homeostatic levels in circulation could exceed the catabolic capacity of ADA2 enzyme activity, pushing cellular equilibrium in favor of nucleoside uptake by the vascular endothelium and other cells. Both Ado and dAdo typically display a half-life of seconds in circulation due to rapid and efficient uptake by vascular endothelial cells (*16*), which express relatively high levels of nucleoside transporters (Fig. 3B) and may confer particular sensitivity to the functional effects of ADA2 deficiency and extracellular dAdo.

Acute elevations of extracellular adenosine triphosphate (ATP) and Ado up to micromolar concentrations have been quantified in vivo upon ischemia in the brain (*29*) and heart (*30*). Extracellular accumulation of dAdo is less well documented; however, a recent unbiased metabolic characterization of myocardial infarction in humans showed that sustained plasma elevations of dAdo occurred during the reperfusion phase of ischemic tissue (*31*), when the bulk of vascular injury and inflammation occurs. Ischemia/reperfusion injury provokes a robust innate type I IFN response (*32*), which is likely to be affected by the dAdo-dependent mechanism reported here. In the solid tumor microenvironment, chronic elevation of Ado at micromolar concentrations plays a critical role in modulating tumor-infiltrating immune cells (*33*). Notably, induction of ERV in tumors drives a protective antitumor IFN-β response (*21*) and is associated with increased survival (*22*), suggesting that enhanced dAdo uptake may play a protective role in the tumor microenvironment by boosting tumor immune surveillance.

Cell-free DNA also accumulates in circulation during pathological inflammation (*4*), and dAdo may be produced upon breakdown of extracellular DNA by circulating DNAse I (*34*). This is particularly relevant for infections and other pathological processes that stimulate the formation of neutrophil extracellular traps (NETs), a unique form of cell death in which decondensed chromatin is released into the extracellular environment by activated neutrophils. A recent study reported enhanced NETosis in cells from DADA2 patients, which was linked to increased Ado signaling via adenosine receptors expressed on neutrophils (*35*). Thus, it is likely that increased NETosis, subsequent release of cellular DNA coupled, and dysregulated dAdo metabolism in ADA2 patients could synergize to drive innate cytokine responses in DADA2 patients. This is of relevance for microbial infections in which excessive neutrophil activation drives pathogenic inflammation, such as the novel coronavirus disease 2019 (COVID-19), which drives excessive NETosis and cytokine-driven pathogenic inflammation in the human lung (*36*).

Clarifying the cellular and molecular mechanisms governing extracellular DNA-DAMP signaling will be necessary to understand how protective or pathogenic inflammatory signals are initiated, propagated, and amplified during the course of inflammatory episodes. In this regard, specialized myeloid cells can sense and respond to immune-stimulatory DNA triggers using both cell-intrinsic and cell-extrinsic mechanisms. This signaling diversity enables these specialized cells to report intrinsic inflammatory provocations or amplify extrinsic signals. Accordingly, here, we show that monocytic cells respond to paracrine effects of ADA2 and dAdo by triggering IFN-β. Whether nonphagocytic or stromal cells can sense external DNA-DAMP triggers, allowing them to amplify inflammation in response to external innate ligands, is unclear. Our results in HUVEC establish purine nucleosides as broadly acting immune-stimulatory agents of sterile inflammation and implicate cellular sensing of extracellular purines as a unique modality, enabling non-specialized cells to fulfill a specialized innate immune function by responding to DNA-DAMP signals originating from distal cell damage. This is relevant to the as-yet incomplete understanding of the origins of innate inflammatory signals in human autoimmune disease and vasculitis, and our data suggest that stromal cells of the vascular endothelium are initiators of inflammatory signals rather than bystanders.

ADA2 deficiency disease

Our previously unknown results linking reduced ADA2 ectoenzyme activity and dysregulated purine nucleoside signaling to spontaneous IFN-β production expands our understanding of the complex and seemingly conflicting pathobiology of DADA2, a poorly understood immunological disease affecting several hundred patients worldwide, primarily children. Many patients with DADA2, whose clinical diagnosis is contingent upon specific reduction of serum ADA2 enzyme activity or loss-of-function mutations in the *ADA2* gene (*11*, *37*), experience clinical symptoms ranging from multiorgan vascular inflammation of the brain, skin, and kidneys to primary immunodeficiency due to bone marrow failure. Our data show that ADA2 deficiency exerts paracrine effects on the immune response that directly relate to dysregulated dAdo/dIno metabolism. Notably, extracellular accumulation of dAdo neither was observed in ADA2-deficient HUVEC nor was it required for spontaneous IFN-β induction (Fig. 3B). Rather, ENT- and ADA1-dependent production of dIno, dysregulation of genomic DNA methylation, and *ERV* induction drove cellular IFN-β production. DADA2 patients do not exhibit steady-state elevations of dAdo in the blood (*11*); however, intracellular measurements of purine metabolites and *ERV* have not been reported in myeloid and stromal cells from these individuals. Furthermore, since steady-state levels of dAdo measured in systemic circulation do not necessarily reflect localized

or dynamic changes in dAdo levels in inflamed tissues or organs, development of an appropriate in vivo or multicellular system will be required to accurately assess effects of ADA2 deficiency on dynamic changes in dAdo levels.

Notable similarities exist between DADA2 and other metabolic syndromes of purines in humans; approximately one-third of patients harboring loss-of-function mutations in PNP, the enzyme that catabolizes dIno, present with elevated levels of dIno and autoimmune manifestations (*38*). These include systemic lupus erythematosus (SLE), a disease of multiorgan autoinflammation and vasculitis, in which elevated type I IFN production is causally implicated. Recently, an unbiased genetic analysis of human SLE identified missense single-nucleotide polymorphisms in the coding region of the *PNP* locus, which correlated in a dose-dependent manner to low PNP expression, low PNP enzyme activity, and high circulating levels of IFN (*39*). The mechanism of IFN up-regulation is unknown but may interface with the novel mechanism described here. The overlapping immunological manifestations of DADA2 and PNP deficiency disease, which run the spectrum of immune deficiency to autoimmunity, support overlapping immune-modulatory mechanisms involving purine nucleosides.

The functional significance of IFN-β in the pathobiology of DADA2 or PNP deficiency disease is not well understood. Notably, IFN-β–driven signaling promotes immune cell activation and drives additional inflammatory cytokines, including tumor necrosis factor (TNF). Notably, elevated TNF is not detectable in the serum in DADA2 patients; however, its functional relevance in driving vascular inflammation in DADA2 arises from the clinical observation that inhibitors of TNF effectively control symptoms of inflammatory vasculopathy and prevent strokes (*40*). Both IFN-β and TNF are functional drivers of pathology in SLE (*41*). In macrophages, TNF induces low levels of IFN-β, and autocrine TNF signaling synergizes with autocrine IFN-β signaling to enhance the expression of proinflammatory genes. Notably, neither control nor ADA2-depleted HUVEC expresses detectable levels of TNF, either at baseline or in response to innate ligands and virus infection. However, our data support a paracrine effect for IFN-β produced by vascular endothelium in driving TNF expression in macrophages. Imbalanced macrophage polarization toward the proinflammatory M1 macrophage phenotype is characteristic of DADA2 disease (*11*), and M1 macrophages are a primary source of TNF in vivo. IFN-β–driven signaling actually drives M1 polarization through activation of the transcription factor signal transducers and activators of transcription 1 and up-regulation of M1-promoting cytokines CXCL10 and CXCL9 (*42*). Thus, our data support a model whereby spontaneous IFN-β produced by vascular endothelium can drive down TNF production by M1 macrophages.

A primary limitation of our study is an absence of in vivo corroboration that ADA2 deficiency or extracellular purine nucleosides drive IFN-β production in stromal and immune cells. This is compounded by the lack of conservation of ADA2 in the mouse and the heterogenous nature of DADA2 disease, which likely affects many different cell types due to autocrine/paracrine activities of ADA2 and purine nucleosides and the variable nature of extracellular purine accumulation. Future studies centered on studying other regulatory enzymes of purine nucleoside metabolism using mouse models of tumor growth, infection, or ischemia, as well as studies of both stromal and immune cells isolated from DADA2 patients undergoing acute manifestations of inflammatory disease, will be required to further understand the complex balance between the proinflammatory and anti-inflammatory activities of purine nucleosides in human disease.

MATERIALS AND METHODS

Experimental design

The objective of this research is to examine novel molecular mechanisms regulating the innate immune IFN-β response using controlled laboratory experiments. To identify novel gene and protein regulators of IFN-β, we optimized and implemented loss-of-function RNA interference screening for activation of the transcription factor IRF3 in primary human vascular endothelial cells. Among 38 candidates associated with systemic inflammation and vasculitis in humans, the metabolic enzyme ADA2 was identified as a novel negative regulator of spontaneous IFN-β production. Subsequent transcriptomic, metabolic, and functional experiments were performed to examine the effects of ADA2 deficiency and purine nucleoside supplementation in triggering IFN-β via activation of cytosolic nucleic acid sensing pathways.

Study design

Sample size (*n*) for each experimental group is described in each figure legend and was determined on the basis of the optimal number to generate statistically significant data. No data were excluded. All experimental findings were independently replicated three times. Outliers were identified at the beginning of the study. Samples were randomized by position on plates, microplates, and flow cells or temporally randomized on the mass spectrometer. For siRNA screening, qRT-PCR, RNA-seq, and MS, research subjects (e.g., cell culture supernatants or lysates) were analyzed in a blinded manner using numerical keys.

Antibodies and reagents

rADA1 (93985) was purchased from Sigma-Aldrich, and rADA2 (7518-AD) was from R&D Systems. ADA2 antibody (HPA007888) was purchased from Sigma-Aldrich, and β-actin antibody (MAB8929) was from R&D. Toll-like receptor 3 antibody (sc-32232) was from Santa Cruz Biotechnology, cGAS antibody (AP10510c) was from Abgent, and DDX58/RIG-I (ab45428), MAVS (ab31334), and IFI16 (ab55328) antibodies were from Abcam. IFIH1/MDA5 antibody (21775-1-AP) was purchased from Proteintech, and phospho-IRF3 (Ser^{396}) (4947S), phospho-TBK1 (Ser^{172}) (5483S), TBK1 (3013S), and IRF3 (4302S) antibodies were from Cell Signaling Technologies. Human IFN-β–neutralizing antibody (31401-1) was purchased from PBL IFN source, and 2′-deoxyadenosine, adenosine, DPM, EHNA, 5-azacytidine, and 9-deazaguanine were from Sigma. Pentostatin was purchased from Cayman Chemicals, cycloleucine was from MP Biomedicals, and 2′-deoxyinosine, inosine, and hypoxanthine were from Alfa Aeser. ELISA kit for ADA2 quantification in cell culture supernatants was purchased from Cloud Clone Corp. poly (dA:dT) DNA was purchased from Invivogen. poly (dA:dT) DNA (P0883) was purchased from Sigma.

Cells

Primary human cells were obtained from Lonza as single-donor aliquots. HUVEC and aortic endothelial cells were maintained in EGM-2 BulletKit growth media (Lonza) supplemented with 2%

hi-FBS (fetal bovine serum). Dermal microvascular endothelial cells, brain microvascular endothelial cells, and kidney microvascular endothelial cells were maintained in EBM-2MV BulletKit growth media (Lonza) supplemented with hi-FBS. Normal bronchial epithelial cells were maintained in BEGM BulletKit growth media (Lonza). U937 monocytes were obtained from American Type Culture Collection and maintained in RPMI supplemented with hi-FBS. Purity of cell populations was verified >95% by flow cytometry using the following markers: CD31 for endothelial cells (BioLegend, 303121), epithelial cell adhesion molecule for epithelial cells (BioLegend, 324207), and CD45 for U937 monocytes (BioLegend, 103115). For activation of IRF3-driven responses, cells were transfected with poly dA:dT for 3 hours (200 ng/ml) using LyoVec transfection reagent (Invivogen) or infected for 3 hours with hCMV strain MOLD [multiplicity of infection (MOI) = 1]. For IFN-β neutralizing experiments, cells were treated with anti–IFN-β–neutralizing antibody (10 to 40 U/ml) for 72 hours.

Virus

hCMV clinical strain MOLD (a gift from M. Raftery) was used. Stocks were prepared as previously described (*6*).

siRNA screening and transfections

siGenome siRNA oligonucleotide pools were purchased from Dharmacon, and transfections were performed as previously described (*6*). siRNA transfection in U937 was performed with a Neon Transfection System 100 μl kit (Thermo Fisher Scientific). Cells were resuspended with R buffer (5×10^6 cells/ml) and electroporated with 100 nM siRNA at 1400 V/10 ms/3 pulse. Following transfection with siRNA (30 to 40 nM), growth media was changed every 24 hours, and cells were harvested after 72 hours unless otherwise specified. For the IRF3 screen, *z* score was calculated for each gene-specific siRNA pool using mean and SD values from triplicate wells compared to control, nontargeting siRNA (Dharmacon). Individual *z* scores are listed in table S1. siRNA sequences are listed in table S2.

IRF3 immunostaining

IRF3 immunostaining was performed as described (*6*). IRF3 antibody (sc-9082) was purchased from Santa Cruz Biotechnology, and Alexa Fluor 647–conjugated anti-rabbit secondary antibody (111-604-144) was from Jackson ImmunoResearch. DAPI (4′,6-diamidino-2-phenylindole) was used to counter stain nuclei. Images were acquired at ×10 magnification on ImageXpress Micro (Molecular Devices), and images were analyzed using the enhanced translocation module of MetaXpress (Molecular Devices). Individual cells were scored as positive for nuclear translocation if >70% of IRF3 fluorescence were spatially correlated with nuclear probe. Each data point is representative of >200 cells.

Quantitative real-time PCR

Total RNA was extracted using a Quick-RNA MiniPrep Plus kit (Zymo Research). RNA (500 ng) was used to synthesize complementary DNA (cDNA) using a qScript cDNA synthesis kit (Quanta). Samples were diluted 10-fold, and gene expression was analyzed on a CFX96 Touch Detection System (Bio-Rad) using FastStart SYBR Green Master Mix (Roche) or TaqMan Universal PCR Master Mix (Thermo Fisher Scientific). Primer pair sequences and catalog numbers for cellular mRNAs are listed in table S3. Primer pair and catalog numbers for ERV genes are listed in table S4.

Western blotting

For Western blotting of whole-cell lysates, cells were lysed in 1X radioimmunoprecipitation assay buffer (Cell Signaling Technology), incubated on ice for 30 min, and centrifuged at 10,000 rpm for 10 min. Protein concentration was quantified from cleared supernatants using bicinchoninic acid protein assay kit (Thermo Fisher Scientific). Equivalent amounts of lysate (30 to 50 μg) were boiled with 5× SDS sample buffer and resolved by 10% SDS–polyacrylamide gel electrophoresis (SDS-PAGE), transferred to polyvinylidene difluoride membrane (Thermo Fisher Scientific), blocked with 5% milk, and incubated with primary antibody (1:1000) at 4°C overnight. Blots were washed with tris buffered saline-0.1% tween 20 and incubated with horseradish peroxidase–conjugated secondary antibodies (1:7000) at room temperature for 1 hour. Signal was detected using Amersham ECL Prime (GE Healthcare). For Western blotting of cultured supernatants to detect ADA2, HUVEC or U937 cells plated at 5×10^5 cells per well in six-well plates in serum-free media were cultured for 48 hours before collecting supernatants. Supernatant (2 to 6 ml) was filtered (0.45-μm filter, 28413-312, VWR) and concentrated with Pierce Protein Concentrator polyethersulfone 10K molecular weight cutoff (88527, Thermo Fisher Scientific) at 5000 rpm for 15 min to obtain 100 to 150 μl of samples. Concentrated supernatants were incubated with 5x SDS sample buffer at 70°C for 10 min and resolved by 10% SDS-PAGE as described above. ADA2 antibody was obtained from Sigma (HPA007888).

RNA sequencing

Total RNA was extracted from 2×10^5 cultured cells using the Quick-RNA MiniPrep samples and samples were checked for RNA integrity number equivalent (RIN^e) scores above 9.0. For each sample, 500 to 1000 ng of total RNA were then prepared into an mRNA library using the TruSeq Stranded mRNA Library Prep Kit (Illumina). Resulting libraries were then pooled at equimolar concentrations using the Quant-iT PicoGreen double-stranded DNA Assay Kit (Life Technologies) and were deep sequenced on an Illumina 2500 in Rapid Run Mode, producing between 10 and 90 M single-end reads with lengths of 50 nucleotides per sample. The single-end reads that passed Illumina filters were filtered for reads aligning to transfer RNA (tRNA), ribosomal RNA (rRNA), adapter sequences, and spike-in controls. The reads were then aligned to University of California, Santa Cruz hg38 reference genome using TopHat (v1.4.1). DUST scores were calculated with PRINSEQ Lite (v0.20.3), and low-complexity reads (DUST of >4) were removed from the BAM files. The alignment results were parsed via SAMtools to generate SAM files. Read counts to each genomic feature were obtained with the htseq-count program (v0.6.0) using the "union" option. After removing absent features (zero counts in all samples), the raw counts were then imported to R/Bioconductor package DESeq2 to identify differentially expressed genes among samples. DESeq2 normalizes counts by dividing each column of the count table (samples) by the size factor of this column. The size factor is calculated by dividing the samples by geometric means of the genes. This brings the count values to a common scale suitable for comparison. *P* values for differential expression are calculated using binomial test for differences between the base means of two conditions. These *P* values are then adjusted for multiple test correction using Benjamini-Hochberg algorithm to control the false discovery rate. We considered genes differentially expressed between two groups of samples when the DESeq2 analysis resulted in an adjusted *P* value of <0.05,

and the fold change in gene expression was twofold. Cluster analyses including principal components analysis and hierarchical clustering were performed using standard algorithms and metrics. Hierarchical clustering was performed using complete linkage with Euclidean metric.

For analysis of ERV expression, the single-end reads that passed Illumina filters were filtered for reads aligning to tRNA, rRNA, adapter sequences, and spike-in controls. The reads were then aligned to the reference genome using TopHat (v1.4.1) and alignment results parsed via SAMtools to generate SAM files. The uniquely and multi-mapped reads were filtered from the BAM files using an MAPQ (mapping quality) of 255 to filter uniquely mapped reads. Read counts to each repeat element were obtained by mapping the multi-mapped reads to each of the repeat classes and counting alignment for each repeat class (all the steps are part of the RepEnrich pipeline). After removing absent repeat elements (zero counts in all samples), the raw counts were then imported to R/Bioconductor package EdgeR differentially expressed genes among samples. The generalized linear model method within the EdgeR package was used to identify differentially expressed genes. We considered genes differentially expressed between two groups of samples when the EdgeR analysis resulted in an adjusted P value of <0.05.

rADA supplementation

HUVECs were seeded at 5×10^5 cells per well in six-well plates and incubated for 24 hours at 37°C with 5% CO_2. Twenty-four hours after transfection with siRNA, fresh growth media containing rADA1 (Sigma, 93985) or ADA2 (R&D Systems, 7518-AD) was added to wells (50 to 1000 ng/ml). Where indicated, pentostatin (10 μM) was added 1 hour before rADA2. Seventy-two hours after transfection, cells were harvested, and gene expression was analyzed by qRT-PCR.

Coculture of HUVEC and U937

HUVECs (1.4×10^5) were plated in 24-well plates. Twenty-four hours after transfection with siRNA, media was changed into fresh EGM-2, and wild-type U937 cells (1×10^5 cells/200 μl of RPMI, 2% FBS) were plated on polyester membrane Transwell clear inserts (0.4 μm) and cocultured with HUVECs for an additional 48 hours before analysis of gene expression by qRT-PCR.

Metabolite extraction and LC-MS/MS

For cell sample preparation, 3×10^6 HUVECs were lysed by aspirating in 80% cold methanol, followed by three freeze-thaw cycles, where samples were alternated between 40° and −80°C baths in 30-s intervals. Lysates were then centrifuged at 4°C at 14,000 rpm for 10 min to remove cellular debris, and supernatants were dried using a Speed Vacuum system for 2 hours at 30°C before resuspension in an 80:20 ratio of methanol and water. For spent media samples, media was centrifuged at 14,000 rpm for 1 min to remove cell debris before adding cold methanol at a ratio of 80:20 methanol and water. Cell and media samples were placed at −20°C for 30 min to precipitate protein before centrifugation at 14,000 rpm for 10 min at 4°C. Each injection for liquid chromatography–tandem MS (LC-MS/MS) analysis was equivalent to 80,000 cells or 0.5 μl of media. LC was performed with a Vanquish UHPLC system (Thermo Fischer Scientific) using a ZIC-pHILIC polymeric column (EMD Millipore) (150 mm by 2.1 mm, 5 μm). Mobile phases used were (A) 20 mM ammonium bicarbonate in water (pH 9.6) and (B) acetonitrile. Mobile phase gradients were as follows: 90% B for first two 2 min followed by a linear gradient to 55% B at 16 min, sustained for an additional 3 min before final re-equilibration for 11 min at 90% B. Column temperature was 25°C, and mobile phase flow rate was 0.3 ml/min. Detection was performed using a Q Exactive Orbitrap mass spectrometer with heated electrospray ionization source (Thermo Scientific). Sheath and auxiliary gas were high purity nitrogen at 40 au (arbitrary units) and 20 au, respectively. MS/MS were collected in both positive and negative ionization polarities. Ion transfer tube was set at 275°C, and vaporizer temperature was set to 350°C. Collision-induced dissociation (CID) MS2 was collected using stepped normalized collision energies of 15, 30, and 45 au and isolation widths of ±0.5 mass/charge ratio (m/z). Peak heights from chromatograms corresponding to metabolites of interest were used for quantification. Indicated metabolites were detected as protonated/deprotonated ions of their monoisotopic masses, and identities were confirmed by comparison to the MS1 and MS2 patterning of known standards under this LC-MS method.

Methionine labeling and SAM flux

EGM-2 medium with dialyzed serum containing 64 μM unlabeled methionine was supplemented with either 64 μM U-^{13}C methionine (Cambridge Isotope Laboratories Inc.; ^{13}C medium) or 64 μM unlabeled methionine (^{12}C medium). At time zero, cells cultured in ^{12}C medium were switched to ^{13}C medium, incubated for 10 or 30 min at 37°C, washed twice with ice-cold phosphate-buffered saline (PBS), extracted in 80% cold methanol, and analyzed with LC-MS as described above to obtain mass isotopomer distributions. Because of lack of remethylation of homocysteine, SAM metabolism was modeled as a linear pathway from methionine to SAM to SAH, governed by the differential equations

$$\dot{x}_{SAM}/\mu_{SAM} = x_{met} - x_{SAM}$$

$$\dot{x}_{SAH}/\mu_{SAH} = x_{SAM} - x_{SAH}$$

where x_{SAM} and x_{SAH} are time-dependent mass isotopomer fractions with time derivatives $\dot{x}$, while x_{met} was considered constant; v is the flux through the linear pathway; and μ is the rate constant, related to the flux as $\mu \cdot c = v$, where c is the pool size (concentration) (*19*). This model was fit to the $^{13}C_5$-methionine, $^{13}C_5$-SAM, and $^{13}C_4$-SAH mass isotopomers at the 10- and 30-min time points by minimizing the variance-weighted sum of square errors. The resulting observed x^2 error was 3.95, which was acceptable at the 95% level. CIs for the parameters μ were obtained as previously described (*43*). Change in flux was obtained as the ratio

$$v^{siADA2}/v^{siCtrl} = \mu^{siADA2} c^{siADA2}/(\mu^{siCtrl} c^{siCtrl})$$

and 95% CIs for this ratio were obtained from the corresponding CIs for μ.

ADA and MAT activity

In vitro dAdo deamination assays were adapted from previously described methodology (*11*) by measuring de novo dIno and hypoxanthine production using LC-MS/MS. For extracellular activity, cultured HUVEC-conditioned EGM-2 supernatant was centrifuged at 1500 rpm for 5 min to remove cellular debris. For intracellular activity, whole-cell lysates were diluted to 10×10^5 cell equivalents per milliliter in cold PBS. Supernatant or lysate was incubated with

100 nM to 1 mM isotopically labeled 2′-deoxyadenosine ($^{15}N_5$, Cambridge Isotope Laboratories) for 30 and 120 min at 37°C. Reactions were quenched using ice-cold methanol at a ratio of 80:20 methanol: reaction volume and centrifuged at 14,000 rpm for 10 min to remove protein. Supernatant was isolated, from which 2 μl was injected for LC-MS/MS analysis. Quantification of deaminase activity was measured as nanomoles of isotopically labeled dAdo and hypoxanthine produced per liter per minute. For quantification of extracellular, HUVEC-derived ADA activity, reaction rate in cell-free EGM-2 growth media was subtracted from rates in culture supernatant, and data were normalized to cell number and protein concentration. In vitro methionine adenosyl-transferase assays were performed by detection of isotopically labeled *S*-adenosyl methionine by LC-MS/MS, as described above. Fresh cytoplasmic protein lysates were isolated from U937 monocytes. In tissue culture–treated, 96-well low-evaporation plates, 10 μg of lysate was incubated with 1 mM U-^{13}C methionine, 1 mM ATP, and nucleoside at indicated concentration. Reactions were buffered in 20 mM $MgSO_4$, 150 mM KCl, and 100 mM tris (pH 7.4). After incubation for 30 min at 37°C, reactions were quenched using ice-cold methanol at a ratio of 80:20 methanol:reaction volume. Samples were centrifuged at 14,000 rpm for 10 min to remove protein, and supernatants were dried using a Speed Vacuum concentrator system for 2 hours at 30°C. Samples were resuspended in 50 μl of 80:20 methanol and water, with 2 μl injected for LC-MS/MS. Quantification of MAT activity was measured via *m/z* 404.1611 chromatogram peak intensity, corresponding to the $M+H^+$ ion of 5-carbon U-^{13}C–labeled *S*-adenosyl methionine and confirmed by diagnostic tandem mass spectra.

Quantification of labeled methyl-cytosine and labeled deoxyinosine

For labeled deoxyinosine, HUVECs were transfected with siRNA against the indicated target for 72 hours, followed by treatment with 85 μM U-^{15}N-2′-deoxyadenosine (Cambridge Isotope Laboratories) for 5 min. Cells were lysed, and metabolite extraction with LC-MS/MS was performed as described. 2′-Deoxyinosine with full label incorporation (4^{15}N-2′-deoxyinosine) was identified by diagnostic MS/MS fragmentation pattern, and relative abundance was quantified by peak height. For labeled methyl-cytosine, HUVECs were transfected with siRNA against the indicated target for 72 hours. During this time, cells were incubated with U-^{13}C methionine (Cambridge Isotope Laboratories) at 64 μM, the equivalent concentration to endogenous levels in growth medium, for the indicated times. Following lysing, genomic DNA was isolated using the PureLink Genomic DNA Mini Kit (Invitrogen) and quantified before digestion with a DNA Degradase I (Zymogen) kit according to the manufacturer's protocol. Product deoxynucleosides were mixed 1:2 with a 50% acetonitrile and 50% water and 40 mM ammonium bicarbonate solution. Each injection of digested genomic isolate was equivalent to 5 ng of DNA. LC was performed with a Vanquish UHPLC system (Thermo Scientific) using a ZIC-HILIC column (EMD Millipore) (150 mm × 1 mm, 5 μM). Metabolites were eluted following a gradient mobile phase of (A) 40 mM ammonium bicarbonate and (B) acetonitrile. The gradients were used as follows: 90% B at start, immediately following a linear gradient to 0% B over 2 min, and sustaining 0% B for two additional minutes of detection before a final equilibration for 1 min at 90% B. Column temperature was 25°C, and mobile phase flow rate was 0.175 ml/min. Detection was performed using a Q Exactve Orbitrap mass spectrometer (Thermo Scientific) as described above, with CID MS2 collected using stepped normalized collision energies of 30 and 45 au. Peak heights from chromatograms corresponding to the M+1 isotopomer of 2′-deoxy-5-methylcytidine were used for quantification, subtracting peak height attributable to naturally occurring ^{13}C isotope.

ERV cloning

cDNAs for ERVK28 transcripts 1 and 3 were ordered from IDT. cDNAs were amplified using the Phusion polymerase enzyme and specific primers that are flanked by Eco RI (forward: 5′-GCT**GAATTC**-CAGGTATTGTAGGGG-3′ for transcript 1; forward: 5′-GCT-**GAATTC**ATGAACCCATCGGAG-3′ for transcript 3; Eco RI sites in bold) and Apa I or Xho I restriction sites (reverse: 5′-GT**GGGCCC**T-GCAAAATGGAGT-3′ for transcript 1; Apa I site in bold; reverse: 5′-G**CTCGAG**GCGCAGAATTTTTC-3′ for transcript 3; Xho I site in bold), respectively. The amplicons [308 base pair (bp) for transcript 1 and 457 bp for transcript 3] were digested with Eco RI/Apa I and Eco RI/Xho I, respectively, and were cloned in frame with Flag between the Eco RI and Apa I sites or Eco RI and Xho I sites, respectively, in pcDNA3-Flag-TRAF6 (Addgene, no. 66929). Plasmids were produced in One Shot TOP10 bacteria, and constructs sequence was verified.

ERV overexpression

For electroporation, 2×10^5 HUVECs were transferred to a sterile 1.5-ml microcentrifuge tube, resuspended in 10 μl of buffer R, and mixed with 500 ng of pcDNA3-Flag vector. DNA electroporation was performed with a Neon Transfection System 10 μl kit (Thermo Fisher Scientific) using the following settings: 1350 V, 30 ms, and 2 pulses. After electroporation, cells from five 10-μl transfection reactions were added to the same well of a six-well plate containing 2 ml of prewarmed supplemented EGM-2 BulletKit growth media (Lonza) without antibiotics. Twenty-four hours after transfection, new media with antibiotics was added to the wells. Forty-eight hours after transfection, total RNA was extracted using a Quick-RNA MiniPrep Plus kit (Zymo Research). RNA (500 ng) was used to synthesize cDNA using a qScript cDNA synthesis kit (Quanta). Samples were diluted sixfold, and gene expression was analyzed on a CFX96 Touch Detection System (Bio-Rad) using TaqMan Universal PCR Master Mix (Thermo Fisher Scientific).

Statistical analysis

Data values are reported as means ± SD, with differences determined using two-tailed Student's *t* test, with Benjamini-Hochberg adjustment for multiple comparisons. Regarding statistical calculations, normal data distribution was confirmed using controls or previously generated large-scale datasets ($n > 384$) before analysis of experimental samples. All reported results were independently replicated using three independent experiments.

SUPPLEMENTARY MATERIALS

Supplementary material for this article is available at http://advances.sciencemag.org/cgi/content/full/6/30/eaba3688/DC1

View/request a protocol for this paper from *Bio-protocol*.

REFERENCES AND NOTES

1. G. R. Stark, How cells respond to interferons revisited: from early history to current complexity. *Cytokine Growth Factor Rev.* **18**, 419–423 (2007).

2. J. Wu, Z. J. Chen, Innate immune sensing and signaling of cytosolic nucleic acids. *Annu. Rev. Immunol.* **32**, 461–488 (2014).
3. R. Dhanwani, M. Takahashi, S. Sharma, Cytosolic sensing of immuno-stimulatory DNA, the enemy within. *Curr. Opin. Immunol.* **50**, 82–87 (2018).
4. D. S. Pisetsky, The origin and properties of extracellular DNA: from PAMP to DAMP. *Clin. Immunol.* **144**, 32–40 (2012).
5. M. K. Crow, M. Olferiev, K. A. Kirou, Type I Interferons in Autoimmune Disease. *Annu. Rev. Pathol.* **14**, 369–393 (2019).
6. C.-W. Lio, B. McDonald, M. Takahashi, R. Dhanwani, N. Sharma, J. Huang, E. Pham, C. A. Benedict, S. Sharma, cGAS-STING Signaling Regulates Initial Innate Control of Cytomegalovirus Infection. *J. Virol.* **90**, 7789–7797 (2016).
7. A. Belot, E. Wassmer, M. Twilt, J. C. Lega, L. A. Zeef, A. Oojageer, P. R. Kasher, A. L. Mathieu, C. Malcus, J. Demaret, N. Fabien, S. Collardeau-Frachon, L. Mechtouff, L. Derex, T. Walzer, G. I. Rice, I. Durieu, Y. J. Crow, Mutations in CECR1 associated with a neutrophil signature in peripheral blood. *Pediatr. Rheumatol. Online J.* **12**, 44 (2014).
8. A. V. Zavialov, A. Engstrom, Human ADA2 belongs to a new family of growth factors with adenosine deaminase activity. *Biochem. J.* **391**, 51–57 (2005).
9. N. A. Andreasyan, H. L. Hairapetyan, Y. G. Sargisova, S. S. Mardanyan, ADA2 isoform of adenosine deaminase from pleural fluid. *FEBS Lett.* **579**, 643–647 (2005).
10. N. Novershtern, A. Subramanian, L. N. Lawton, R. H. Mak, W. N. Haining, M. E. McConkey, N. Habib, N. Yosef, C. Y. Chang, T. Shay, G. M. Frampton, A. C. Drake, I. Leskov, B. Nilsson, F. Preffer, D. Dombkowski, J. W. Evans, T. Liefeld, J. S. Smutko, J. Chen, N. Friedman, R. A. Young, T. R. Golub, A. Regev, B. L. Ebert, Densely interconnected transcriptional circuits control cell states in human hematopoiesis. *Cell* **144**, 296–309 (2011).
11. Q. Zhou, D. Yang, A. K. Ombrello, A. V. Zavialov, C. Toro, A. V. Zavialov, D. L. Stone, J. J. Chae, S. D. Rosenzweig, K. Bishop, K. S. Barron, H. S. Kuehn, P. Hoffmann, A. Negro, W. L. Tsai, E. W. Cowen, W. Pei, J. D. Milner, C. Silvin, T. Heller, D. T. Chin, N. J. Patronas, J. S. Barber, C. C. Lee, G. M. Wood, A. Ling, S. J. Kelly, D. E. Kleiner, J. C. Mullikin, N. J. Ganson, H. H. Kong, S. Hambleton, F. Candotti, M. M. Quezado, K. R. Calvo, H. Alao, B. K. Barham, A. Jones, J. F. Meschia, B. B. Worrall, S. E. Kasner, S. S. Rich, R. Goldbach-Mansky, M. Abinun, E. Chalom, A. C. Gotte, M. Punaro, V. Pascual, J. W. Verbsky, T. R. Torgerson, N. G. Singer, T. R. Gershon, S. Ozen, O. Karadag, T. A. Fleisher, E. F. Remmers, S. M. Burgess, S. L. Moir, M. Gadina, R. Sood, M. S. Hershfield, M. Boehm, D. L. Kastner, I. Aksentijevich, Early-onset stroke and vasculopathy associated with mutations in ADA2. *N. Engl. J. Med.* **370**, 911–920 (2014).
12. A. V. Zavialov, X. Yu, D. Spillmann, G. Lauvau, A. V. Zavialov, Structural basis for the growth factor activity of human adenosine deaminase ADA2. *J. Biol. Chem.* **285**, 12367–12377 (2010).
13. P. Y. Lee, Z. Huang, M. S. Hershfield, P. A. Nigrovic, Analysis of peripheral blood ADA1 and ADA2 levels in children and adults. Response to: 'Total adenosine deaminase highly correlated with adenosine deaminase 2 activity in serum' by Gao et al. *Ann. Rheum. Dis.* annrheumdis-2020-217055 (2020).
14. J. Linden, F. Koch-Nolte, G. Dahl, Purine Release, Metabolism, and Signaling in the Inflammatory Response. *Annu. Rev. Immunol.* **37**, 325–347 (2019).
15. E. Szabados, R. G. Duggleby, R. I. Christopherson, Metabolism of adenosine and deoxyadenosine by human erythrocytes and CCRF-CEM leukemia cells. *Int. J. Biochem. Cell Biol.* **28**, 1405–1415 (1996).
16. R. T. Smolenski, Z. Kochan, R. McDouall, A. M. Seymour, M. H. Yacoub, Adenosine uptake and metabolism in human endothelial cells. *Adv. Exp. Med. Biol.* **370**, 435–438 (1994).
17. M. Jain, R. Nilsson, S. Sharma, N. Madhusudhan, T. Kitami, A. L. Souza, R. Kafri, M. W. Kirschner, C. B. Clish, V. K. Mootha, Metabolite profiling identifies a key role for glycine in rapid cancer cell proliferation. *Science* **336**, 1040–1044 (2012).
18. R. L. Williams-Karnesky, U. S. Sandau, T. A. Lusardi, N. K. Lytle, J. M. Farrell, E. M. Pritchard, D. L. Kaplan, D. Boison, Epigenetic changes induced by adenosine augmentation therapy prevent epileptogenesis. *J. Clin. Invest.* **123**, 3552–3563 (2013).
19. R. Nilsson, M. Jain, Simultaneous tracing of carbon and nitrogen isotopes in human cells. *Mol. BioSyst.* **12**, 1929–1937 (2016).
20. L. Huang, D. Kim, X. Liu, C. R. Myers, J. W. Locasale, Estimating relative changes of metabolic fluxes. *PLOS Comput. Biol.* **10**, e1003958 (2014).
21. K. B. Chiappinelli, P. L. Strissel, A. Desrichard, H. Li, C. Henke, B. Akman, A. Hein, N. S. Rote, L. M. Cope, A. Snyder, V. Makarov, S. Budhu, D. J. Slamon, J. D. Wolchok, D. M. Pardoll, M. W. Beckmann, C. A. Zahnow, T. Merghoub, T. A. Chan, S. B. Baylin, R. Strick, Inhibiting DNA Methylation Causes an Interferon Response in Cancer via dsRNA Including Endogenous Retroviruses. *Cell* **164**, 1073 (2016).
22. I. Canadas, R. Thummalapalli, J. W. Kim, S. Kitajima, R. W. Jenkins, C. L. Christensen, M. Campisi, Y. Kuang, Y. Zhang, E. Gjini, G. Zhang, T. Tian, D. R. Sen, D. Miao, Y. Imamura, T. Thai, B. Piel, H. Terai, A. R. Aref, T. Hagan, S. Koyama, M. Watanabe, H. Baba, A. E. Adeni, C. A. Lydon, P. Tamayo, Z. Wei, M. Herlyn, T. U. Barbie, R. Uppaluri, L. M. Sholl, E. Sicinska, J. Sands, S. Rodig, K. K. Wong, C. P. Paweletz, H. Watanabe, D. A. Barbie, Tumor innate immunity primed by specific interferon-stimulated endogenous retroviruses. *Nat. Med.* **24**, 1143–1150 (2018).
23. Z.-W. Gao, X. Wang, F. Lin, K. Dong, Total adenosine deaminase highly correlated with adenosine deaminase 2 activity in serum. *Ann. Rheum. Dis.* annrheumdis-2020-217055 (2020).
24. I. Khodadadi, M. Abdi, A. Ahmadi, M. S. Wahedi, S. Menbari, F. Lahoorpour, R. Rahbari, Analysis of serum adenosine deaminase (ADA) and ADA1 and ADA2 isoenzyme activities in HIV positive and HIV-HBV co-infected patients. *Clin. Biochem.* **44**, 980–983 (2011).
25. E. Fernandez, L. Rodrigo, S. Riestra, S. Carcia, F. Gutierrez, G. Ocio, Adenosine deaminase isoenzymes and neopterin in liver cirrhosis. *J. Clin. Gastroenterol.* **30**, 181–186 (2000).
26. S. Valadbeigi, R. Saghiri, M. Ebrahimi-Rad, S. Khatami, H. Akhbari, Adenosine Deaminase Activity and HLA-DRB as Diagnostic Markers for Rheumatoid Arthritis. *Curr. Rheumatol. Rev.* **15**, 44–49 (2019).
27. J. P. Ungerer, H. M. Oosthuizen, J. H. Retief, S. H. Bissbort, Significance of adenosine deaminase activity and its isoenzymes in tuberculous effusions. *Chest* **106**, 33–37 (1994).
28. S. Iwaki-Egawa, T. Yamamoto, Y. Watanabe, Human plasma adenosine deaminase 2 is secreted by activated monocytes. *Biol. Chem.* **387**, 319–321 (2006).
29. N. Dale, B. G. Frenguelli, Release of adenosine and ATP during ischemia and epilepsy. *Curr. Neuropharmacol.* **7**, 160–179 (2009).
30. A. Sollevi, Cardiovascular effects of adenosine in man; possible clinical implications. *Prog. Neurobiol.* **27**, 319–349 (1986).
31. A. Surendran, M. Aliani, A. Ravandi, Metabolomic characterization of myocardial ischemia-reperfusion injury in ST-segment elevation myocardial infarction patients undergoing percutaneous coronary intervention. *Sci. Rep.* **9**, 11742 (2019).
32. A. McDonough, R. V. Lee, S. Noor, C. Lee, T. Le, M. Iorga, J. L. H. Phillips, S. Murphy, T. Moller, J. R. Weinstein, Ischemia/Reperfusion Induces Interferon-Stimulated Gene Expression in Microglia. *J. Neurosci.* **37**, 8292–8308 (2017).
33. P. Vaupel, A. Mayer, Hypoxia-Driven Adenosine Accumulation: A Crucial Microenvironmental Factor Promoting Tumor Progression. *Adv. Exp. Med. Biol.* **876**, 177–183 (2016).
34. P. A. Keyel, Dnases in health and disease. *Dev. Biol.* **429**, 1–11 (2017).
35. C. Carmona-Rivera, S. S. Khaznadar, K. W. Shwin, J. A. Irizarry-Caro, L. J. O'Neil, Y. Liu, K. A. Jacobson, A. K. Ombrello, D. L. Stone, W. L. Tsai, D. L. Kastner, I. Aksentijevich, M. J. Kaplan, P. C. Grayson, Deficiency of adenosine deaminase 2 triggers adenosine-mediated NETosis and TNF production in patients with DADA2. *Blood* **134**, 395–406 (2019).
36. B. J. Barnes, J. M. Adrover, A. Baxter-Stoltzfus, A. Borczuk, J. Cools-Lartigue, J. M. Crawford, J. Dassler-Plenker, P. Guerci, C. Huynh, J. S. Knight, M. Loda, M. R. Looney, F. McAllister, R. Rayes, S. Renaud, S. Rousseau, S. Salvatore, R. E. Schwartz, J. D. Spicer, C. C. Yost, A. Weber, Y. Zuo, M. Egeblad, Targeting potential drivers of COVID-19: Neutrophil extracellular traps. *J. Exp. Med.* **217**, e20200652 (2020).
37. P. Navon Elkan, S. B. Pierce, R. Segel, T. Walsh, J. Barash, S. Padeh, A. Zlotogorski, Y. Berkun, J. J. Press, M. Mukamel, I. Voth, P. J. Hashkes, L. Harel, V. Hoffer, E. Ling, F. Yalcinkaya, O. Kasapcopur, M. K. Lee, R. E. Klevit, P. Renbaum, A. Weinberg-Shukron, E. F. Sener, B. Schormair, S. Zeligson, D. Marek-Yagel, T. M. Strom, M. Shohat, A. Singer, A. Rubinow, E. Pras, J. Winkelmann, M. Tekin, Y. Anikster, M. C. King, E. Levy-Lahad, Mutant adenosine deaminase 2 in a polyarteritis nodosa vasculopathy. *N. Engl. J. Med.* **370**, 921–931 (2014).
38. M. L. Markert, Purine nucleoside phosphorylase deficiency. *Immunodefic. Rev.* **3**, 45–81 (1991).
39. Y. Ghodke-Puranik, J. M. Dorschner, D. M. Vsetecka, S. Amin, A. Makol, F. Ernste, T. Osborn, K. Moder, V. Chowdhary, E. Eliopoulos, M. I. Zervou, G. N. Goulielmos, M. A. Jensen, T. B. Niewold, Lupus-Associated Functional Polymorphism in PNP Causes Cell Cycle Abnormalities and Interferon Pathway Activation in Human Immune Cells. *Arthritis Rheumatol.* **69**, 2328–2337 (2017).
40. A. K. Ombrello, J. Qin, P. M. Hoffmann, P. Kumar, D. Stone, A. Jones, T. Romeo, B. Barham, G. Pinto-Patarroyo, C. Toro, A. Soldatos, Q. Zhou, N. Deuitch, I. Aksentijevich, S. L. Sheldon, S. Kelly, A. Man, K. Barron, M. Hershfield, W. A. Flegel, D. L. Kastner, Treatment Strategies for Deficiency of Adenosine Deaminase 2. *N. Engl. J. Med.* **380**, 1582–1584 (2019).
41. A. Yarilina, L. B. Ivashkiv, Type I interferon: a new player in TNF signaling. *Curr. Dir. Autoimmun.* **11**, 94–104 (2010).
42. N. Wang, H. Liang, K. Zen, Molecular mechanisms that influence the macrophage m1-m2 polarization balance. *Front. Immunol.* **5**, 614 (2014).
43. M. R. Antoniewicz, J. K. Kelleher, G. Stephanopoulos, Determination of confidence intervals of metabolic fluxes estimated from stable isotope measurements. *Metab. Eng.* **8**, 324–337 (2006).

Acknowledgments: We thank D. Araujo for editing and formatting the manuscript; Y. Kushnareva and C. Oh at the LJI Functional Genomics Center for screening; J. Huang for optimization of IRF3/IFNβ assays in primary human endothelial cells; D. Freeman, B. McDonald, and R. El Morabiti for hCMV propagation and infections; J. Day for RNA-seq; K. Lehman for metabolomics; G. Guzzo for optimization of SAM flux; J. Greenbaum and the LJI Bioinformatics

Core for statistical analyses; and E. Chambers and the DADA2 Foundation for helpful discussions. **Funding:** This work was supported by NIH grants R01CA199376 and U01DE028227 and a research grant from the Vasculitis Foundation awarded to S.S., NIH S10OD020025 and R01ES027595 to M.J., and RO1 AI101423 to C.A.B. I.T.M. is supported by NIH T32GM007752, F31CA236405, and a SPARK award from the LJI Board of Directors. **Author contributions:** The functional screen was optimized and performed by R.D. who validated ADA2 and purine nucleosides as regulators of IFNβ in endothelial cells and monocytes. M.T. contributed significantly in defining the specific mechanism linking ADA2 and IFNβ by designing, optimizing, and performing in vitro experiments. I.T.M. performed MS metabolomics. C.L. performed rescue experiments using rADA2 and rADA1 and ERV cloning and overexpression. A.R. performed ADA2 Western blot. J.D.W. optimized the labeled methyl-cytosine assay. B.P. compiled the list of disease genes. S.S. provided overall direction and supervised project planning and execution, with input from C.A.B. (hCMV), C.C.H. (endothelial and monocyte biology), J.L. (purine signaling), R.N. (SAM flux), and M.J. (metabolomics). S.S., M.J., R.D., M.T., and I.T.M. wrote the manuscript with input from co-authors. **Competing interests:** S.S., R.D., and M.T. are inventors on a patent application related to this work filed by La Jolla Institute for Immunology (no. US2019/0282580 A1, filed 19 September 2019). The authors declare that they have no other competing interests. **Data and materials availability:** All materials used in this study will be made available to the scientific research community, conditional upon execution of a materials transfer agreement between LJI and the recipient institution. All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. Raw RNA-seq data for Figs. 1D and 3D can be accessed at the Gene Expression Omnibus (GSE150540 and sGSE151924). Additional data related to this paper may be requested from the authors.

Submitted 27 November 2019
Accepted 9 June 2020
Published 22 July 2020
10.1126/sciadv.aba3688

Citation: R. Dhanwani, M. Takahashi, I. T. Mathews, C. Lenzi, A. Romanov, J. D. Watrous, B. Pieters, C. C. Hedrick, C. A. Benedict, J. Linden, R. Nilsson, M. Jain, S. Sharma, Cellular sensing of extracellular purine nucleosides triggers an innate IFN-β response. *Sci. Adv.* **6**, eaba3688 (2020).

ScienceAdvances

Cellular sensing of extracellular purine nucleosides triggers an innate IFN-β response

Rekha Dhanwani, Mariko Takahashi, Ian T. Mathews, Camille Lenzi, Artem Romanov, Jeramie D. Watrous, Bartijn Pieters, Catherine C. Hedrick, Chris A. Benedict, Joel Linden, Roland Nilsson, Mohit Jain and Sonia Sharma

Sci Adv **6** (30), eaba3688.
DOI: 10.1126/sciadv.aba3688

IMMUNOLOGY

The scent of atherosclerosis

Vascular macrophages sense an odorant to induce atherosclerotic plaque formation

By **Katey J. Rayner**[1,2] *and* **Adil Rasheed**[1,2]

In the blood vessel wall, atherosclerotic plaque development (atherogenesis) occurs over several decades and involves the coordinated dysfunction of multiple cell types and organ systems. Although treatments targeting hypercholesterolemia and hypertension, the major risk factors of atherosclerosis, are effective at preventing fatal complications such as heart attack and stroke, there remains a major burden of atherosclerotic disease. Therefore, a more comprehensive understanding of the cells that drive disease, and how to therapeutically target them, is urgently needed. On page 214 of this issue, Orecchioni *et al.* (*1*) reveal an important role for olfactory receptor 2 (OLFR2) in atherosclerotic plaque development, identifying an additional target for intervention during the progression of atherosclerosis.

Olfactory receptors are a structurally diverse family of proteins that sense a wide range of odorants. The human genome comprises more than 400 olfactory receptors, which are expressed on the surface of olfactory sensory neurons in the nasal epithelium (*2*, *3*). However, olfactory receptors are also expressed in sperm, skin, and the intestine, suggesting that they play a role outside of the olfactory system. Previous work identified the expression of *Olfr2* and its downstream signaling components in mouse vascular macrophages during atherosclerosis (*4*). Orecchioni *et al.* extend these observations and demonstrate that octanal, the endogenous ligand of the olfactory receptors, triggers calcium-dependent induction of the NLRP3 (NACHT, LRR- and PYD domains-containing protein 3) inflammasome and the production of proinflammatory interleukin-1β (IL-1β) to exacerbate atherosclerosis (see the figure).

Macrophages are present and maintain homeostasis in virtually every tissue, including the arterial wall, by promoting host defense and clearing deposited lipids and apoptotic cells. However, when these functions become disabled or overwhelmed, macrophages directly promote atherogenesis. The secretion of IL-1β by macrophages in the atherosclerotic plaque is a key process during atherosclerosis (*5*, *6*), and there has been increasing clinical interest in blocking IL-1β in patients using monoclonal antibodies (*7*). However, there are numerous issues associated with inactivating IL-1β, including increased risk of death from opportunistic infections. Given that octanal and OLFR2 act upstream of IL-1β, this could represent a therapeutic opportunity to block lipid-driven inflammation in the vascular wall without disarming host defense.

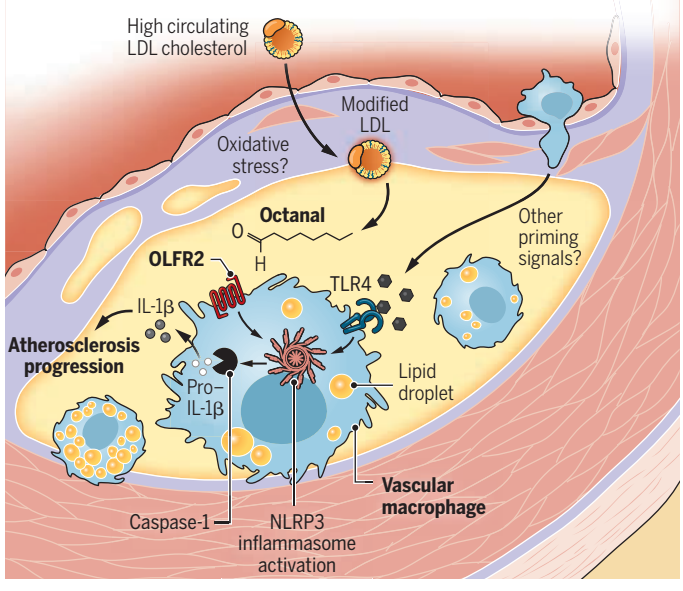
Octanal is a product of the peroxidation of fatty acids and occurs primarily as a result of oxidative stress in local vascular cells in the arterial wall (*8*). Orecchioni *et al.* posit that octanal is derived from oxidized low-density lipoprotein (oxLDL), a modified form of LDL cholesterol, the so-called "bad" cholesterol that directly accumulates in the vessel wall and causes atherosclerosis (*5*, *6*). The authors show that octanal signals through OLFR2 in macrophages, triggering downstream inflammation. Using aortic explants from mice with apolipoprotein E (*Apoe*) deletion (which increases production of cholesterol and atherogenic accumulation within macrophages), they showed that octanal could be produced by the aorta from oleic acid, a fatty acid prevalent in our diets. Notably, in mouse models of atherosclerosis and in humans, the circulating concentrations of octanal (~2 μM) are similar, suggesting that these mechanisms may translate to human disease. However, the precise source of octanal in the vessel wall is unknown.

Orecchioni *et al.* demonstrate that octanal is not sufficient to induce the activation of IL-1β by macrophages and that this induction of IL-1β only occurs with concomitant Toll-like receptor 4 (TLR4) agonism. TLR4 senses extracellular inflammatory molecules, leading to the transcriptional induction of NLRP3. This requirement for TLR4 ligation prior to octanal stimulation is reminiscent of previous studies demonstrating the two-hit activation of the inflammasome—priming and then activation—which is first driven by oxLDL engagement of TLR4 followed by signaling induced by cholesterol crystals that accumulate below the endothelial lining of the vessel wall during atherosclerosis (*9*). This raises the possibility that additional TLR priming, such as by other atherogenic ligands or by pathogen-associated molecular patterns (PAMPs), may act as a priming signal to activate the inflammasome in vascular macrophages. Thus, this study provides further evidence of the interplay of lipids and inflammation that occurs in macrophages and the role of macrophages as major immunometabolic hubs in atherosclerosis and other cardiometabolic diseases, including diabetes and obesity (*10*).

257

Atherogenesis involving octanal

Macrophages in the arterial wall are a hub of atherogenesis. Toll-like receptor 4 (TLR4) activation primes the NLRP3 (NACHT, LRR- and PYD domains-containing protein 3) inflammasome, which is fully activated by olfactory receptor 2 (OLFR2). Octanal, the OLFR2 ligand, is produced from oxidized low-density lipoprotein (oxLDL) cholesterol. This two-hit NLRP3 activation leads to interleukin-1β (IL-1β) production and exacerbation of atherosclerosis.



GRAPHIC: KELLIE HOLOSKI/*SCIENCE*

[1]University of Ottawa Heart Institute, Ottawa, ON, Canada. [2]Department of Biochemistry, Microbiology and Immunology, Faculty of Medicine, University of Ottawa, Ottawa, ON, Canada. Email: krayner@ottawaheart.ca

OLFR2 is expressed by multiple cell types within the aorta but was found to be most highly expressed in vascular macrophages (*1*, *4*). Although only ~30% of vascular macrophages expressed OLFR2 in mice, deletion of *Olfr2* or its downstream signaling mediators *Adcy3* (adenylate cyclase 3) and *Rtp1* and *Rtp2* (receptor transporter proteins 1 and 2) in hematopoietic cells led to a reduction in markers of plaque vulnerability (e.g., necrotic core in the vessel wall) that predispose to fatal complications, and an increase in collagen content (linked to plaque stability), indicating the potent role of OLFR2 signaling in macrophages on the atherosclerotic plaque. Moreover, despite the relatively low frequency of OLFR2$^+$ vascular macrophages, delivery of exogenous octanal to atherosclerosis-prone *Apoe*-deficient mice led to about a twofold increase in plaque area. Together, these data indicate that octanal may signal through other olfactory receptors expressed by lesional cells and/or that OLFR2-expressing macrophages release proatherogenic signals that cross-talk with other cells in the plaque. This study, along with high-resolution single-cell sequencing of atherosclerotic plaques (*11*), further contributes to our understanding of the function of macrophage subpopulations and their influence on atherogenesis.

There remain some key unknowns regarding how octanal and its receptors function in the vessel wall. Is there another role for octanal, in addition to promoting inflammasome activation? It is possible that, similar to how odorants act as attraction or repulsion signals in the olfactory epithelium or sperm cells, octanal acts as a chemoattractant in the vessel wall. Are there additional sources of octanal, perhaps derived from diet or the environment? As large-scale omics studies continue to be used to identify additional genes associated with risk of cardiovascular-related death, and as their functions in the vessel wall are characterized, there is an opportunity to better understand atherosclerosis progression and derive new treatments. ■

REFERENCES AND NOTES

1. M. Orecchioni *et al.*, *Science* **375**, 214 (2022).
2. C. Trimmer *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **116**, 9475 (2019).
3. B. Malnic *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **101**, 2584 (2004).
4. S. McArdle *et al.*, *Circ. Res.* **125**, 1038 (2019).
5. K. J. Moore *et al.*, *Nat. Rev. Immunol.* **13**, 709 (2013).
6. P. Libby, *Nature* **592**, 524 (2021).
7. P. M. Ridker *et al.*, *N. Engl. J. Med.* **377**, 1119 (2017).
8. S. Janfaza *et al.*, *Biol. Methods Protoc.* **4**, bpz014 (2019).
9. A. Grebe *et al.*, *Circ. Res.* **122**, 1722 (2018).
10. A. Rasheed, K. J. Rayner, *Endocr. Rev.* **42**, 407 (2021).
11. A. Zernecke *et al.*, *Circ. Res.* **127**, 402 (2020).

ACKNOWLEDGMENTS

A.R. is supported by the University of Ottawa Cardiac Endowment Fund, and K.J.R. is supported by the Canadian Institutes for Health Research.

10.1126/science.abn4708

PLANETARY SCIENCE

Molten iron in Earth-like exoplanet cores

Iron crystallization in super-Earth interiors plays a key role in their habitability

By **Youjun Zhang**[1,2] *and* **Jung-Fu Lin**[3]

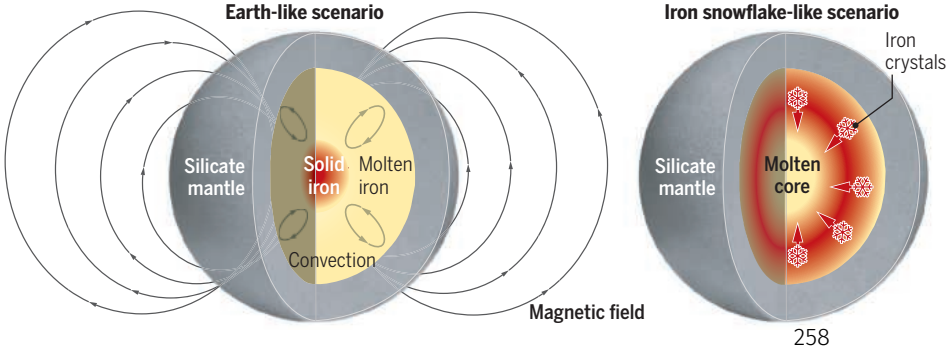
Earth, the only known habitable planet in the Universe, has a magnetic field that shields organic life-forms from harmful radiation coming from the Sun and beyond. This magnetic field is generated by the churning of molten iron in its outer core. The habitability of exoplanets orbiting other stars could be gleaned through better understanding of their iron cores and magnetic fields (*1*). However, extreme pressure and temperature conditions inside exoplanets that are much heavier than Earth may mean that their cores behave differently. On page 202 of this issue, Kraus *et al.* (*2*) used a powerful laser to generate conditions similar to those inside the cores of such "super-Earths" and reveal that even under extreme conditions, molten iron can crystallize similarly to that found at the base of Earth's outer core.

To date, more than 4500 exoplanets have been discovered, with approximately one-third of them categorized as Earth-like exoplanets (*3*). The discoveries of these exoplanets have raised hopes about finding habitable conditions beyond the Solar System and that exoplanetary habitability could be quite diverse in the Universe. Although surface water in a star's habitable zone has always been used as a qualifying condition for habitability, other key factors for habitability lie beneath the surface of the exoplanet, such as the property of its dynamo, a self-sustaining mechanism that generates a magnetic field (*4*).

Similar to Earth, super-Earths are thought to have formed through collisions and then differentiated into light silicate mantles and heavy iron cores. The iron cores were initially hot and molten but slowly lost heat to the silicate mantles. If core cooling is efficient, it can lead to iron crystallization, which releases energy. The cooling and solidification processes are thought to be the main sources of power that drives the convection of molten iron in the liquid core, generating magnetic fields through dynamo action, also known as magnetospheres. The pressure-temperature condition in which convection occurs is close to adiabatic, meaning that hot upwelling fluid follows a predictable temperature profile without heat gain or loss to the surroundings. Depending on the intersection relation between the iron melting temperature and the adiabatic profile under compression in a super-Earth's core, the molten cores can crystallize in two possible scenarios: either in an Earth-like "bottom-up" iron crystallization scenario or in an iron snowflake–like "top-down" scenario (see the figure). Bottom-up crystallization happens in the case of an iron melting curve steeper than the adiabatic pro-

Iron crystallization in super-Earth cores

Exoplanets with an Earth-like iron crystallization in their cores are more likely to possess and sustain a magnetic shield necessary for organic life-forms to exist. However, exoplanets with a higher content of light elements in the core may not have the internal condition necessary to sustain a solid core in the center and subsequently to sustain a magnetic shield over a long period.



GRAPHIC: N. DESAI/SCIENCE

The scent of atherosclerosis

Katey J. RaynerAdil Rasheed

Science, 375 (6577),

Methods in Molecular Biology 2419

Springer Protocols

Dipak P. Ramji *Editor*

Atherosclerosis

Methods and Protocols

Humana Press

261

METHODS IN MOLECULAR BIOLOGY

Series Editor
John M. Walker
School of Life and Medical Sciences
University of Hertfordshire
Hatfield, Hertfordshire, UK

For further volumes:
http://www.springer.com/series/7651

For over 35 years, biological scientists have come to rely on the research protocols and methodologies in the critically acclaimed *Methods in Molecular Biology* series. The series was the first to introduce the step-by-step protocols approach that has become the standard in all biomedical protocol publishing. Each protocol is provided in readily-reproducible step-by-step fashion, opening with an introductory overview, a list of the materials and reagents needed to complete the experiment, and followed by a detailed procedure that is supported with a helpful notes section offering tips and tricks of the trade as well as troubleshooting advice. These hallmark features were introduced by series editor Dr. John Walker and constitute the key ingredient in each and every volume of the *Methods in Molecular Biology* series. Tested and trusted, comprehensive and reliable, all protocols from the series are indexed in PubMed.

Atherosclerosis

Methods and Protocols

Edited by

Dipak P. Ramji

Cardiff School of Biosciences, Cardiff University, Cardiff, UK

Humana Press

Editor
Dipak P. Ramji
Cardiff School of Biosciences
Cardiff University
Cardiff, UK

ISSN 1064-3745 ISSN 1940-6029 (electronic)
Methods in Molecular Biology
ISBN 978-1-0716-1923-0 ISBN 978-1-0716-1924-7 (eBook)
https://doi.org/10.1007/978-1-0716-1924-7

This Humana imprint is published by the registered company Springer Science+Business Media, LLC, part of Springer Nature.
The registered company address is: 1 New York Plaza, New York, NY 10004, U.S.A.

Dedication

This book is dedicated to Professor Sampath Parthasarathy who passed away on 1 December 2020. Prof. Parthasarathy published over 250 articles on atherosclerosis and other diseases. His research was instrumental in the development of a key concept that oxidized low-density lipoprotein is involved in the initiation and progression of atherosclerosis. The world has lost not only a dedicated scientist of great intellect who made tremendous contributions in the field of cardiovascular research but also a kind and considerate human being.

Preface

Atherosclerosis and its clinical complications, such as myocardial infarction and cerebrovascular accident, are the major cause of global morbidity and mortality. The disease is an enormous economic burden, and it is anticipated that this will worsen in the future because of the prevalence of risk factors such as hypercholesterolemia, hypertension, diet rich in saturated fats, sedentary lifestyle, obesity, and diabetes. Current therapies are not fully effective and associated with various adverse side effects. It is therefore essential to increase our understanding of the molecular basis of atherosclerosis as this will not only reveal novel targets for therapeutic intervention but ultimately improve diagnosis, prevention, and treatment of this disease. The use of *in vitro* and *in vivo* model systems coupled with techniques that probe atherosclerosis-associated processes in such models has been instrumental in advancing our knowledge of the genetic, molecular, and cellular mechanisms underlying the disease and will continue to do so in the future.

This book provides a comprehensive and up-to-date coverage of methods used in atherosclerosis research. The first part of the book addresses the pathogenesis of atherosclerosis, model systems employed to probe the disease, and emerging concepts in the field. This is followed by in-depth coverage of *in vitro and in vivo* model systems together with current and emerging methods in the field. Each methods chapter includes introduction to the topic, list of materials and reagents required, detailed step-by-step laboratory protocols, and invaluable notes based on the authors' personal experiences. The book will be an invaluable resource for those engaged in research on atherosclerosis and cardiovascular disease, those that are planning to move into the field, those working on other inflammatory disorders that predispose to atherosclerosis (e.g., obesity, diabetes, arthritis, etc.), and students at both the undergraduate and postgraduate levels.

Cardiff, UK ***Dipak P. Ramji***

Contents

Part III In Vivo Model Systems for Atherosclerosis Research

Contributors

Ali Abdelfattah • *Department of Medical Laboratory Sciences, Faculty of Applied Medical Sciences, Hashemite University, Zarqa, Jordan*

Aaron D. Aguirre • *Wellman Center for Photomedicine, Harvard Medical School and Massachusetts General Hospital, Boston, MA, USA; Cardiovascular Research Center, Division of Cardiology, Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA*

Ramzi Ajjan • *Leeds Institute of Cellular & Metabolic Medicine, Faculty of Medicine & Health, University of Leeds, Leeds, UK*

Sulaiman Al Alawi • *Cardiff School of Biosciences, Cardiff University, Cardiff, UK*

Ahmed Al-Aufi • *Leeds Vascular Institute, Faculty of Medicine & Health, University of Leeds, Leeds, UK*

Basmah Al-Sayejh • *Leeds Institute of Cellular & Metabolic Medicine, Faculty of Medicine & Health, University of Leeds, Leeds, UK*

Alaa Alahmadi • *Cardiff School of Biosciences, Cardiff University, Cardiff, UK*

Mazen S. Albaghdadi • *Cardiovascular Research Center, Division of Cardiology, Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA*

Alhomidi Almotiri • *College of Applied Medical Sciences-Dawadmi, Shaqra University, Dawadmi, Saudi Arabia*

Reem Alotibi • *Cardiff School of Biosciences, Cardiff University, Cardiff, UK*

Fahad Alradi • *Cardiff School of Biosciences, Cardiff University, Cardiff, UK*

Nouf Alshehri • *Cardiff School of Biosciences, Cardiff University, Cardiff, UK*

Chandrakala Aluganti Narasimhulu • *Burnett School of Biomedical Sciences, College of Medicine, University of Central Florida, Orlando, FL, USA*

G. M. Anantharamaiah • *Department of Medicine, UAB Medical Centre, Birmingham, AL, USA*

Vicente Andrés • *Laboratory of Molecular and Genetic Cardiovascular Pathophysiology, Vascular Biology Area, Centro Nacional de Investigaciones Cardiovasculares (CNIC), Madrid, Spain; Centro de Investigación Biomédica en Red de Enfermedades Cardiovasculares (CIBERCV), Madrid, Spain*

Ananth Annapragada • *E.B. Singleton Department of Radiology, Texas Children's Hospital & Baylor College of Medicine, Houston, TX, USA*

Andrew Badachhape • *E.B. Singleton Department of Radiology, Texas Children's Hospital & Baylor College of Medicine, Houston, TX, USA*

Andrew H. Baker • *BHF Centre for Cardiovascular Sciences (CVS), University of Edinburgh, Edinburgh, UK; Department of Pathology, Maastricht University Medical Center (MUMC), Cardiovascular Research Institute Maastricht (CARIM), Maastricht, The Netherlands*

Ana Barettino • *Laboratory of Molecular and Genetic Cardiovascular Pathophysiology, Vascular Biology Area, Centro Nacional de Investigaciones Cardiovasculares (CNIC), Madrid, Spain; Centro de Investigación Biomédica en Red de Enfermedades Cardiovasculares (CIBERCV), Madrid, Spain*

Tobias Becher • *Laboratory of Molecular Metabolism, The Rockefeller University, New York, NY, USA; First Department of Medicine (Division of Cardiology), University Medical*

Center Mannheim, Mannheim, Germany; DZHK (German Centre for Cardiovascular Research), Partner Site Heidelberg/Mannheim, Mannheim, Germany

Ignacio Benedicto • *Laboratory of Molecular and Genetic Cardiovascular Pathophysiology, Vascular Biology Area, Centro Nacional de Investigaciones Cardiovasculares (CNIC), Madrid, Spain*

Jacob F. Bentzon • *Department of Clinical Medicine, Heart Diseases, Aarhus University, Aarhus, Denmark; Centro Nacional de Investigaciones Cardiovasculares (CNIC), Madrid, Spain*

Juan A. Bernal • *Centro Nacional de Investigaciones Cardiovasculares (CNIC), Madrid, Spain*

Rohan Bhavane • *E.B. Singleton Department of Radiology, Texas Children's Hospital & Baylor College of Medicine, Houston, TX, USA*

Martin Mæng Bjørklund • *Department of Clinical Medicine, Heart Diseases, Aarhus University, Aarhus, Denmark; Centro Nacional de Investigaciones Cardiovasculares (CNIC), Madrid, Spain*

Bethan A. Brown • *Department of Translational Health Sciences, Bristol Medical School, University of Bristol, Bristol, UK*

Úrsula Catalán Santos • *Facultat de Medicina i Ciències de la Salut, Departament de Medicina i Cirurgia, Grup Nutrició Funcional, Oxidació i Malalties Cardiovasculars (NFOC-Salut), Universitat Rovira i Virgili, Reus, Spain; Eurecat, Centre Tecnològic de Catalunya, Unitat de Nutrició i Salut, Reus, Spain*

Yee-Hung Chan • *Cardiff School of Biosciences, Cardiff University, Cardiff, UK*

Jing Chen • *Cardiff School of Biosciences, Cardiff University, Cardiff, UK*

Yajie Chen • *Department of Molecular Pathology, Faculty of Medicine, Interdisciplinary Graduate School of Medicine, University of Yamanashi, Chuo, Japan; School of Biotechnology and Health Sciences, Wuyi University, Jiangmen, China*

Jingyi Chi • *Laboratory of Molecular Metabolism, The Rockefeller University, New York, NY, USA*

Mohammed M. Chowdhury • *Cardiovascular Research Center, Division of Cardiology, Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA; Division of Vascular Surgery, Department of Surgery, Addenbrooke's Hospital, University of Cambridge, Cambridge, UK*

Paul Cohen • *Laboratory of Molecular Metabolism, The Rockefeller University, New York, NY, USA*

Patrick A. Coughlin • *Division of Vascular Surgery, Department of Surgery, Addenbrooke's Hospital, University of Cambridge, Cambridge, UK*

Stephen J. Cross • *Department of Translational Health Sciences, Bristol Medical School, University of Bristol, Bristol, UK*

Gary A. Cuthbert • *Leeds Vascular Institute, Faculty of Medicine & Health, University of Leeds, Leeds, UK*

Thomas Q. de Aguiar Vallim • *Department of Medicine, Cardiology, University of California Los Angeles, Los Angeles, CA, USA; Department of Biological Chemistry, David Geffen School of Medicine at UCLA, University of California Los Angeles, Los Angeles, CA, USA; Molecular Biology Institute, David Geffen School of Medicine at UCLA, University of California Los Angeles, Los Angeles, CA, USA; Jonsson Comprehensive Cancer Center, David Geffen School of Medicine at UCLA, University of California Los Angeles, Los Angeles, CA, USA*

Marco De Giorgi • *Department of Molecular Physiology and Biophysics, Baylor College of Medicine, Houston, TX, USA*
Alba de Juan • *Institut Curie, PSL Research University, INSERM, U932, Paris, France*
Anuradha Dhanasekaran • *Centre for Biotechnology, Anna University, Chennai, Tamil Nadu, India*
Charles Drummer IV • *Cardiovascular Research Center, Lewis Katz School of Medicine at Temple University, Philadelphia, PA, USA*
Christopher P. Durant • *La Jolla Institute for Immunology, La Jolla, CA, USA*
Ali Ertürk • *Institute for Stroke and Dementia Research, Klinikum der Universität München, Ludwig-Maximilians-University (LMU), Munich, Germany; Helmholtz Zentrum München, Deutsches Forschungszentrum für Gesundheit und Umwelt, Neuherberg, Germany*
Joan Carles Escolà-Gil • *Institut de Recerca de l'Hospital de la Santa Creu i Sant Pau, Institut d'Investigacions Biomèdiques (IIB) Sant Pau, Barcelona, Spain; CIBER de Diabetes y Enfermedades Metabólicas Asociadas, CIBERDEM, Madrid, Spain; Departament de Bioquímica i Biologia Molecular, Universitat Autònoma de Barcelona, Barcelona, Spain*
Jianglin Fan • *Department of Molecular Pathology, Faculty of Medicine, Interdisciplinary Graduate School of Medicine, University of Yamanashi, Chuo, Japan; School of Biotechnology and Health Sciences, Wuyi University, Jiangmen, China*
Sara Fernández-Castillejo • *Facultat de Medicina i Ciències de la Salut, Departament de Medicina i Cirurgia, Grup Nutrició Funcional, Oxidació i Malalties Cardiovasculars (NFOC-Salut), Universitat Rovira i Virgili, Reus, Spain; Eurecat, Centre Tecnològic de Catalunya, Unitat de Nutrició i Salut, Reus, Spain*
Benjamin Förstera • *Institute for Stroke and Dementia Research, Klinikum der Universität München, Ludwig-Maximilians-University (LMU), Munich, Germany; Helmholtz Zentrum München, Deutsches Forschungszentrum für Gesundheit und Umwelt, Neuherberg, Germany*
Sarah J. George • *Department of Translational Health Sciences, Bristol Medical School, University of Bristol, Bristol, UK*
Godfrey S. Getz • *Department of Pathology, The University of Chicago, Chicago, IL, USA*
Ketan B. Ghaghada • *E.B. Singleton Department of Radiology, Texas Children's Hospital & Baylor College of Medicine, Houston, TX, USA*
Kayleigh Griffiths • *Institute of Cardiovascular Sciences, College of Medical and Dental Sciences, University of Birmingham, Birmingham, UK*
Andreas Habenicht • *Institute for Cardiovascular Prevention (IPEK), Klinikum of the University of Munich (KUM), Ludwig-Maximilians-University (LMU), Munich, Germany; German Center for Cardiovascular Research (DZHK), Partner site Munich Heart Alliance, Munich, Germany*
Magda R. Hamczyk • *Centro de Investigación Biomédica en Red de Enfermedades Cardiovasculares (CIBERCV), Madrid, Spain; Departamento de Bioquímica y Biología Molecular, Instituto Universitario de Oncología (IUOPA), Universidad de Oviedo, Oviedo, Spain*
Michael A. Harrison • *School of Biomedical Sciences, University of Leeds, Leeds, UK*
Catherine C. Hedrick • *La Jolla Institute for Immunology, La Jolla, CA, USA*
Shervanthi Homer-Vanniasinkam • *Leeds Vascular Institute, Faculty of Medicine & Health, University of Leeds, Leeds, UK*
Alaa Ismail • *Cardiff School of Biosciences, Cardiff University, Cardiff, UK*

Farouc A. Jaffer • *Cardiovascular Research Center, Division of Cardiology, Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA; Wellman Center for Photomedicine, Harvard Medical School and Massachusetts General Hospital, Boston, MA, USA*

Philip E. James • *Cardiff School of Sport and Health Sciences, Cardiff Metropolitan University, Cardiff, UK*

Kelsey E. Jarrett • *Department of Medicine, Cardiology, University of California Los Angeles, Los Angeles, CA, USA*

Xiaohua Jiang • *Cardiovascular Research Center, Lewis Katz School of Medicine at Temple University, Philadelphia, PA, USA; Metabolic Disease Research, Departments of Pharmacology, Microbiology and Immunology, Lewis Katz School of Medicine at Temple University, Philadelphia, PA, USA*

Jason L. Johnson • *Department of Translational Health Sciences, Bristol Medical School, University of Bristol, Bristol, UK*

Josep Julve • *Institut de Recerca de l'Hospital de la Santa Creu i Sant Pau, Institut d'Investigacions Biomèdiques (IIB) Sant Pau, Barcelona, Spain; CIBER de Diabetes y Enfermedades Metabólicas Asociadas, CIBERDEM, Madrid, Spain; Departament de Bioquímica i Biologia Molecular, Universitat Autònoma de Barcelona, Barcelona, Spain*

Mohamad B. Kassab • *Cardiovascular Research Center, Division of Cardiology, Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA*

Yoichiro Kawamura • *Cardiovascular Research Center, Division of Cardiology, Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA*

Katie Lacey • *School of Molecular & Cellular Biology, University of Leeds, Leeds, UK*

William R. Lagor • *Department of Molecular Physiology and Biophysics, Baylor College of Medicine, Houston, TX, USA*

Begoña Lavin Plaza • *School of Biomedical Engineering and Imaging Sciences, King's College London, London, UK; Department of Biochemistry and Molecular Biology, School of Chemistry, Complutense University, Madrid, Spain*

Klaus Ley • *La Jolla Institute for Immunology, La Jolla, CA, USA; Department of Bioengineering, University of California San Diego, La Jolla, CA, USA*

Yuanfang Li • *Institute for Cardiovascular Prevention (IPEK), Klinikum of the University of Munich (KUM), Ludwig-Maximilians-University (LMU), Munich, Germany*

Yutao Li • *Institute for Cardiovascular Prevention (IPEK), Klinikum of the University of Munich (KUM), Ludwig-Maximilians-University (LMU), Munich, Germany*

Jonathan R. Lindner • *Oregon Health & Science University, Portland, OR, USA*

Enqi Liu • *Research Institute of Atherosclerotic Disease and Laboratory Animal Center, Xi'an Jiaotong University School of Medicine, Xi'an, China*

George Liu • *Institute of Cardiovascular Sciences, Key Laboratory of Molecular Cardiovascular Sciences, Ministry of Education, School of Basic Medical Sciences, Peking University, Beijing, China*

Ronan Lordan • *Department of Biological Sciences, University of Limerick, Limerick, Ireland; Health Research Institute, University of Limerick, Limerick, Ireland*

Shu Lu • *Institute for Cardiovascular Prevention (IPEK), Klinikum of the University of Munich (KUM), Ludwig-Maximilians-University (LMU), Munich, Germany*

Yifan Lu • *Cardiovascular Research Center, Lewis Katz School of Medicine at Temple University, Philadelphia, PA, USA*

Melanie Madhani • *Institute of Cardiovascular Sciences, College of Medical and Dental Sciences, University of Birmingham, Birmingham, UK*

Manousos Makridakis • *Center of Systems Biology, Biomedical Research Foundation, Academy of Athens (BRFAA), Athens, Greece*
Keiichiro Matoba • *Division of Diabetes, Endocrinology and Metabolism, Department of Internal Medicine, The Jikei University School of Medicine, Tokyo, Japan*
Remco T. A. Megens • *Institute for Cardiovascular Prevention (IPEK), Klinikum of the University of Munich (KUM), Ludwig-Maximilians-University (LMU), Munich, Germany; Cardiovascular Research Institute Maastricht (CARIM), Maastricht University, Maastricht, The Netherlands*
Juan Bautista Menendez-Gonzalez • *Department of Stem Cell and Regenerative Biology, Harvard Stem Cell Institute, Harvard University, Cambridge, MA, USA; Center for Regenerative Medicine, Massachusetts General Hospital, Boston, MA, USA*
Melissa A. Meyer • *La Jolla Institute for Immunology, La Jolla, CA, USA*
Sarajo K. Mohanta • *Institute for Cardiovascular Prevention (IPEK), Klinikum of the University of Munich (KUM), Ludwig-Maximilians-University (LMU), Munich, Germany; German Center for Cardiovascular Research (DZHK), Partner site Munich Heart Alliance, Munich, Germany; Munich Cluster of Systems Neurology (SyNergy), Ludwig-Maximilians-University Munich, Munich, Germany*
Claude Motta • *Facultat de Medicina i Ciències de la Salut, Departament de Medicina i Cirurgia, Grup Nutrició Funcional, Oxidació i Malalties Cardiovasculars (NFOC-Salut), Universitat Rovira i Virgili, Reus, Spain*
Jan Nagenborg • *Pathology Department, Cardiovascular Research Institute Maastricht (CARIM), Maastricht University Medical Center (MUMC), Maastricht, The Netherlands*
Jamie Nash • *Cardiff School of Sport and Health Sciences, Cardiff Metropolitan University, Cardiff, UK*
Rosa M. Nevado • *Centro Nacional de Investigaciones Cardiovasculares (CNIC), Madrid, Spain; Centro de Investigación Biomédica en Red de Enfermedades Cardiovasculares (CIBERCV), Madrid, Spain*
The Anh Nguyen • *Oregon Health & Science University, Portland, OR, USA*
Manabu Niimi • *Department of Molecular Pathology, Faculty of Medicine, Interdisciplinary Graduate School of Medicine, University of Yamanashi, Chuo, Japan*
Marco Orecchioni • *La Jolla Institute for Immunology, La Jolla, CA, USA*
Mayakonda Palgunachari • *Department of Medicine, UAB Medical Centre, Birmingham, AL, USA*
Sampath Parthasarathy • *Burnett School of Biomedical Sciences, College of Medicine, University of Central Florida, Orlando, FL, USA*
Anna Pedret Figuerola • *Facultat de Medicina i Ciències de la Salut, Departament de Medicina i Cirurgia, Grup Nutrició Funcional, Oxidació i Malalties Cardiovasculars (NFOC-Salut), Universitat Rovira i Virgili, Reus, Spain; Eurecat, Centre Tecnològic de Catalunya, Unitat de Nutrició i Salut, Reus, Spain*
Zhonglie Piao • *Wellman Center for Photomedicine, Harvard Medical School and Massachusetts General Hospital, Boston, MA, USA*
Sreenivasan Ponnambalam • *School of Molecular & Cellular Biology, University of Leeds, Leeds, UK*
Dipak P. Ramji • *Cardiff School of Biosciences, Cardiff University, Cardiff, UK*
Catherine A. Reardon • *Ben May Department, The University of Chicago, Chicago, IL, USA*
Lien M. Reolizo • *Department of Translational Health Sciences, Bristol Medical School, University of Bristol, Bristol, UK*

Dario F. Riascos-Bernal • *Department of Medicine (Cardiology Division), Wilf Family Cardiovascular Research Institute, Albert Einstein College of Medicine, Bronx, NY, USA; Department of Developmental and Molecular Biology, Wilf Family Cardiovascular Research Institute, Albert Einstein College of Medicine, Bronx, NY, USA*

Andrea Rivas-Urbina • *Institut de Recerca de l'Hospital de la Santa Creu i Sant Pau, Institut d'Investigacions Biomèdiques (IIB) Sant Pau, Barcelona, Spain*

Neil P. Rodrigues • *European Cancer Stem Cell Research Institute, School of Biosciences, Cardiff University, Cardiff, UK*

Anne Roivainen • *Turku PET Centre, University of Turku and Turku University Hospital, Turku, Finland; Turku Center for Disease Modeling, University of Turku, Turku, Finland*

Barnaby W. R. Roper • *School of Molecular & Cellular Biology, University of Leeds, Leeds, UK*

Noemi Rotllan • *Institut de Recerca de l'Hospital de la Santa Creu i Sant Pau, Institut d'Investigacions Biomèdiques (IIB) Sant Pau, Barcelona, Spain; CIBER de Diabetes y Enfermedades Metabólicas Asociadas, CIBERDEM, Madrid, Spain*

James H. F. Rudd • *Division of Cardiovascular Medicine, Addenbrooke's Hospital, University of Cambridge, Cambridge, UK*

Fatma Saaoud • *Cardiovascular Research Center, Lewis Katz School of Medicine at Temple University, Philadelphia, PA, USA*

Ryosuke Saigusa • *La Jolla Institute for Immunology, La Jolla, CA, USA*

Jose Luis Sanchez-Quesada • *Institut de Recerca de l'Hospital de la Santa Creu i Sant Pau, Institut d'Investigacions Biomèdiques (IIB) Sant Pau, Barcelona, Spain; CIBER de Diabetes y Enfermedades Metabólicas Asociadas, CIBERDEM, Madrid, Spain; Departament de Bioquímica i Biologia Molecular, Universitat Autònoma de Barcelona, Barcelona, Spain*

David Santos • *Institut de Recerca de l'Hospital de la Santa Creu i Sant Pau, Institut d'Investigacions Biomèdiques (IIB) Sant Pau, Barcelona, Spain; CIBER de Diabetes y Enfermedades Metabólicas Asociadas, CIBERDEM, Madrid, Spain*

Antti Saraste • *Turku PET Centre, University of Turku and Turku University Hospital, Turku, Finland; Heart Center, Turku University Hospital and University of Turku, Turku, Finland*

Jessica P. Scanlon • *BHF Centre for Cardiovascular Sciences (CVS), University of Edinburgh, Edinburgh, UK*

Ying Shao • *Cardiovascular Research Center, Lewis Katz School of Medicine at Temple University, Philadelphia, PA, USA*

Nicholas E. S. Sibinga • *Department of Medicine (Cardiology Division), Wilf Family Cardiovascular Research Institute, Albert Einstein College of Medicine, Bronx, NY, USA; Department of Developmental and Molecular Biology, Wilf Family Cardiovascular Research Institute, Albert Einstein College of Medicine, Bronx, NY, USA*

Judith C. Sluimer • *Pathology Department, Cardiovascular Research Institute Maastricht (CARIM), Maastricht University Medical Center (MUMC), Maastricht, The Netherlands; BHF Centre for Cardiovascular Sciences (CVS), University of Edinburgh, Edinburgh, UK*

Rosa Solà • *Facultat de Medicina i Ciències de la Salut, Departament de Medicina i Cirurgia, Grup Nutrició Funcional, Oxidació i Malalties Cardiovasculars (NFOC-Salut), Universitat Rovira i Virgili, Reus, Spain; Hospital Universitari Sant Joan de Reus, Reus, Spain*

Alexandros Somos • *Department of Translational Health Sciences, Bristol Medical School, University of Bristol, Bristol, UK*
Mia Ståhle • *Turku PET Centre, University of Turku and Turku University Hospital, Turku, Finland*
Karen Stanic • *Institute for Stroke and Dementia Research, Klinikum der Universität München, Ludwig-Maximilians-University (LMU), Munich, Germany; Helmholtz Zentrum München, Deutsches Forschungszentrum für Gesundheit und Umwelt, Neuherberg, Germany*
Sabine Steffens • *Institute for Cardiovascular Prevention (IPEK), Klinikum of the University of Munich (KUM), Ludwig-Maximilians-University (LMU), Munich, Germany; German Center for Cardiovascular Research (DZHK), Partner site Munich Heart Alliance, Munich, Germany*
Jianxin Sun • *Department of Medicine, Sidney Kimmel Medical College, Thomas Jefferson University, Philadelphia, PA, USA*
Ting Sun • *Institute for Cardiovascular Prevention (IPEK), Klinikum of the University of Munich (KUM), Ludwig-Maximilians-University (LMU), Munich, Germany*
Yu Sun • *Cardiovascular Research Center, Lewis Katz School of Medicine at Temple University, Philadelphia, PA, USA*
Vasantika Suryawanshi • *La Jolla Institute for Immunology, La Jolla, CA, USA*
Ritsuko Suzuki • *Admission Center, University of Yamanashi, Kofu, Japan*
Eric Tanifum • *E.B. Singleton Department of Radiology, Texas Children's Hospital & Baylor College of Medicine, Houston, TX, USA*
Guillermo J. Tearney • *Wellman Center for Photomedicine, Harvard Medical School and Massachusetts General Hospital, Boston, MA, USA; Department of Pathology, Harvard Medical School and Massachusetts General Hospital, Boston, MA, USA*
Darren C. Tomlinson • *School of Molecular & Cellular Biology, University of Leeds, Leeds, UK*
Alexandros Tsoupras • *Department of Biological Sciences, University of Limerick, Limerick, Ireland; Health Research Institute, University of Limerick, Limerick, Ireland; Bernal Institute, University of Limerick, Limerick, Ireland*
Kim van Kuijk • *Pathology Department, Cardiovascular Research Institute Maastricht (CARIM), Maastricht University Medical Center (MUMC), Maastricht, The Netherlands; Institute of Experimental Medicine and Systems Biology, Faculty of Medicine, RWTH Aachen University, Aachent, Germany*
Selvaraj Vimalraj • *Centre for Biotechnology, Anna University, Chennai, Tamil Nadu, India*
Antonia Vlahou • *Center of Systems Biology, Biomedical Research Foundation, Academy of Athens (BRFAA), Athens, Greece*
Kerry S. Wadey • *Department of Translational Health Sciences, Bristol Medical School, University of Bristol, Bristol, UK*
Hong Wang • *Metabolic Disease Research, Departments of Pharmacology, Microbiology and Immunology, Lewis Katz School of Medicine at Temple University, Philadelphia, PA, USA*
Yuhui Wang • *Institute of Cardiovascular Sciences, Key Laboratory of Molecular Cardiovascular Sciences, Ministry of Education, School of Basic Medical Sciences, Peking University, Beijing, China*
Zhihua Wang • *Institute for Cardiovascular Prevention (IPEK), Klinikum of the University of Munich (KUM), Ludwig-Maximilians-University (LMU), Munich, Germany*

Christian Weber • *Institute for Cardiovascular Prevention (IPEK), Klinikum of the University of Munich (KUM), Ludwig-Maximilians-University (LMU), Munich, Germany; German Center for Cardiovascular Research (DZHK), Partner site Munich Heart Alliance, Munich, Germany; Cardiovascular Research Institute Maastricht (CARIM), Maastricht University, Maastricht, The Netherlands; Munich Cluster of Systems Neurology (SyNergy), Ludwig-Maximilians-University Munich, Munich, Germany*

Cass Whelan • *Cardiff School of Sport and Health Sciences, Cardiff Metropolitan University, Cardiff, UK*

C. Roger White • *Department of Medicine, UAB Medical Centre, Birmingham, AL, USA*

Helen Williams • *Department of Translational Health Sciences, Bristol Medical School, University of Bristol, Bristol, UK*

Jessica O. Williams • *Cardiff School of Sport and Health Sciences, Cardiff Metropolitan University, Cardiff, UK*

Georg Wissmeyer • *Cardiovascular Research Center, Division of Cardiology, Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA*

Paul Wolkowicz • *Department of Medicine, UAB Medical Centre, Birmingham, AL, USA*

Xunde Xian • *Institute of Cardiovascular Sciences, Key Laboratory of Molecular Cardiovascular Sciences, Ministry of Education, School of Basic Medical Sciences, Peking University, Beijing, China*

Keman Xu • *Cardiovascular Research Center, Lewis Katz School of Medicine at Temple University, Philadelphia, PA, USA*

Xiaofeng Yang • *Cardiovascular Research Center, Lewis Katz School of Medicine at Temple University, Philadelphia, PA, USA; Metabolic Disease Research, Departments of Pharmacology, Microbiology and Immunology, Lewis Katz School of Medicine at Temple University, Philadelphia, PA, USA; Centers for Cardiovascular Research, Inflammation & Lung Research, Metabolic Disease Research, and Thrombosis Research, Lewis Katz School of Medicine at Temple University, Philadelphia, PA, USA*

Changjun Yin • *Institute for Cardiovascular Prevention (IPEK), Klinikum of the University of Munich (KUM), Ludwig-Maximilians-University (LMU), Munich, Germany; German Center for Cardiovascular Research (DZHK), Partner site Munich Heart Alliance, Munich, Germany; Munich Cluster of Systems Neurology (SyNergy), Ludwig-Maximilians-University Munich, Munich, Germany*

Jun Yu • *Metabolic Disease Research, Departments of Pharmacology, Microbiology and Immunology, Lewis Katz School of Medicine at Temple University, Philadelphia, PA, USA*

Stephanie Yu • *Northeastern University, Boston, MA, USA*

Arif Yurdagul Jr • *Department of Molecular and Cellular Physiology, LSU Health Shreveport, Shreveport, LA, USA*

Ioannis Zabetakis • *Department of Biological Sciences, University of Limerick, Limerick, Ireland; Health Research Institute, University of Limerick, Limerick, Ireland*

Chuankai Zhang • *Institute for Cardiovascular Prevention (IPEK), Klinikum of the University of Munich (KUM), Ludwig-Maximilians-University (LMU), Munich, Germany*

Xi Zhang • *Institute for Cardiovascular Prevention (IPEK), Klinikum of the University of Munich (KUM), Ludwig-Maximilians-University (LMU), Munich, Germany*

Chapter 47

Flow Cytometry and Mass Cytometry for Measuring the Immune Cell Infiltrate in Atherosclerotic Arteries

Marco Orecchioni, Melissa A. Meyer, Catherine C. Hedrick, and Klaus Ley

Abstract

Atherosclerosis is characterized by the abundant infiltration of immune cells starting at early stages and progressing to late stages of the disease. The study and characterization of immune cells infiltrating and residing in the aorta has being tackled by several methodologies such as flow cytometry and mass cytometry (CyTOF). Flow cytometry has been primarily used to address the aortic leukocyte composition; however, only a limited number of markers can be analyzed simultaneously. CyTOF started to overcome these limitations by employing rare element-tagged antibodies and combines mass spectrometry with the ease and precision of flow cytometry. CyTOF currently allows for the simultaneous measurement of more than 40 cellular parameters at single-cell resolution.

In this chapter, we describe the methodology used to isolate single immune cells from mouse aortas, followed by protocols for flow cytometry and CyTOF for aortic immune cell characterization.

Key words Atherosclerosis, Aorta, Immune cells, Leukocytes, Flow cytometry, Mass cytometry, CyTOF

1 Introduction

Atherosclerosis is a chronic inflammatory disease of the arterial wall characterized by lipid deposition. Innate and adaptive immune cells infiltrate the aorta of atherosclerotic mice during the development of the disease [1]. Common animal models used in cardiovascular research include genetically modified mice with altered lipid levels such as apolipoprotein E gene knockout and low-density lipoprotein receptor gene knockout mice [2].

In early studies, the immune cell infiltrate was characterized by immunohistochemistry [3]. Later, methods were developed to generate single-cell suspensions from the mouse aorta for flow cytometry analysis [4]. Recently, improvement in flow cytometry

Marco Orecchioni and Melissa A. Meyer contributed equally to this work.

Dipak P. Ramji (ed.), *Atherosclerosis: Methods and Protocols*, Methods in Molecular Biology, vol. 2419,
https://doi.org/10.1007/978-1-0716-1924-7_47,

allowed the detection of 16 to 20 markers simultaneously. Mass cytometry (CyTOF) uses rare earth metal tags instead of fluorochromes, further expanding the range of markers up to 50 [5].

To study the immune cell composition of the artery wall by flow cytometry or CyTOF, single-cell suspensions must be made. Most of our knowledge is derived from animals, usually mouse models. The majority of methods to evaluate mechanisms of disease and efficacy of therapy would include the isolation and excision of the aorta. Analogous protocols can be used for other arteries (not described here).

The yield, viability, and composition of the single-cell suspensions depend on the diet, age and sex of the animals, harvesting practices, and the protocol used to obtain single-cell suspensions [4, 6–12]. The number of immune cells in the aorta differs significantly using different protocols [12–16].

This chapter provides an overview of the methods used for aortic immune cell characterization, from aortic leukocyte harvesting to flow cytometry and CyTOF analysis.

2 Materials

2.1 Mouse aorta Harvesting

1. Microdissection tools: two blunt curved forceps, iris scissors.
2. Surgical dissecting microscope.
3. Foam plate.
4. 1× Phosphate buffered saline (PBS): 137 mM NaCl, 2.7 mM KCl, 8 mM Na_2HPO_4, and 2 mM KH_2PO_4.
5. 1 mL syringes with 25G Needle.
6. 10 mL syringes with 25G Needle and Luer-Lok Tip.
7. 70% ethanol.
8. CO_2.
9. Ketamine.
10. Perfusion Buffer: 2% heparin diluted in PBS.
11. Fetal bovine serum (FBS).
12. RPMI 1640 Media.
13. Penicillin–streptomycin solution.
14. 0.5 M sterile ethylenediaminetetraacetic acid (EDTA).
15. 15 mL conical tubes.

2.2 Aorta Digestion

1. Microdissection tools: two blunt curved forceps, iris scissors.
2. 1× PBS.
3. FBS.
4. 1× HBSS with Ca^{2+} and Mg^{2+}.

5. Collagenase, Type I.
6. Collagenase, Type XI.
7. Deoxyribonuclease I (DNase) from bovine pancreas, type I.
8. Hyaluronidase from bovine testes, type 1-s.
9. Penicillin–streptomycin solution.
10. Completed RPMI Media: RPMI 1640 containing 10% FBS and 1% penicillin–streptomycin solution.
11. 15 mL conical tubes.
12. 70 μm nylon mesh strainers.
13. Thermally controlled circular shaker.
14. Thermally controlled 15/50 mL conical tubes centrifuge.
15. Ice.

2.3 Flow Cytometry (FACS)

1. Cell counting tool (e.g., hemocytometer or automatic cell counter).
2. Fc blocking antibodies (CD16/CD32 solution).
3. Antibodies as per panel (*see* Table 1).
4. Fixable viability dye.
5. 2% paraformaldehyde (PFA).
6. 1× PBS.
7. FBS.
8. 5% sodium azide (NaN_3).
9. FACS Buffer: PBS, 0.5–1% BSA or 4% FBS, 0.1% NaN_3.
10. Fc blocking buffer: FACS buffer, Fc blocking antibodies.
11. Flow cytometry acquisition instrument (e.g., BD LSR 2 II, BD Fortessa).
12. Compensation beads.
13. FACS tubes.
14. Table top centrifuge able to accommodate 15 mL conical tubes.
15. Flow cytometry analysis software (e.g., FlowJo).

2.4 Mass Cytometry (CyTOF)

1. CyPBS: 1× PBS free from heavy metal contaminates.
2. Cell Freezing Media: 5% DMSO in FBS.
3. CyFACS: 2 mM EDTA, 0.1% BSA, 0.05% NaN_3 in CyPBS. Filter-sterilize and store at 4 °C.
4. Purified CD16/32 Antibody.
5. Liquid nitrogen.

Table 1
Aortic panleukocyte panel example for flow cytometry

Antigen	Color	Dilution	Clone	Source
CD5	PerCP Cy5.5	1 to 200	54-7.3	BioLegend
CD11b	PE-Cy7	1 to 200	M1/70	BioLegend
CD64	APC	1 to 200	X54-5/7.1	BioLegend
CD19	APC-Cy7	1 to 200	6D5	BioLegend
CD45	AF700	1 to 400	30-F11	BioLegend
CD11c	BV605	1 to 200	N418	BioLegend
CD127	BV570	1 to 200	A7R34	BioLegend
TCR-y/d	FITC	1 to 200	GL3	BioLegend
CD8a	BV421	1 to 400	53-6.7	BioLegend
NK1.1	BV650	1 to 200	PK139	BD Biosciences
TCR-β	BV711	1 to 400	H57-597	BioLegend
Ly6C	BV786	1 to 200	RB6-8C5	BD Biosciences
CD4	PE/Dazzle 594	1 to 400	RM4-5	BioLegend
CD25	PE	1 to 200	PC61	BioLegend
LD	AQUA	1 to 1000	N/A	Thermo Scientific

6. Conjugated antibody: extracellular, intracellular, spike-in (will vary depending on designed panel, *see* Table 2).
7. Cell counting tool (e.g., hemocytometer or automatic cell counter).
8. Table top centrifuge able to accommodate 15 mL conical tubes.
9. Live/Dead Stain: 5 μM Cell ID Cisplatin (Fluidigm) in CyPBS. Prepare fresh each day.
10. 2% PFA: 16% methanol-free formaldehyde diluted to 2% in CyPBS. Store aliquots at −20 °C, protected from light.
11. Fixation/Permeabilization Buffer from FoxP3/Transcription Buffer Set (eBioscience). Working solution as described in the manufacturer's protocol: 4× Fixation/Permeabilization Concentrate diluted to 1× in Fixation/Perm Diluent. Prepare fresh each day.
12. Permeabilization Buffer from FoxP3/Transcription Buffer Set (eBioscience): 10× Permeabilization Buffer diluted to 1× with Milli-Q water. Prepare fresh each day.
13. Intercalator Solution: Cell ID Intercalator-IR, which contains natural-abundance iridium, 191Ir and 193Ir (Fluidigm)

Table 2
CyTOF panel as published in Winkels et al. [9]

Antigen	Label	Clone	Source	Catalog number	Conjugation
CD45	89Y	30-F11	Fluidigm	3089005B	Commercial
Ly-6C	141Pr	HK1.4	BioLegend	128002	In-house
CD11c	142Nd	N418	Fluidigm	3142003B	Commercial
TCR-β	143Nd	H57-597	Fluidigm	3143010B	Commercial
CD41	144Nd	MWReg30	BioLegend	133902	In-house
CD45 (Spike In)	145Nd	A20	BioLegend	110702	In-house
CD5	146Nd	53-7.3	Fluidigm	3146012B	Commercial
CD11b	148Nd	M1/70	Fluidigm	3148003B	Commercial
CD19	149Sm	6D5	Fluidigm	3149002B	Commercial
CD64	150Nd	X54-5/7.1	BioLegend	139302	In-house
CD25	151Eu	3C7	Fluidigm	3151007B	Commercial
Siglec-F	152Sm	1RNM44N	eBioscience	14-1702-82	In-house
CD8a	153Eu	53-6.7	Fluidigm	3153012B	Commercial
TER-119	154Sm	TER-119	Fluidigm	3154005B	Commercial
FR4	155Gd	TH6	BioLegend	125102	In-house
CD103	156Gd	2E7	BioLegend	121402	In-house
TCR-γ/δ	159Tb	GL3	Fluidigm	3159012B	Commercial
B220	160Gd	RA3-6B2	Fluidigm	3160012B	Commercial

diluted 1:2000 in permeabilization buffer. Prepare fresh each day.

14. Cell Acquisition Solution (Fluidigm).
15. Washing Solution (Fluidigm).
16. Cell Acquisition Solution and Beads: EQ Four Element Calibration Beads (Fluidigm) diluted 1:10 in Cell Acquisition Solution (Fluidigm). Prepare fresh each day.
17. 70 um filters.
18. FACS tubes.
19. CyTOF Helios (Fluidigm).
20. Normalizer (https://github.com/nolanlab/bead-normalization/releases, [17]).
21. CyTOF analysis software (e.g., FlowJo, Cytobank, Omiq).

3 Methods

3.1 Mouse Aorta Harvesting

1. Euthanize mouse by CO_2 inhalation or by exposing the animal to a supratherapeutic dose of anesthetics, such as 100–200 mg/kg of ketamine by intraperitoneal injection. The method used needs to be in accordance with the approved protocol of the appropriate Animal Care and Use Committee.
2. Verify primary euthanasia via toe pinch as a noxious stimulus.
3. Use a secondary method of euthanasia for confirmation (diaphragm cutting or heart puncture). Avoid cervical dislocation due to possible mechanical aortic damage.
4. After euthanasia, spray the mice with 70% ethanol and place in supine position with all 4 legs fixed onto a foam plate.
5. Puncture the heart with a 1 mL syringe with 25G needle to harvest blood, if desired.
6. To remove the remaining blood, use a 10 mL syringe with 25G needle to perfuse the aorta by puncture of the left ventricle with 5 to 10 mL of sterile cold (4 °C) perfusion buffer to avoid blood clots. The aorta will appear white after perfusion.
7. Using microdissection tools, remove the intestines, spleen, liver, and lung, keeping the heart, kidneys, and aorta in situ (*see* **Note 1**).
8. Remove the aorta under a dissecting microscope (*see* **Note 2**), carefully cleaning it from the periaortic fat.
9. Leave the main artery branches intact (0.5 cm from their branching location): Innominate artery, left carotid artery, left subclavian artery, renal arteries, celiac artery, and the common iliac artery.
10. Place the harvested aorta in a 15 mL conical tube containing PBS supplemented with 2% FBS and 2 mM EDTA or RPMI 1640 media supplemented with 10% FBS.

3.2 Aorta Digestion

1. Prepare enzyme digestion solution (diluted in HBSS containing Ca^{2+} and Mg^{2+}) with the following enzymes: 675–450 U/mL collagenase type I; 250–125 U/mL collagenase type XI; 120–60 U/mL hyaluronidase type I-s; and 60 U/mL DNase-I (to establish the exact enzyme concentration for the specific experimental settings *see* **Notes 3–6**).
2. Place 5 mL of the enzyme digestion solution in a 15 mL conical tube containing the aorta. Aortas can be cut into small pieces to improve digestion of the tissues.
3. Incubate at 37 °C for 40–70 min in agitation at 160–180 rpm In a thermally controlled circular shaker (*see* **Note 3**).

4. After incubation, filter through 70 μm nylon mesh strainers, and add 10 mL of cold PBS.
5. Wash by centrifugation at 400 × *g* at 4 °C for 5 min and then decant supernatant.
6. Resuspend in 2 mL of completed RPMI media and incubate at 37 °C for antigen recovery (*see* **Note 5**).
7. Wash by centrifugation at 400 × *g* at 4 °C for 5 min and then decant supernatant.
8. Add 1 mL of FACS buffer and place on ice and proceed with the chosen next steps (Flow Cytometry or CyTOF).

3.3 Flow Cytometry (FACS)

3.3.1 FACS Panel Design

Panel design is the most important step in designing a flow cytometry experiment. One of the most commonly available Flow Cytometers are the BD LSR II and the BD Fortessa. Both are capable of detecting from 15 to 20 markers including nuclear markers. For the design and optimization of an effective multi colorimetric flow cytometric panel, several steps are required such as correct fluorophore selection (*see* **Notes 7** and **8**), titration (*see* **Note 9**), live–dead cell staining (*see* **Note 10**), and validation with Fluorescence Minus One (FMO) control (*see* **Note 11**).

3.3.2 FACS Extracellular Staining

1. Perform extracellular marker staining on aortic cell suspension using a standard flow cytometry protocol as described below.
2. Count the cells with the desired method (e.g., hemocytometer or automatic cell counter) and incubate 1.0–1.5 × 10^6 of aortic cell suspension with 50 μL of 10 μg/mL of Fc blocking buffer for 10 min at room temperature to inhibit nonspecific binding of antibodies.
3. Add 50 μL of antibodies as per panel (*see* **Note 12** and Table 1) prepared in 10 μg/mL of Fc blocking buffer.
4. The extracellular antibody staining can be combined with the viability stain (*see* **Note 10**).
5. Incubate samples in the dark for 20–30 min at room temperature or 40 min on ice.
6. After the antibody incubation, add 1 mL of PBS to wash.
7. Centrifuge at 400 × *g* at 4 °C for 5 min and decant the supernatant.
8. Repeat **steps 6** and **7**.
9. After washing, fix the cells in 100 μL of 2% PFA for 30 min.
10. Repeat **steps 6** and **7**.
11. Resuspend the cell pellet in 200 to 300 μL of FACS buffer.
12. Proceed to data acquisition using the flow cytometer or store overnight at 4 °C in the dark.

3.3.3 FACS Sample Acquisition

1. Before sample acquisition, set up the flow cytometer and settings according to the manufacturer's protocol.
2. Set up voltages and run compensation controls for each fluorophore by using compensation beads or single fluorophore stained cells and apply to the experiment (*see* **Note 13**).
3. Use unstained and FMO controls (*see* **Note 11**) to set the appropriate forward (FSC) and side (SSC) scatter as well as gates for the fluorescent channels to depict clear negative and positive populations for each fluorophore used. Then, begin collecting control samples.
4. Run samples from Subheading 3.3.2, **step 12**.

3.3.4 FACS Data Analysis

For multiparametric flow cytometry analysis, often the software installed in the Flow Cytometer are not the best tools, being inherently slow and poorly intuitive. To overcome these problems, other software can be used (e.g., FlowJo and Cytobank). Both FlowJo and Cytobank allow processing of data from initial compensation and doublet removal (*see* **Notes 13** and **14**) to figure editing, providing all the tools required. Recently, these programs also include a user-friendly interface for biaxial analysis such as viSNE that allows the two dimensional visualization of high-dimensional single-cell data based on the t-Distributed Stochastic Neighbor Embedding (t-SNE) [18] and Uniform Manifold Approximation and Projection method (UMAPs) allowing a full characterization and visualization of multiparametric data as recently shown by McArdle and colleagues [15].

3.4 Mass Cytometry (CyTOF)

3.4.1 CyTOF Panel Design

The most widely available CyTOF, the Helios, is capable of detecting ions atomic masses from 75 to 209 Da [19]. *See* **Notes 15** and **16** for extensive description of the procedures needed for panel design, antibody conjugation, antibody titration, and required controls.

*3.4.2 CyTOF Internal Controls (See **Notes 17–20**)*

1. Prepare aliquots of spike-in control by isolating a single cell suspension of aorta or spleen.
2. Wash with CyPBS, resuspend in Cell Freezing Media, and slow freeze at −80 °C. Store at −80 °C or in liquid nitrogen until use. Prepare at least one vial per day of staining.

*3.4.3 CyTOF Antibody Cocktails (See **Note 21**)*

1. Prepare IgG FC Receptor block by diluting purified CD16/32 in CyFACS at 1:200 (*see* **Note 22**).
2. Prepare CD45 Sample Cocktail by diluting sample-specific antibody (89Y-CD45 in example panel, Table 2) in CyFACS at determined titration.

3. Prepare CD45 Spike-In Cocktail by diluting spike-in-specific antibody (145Nd-CD45 in example panel, Table 2) in CyFACS at determined titration.
4. Prepare Extracellular Antibody Cocktail by diluting all extracellular antibodies except CD45 in CyFACS at determined titrations (*see* **Note 23**).
5. Prepare Intracellular Antibody Cocktail by diluting all intracellular antibodies in Permeabilization Buffer at determined titrations (*see* **Note 24**).

3.4.4 CyTOF Live/Dead Staining

1. Proceed from the isolation of single cells from the aorta (Subheading 3.2, **step 8**). Use the entire aorta digestion for each sample. Note a cell count for each sample using desired method.
2. Thaw a vial of spike-in control into 5 mL of CyPBS. Centrifuge 350 × *g* at 4 °C for 5 min. Aspirate (*see* **Note 25**).
3. For each sample and the spike-in control, wash cells with 5 mL of CyPBS.
4. Centrifuge 350 × *g* at 4 °C for 5 min. Aspirate.
5. Resuspend in 1 mL of Live/Dead stain (*see* **Note 26**).
6. Incubate at 4 °C for 5 min.
7. Quench with 1.5 mL of CyFACS.
8. Centrifuge 350 × *g* at 4 °C for 5 min. Aspirate.
9. Resuspend in 1 mL of CyFACS.
10. Centrifuge 350 × *g* at 4 °C for 5 min. Aspirate.

3.4.5 CyTOF Cell Staining

1. Resuspend in 100 μL of IgG FC Receptor Block. Incubate at 4 °C for 10 min.
2. Centrifuge 350 × *g* at 4 °C for 5 min. Aspirate.
3. Resuspend each sample in 100 μL of CD45 Sample Cocktail. Resuspend spike-in in 100 μL of CD45 Spike-In Cocktail (*see* **Note 17**). Incubate at 4 °C for 30 min.
4. Add 1 mL of CyFACS. Resuspend samples and spike-in in 100 μL extracellular antibody cocktail (*see* **Note 16**). Add appropriate volume of spike-in control suspension to each sample.
5. Incubate at 4 °C for 30 min.
6. Add 1 mL of CyFACS.
7. Centrifuge 350 × *g* at 4 °C for 5 min. Aspirate.

3.4.6 CyTOF Cell Fixation

1. Resuspend in 100 μL of 2% PFA. Incubate overnight at 4 °C (*see* **Note 27**).

2. Add 1 mL of CyFACS.
3. Centrifuge 800 × *g* at 4 °C for 5 min. Aspirate (*see* **Note 28**).

3.4.7 CyTOF Intercalator Stain and Intracellular Staining

Intracellular staining is not necessary for all panels. If no intracellular antibodies are used, skip **steps 6–11**.

1. Resuspend in 1 mL of Fixation/Permeabilization Buffer. Incubate at 4 °C for 30 min.
2. Add 1 mL of Permeabilization Buffer.
3. Centrifuge 800 × *g* at 4 °C for 5 min. Aspirate.
4. Resuspend in 1 mL of Permeabilization Buffer.
5. Centrifuge 800 × *g* at 4 °C for 5 min. Aspirate.
6. Resuspend in 100 μL of Intracellular Antibody Cocktail (*see* **Note 16**).
7. Incubate at 4 °C for 30 min.
8. Add 1 mL of CyFACS.
9. Centrifuge 800 × *g* at 4 °C for 5 min. Aspirate.
10. Resuspend in 1 mL of CyFACS.
11. Centrifuge 800 × *g* at 4 °C for 5 min. Aspirate.
12. Resuspend in 100 μL of intercalator solution (*see* **Note 29**).
13. Incubate at room temperature for 30 min.
14. Add 1 mL of CyFACS.
15. Centrifuge 800 × *g* at 4 °C for 5 min. Aspirate.
16. Resuspend in 1 mL of CyFACS.
17. Store at 4 °C until prepared to acquire samples (*see* following sections and **Note 30**).

3.4.8 CyTOF Acquisition

1. Warm up, tune and inspect make-up gas each day of use of CyTOF Helios in addition to regular maintenance (*see* **Note 31**).
2. Prepare for acquisition in the software by input of your panel and setting the run time or event count for each sample.
3. Clean the CyTOF Helios by running 5 min of Washing Solution followed by 5 min of Cell Acquisition Solution.
4. Prepare sample by first centrifugation at 800 × *g* at 4 °C for 5 min. Aspirate (*see* **Note 32**).
5. Resuspend in 1 mL of Cell Acquisition Solution.
6. Centrifuge 800 × *g* at 4 °C for 5 min. Aspirate.
7. Resuspend in 1 mL Cell Acquisition Solution + Beads.
8. Filter through 70 μm filter into a FACS tube (*see* **Note 33**).

9. Analyze sample on CyTOF Helios. Acquire one million cells or the entire samples, whichever occurs first. The event rate should be no greater than 500 events/s.

3.4.9 CyTOF Cleaning Data

The following process will normalize data using the four element beads and remove debris, doublet, and dead cells using FlowJo. Further gating in FlowJo may be necessary to select for subpopulations of cells or to separate out the spike-in control. Many computational pipelines or user-friendly software packages are capable of cleaning the data and/or normalizing with beads, spike-in controls and barcoding independent of FlowJo manipulation.

1. Normalize to and remove beads using Normalizer (https://github.com/nolanlab/bead-normalization/releases [17]).
2. Import normalized files into FlowJo.
3. Select singlets (*see* **Note 34**).
4. Select live cells.
5. Export live cells as a new FCS file (*see* **Notes 35** and **36**).

3.4.10 CyTOF Analysis

CyTOF analysis can be accomplished in many different ways and is dependent on the skill sets available and the biological question. Many user-friendly platforms for high-dimensional cytometry analysis are now available, including Cytobank, Omiq, and FlowJo. Pay for service analyses are also available including the Astrolabe Cytometry Platform. The most flexible type of analysis is the use of computational pipelines, though they require computational expertise.

4 Notes

1. To avoid contamination of the aorta, all tools (scissors, forceps) must be clean and sterile by autoclaving.
2. To avoid cell death, aorta harvesting steps need to be performed in a timely manner, avoiding leaving the aorta sitting for a long time. The average time for aorta harvest from mouse euthanasia should not exceed 15 min. Moreover, during harvesting, aorta should be always kept moist with cold perfusion solution.
3. Enzyme concentration, as well as incubation time, are the key factor in the quality and quantity of the aortic single-cell recovered. Lower concentration of collagenase XI and hyaluronidase (e.g., 125 U/mL, and 60 U/mL respectively), followed by a reduced incubation time (e.g., 40-min), is preferred for a good representation of myeloid cells in the single-cell pool. Higher enzyme concentration (e.g., 250 U/mL of collagenase XI, and

120 U/mL of hyaluronidase) followed by at least 60-min incubation provides a better representation of T cells and B cells, with fewer macrophages and foam cells [12]. The best recovery of myeloid cells and foam cells so far obtained has been achieved by Kim and Lin et al. [11, 20]. Kim and colleagues applied a double digestion method using an initial short incubation time with Collagenase type II followed by a long 70-min incubation time that included all classical enzymes [11]. This approach provided a good amount of foam cells but lost all the other major cell populations. Similarly, McArdle et al. used a 45 + 45 min incubation with the same enzymatic pool described above to achieve macrophage isolation that was used for bulk RNA seq analysis [15].

4. Enzymatic digestion is necessary to retrieve single-cell suspension especially for studying tissue cell composition and describing new players in diseases like atherosclerosis. The methods used until now are far from being perfect, the high variability in the cell quality and quantity still require adjustment based on the focus of the study. Moreover, collagenase-based enzymes are known to induce changes in the cell transcriptomic profile. Often, the expression of stress response and cell death genes are induced. Due to continuous advancements in the sensitivity of high throughput technologies, small changes in gene expression can dramatically influence the interpretation of biological data. Transcription is active at 37 °C, and extended incubation at high temperatures may introduce gene expression artifacts, independent of the biology at the time of harvest [21].
5. Another example of the effect of digestion on the cells is the loss of some antigen expression like CD4. Rescue methods can be applied to retrieve the expression of some such markers. This requires incubation in complete media like RPMI containing 10% fetal bovine serum at 37 °C for 30–40 min [9].
6. Recently, a serine protease (subtilisin A) isolated from a Himalayan glacier–resident bacterium, *Bacillus licheniformis*, has been shown suitable for dissociation of tissues such as kidney and solid tumors at 4–6 °C. Its application reduced artifacts, including global and single-cell gene expression changes compared to classical collagenase-based digestion [22]. Cold enzymatic digestion has not been applied yet for the aorta; however, its application could provide improvement in cell viability with the cold incubation allowing the maintaining of the cell pool diversity in tissues, finally achieving a complete overview of the population habiting tissues in a healthy or disease state.
7. Not all fluorophores work the same. It is critical to study the absorption spectra (to select the laser) and the emission spectra. You must select the right fluorophore for the right marker. You

may use a less efficient marker for the dump channel (cell types you are not studying). The number of fluorophores available has increased significantly thanks to the development of new tandem dyes and dyes with large Stokes's shift. Each fluorophore, other than a specific excitation and emission wavelength, can also be ranked for its brightness and stability in the buffer. The relative abundance of a marker in the cells of interest is what should guide the selection of the fluorophore; CD45, for example, is abundant in leukocytes and should be always used in dim fluorophores such as FITC or PerCP. Conversely, rare markers will preferentially require bright fluorophores like PE or Alexa Fluor 647.

8. Tandem dyes contain chemically linked donor and acceptor molecules. The donor molecule is generally a dye with a large extinction coefficient, and the acceptor molecule absorbs energy emitted by the donor molecule through a phenomenon called fluorescence-resonance energy transfer (FRET). Tandem dyes have substantially increased the fluorophore choice. However, the stability of these dyes is limited. The decoupling of the donor from the acceptor can change substantially the emission wavelength, causing emission to erroneously fall into the channel of another fluorophore. For example, when PE-Cy7 disintegrates over time, the signal will fall in the PE channels giving false positive signals. Tandem dyes must be stored strictly according to the manufacturer's specifications.
9. Titration is required for optimizing the resolution and obtaining robust results for population identification and expression level measurements. It requires little effort and can save time, money, and problems in the long run. Titration means to test different concentrations of an antibody, for example, from 1:50 to 1:1000, and check which concentration will provide the best separation index between positive and negative events. When this is set, that concentration can be used among all the experiments.
10. Due to the elevated cell death caused by the digestion protocols, effective and correct aortic leukocyte characterization require the use of specific live–dead markers to assess cell viability and avoid false-positive signals. When cells die, they lose their membrane integrity and become "leaky." This is what causes the nonspecific binding. In our laboratory, we reject samples with leukocyte viability below 60%. Live–dead markers can be added to the staining mix or used before the antibody staining by incubation for 15 min in the dark. There are two main live–dead markers used in flow cytometry: DNA binding dyes and protein binding dyes. DNA binding dyes, such as propidium iodide (PI), DAPI, or 7-AAD only enter the cells when the membranes have been compromised. Once they do,

they bind the DNA and fluoresce. A range of these dyes with different excitation and emission characteristics exists. However, they should be avoided if the protocol requires a final fixation of the cells. Protein binding dyes, also known as amine-reactive dyes or live/dead fixable dyes, are still based on the principle that dead cells have compromised membranes but instead of binding DNA, they bind specific intracellular proteins. Moreover, this dye can be used also in protocols that require cell fixation.

11. Fluorescence minus one (FMO) controls are important when optimizing multicolor flow cytometry panels as they will help to determine where the gates delineating negative with positive events should be set. They are needed because of the so-called fluorescent spread, which is most pronounced with the brighter fluorophores. That is why the unstained control is not sufficient to correctly set the gates. FMOs consume a lot of cells, because you would have 20 different FMO tubes in a 20-flurophore panel. FMO controls are highly recommended when accurate identification between a positive and negative population is vital in the correct characterization of such population. FMOs can be omitted for abundant markers such as CD45, CD4, CD8, and CD19.
12. Aortic leukocyte panel design: In 2018, Winkels et al. applied a 13-marker panel for the discrimination of aortic leukocytes including: CD45, CD11b, F4/80, TCRyδ, CD117 (c-kit), CD19, FceR1, CD11c, NK1.1, TCR-b, Ly6C/G (Gr-1), CD4, and CD8 [9]. With this set of markers, they were able to successfully identify 12 leukocyte populations from the aorta of $ApoE^{-/-}$ mice fed 12 weeks western diet (WD), including mast cells, γδ T Cells, CD4+ T cells, CD8+ T cells, NK cells, NK T cells, macrophages, CD11c+ macrophages, DCs and myeloid DCs, and CD11b+ and CD11b- B cells acquired with an LSRII flow cytometer. However, 17% of cells remained unidentified, suggesting that other leukocyte populations are present in the aorta and the markers used are not sufficient for their detection. The panleukocyte panel currently used and validated in our laboratory did change with the addition of few specific markers, increasing the limit to currently 15 or 16 markers designed for the LSRII configuration (panel example in Table 1) and includes: (1) CD127 and CD25 for the identification of CD4+ T regs (TCRβ+, CD4+, CD25+ CD127 low); (2) Ly6C (instead of GR1) for the identification of classical and nonclassical monocytes; (3) CD5 recently used together with CD19 and CD11b to identify B1a (CD19+, CD11b+, CD5+) and B1b (CD19+, CD11b+, CD5+) B Cells subpopulations; and (4) CD64 to improve macrophage detection given that some DCs also express F4/80

[23]. Thanks to the new single cell RNA sequencing and CyTOF studies [9, 8, 12], new knowledge on the leukocytes subsets such as macrophages inhabiting the aorta are available. New panels are currently being developed with a combination of markers like CD45, CD64, CD206, CD169, CD11c, CD11b, MHCII, F4/80, Ly6C, and CD9/TREM2 for the identification of the newly defined subsets.

13. Compensation is always required because a fluorophore can bleed into a channel applied for the detection of another dye. This spillover can be extensive for fluorophore with close emission spectra such as FITC and PE resulting in false double-positive events. The spillover fluorescence can be defined as a linear function of the fluorescence intensity at the same photomultiplier tube (PMTs) voltages and filter sets. Introducing a spillover coefficient for each channel is sufficient for the flow cytometry software to calculate the compensated data [24]. Thus, voltage and spillovers among fluorophores in the same PMT should be carefully checked before the acquisition, aiming to keep the spillover below 30%. When using a large multicolor panel, further compensation might be required with third-party software like FlowJo, which allows manual correction of the automatic compensation generated by the flow cytometry software. This will help in improving the quality of the final data.
14. Another crucial aspect in the flow cytometric data analysis is the removal of doublet events (two cells passing the laser at the same time) as this can lead to false discovery especially when, as in atherosclerosis research, the goal is to define new populations inhabiting the aorta. Any "new" cell type is initially suspect of being a doublet unless proven otherwise. Several strategies are applied to exclude cell doublets from the data: (1) The most common is to exclude the events that deviate from the linear correlation between the forward scatter area (FSC-A) and the FSC height (FSC-H) parameters; and (2) To be more restrictive, two more filtering steps can be done with FSC-A by FSC width (FSC-W) and side scatter area (SSC-A) by SSC width (SSC-W) utilizing the low pulse width signal indicative of single cells [25]. The method indicated will remove the majority of cell doublets; however, some are not fully removed, which needs to be always considered when performing the analysis. Not all doublets are technical. There is mounting evidence that some doublets may have functional and biological meaning suggesting true cell-cell interactions [26].
15. CyTOF allows for the analysis of many cell populations and phenotypes in a single experiment. It is important to consider each cell type you would like to identify, as well as the cell types that need to be excluded. Markers for each should be included

in the panel. It is acceptable to use "dump" channels, where multiple antibodies are stained in the same channel and all positive cells are excluded from the analysis. Table 2 displays a CyTOF panel similar to the one used in Winkels et al. [9]. This panel allowed for the identification of 28 clusters or cell populations in the mouse aorta across healthy and atherosclerotic mice. Generally, panel design for CyTOF is more forgiving than panel design for flow cytometry. Because heavy metals are utilized rather than fluorophores, spectral overlap and compensation are not an issue like in flow cytometry. Yet there are a few considerations to be aware of in assigning antibodies against specific antigens to certain heavy metals. Like flow cytometry fluorophores, some heavy metals are "brighter" than others. Metals near the middle of the detection range (~147Gd to 170Er) provide a stronger signal than those metals on the ends of the range [27]. Low abundance markers should be assigned to the "brighter" metals to ensure detection. Next, a small amount of metal impurity can occur. This leads to spillover in the channel the next mass higher or the next mass lower (M + 1 or M − 1). The Maxpar labeling kits and commercially conjugated antibodies guarantee the purity of 98% or higher and rarely lead to enough rogue signal to require compensation [28]. Another source of spillover is oxidation. Oxidation can occur during the acquisition of the sample on the Helios leading to additional signals in the M + 16 channel. Careful warm-up, tuning, and adjusting of the make-up gas on the Helios each day can keep oxidation to a minimum (<3% of the signal) [28]. If issues from spillover arise, it would be prudent to consider using antibodies against antigens that are relatively low in abundance for the channel in question, to reduce the magnitude of the spillover.

16. Many antibodies can be purchased commercially from Fluidigm already conjugated to heavy metal isotopes. In the case that a commercial reagent is not available, purified antibodies can be conjugated to the heavy metal isotope of choice using the Maxpar antibody labeling kits available from Fluidigm. Purified antibody preparations should be free of carrier proteins. The antibody labeling kits work best with IgG isotypes and also work with affinity-purified polyclonal antibodies, but are not useful for conjugation to IgM isotypes. Once antibodies are obtained from Fluidigm and/or conjugated using the Maxpar antibody labeling kit, it is prudent to validate the antibodies using titration. The best method is to use the tissue of interest to perform titrations, but another tissue (such as the spleen) is acceptable if the material is limited (as it is for the aorta). Titrations should be performed on a similar number of cells to match expected staining conditions, or on ten million

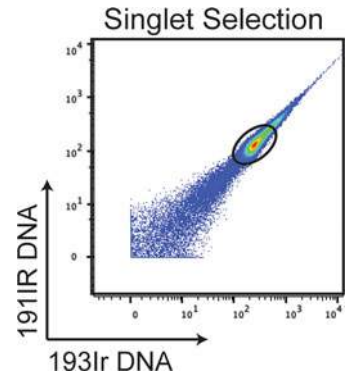


Fig. 1 Example of FlowJo-based sample cleaning. Singlet selection is preformed first by inspecting 191IR vs. 193IR to look at cellular DNA content. An oval shaped, concentrated bulge is selected as shown. Selected singlets are suitable for export and use in further downstream analysis

cells if the staining conditions are unknown. The full CyTOF staining protocol should be performed in the same way as is expected for the final experiment. Titration of 1:400, 1:200, 1:100, and 1:50 are suggested but could be adjusted based on the user's experience. An unstained control is necessary for comparison. Unlike flow cytometry, titration for multiple antibodies can be performed using a single tube for a single dilution (i.e., if four antibodies need to be titrated, they could be combined at 1:400, a second tube at 1:200, a third tube at 1:100, a fourth tube at 1:50, and an unstained control for a total of only 5 samples). Titration can be analyzed simply using Flowjo. After the selection of single (Fig. 1) and live cells, titrations can be inspected as shown in Fig. 2. In the case of Fig. 2, a dilution of 1:100 would be chosen because it shows distinct staining and higher dilutions do not provide additional resolution. All antibodies in a "dump" channel need to be titrated independently to ensure each is staining. In this case, more than five tubes would be necessary. In CyTOF, we generally do not perform control similar to fluorescence minus one. Instead, it is prudent to use knowledge of biology to inspect for artifacts or nonspecific staining. For example, CD4 and CD8 are not expected to be coexpressed on T cells outside of the thymus, so coexpression might suggest a problem with staining. It is important to remember that CyTOF data may need to be validated using an additional method like FACS if a novel or unexpected population of cells is identified.

17. One limitation of CyTOF is the degradation of the signal on the CyTOF Helios across a day of acquisitions and the inconsistency in detection and staining from day to day, or month to month for longitudinal studies. Two methods of internal

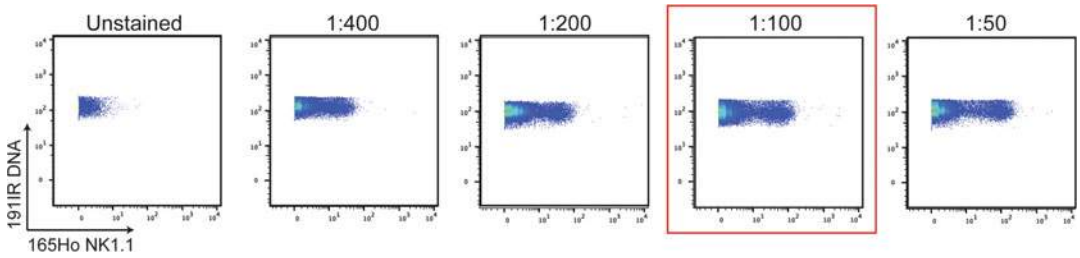


Fig. 2 Example of antibody titration. Blood cells were stained with anti-mouse NK1.1 tagged with 165Ho at dilutions of 1:400, 1:200, 1:100, and 1:50. An unstained sample was also collected. Decreasing dilutions of antibody showed increased staining and better separation between the positive and negative populations. Separation and staining plateau at 1:100. Therefore, this dilution is chosen

control can be used to ensure consistency after normalization. The first is a spike-in control. In this case, a control sample of the tissue in question or another tissue that also expresses all of the incorporated markers such as the spleen, is collected, processed into single cells, and frozen using cell freezing media, aliquoting enough vials to have sufficient supply to thaw one vial each time staining occurs. Each staining day, a vial is thawed and first stained with a CD45 antibody tagged with a unique metal. This is considered the spike-in control. In parallel, the sample of interest is stained with a CD45 antibody tagged with a different metal. Once washed, the cells are mixed before proceeding with additional staining and acquisition. During the analysis, the spike-in is separated and used as a reference to normalize all samples.

If samples are stained all on the same day, barcoding is the best internal control method to use. Recent advances have provided a cadmium-based barcoding system available from Fluidigm [29]. In this protocol, each sample is labeled with a different combination of the seven available cadmium isotopes tagged to a CD45 antibody, before mixing and proceeding with further staining and acquisition. Before analysis, samples are de-barcoded using available open-source software packages or user-friendly software programs such as Cytobank. In addition to these optional internal controls, all samples are acquired on the Helios with EQ 4 Element Beads from Fluidigm. These polystyrene beads are tagged with 140/142Ce, 151/153Eu, 165Ho and 175/176Lu and are added to the sample during the final resuspension before the acquisition. Like a spike-in control, these beads provide a consistent signal that is used to normalize signal variation that can occur over time especially if the machine is run for hours or across days. It is important to make sure that signals from these channels are collected, but do not exclude their use for other markers. Beads are easily distinguished from cells due to their coexpression of all four metals

and the bright signal. Normalizing software is available to correct the data and remove the beads [17]. Normalizer v0.3 from the Nolan laboratory is available for download from github (https://github.com/nolanlab/bead-normalization/releases).

18. Reagents and samples should be kept cool throughout the procedure unless otherwise noted. Samples and reagents can be incubated, centrifuged, and stored at 4 °C. Samples can be quickly manipulated at the bench before returning to 4 °C or work can be done on ice.
19. Reagents need to be clean of heavy metal contamination. Common contaminants include barium, lead and iodine. Barium can be present on some laboratory materials and in laboratory soap. To avoid barium, it is suggested to use new plastics rather than washed reusable containers. Iodine can be present on some laboratory materials. Lead can be found in some municipal water supplies. To avoid lead, it is suggested that Milli-Q water be used for diluting reagents. To test for contamination, process a sample without antibody staining and look for contamination.
20. The protocol details a spike-in internal control. For protocols defining barcoding, *see* reference of Zivanovic et al. [30].
21. Prepare antibody cocktails before beginning an experiment. Antibody cocktails can be prepared a day prior to an experiment and stored at 4 °C.
22. If CD16/32 is part of the panel design, use the conjugated CD16/32 antibody at the tested dilution in the IgG block. This step will block any nonspecific antibody interactions.
23. Do account for antibody volume when preparing dilutions of the extracellular antibody cocktail.
24. If all antibodies are extracellular, skip intracellular staining. Prepare the intracellular antibody cocktail the day it is used.
25. Spike-in should account for 10–20% of each sample. Make sure enough spike-in is processed to cover all samples.
26. Cisplatin is used to detect dead cells. Cisplatin is excluded from live cells but penetrates dead cells during a short incubation.
27. Cells are acquired in a hypotonic solution. Fixation is imperative to stabilize cells.
28. After fixation, increase centrifuge speed to 800 × *g* to increase cell recovery.
29. The cell structure is not maintained or recorded as cells pass through the CyTOF. Instead, a DNA intercalator is used to identify cells and exclude doublets.

30. Sample can be stored up to 2 days before the acquisition. Cell integrity will be maintained better the sooner samples are acquired.
31. This is usually performed by staff in a core facility. If you are to operate the CyTOF independently, ask for training from your institution.
32. Prepare a single sample at a time for acquisition. Once a sample is washed into Cell Acquisition Solution, it cannot go back into CyFACS. If problems with the CyTOF Helios arise, the washed but unacquired sample will be lost.
33. The CyTOF Helios is prone to clogs. It is prudent to filter the sample before the acquisition. It is also useful to add an in-line filter to the CyTOF sampler before applying a sample.
34. The singlet gate may move up and down the plot along the diagonal day to day and even sample to sample. The gate should encompass a single cluster and, usually, the cluster that is the largest.
35. If two acquisitions were used for a single sample (i.e., after the machine clogs), concatenate for a single FCS file at the point of export after cleaning the data. Export and concatenation can both be accomplished by right-clicking on the population you would like to export in the workspace and selecting "export/concatenate." Many naming and file location selection options are available.
36. For many analyses, the spike-in control should be separated from the sample at this point. To do this in FlowJo, select each population using the two CD45-labeled channels. Export each population separately.

Acknowledgments

This work was supported by grants to M.O. from the American Heart Association (AHA18POST34060251), and from The Conrad Prebys Foundation Award. M.A.M acknowledges support from the American Cancer Society Postdoctoral Fellowship (PF-20-132-01-LIB) and the National Institute of Health (2T32AR064194). C. C. H. Acknowledges support from NIH grants U01 CA224766, R01 CA202987, R01 HL134236, and P01 HL136275. K. L. was supported by grants NIH HL 115232, 145241, and HL088093.

References

1. Weber C, Zernecke A, Libby P (2008) The multifaceted contributions of leukocyte subsets to atherosclerosis: lessons from mouse models. Nat Rev Immunol 8(10):802–815. https://doi.org/10.1038/nri2415
2. Getz GS, Reardon CA (2012) Animal models of atherosclerosis. Arterioscler Thromb Vasc Biol 32(5):1104–1115. https://doi.org/10.1161/ATVBAHA.111.237693
3. Jonasson L, Holm J, Skalli O, Bondjers G, Hansson GK (1986) Regional accumulations of T cells, macrophages, and smooth muscle cells in the human atherosclerotic plaque. Arteriosclerosis 6(2):131–138. https://doi.org/10.1161/01.atv.6.2.131
4. Galkina E, Kadl A, Sanders J, Varughese D, Sarembock IJ, Ley K (2006) Lymphocyte recruitment into the aortic wall before and during development of atherosclerosis is partially L-selectin dependent. J Exp Med 203(5):1273–1282. https://doi.org/10.1084/jem.20052205
5. Hamers AAJ, Dinh HQ, Thomas GD, Marcovecchio P, Blatchley A, Nakao CS, Kim C, McSkimming C, Taylor AM, Nguyen AT, McNamara CA, Hedrick CC (2019) Human monocyte heterogeneity as revealed by high-dimensional mass cytometry. Arterioscler Thromb Vasc Biol 39(1):25–36. https://doi.org/10.1161/ATVBAHA.118.311022
6. Clement M, Raffort J, Lareyre F, Tsiantoulas D, Newland S, Lu Y, Masters L, Harrison J, Saveljeva S, Ma MKL, Ozsvar-Kozma M, Lam BYH, Yeo GSH, Binder CJ, Kaser A, Mallat Z (2019) Impaired autophagy in CD11b(+) dendritic cells expands CD4(+) regulatory T cells and limits atherosclerosis in mice. Circ Res 125(11):1019–1034. https://doi.org/10.1161/CIRCRESAHA.119.315248
7. Campbell LA, Rosenfeld ME (2015) Infection and atherosclerosis development. Arch Med Res 46(5):339–350. https://doi.org/10.1016/j.arcmed.2015.05.006
8. Cole JE, Park I, Ahern DJ, Kassiteridi C, Danso Abeam D, Goddard ME, Green P, Maffia P, Monaco C (2018) Immune cell census in murine atherosclerosis: cytometry by time of flight illuminates vascular myeloid cell diversity. Cardiovasc Res 114(10):1360–1371. https://doi.org/10.1093/cvr/cvy109
9. Winkels H, Ehinger E, Vassallo M, Buscher K, Dinh HQ, Kobiyama K, Hamers AAJ, Cochain C, Vafadarnejad E, Saliba AE, Zernecke A, Pramod AB, Ghosh AK, Anto Michel N, Hoppe N, Hilgendorf I, Zirlik A, Hedrick CC, Ley K, Wolf D (2018) Atlas of the immune cell repertoire in mouse atherosclerosis defined by single-cell RNA-sequencing and mass cytometry. Circ Res 122(12):1675–1688. https://doi.org/10.1161/CIRCRESAHA.117.312513
10. Cochain C, Vafadarnejad E, Arampatzi P, Pelisek J, Winkels H, Ley K, Wolf D, Saliba AE, Zernecke A (2018) Single-cell RNA-seq reveals the transcriptional landscape and heterogeneity of aortic macrophages in murine atherosclerosis. Circ Res 122(12):1661–1674. https://doi.org/10.1161/CIRCRESAHA.117.312509
11. Kim K, Shim D, Lee JS, Zaitsev K, Williams JW, Kim KW, Jang MY, Seok Jang H, Yun TJ, Lee SH, Yoon WK, Prat A, Seidah NG, Choi J, Lee SP, Yoon SH, Nam JW, Seong JK, Oh GT, Randolph GJ, Artyomov MN, Cheong C, Choi JH (2018) Transcriptome analysis reveals nonfoamy rather than foamy plaque macrophages are proinflammatory in atherosclerotic murine models. Circ Res 123(10):1127–1142. https://doi.org/10.1161/CIRCRESAHA.118.312804
12. Zernecke A, Winkels H, Cochain C, Williams JW, Wolf D, Soehnlein O, Robbins CS, Monaco C, Park I, McNamara CA, Binder CJ, Cybulsky MI, Scipione CA, Hedrick CC, Galkina EV, Kyaw T, Ghosheh Y, Dinh HQ, Ley K (2020) Meta-analysis of leukocyte diversity in atherosclerotic mouse aortas. Circ Res 127(3):402–426. https://doi.org/10.1161/CIRCRESAHA.120.316903
13. Butcher MJ, Galkina EV (2012) Phenotypic and functional heterogeneity of macrophages and dendritic cell subsets in the healthy and atherosclerosis-prone aorta. Front Physiol 3:44. https://doi.org/10.3389/fphys.2012.00044
14. Gjurich BN, Taghavie-Moghadam PL, Galkina EV (2015) Flow cytometric analysis of immune cells within murine aorta. Methods Mol Biol 1339:161–175. https://doi.org/10.1007/978-1-4939-2929-0_11
15. McArdle S, Buscher K, Ghosheh Y, Pramod AB, Miller J, Winkels H, Wolf D, Ley K (2019) Migratory and dancing macrophage subsets in atherosclerotic lesions. Circ Res 125(12):1038–1051. https://doi.org/10.1161/CIRCRESAHA.119.315175
16. Hu D, Mohanta SK, Yin C, Peng L, Ma Z, Srikakulapu P, Grassia G, MacRitchie N, Dever G, Gordon P, Burton FL, Ialenti A, Sabir SR, McInnes IB, Brewer JM, Garside P, Weber C, Lehmann T, Teupser D,

Habenicht L, Beer M, Grabner R, Maffia P, Weih F, Habenicht AJ (2015) Artery tertiary lymphoid organs control aorta immunity and protect against atherosclerosis via vascular smooth muscle cell lymphotoxin beta receptors. Immunity 42(6):1100–1115. https://doi.org/10.1016/j.immuni.2015.05.015

17. Finck R, Simonds EF, Jager A, Krishnaswamy S, Sachs K, Fantl W, Pe'er D, Nolan GP, Bendall SC (2013) Normalization of mass cytometry data with bead standards. Cytometry A 83(5):483–494. https://doi.org/10.1002/cyto.a.22271
18. el AD A, Davis KL, Tadmor MD, Simonds EF, Levine JH, Bendall SC, Shenfeld DK, Krishnaswamy S, Nolan GP, Pe'er D (2013) viSNE enables visualization of high dimensional single-cell data and reveals phenotypic heterogeneity of leukemia. Nat Biotechnol 31(6):545–552. https://doi.org/10.1038/nbt.2594
19. Hartmann FJ, Simonds EF, Vivanco N, Bruce T, Borges L, Nolan GP, Spitzer MH, Bendall SC (2019) Scalable conjugation and characterization of immunoglobulins with stable mass isotope reporters for single-cell mass cytometry analysis. Methods Mol Biol 1989:55–81. https://doi.org/10.1007/978-1-4939-9454-0_5
20. Lin JD, Nishi H, Poles J, Niu X, McCauley C, Rahman K, Brown EJ, Yeung ST, Vozhilla N, Weinstock A, Ramsey SA, Fisher EA, Loke P (2019) Single-cell analysis of fate-mapped macrophages reveals heterogeneity, including stem-like properties, during atherosclerosis progression and regression. JCI Insight 4(4). https://doi.org/10.1172/jci.insight.124574
21. O'Flanagan CH, Campbell KR, Zhang AW, Kabeer F, Lim JLP, Biele J, Eirew P, Lai D, McPherson A, Kong E, Bates C, Borkowski K, Wiens M, Hewitson B, Hopkins J, Pham J, Ceglia N, Moore R, Mungall AJ, McAlpine JN, Team CIGC, Shah SP, Aparicio S (2019) Dissociation of solid tumor tissues with cold active protease for single-cell RNA-seq minimizes conserved collagenase-associated stress responses. Genome Biol 20(1):210. https://doi.org/10.1186/s13059-019-1830-0
22. Adam M, Potter AS, Potter SS (2017) Psychrophilic proteases dramatically reduce single-cell RNA-seq artifacts: a molecular atlas of kidney development. Development 144(19):3625–3632. https://doi.org/10.1242/dev.151142
23. Randolph GJ, Inaba K, Robbiani DF, Steinman RM, Muller WA (1999) Differentiation of phagocytic monocytes into lymph node dendritic cells in vivo. Immunity 11(6):753–761. https://doi.org/10.1016/s1074-7613(00)80149-1
24. Szaloki G, Goda K (2015) Compensation in multicolor flow cytometry. Cytometry A 87(11):982–985. https://doi.org/10.1002/cyto.a.22736
25. Stadinski BD, Huseby ES (2020) How to prevent yourself from seeing double. Cytometry A 97(11):1102–1104. https://doi.org/10.1002/cyto.a.24045
26. Burel JG, Pomaznoy M, Lindestam Arlehamn CS, Weiskopf D, da Silva AR, Jung Y, Babor M, Schulten V, Seumois G, Greenbaum JA, Premawansa S, Premawansa G, Wijewickrama A, Vidanagama D, Gunasena B, Tippalagama R, deSilva AD, Gilman RH, Saito M, Taplitz R, Ley K, Vijayanand P, Sette A, Peters B (2019) Circulating T cell-monocyte complexes are markers of immune perturbations. Elife 8:e46045. https://doi.org/10.7554/eLife.46045
27. Gonder S, Fernandez Botana I, Wierz M, Pagano G, Gargiulo E, Cosma A, Moussay E, Paggetti J, Largeot A (2020) Method for the analysis of the tumor microenvironment by mass cytometry: application to chronic lymphocytic leukemia. Front Immunol 11:578176. https://doi.org/10.3389/fimmu.2020.578176
28. Leipold MD, Newell EW, Maecker HT (2015) Multiparameter phenotyping of human PBMCs using mass cytometry. Methods Mol Biol 1343:81–95. https://doi.org/10.1007/978-1-4939-2963-4_7
29. Lai L, Ong R, Li J, Albani S (2015) A CD45-based barcoding approach to multiplex mass-cytometry (CyTOF). Cytometry A 87(4):369–374. https://doi.org/10.1002/cyto.a.22640
30. Zivanovic N, Jacobs A, Bodenmiller B (2014) A practical guide to multiplexed mass cytometry. Curr Top Microbiol Immunol 377:95–109. https://doi.org/10.1007/82_2013_335

How the immune system shapes atherosclerosis: roles of innate and adaptive immunity

Payel Roy[1], *Marco Orecchioni*[1] *and Klaus Ley*[1,2]✉

Abstract | Atherosclerosis is the root cause of many cardiovascular diseases. Extensive research in preclinical models and emerging evidence in humans have established the crucial roles of the innate and adaptive immune systems in driving atherosclerosis-associated chronic inflammation in arterial blood vessels. New techniques have highlighted the enormous heterogeneity of leukocyte subsets in the arterial wall that have pro-inflammatory or regulatory roles in atherogenesis. Understanding the homing and activation pathways of these immune cells, their disease-associated dynamics and their regulation by microbial and metabolic factors will be crucial for the development of clinical interventions for atherosclerosis, including potentially vaccination-based therapeutic strategies. Here, we review key molecular mechanisms of immune cell activation implicated in modulating atherogenesis and provide an update on the contributions of innate and adaptive immune cell subsets in atherosclerosis.

Major adverse cardiovascular events (MACEs). A composite end point frequently used in cardiovascular research that may include clinical complications such as myocardial infarction, stroke, heart failure, need for coronary revascularization, recurrent angina and death from cardiovascular disease.

Low-density lipoprotein (LDL). A group of apolipoprotein B-containing macromolecular lipid carriers in the blood that transport cholesterol and triglycerides from the liver to other body tissues.

[1]*Division of Inflammation Biology, La Jolla Institute for Immunology, La Jolla, CA, USA.*

[2]*Department of Bioengineering, University of California, San Diego, San Diego, CA, USA.*

✉*e-mail: klaus@lji.org*

https://doi.org/10.1038/s41577-021-00584-1

The major cardiovascular diseases, including coronary artery disease, myocardial infarction, stroke and peripheral artery disease, dominate death and disability statistics globally. Atherosclerosis, which is a common pathology underlying many cardiovascular diseases, is characterized by the accumulation of lipid-laden, immune cell-rich plaques known as atheromata in large and medium-sized arteries. Advanced plaques can rupture or erode, causing thromboses that occlude arteries and obstruct blood flow, leading to an array of life-threatening clinical manifestations known as major adverse cardiovascular events (MACEs). Disease pathogenesis is, in part, linked to dyslipidaemia and hypercholesterolaemia, which can be triggered by genetic factors, diet and lifestyle choices, and to metabolic disorders such as obesity and type 2 diabetes. Current therapies for atherosclerosis, such as statins or inhibitors of the serine protease PCSK9, aim to control levels of low-density lipoprotein (LDL) cholesterol in the blood[1]. However, despite the effectiveness of these therapies in reducing LDL cholesterol levels to guideline recommendations, MACEs are reduced by less than 50%[2]. Some of the residual risk is thought to be inflammatory in nature[3]. The Canakinumab Anti-inflammatory Thrombosis Outcome Study (CANTOS), which tested the effects of immunomodulation through IL-1β inhibition, led to successful reduction of MACEs and provided the first major evidence for the feasibility of treating the inflammatory component of atherothrombosis[4]. Colchicine, as given in the Colchicine Cardiovascular Outcomes Trial (COLCOT)[5], also reduced MACEs. In both studies, however, systemic immune suppression resulted in a higher incidence of infections in the treatment arm than in the placebo group, which led to increased serious adverse events and even fatality. This underlines the need to develop immune-targeted interventions that can modulate atherosclerosis more precisely, and are safe, durable and efficacious. Selectively targeting specific components of the immune network that promote atherogenesis will require an in-depth mechanistic knowledge of the cellular and molecular mechanisms that drive atherosclerosis.

Role of immune cells in atherogenesis

The complex and multifactorial aetiology of atherosclerosis involves contributions from both metabolic and immune mediators. Atherogenesis begins with endothelial dysfunction in susceptible regions of arteries that are characterized by disturbed blood flow (predilection sites). This allows the entry of cholesterol-rich, apolipoprotein B (APOB)-containing lipoproteins into the subendothelial space, where they are retained through interactions with intramural extracellular matrix components. Oxidative, enzymatic and chemical modifications of the trapped lipoproteins activate endothelial cells and vascular smooth muscle cells (VSMCs). Chemokine signalling and expression of adhesion molecules trigger a cascade of immune cell influx, leading to the nucleation of a lesion that can develop to form an atherosclerotic plaque. Both innate

Atherothrombosis
The state of pathological complication when erosion or rupture of atherosclerotic plaques leads to thrombosis or clot formation, resulting in stroke or myocardial infarction.

Atherogenesis
The process of atherosclerotic plaque formation in the intimal layers of the artery, which is mediated by chronic inflammation and lipid deposition in the vessel walls.

Apolipoprotein B
(APOB). The amphipathic protein backbone of most lipid-transport particles in the plasma, including very-low-density lipoprotein, low-density lipoprotein and chylomicrons.

Efferocytosis
The process by which apoptotic cells are engulfed and cleared by phagocytic cells.

Classical monocytes
Ly-6C$^+$ monocytes in mice and the CD14^{hi}CD16$^-$ subset in humans. They are CCR2^{hi}CX3CR1^{low} and are important mediators of tissue inflammation.

Non-classical monocytes
Ly-6C$^-$ monocytes in mice and the CD14^{low}CD16$^+$ subset in humans. They are CCR2^{low}CX3CR1^{hi} and patrol the vessel walls to maintain endothelial integrity and homeostasis.

and adaptive immune cells, particularly dendritic cells (DCs), macrophages and T cells, dominate the cellular landscape of the evolving plaque.

Early lesions known as fatty streaks are characterized by cholesteryl ester-laden 'foam cells', which accumulate owing to unregulated uptake of native and modified lipoproteins, predominantly by macrophages and VSMCs. More-advanced lesions contain a core region of activated immune cells, cholesterol crystals and extracellular lipids, surrounded by a fibrous cap of VSMCs and collagen. Recent studies using lineage-tracing and transcriptomic analyses have revealed that the role of VSMCs in atherogenesis is not limited to their ability to produce extracellular matrix for cap formation[6,7]. Cholesterol accumulation in VSMCs has been demonstrated in sections of human coronary arteries and mouse aortas[6]. Single-cell transcriptomic studies in mouse models suggest that VSMC-derived foam cells constitute about half of all foam cells[8]. VSMCs exhibit enormous phenotypic plasticity[9,10] and can acquire lesion-promoting macrophage-like features[11], as well as plaque-stabilizing fibroblast-like features[10].

Over time, excessive death of immune cells in the lesion core puts strain on their removal by efferocytosis. Insufficient efferocytosis results in secondary necrosis and the formation of a necrotic core, in addition to primary necrosis[12]. Pro-inflammatory forms of programmed cell death such as necroptosis and pyroptosis also occur in the vessel wall[12] and may contribute to atherosclerosis progression. Both growth and stability of the progressing plaques can differentially affect disease outcome. Plaque growth is driven by a net positive balance between continued recruitment of immune cells and macrophage proliferation versus apoptosis and efflux. Plaques are rendered unstable by increased inflammatory activity, a growing necrotic core and thinning of the fibrous cap, owing to decreased collagen synthesis and increased secretion of matrix metalloproteinases. Calcifications in the fibrous cap concentrate mechanical stresses[13], promoting rupture or erosion of these plaques. Exposure of the thrombogenic material in the plaque to blood platelets and coagulation factors leads to the formation of a thrombus that can immediately block blood flow, resulting in end organ damage.

Flow cytometry, mass cytometry and sequencing-based multidimensional single-cell molecular profiling approaches have resolved the cellular compositions of atherosclerotic plaques and of the arterial wall at unprecedented resolution[8,14,15]. An anti-inflammatory network of immune cells and resolving factors attempts to mitigate the collateral damage associated with immune cell activation[16,17], but chronic inflammation prevails during plaque progression. Thus, interventions that can selectively inhibit atherosclerosis-associated maladaptive inflammation or boost anti-atherogenic regulatory processes have the potential to improve patient outcome through decreased progression or enhanced resolution of atherosclerosis. For example, pioneering work in mouse models of atherosclerosis has established the prophylactic potential of peptide vaccines to boost regulatory T (T_{reg}) cell responses and prevent atherosclerosis[18]. Tolerogenic vaccination holds promise as an efficient, durable and relatively inexpensive approach to induce protective adaptive immunity in patients with atherosclerosis[19]. If successful, it could save the lives of millions of at-risk individuals around the world.

Here, we provide a succinct overview of the key innate and adaptive immune components that orchestrate inflammation in atherosclerosis, and discuss their activation by endogenous danger signals and microbial triggers. The effects of the immune system or immune mediators on non-immune cells in atherosclerotic plaques are outside the scope of this Review. As much of our current mechanistic understanding of the immune-dependent processes in the vessel wall is based on preclinical data, we focus primarily on recent studies in mouse models of atherosclerosis, but note that these have limitations with respect to human disease (BOX 1).

Innate immune cells

A wide range of innate immune cells, including macrophages, DCs, monocytes, mast cells and neutrophils, are relevant to atherosclerosis progression[20]. Recent studies have also highlighted the involvement of natural killer cells and non-cytotoxic innate lymphoid cells (BOX 2).

Monocytes and macrophages. In healthy arteries, monocytes are rare. Hypercholesterolaemia promotes the proliferation of haematopoietic stem and progenitor cells (BOX 3), leading to systemic monocytosis[21]. Subendothelial accumulation of lipoproteins and chemokines triggers an influx of monocytes into the vessel wall[22] (FIG. 1). The chemokine receptor CCR2 and its main ligand CCL2 regulate the recruitment of classical monocytes[23,24], which dominate monocyte influx in atherosclerosis[21,25]. Plaque monocytes differentiate into macrophages and monocyte-derived DCs that can present antigens to T cells[26]. Most non-classical monocytes do not transmigrate, instead having a characteristic 'patrolling' behaviour and a role in maintaining vascular endothelial homeostasis[27,28]. Although some non-classical monocytes can enter the atheroma, their extravascular function remains unclear. In mouse models of atherosclerosis, non-classical monocytes have been shown to have an atheroprotective role[29]. However, in humans, mass cytometry analysis of blood monocyte subtypes showed that a non-classical monocyte subset expressing the carbohydrate marker 6-sulfo LacNAc was enriched in the blood of patients with cardiovascular disease[30].

Monocyte-derived macrophages initiate the uptake and clearance of lipoproteins, resulting in the formation of lipid-rich foam cells[25]. Recent studies involving lipidomic and transcriptomic analyses of lipid-loaded macrophages have revealed an enrichment of anti-inflammatory genes in cells with foamy characteristics, suggesting that foam cell formation by itself is not pro-inflammatory[15,31,32]. However, increasing oxidative stress in the artery wall and an accumulation of modified lipoproteins may overwhelm or rewire macrophage metabolism[25,33] (BOX 4). This triggers an inflammatory cytokine and chemokine cascade that leads to pro-atherogenic immune cell infiltration and activation.

Box 1 | **Limitations of mouse models of atherosclerosis**

Coronary artery disease

The commonly used mouse models of atherosclerosis, *Ldlr*$^{-/-}$ and *Apoe*$^{-/-}$ mice, do not develop coronary artery disease. Spontaneous atherothrombotic events resembling myocardial infarction or stroke do not occur in these mice. However, several mouse strains that have mutations in the high-density lipoprotein receptor SR-BI or its adaptor protein PDZK1 (REFS[153,154]) do develop occlusive coronary arterial atherothrombosis, myocardial infarction, heart failure and premature death. In some of these mice, the disease develops spontaneously when they are fed a standard diet (low fat and low cholesterol), whereas other mice develop the disease only after being fed an atherogenic diet. These mice are available from the Jackson Laboratory (B6;129S-*Scarb1*^{tm1Kri}/J mice and B6;129-*Scarb1*^{tm1Kri}*Apoe*^{tm1Unc}/J mice) but are not widely used.

Lipoproteins

Blood lipoprotein profiles in mice are characterized by high levels of high-density lipoprotein, low levels of low-density lipoprotein, and some very-low-density lipoprotein and intermediate-density lipoprotein[155]. By contrast, humans have high levels of low-density lipoprotein. Total cholesterol levels in *Apoe*$^{-/-}$ mice fed a Western-style (high-fat, high-cholesterol) diet are extremely high (~1,000 mg dl^{-1}), levels that are not observed in humans except in cases of genetic familial hypercholesterolaemia. Also, the genetic background of the mice matters. Most atherosclerosis studies are conducted in a single mouse strain (C57BL/6), which cannot capture the genetic diversity seen in humans. In *Apoe*$^{-/-}$ mice on the C57BL/6 background, atherosclerotic lesions are larger in female mice than in male mice, whereas premenopausal female humans are protected from atherosclerosis[156].

Genetics

In C57BL/6 knockout mice, contamination with 129/Sv DNA near the targeted gene is possible if embryonic stem cells from the 129/Sv mouse strain were used to generate the knockout mice. This 129/Sv genetic contamination can affect the susceptibility to atherosclerosis[157]. Genetic diversity is known to modulate the response to antigens and atherosclerosis-relevant stimuli within a spectrum of responses ranging from pro-inflammatory to anti-inflammatory[158]. For example, a study of more than 100 inbred mouse strains from the hybrid mouse diversity panel found very large differences in atherosclerosis susceptibility and sex distribution[159].

Evolution of the immune system

Some cytokines, chemokines and immune receptors (for example, CXCL8 and CXCR1) are not conserved between mice and humans because the immune systems of both species are under intense evolutionary pressure. Mice represent a simplified model system for antigen presentation and recognition; whereas C57BL/6 mice have just one MHC class II protein with one allele (I-Ab), humans express several alleles of more than 10,000 MHC class II variants.

Diet and gut microbiota

The levels of cholesterol in different diets (chow, Western, high-cholesterol or cholate-containing diets) used for experimental studies of atherosclerosis can directly affect the phenotype and proliferation of T cells[78,160,161]. Furthermore, mice are housed in specific-pathogen-free facilities, which influences immune cell activation, differentiation and formation of their antigenic repertoire in response to microbial ligands[162]. The gut microbiota of mice differs between animal facilities and even between cages in the same facility[162]. The gut microbiota can have large effects on atherosclerosis development and progression[163,164].

Lipid-engorged arterial macrophages that undergo cell death are initially cleared by other macrophages through efferocytosis[34]. When efferocytes become overwhelmed, they release pro-inflammatory cellular and lipid contents[34]. Plaque-resident macrophages cannot migrate much[35,36] and undergo local proliferation[37]. The retention of macrophages in the artery wall is promoted by a combination of local and systemic signals that include adhesion molecules (such as vascular cell adhesion molecule 1 (VCAM1) and platelet endothelial cell adhesion molecule 1)[35] and neural guidance factors such as netrin 1 and semaphorin 3E that block the chemotaxis of macrophages towards CCL19 and CCL21 (REFS[38,39]). Intriguingly, high-density lipoprotein has been reported to promote the egress of macrophages from plaques through a process mediated by CCR7, the receptor for CCL19 and CCL21 (REFS[40,41]). It is controversial whether the egress of plaque macrophages to the lumen or lymphatic system can affect atherosclerosis progression[21,42,43].

In general, plaque macrophages can have pro-inflammatory or anti-inflammatory features, but do not follow a classical M1 or M2 classification[44]. Certain subsets of foam cells and resident macrophages can promote the resolution of inflammation and healing[31,32]. By contrast, inflammatory macrophages have increased levels of mRNA encoding MHC class II molecules, Fcγ receptor I (also known as CD64), the co-stimulatory molecules CD80 and CD86, nitric oxide synthase 2 (NOS2) and pro-inflammatory cytokines such as IL-6, tumour necrosis factor (TNF) and IL-1β. Single-cell studies, involving mass cytometry and single-cell RNA sequencing (scRNA-seq), have defined at least five distinct subsets of macrophages in mouse aortas[8]. The inflammatory macrophage subset, which is derived from circulating monocytes, is the main population involved in atherogenesis and is not found in healthy arteries. Inflammatory macrophages are characterized by increased expression of cytokine and chemokine transcripts as well as inflammasome components[45,46]. A second subset, known as type I interferon-inducible cells[47], is also derived from monocytes, and is defined by increased expression of numerous interferon-inducible genes, including *Ifit3*, *Irf7* and *Isg15* (REF.[48]). This subset is found only during atherosclerosis and has a pro-inflammatory role through type I interferon production[31,45]. Foam cells expressing TREM2 (triggering receptor expressed on myeloid cells 2)[45] have been identified in mice with both early and advanced atherosclerotic lesions but not healthy arteries. TREM2^{hi} macrophages can be derived either from circulating monocytes or from embryonic precursors[8]. In addition to *Trem2*, they express *Cd9*, *Ctsd*, *Fabp4* and *Abcg1*, which have roles in lipid metabolism, cholesterol efflux and oxidative phosphorylation. Labelling with the neutral lipid dye BODIPY showed that lipid-filled macrophage-derived foam cells are TREM2^{+} cells[31]. Although these macrophages are prevalent in atherosclerotic arteries, they are characterized by low levels of expression of inflammatory genes and are postulated to have a role in mitigating vascular inflammation[31,32]. This subset of TREM2^{hi} macrophages resembles the CD11c–YFP^{+} macrophages identified in intact mouse aortas by intravital microscopy[36]. Recently, an aortic intima-resident macrophage (Mac^{AIR}) subset was identified[49]. These macrophages are derived from blood monocytes, which are mostly seeded shortly after birth and then maintained by local proliferation. Under the inflammatory conditions associated with plaque progression, Mac^{AIR} cells also differentiate from recruited monocytes. Fate-mapping studies with Mac^{AIR} depletion models have shown a pro-atherogenic role of these cells in early lesion development. Mac^{AIR} cells constitutively express *Il1b* mRNA but seem to lack other inflammasome-associated genes. They have similarities with the transcriptomic profile of cavity macrophages,

High-density lipoprotein
This is the only group of lipoproteins that do not contain apolipoprotein B. High-density lipoprotein is involved in reverse cholesterol transport, delivering excess cholesterol from tissues to the liver.

Box 2 | **Role of innate lymphoid cells and unconventional T cells in atherosclerosis**

Innate lymphoid cells (ILCs) include the cytotoxic natural killer (NK) cells and the non-cytotoxic group 1, group 2 and group 3 ILCs (ILC1s, ILC2s and ILC3s). Limited numbers of NK cells have been detected in mouse aortas before and after the onset of atherosclerosis[8] and in human lesions[165]. Antibody-based depletion of NK cells and adoptive transfer studies suggest that perforin and granzyme B produced by NK cells promote lesion growth in atherosclerotic mice[166]. However, a direct effect of NK cells on cholesterol-induced atherosclerosis was not confirmed when more precise NK cell loss-of-function genetic models were used[167]. ILC1s, ILC2s and ILC3s share transcription factor and cytokine expression profiles with T helper 1 (T_H1) cells, T_H2 cells and T_H17 cells, respectively. As a result of this overlap, mechanistic studies exploring the relative contributions of ILCs and T_H cells in atherosclerosis are scarce[73]. Recently, genetic depletion of ILC2s and transfer of specific ILC2 subsets have shown their direct role in mediating atheroprotection[89].

Mass cytometry and single-cell RNA-sequencing studies have recently shown that $CD8^+$ T cells accumulate in atherosclerotic mouse aortas[8] and are significantly enriched in human plaques[14]. Interferon-γ, perforin and granzyme B production by $CD8^+$ T cells promote vascular inflammation and lesion growth through apoptotic cell lysis[168,169] and induction of monopoiesis[170]. Although some $CD8^+$ T cells recognize apolipoprotein B-derived epitopes[171], the antigen specificities of most $CD8^+$ T cells in atherosclerotic plaques remain unknown[172]. Minor populations of invariant NK T cells have also been described in atherosclerosis[173]. These cells can be activated by self-antigens or microbial glycolipid antigens presented by CD1d molecules on antigen-presenting cells or by cytokines such as IL-12 and IL-18, which are secreted by Toll-like receptor-stimulated antigen-presenting cells. T_H cell-associated cytokines and cytotoxic molecules produced by activated NK T cells promote inflammation and necrotic core formation[174]. γδ T cells also respond to innate immune stimulants and produce inflammatory cytokines. Although γδ T cells are found in mouse atherosclerotic lesions[175], genetic deficiency of γδ T cells had no effect on the development of diet-induced atherosclerosis in mice[176].

a subset that was found in a single-cell meta-analysis study of atherosclerotic aortas[8]. As both Mac^{AIR} cells and cavity macrophages are defined by the expression of *Itgax*, genes encoding MHC class II, *Cd226*, *Ccr2* and *Retnla*[8,49], we suspect that they represent the same macrophage subset. In both atherosclerotic and healthy arteries, resident macrophages[8] derived embryonically from $CX3CR1^+$ precursor cells are found[50]. They proliferate locally in the adventitia[37,50] and are characterized by the expression of lymphatic vessel endothelial hyaluronan receptor 1 (LYVE1), MHC class II and macrophage mannose receptor C type 1 (MRC1)[50]. Depletion of arterial $LYVE1^+$ macrophages in animal models suggested that they inhibit collagen production by VSMCs and thereby limit arterial stiffness[51]. Most resident macrophages originate from embryonic progenitors, whereas atherosclerotic aortas have a prevalence of monocyte-derived macrophages[8,49,50]. Although these single-cell transcriptomic analyses of mouse aortas have revealed that plaque macrophages are more heterogeneous than was shown by traditional immunophenotyping (FIG. 1), most of these new macrophage populations have been defined by their gene expression profiles and lack comprehensive characterization of protein markers and functional studies.

Efforts to define immune and non-immune cell populations in human atherosclerotic plaques are also under way[52]. Transcriptomes of human macrophage subsets isolated from carotid endarterectomy specimens showed significant overlap with the macrophage clusters defined in mouse aortas, particularly with respect to the inflammatory macrophages and the $TREM2^+$ foam cell-like macrophages[15]. Another study suggests that some human macrophages acquire wound-healing properties after plaque rupture[14]. Identification of the precise role of each of the aortic macrophage subsets in modulating atherosclerosis, together with the characterization of their functions and crosstalk with other immune cells such as T cells and DCs, will be crucial for understanding their relevance to atherosclerosis progression.

Dendritic cells. Both conventional DCs (cDCs) and plasmacytoid DCs (pDCs) have been detected in mouse and human arteries and atherosclerotic lesions[53]. $CD103^+$ type 1 cDCs are derived from $FLT3^+$ migratory pre-cDCs[54] and are similar to the lymphoid-resident $CD8α^+$ DCs[55]. $CD11b^+$ type 2 cDCs are the most abundant aortic DC subset and also include monocyte-derived DCs. An inflammatory $CD11c^+CD11b^+CD8α^-$ CCL17-secreting DC subset, expressing high levels of the co-stimulatory molecules CD40, CD80 and CD86, has been described specifically in atherosclerotic lesions but not in healthy vessels[56].

cDCs take up antigens in the vessel wall and activate antigen-specific naive T cells in draining lymph nodes or the spleen. Although DCs are present in the intimal layers of normal mouse and human arteries, their numbers are greatly increased and their phenotypes are markedly altered under atherosclerotic conditions[53]. Multiphoton live-cell imaging of aortic explants from atherosclerotic mice that were incubated with fluorescently labelled transgenic T cells showed significantly increased productive interactions between aortic CD11c–YFP^+ antigen-presenting cells (APCs) and T cells in the vessel wall only in the presence of the cognate antigen[57]. Polyclonal antigen-experienced $CD44^{hi}CD4^+$ T cells isolated from spleens of atherosclerotic mice, but not control mice, vigorously interacted with APCs in explanted aortas of atherosclerotic mice in the absence of any external antigen stimulation, which suggests that the T cells have a productive recall response to unidentified atherosclerosis epitopes[57]. Incubating ex vivo-isolated CD11c–$EYFP^+$ aortic APCs with ovalbumin-specific $CD8^+$ T cells and $CD4^+$ T cells triggered antigen-specific proliferation of both T cell subsets in the presence of ovalbumin, which indicates that aortic DCs can cross-present MHC class I-restricted antigens to $CD8^+$ T cells[58]. It is likely that under non-inflammatory conditions, DCs from healthy arteries present self-antigens in the absence of adequate co-stimulation, leading to T cell tolerance. However, in atherosclerotic lesions, the maturation of plaque DCs in the presence of Toll-like receptor (TLR) agonists, danger-associated ligands and pro-inflammatory cytokines probably favours a pro-atherogenic activation status. In lesions with necrotic cores, CLEC9A-dependent sensing of necrotic cells by $CD8α^+$ DCs promotes atherosclerosis by restraining anti-inflammatory IL-10 production[59]. Accumulation of plaque $CD11b^-CD103^+IRF8^{hi}$ DCs derived from a $CLEC9A^+$ precursor contributes to atherogenesis by promoting pro-inflammatory T cell activity[60].

T cell activation by DCs is dependent on TLR-mediated maturation of DCs. Bone marrow transplantation from *Cd11c*–Cre^+*Myd88*$^{fl/fl}$ mice — which have $CD11c^+$ DCs deficient in the TLR adaptor MYD88 — into

Inflammasome
A multiprotein cytosolic complex containing members of the NOD-like receptor (NLR) family (such as NLRP3) that integrates signals from several pattern recognition receptors and results in the maturation and secretion of the pro-inflammatory cytokines IL-1β and IL-18.

atherosclerotic mice led to decreased recruitment of both effector T cells and T_{reg} cells to atherosclerotic lesions[61]. The overall effect was larger lesion size, which was attributed to increased production of the monocyte-recruiting chemokine CCL2 as a result of the loss of T_{reg} cells[61]. Different cDC subtypes disparately modulate T_{reg} cell homeostasis. Increased generation of pro-inflammatory cytokine transcripts and increased average lesion size in the aortas of atherosclerotic mice depleted of $CD103^+$ DCs were associated with a decrease in both the number of aortic T_{reg} cells and IL-10 secretion, suggesting that $CD103^+$ DCs may be potent inducers of T_{reg} cells[54]. With use of bone marrow transplants and RNA sequencing, a tolerogenic role of $CD11b^+CD11c^+$ DCs in transforming growth factor-β (TGFβ)-mediated induction of antigen-specific T_{reg} cells was demonstrated in atherosclerotic mice in which autophagy was disrupted through DC-specific abrogation of *Atg16l1* (REF.[62]). Unlike $CD11b^+$ DCs, autophagy disruptions in $CD8α^+$ cDCs or $CD103^+$ cDCs did not induce a tolerogenic phenotype in these subsets. This showed that the previously reported atheroprotective role of $CD103^+$ DCs is not autophagy dependent, highlighting mechanistic differences in the regulatory network that modulates the tolerogenic properties of different DC subsets[62]. Inflammatory $CD11c^+CD11b^+CD8α^-CCL17^+$ DCs inhibit T_{reg} cell development[56].

Atherosclerotic lesions also contain pDCs, albeit in low numbers. Plaque pDCs respond not only to CpG oligonucleotides of microbial origin but also to extracellular self-DNA complexed with cathelicidin-related antimicrobial peptide (CRAMP; encoded by *Cnlp*, also known as *Camp*), which is the mouse orthologue of human cathelicidin antimicrobial peptide (CAMP; also known as LL-37)[63]. Breaching immune tolerance to self-DNA complexed with CRAMP in these atherosclerotic mice increased antibody titres to double-stranded DNA and promoted atherogenesis through interferon-α (IFNα) production. Abrogation of antigen presentation in pDCs through MHC class II deficiency resulted in atheroprotection, which was independent of IFNα and was associated with inhibition of T helper 1 (T_H1) cell responses and T cell infiltration into lesions[64]. Other studies have reported an anti-atherogenic role of pDCs, associated with indoleamine 2,3-dioxygenase 1 (IDO1)-dependent induction of aortic T_{reg} cells[65,66]. These conflicting results are probably owing to differences in pDC identification, depletion strategies and mouse models of atherosclerosis used.

We conclude that both cDCs and pDCs can modulate atherogenesis in mice, in part through regulation of T cell activation and adaptive immune responses, in an antigen-dependent manner. The exact mechanisms by which DCs exert such control require further elucidation. Studies of DCs in human atherosclerosis have been observational so far, and their functional role remains to be characterized.

Adaptive immune cells

Adaptive immunity is a key modulator of atherosclerosis (FIG. 2). Here, we focus on recent preclinical data that highlight the role of MHC class II-mediated activation of $CD4^+$ T cells by various APCs (BOX 5) and the relative contributions of the most widely studied subclasses of $CD4^+$ T_H cells. The roles of $CD8^+$ T cells and lipid antigen-specific natural killer T cells and γδ T cells in atherosclerosis are summarized in BOX 2. The roles of antibodies and B cells in atherosclerosis have been reviewed elsewhere[67].

T cells are present in the adventitia of healthy arteries of wild-type mice[68] and the arteries of mouse models that are genetically prone to develop atherosclerosis[46]. scRNA-seq studies have detected aortic T cells at all stages of atherosclerotic disease[8]. Chemokine receptors such as CCR5 and CXCR6 mediate the infiltration of T cells into the plaque[69,70]. CCL5, the main ligand for CCR5, is found abundantly in atherosclerotic lesions[70]. Circulating activated platelets produce CCL5 and deliver it to the vessel wall through the formation of platelet–monocyte aggregates[71]. Blocking the binding of CCL5 to CCR5 with monoclonal antibodies to the ligand or receptor resulted in reduced infiltration of effector $CD4^+$ T cells into aorta explants[70]. CXCL16, which is expressed as both soluble and membrane-bound proteins in mouse and human lesions, is produced by several types of immune cell, including macrophages, DCs, aortic VSMCs and endothelial cells[72]. The CXCR6-dependent chemotaxis of T cells is likely to be mediated by soluble CXCL16 (REF.[69]).

$CD4^+$ T cell subpopulations can differentially impact atherosclerosis progression through immune activation or immune suppression, or by providing help to B cells for antibody production[73]. Co-stimulatory molecules and immune checkpoint proteins have been reported to have a pivotal role in modulating atherogenesis[74].

Box 3 | Role of haematopoiesis in atherosclerosis

The development of atherosclerosis depends on the haematopoietic supply of immune cells, particularly classical monocytes[25]. In turn, hypercholesterolaemia, stress and inflammation associated with atherosclerosis induce the proliferation and myeloid-biased differentiation of haematopoietic stem and progenitor cells (HSPCs) in the bone marrow, the release of immune cells into the circulation and the diversion of haematopoiesis to extramedullary sites[177]. Mechanistically, cholesterol enrichment in membrane lipid rafts triggers the myeloproliferation of HSPCs through increased sensitivity to IL-3 and granulocyte–macrophage colony-stimulating factor (GM-CSF) and increased RAS–ERK signalling[178]. Proteoglycan-bound apolipoprotein E was shown to regulate cholesterol efflux from HSPCs and inhibit their growth factor-induced hyperproliferative response in a cell-intrinsic manner[179]. Hypercholesterolaemia can also promote a loss of cellular quiescence in HSPCs and skewing towards pro-inflammatory myeloid cell types, a cell-intrinsic effect that was maintained even after return to normocholesterolaemic conditions[180]. During atherosclerosis, myeloid-biased haematopoiesis in the GM-CSF-rich and IL-3-rich environment of the splenic red pulp significantly contributes to the systemic abundance of inflammatory monocytes and neutrophils[181]. Granulocyte colony-stimulating factor-dependent mobilization of HSPCs to the red pulp can be triggered by a high-fat diet and impaired cholesterol efflux, which elicits IL-23 production by splenic dendritic cells and macrophages[182]. Activation of the sympathetic nervous system in acute cardiovascular events, such as myocardial infarction, can also release HSPCs from the bone marrow and drive extramedullary myelopoiesis[183]. Noradrenaline-dependent sympathetic nervous system signalling promotes proliferation of HSPCs and augments plaque inflammation under conditions of chronic stress[184]. Disease-promoting lifestyle choices such as inadequate sleep[185] and insufficient exercise[186] can impact atherosclerosis progression by influencing haematopoiesis[187]. Finally, ageing increases the likelihood of clonal haematopoiesis of indeterminate potential, whereby somatic mutations in *DNMT3A*, *TET2* and *ASXL1* lead to the formation of a genetically distinct subpopulation of haematopoietic cells, which is strongly associated with increased risk of cardiovascular events[188].

The prevailing notion is that regulatory T cells, such as FOXP3$^+$ T_{reg} cells and IL-10$^+$ type 1 regulatory T cells (Tr1 cells), dominate before disease onset but that pro-inflammatory CD4$^+$ effector T cells outcompete them after the onset of atherosclerosis.

T_H1 cells. IFNγ-secreting T_H1 cells, which express the lineage-defining transcription factor T-bet (encoded by *Tbx21*), are the most prominent CD4$^+$ T cell subtype in atherosclerotic plaques. Several lines of evidence, including genetic deficiency of T-bet, IFNγ or its receptor, or

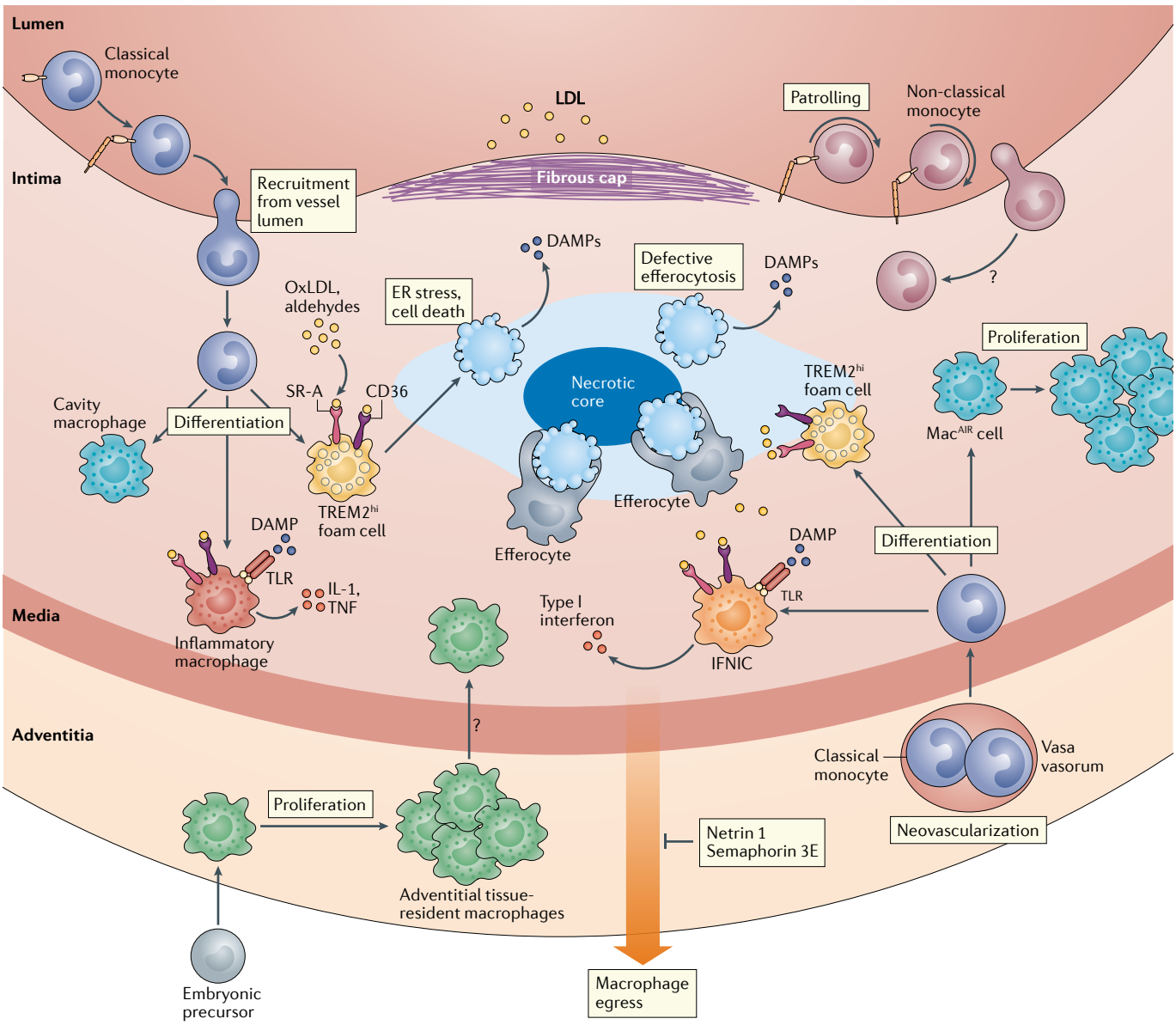


Fig. 1 | **Vascular macrophage populations in mouse atherosclerotic lesions.** The adventitia of the healthy artery wall contains tissue-resident vascular macrophages. The adventitial macrophage subset is derived from embryonic precursors and is maintained by local proliferation. These tissue-resident macrophages help to regulate arterial stiffness, but their role in plaque progression is unknown. The aortic intima-resident macrophage (Mac^{AIR}) subset originates from monocytes that were seeded postnatally and is sustained by local proliferation during the initial phases of atherosclerosis[49]. As disease progresses, Mac^{AIR} cells are replenished from classical monocytes recruited from the artery lumen and, in aged mice, from areas of neovascularization (vasa vasorum)[151,152]. The Mac^{AIR} subset and cavity macrophages[8] are likely to be the same cells. During atherosclerosis, macrophages in the neointima phagocytose apoptotic cells by efferocytosis. Monocyte-derived macrophages can become inflammatory macrophages that produce IL-1, tumour necrosis factor (TNF) and other cytokines and chemokines, which attract more classical monocytes into the intima. Decreased egress and increased retention of macrophages and monocytes in response to neuroimmune guidance cues such as netrin 1 and semaphorin 3E have been proposed to amplify the inflammatory response[21]. Low-density lipoprotein (LDL) from the blood in the vessel lumen can enter the intimal layer, where it can become oxidized (oxLDL) or modified with aldehydes. Modified LDL is taken up by foam cells through scavenger receptors such as CD36 and SR-A. Single-cell studies have defined foam cells as the TREM2^{hi} macrophage subset. Endoplasmic reticulum (ER) stress of efferocytes can lead to cell death, which leads to the release of danger-associated molecular patterns (DAMPs) that are recognized by Toll-like receptors (TLRs) on inflammatory macrophages. If efferocytosis is insufficient, dead macrophages accumulate in the necrotic core. Type I interferon-inducible cells (IFNICs) are likely to produce type I interferons. IFNICs are replenished by differentiation from circulating classical monocytes. Non-classical monocytes patrol the endothelium of the blood vessel and can enter the plaque, but their further fate remains unknown. TREM2, triggering receptor expressed on myeloid cells 2.

Box 4 | **Modulation of atherogenesis by immune cell metabolic pathways**

Hypoxia and lipid imbalances can influence immune cell phenotypes through metabolic rewiring[189]. Pro-inflammatory macrophages and T cells favour catabolic processes such as aerobic glycolysis, fatty acid synthesis, the pentose phosphate pathway and glutaminolysis. Indeed, increased glucose uptake by plaque macrophages forms the basis of ^{18}F-fluorodeoxyglucose positron emission tomography of vascular inflammation[190]. By contrast, anti-inflammatory macrophages preferentially use anabolic processes such as mitochondrial fatty acid β-oxidation, oxidative phosphorylation and tryptophan metabolism. Molecular disruptions that shift the balance from aerobic glycolysis to mitochondrial oxidative respiration can lead to anti-inflammatory polarization of immune cells and suppress plaque inflammation[191]. Leptin, an adipocyte-secreted hormone associated with obesity, causes pro-inflammatory reprogramming of T cells through upregulation of glucose metabolism[192]. Leptin deficiency inhibits atherosclerosis through improved regulatory T (T_{reg}) cell function[193]. A molecular connection between glucose metabolism and monocytic inflammation was also found in human atherosclerosis[194]. Metabolomic profiling of 159 human carotid plaques identified inflammatory metabolic signatures such as increased glycolysis, increased amino acid utilization and impaired fatty acid β-oxidation in those plaques that were prone to rupture[195].

Intracellular cholesterol metabolism can also regulate immune cell activation in atherosclerosis. In mouse models, cholesterol feeding not only had immediate effects such as inflammasome activation but also triggered epigenetic reprogramming in bone marrow myeloid precursors that allowed for the maintenance of metabolic adaptations even after diet reversal[196]. Indeed, mechanistic studies involving pharmacological inhibitors have shown that mevalonate, an intermediate in the cholesterol biosynthesis pathway, can imprint a memory of inflammatory insults in myeloid cells[197], a phenomenon known as trained immunity[198]. So far, the trained immunity hypothesis of atherosclerosis is lacking a specific molecular mechanism that, when suppressed, would have an impact on the disease process. Other reports have suggested an anti-inflammatory effect of cholesterol uptake, whereby inflammatory gene expression in macrophages is suppressed, either through sterol-dependent signalling[32] or through impairment of the pentose phosphate pathway[199]. Similar effects were observed for T cells, in which excess cholesterol was shown to favour T_{reg} cell differentiation[160,200]. By contrast, cholesterol enrichment in lipid rafts may facilitate assembly of the immune synapse and boost T cell receptor signalling in effector T cells, thereby rendering them more resistant to T_{reg} cell-mediated suppression[160]. Cholesterol-induced dysregulation of lipid uptake and efflux pathways in dendritic cells can modulate inflammation and atherogenesis through regulation of T cell activation and polarization[201].

Atherogenesis can also be modulated by amino acid metabolism. Indoleamine 2,3-dioxygenase 1 (IDO1), a key rate-limiting enzyme in tryptophan degradation, mediates the formation of immunomodulatory metabolites such as kynurenine and anthranilic acid[202]. Inhibitor-mediated antagonism[203] or global genetic abrogation[204] of IDO1 activity has been shown to accelerate vascular inflammation and lesion formation in mouse models of atherosclerosis. However, whereas anthranilic acid-induced IL-10 production, particularly from B cells, reduced plaque inflammation[204], kynurenine, which correlates with poor prognosis in patients with cardiovascular disease, was shown to promote atherosclerosis in hypercholesterolaemic mice through inhibition of IL-10 production[205].

administration of exogenous IFNγ, have established a pro-atherogenic role of T_H1 cells in promoting lesion development and plaque instability[75]. TNF, which is expressed by multiple effector T cell types, promotes lesion development[76], in part by facilitating leukocyte transmigration across vascular endothelial cells[77]. IFNγ secreted in the developing lesion induces uptake of oxidized LDL (oxLDL) and foam cell formation, but foam cells have little inflammatory potential[31,32]. IFNγ can also polarize macrophages towards a pro-inflammatory phenotype and promote VSMC proliferation[73,76].

T_{reg} cells. *Foxp3*-expressing CD25$^+$ T_{reg} cells and the anti-inflammatory cytokines IL-10 and TGFβ correlate negatively with atherogenesis[17]. In atherosclerosis-prone *Foxp3*–GFP$^+$ reporter mice, fewer aortic T_{reg} cells were found after prolonged administration of a cholesterol-rich diet than a normal diet[78]. However, the number of GFP$^+$ cells reported in that study is higher than the entire leukocyte content of atherosclerotic mouse aortas, which suggests that there might have been contamination from adjacent adipose tissue. Defective recruitment of T_{reg} cells in atherosclerosis-prone mice fed a cholesterol-rich diet was associated with reduced selectin-mediated binding to the endothelial lining[78]. In transgenic mice, diphtheria toxin-mediated depletion of T_{reg} cells under hypercholesterolaemic settings led to increased lesion size[79]. Surprisingly, T_{reg} cell depletion also induced pro-atherogenic shifts in the plasma cholesterol profile through impaired lipoprotein catabolism in the liver, which was linked to reduced expression of the very low density lipoprotein-binding protein sortilin 1 (REF.[79]). This complicates interpretation of the results, and the direct effect of T_{reg} cell depletion on vascular inflammation remains to be investigated.

Three different mouse models of atherosclerosis regression — delivery of antagonistic *Apob*-antisense oligonucleotide[80], inducible genetic deletion of microsomal triglyceride transfer protein ('Reversa' mice)[41] and surgical transplantation of atherosclerotic aortic arches[81] — suggested that the regression of established plaques is associated with a restoration of plaque T_{reg} cell numbers, mostly peripherally induced T_{reg} cells that do not express neuropilin 1 (NRP1)[82]. It is not clear whether these NRP1$^-$ peripherally induced T_{reg} cells are derived from local activation of naive T cells in the vessel wall or are generated in secondary lymphoid organs upon encountering aorta-derived tolerogenic DCs. However, these data indicate that plaque regression is associated with an anti-inflammatory microenvironment that can support the local expansion of tolerogenic immune cells[82]. T_{reg} cell enrichment in regressing plaques, but not lipid lowering alone, led to macrophage-mediated dampening of inflammation and tissue repair through improved efferocytosis and enhanced synthesis and sensing of pro-resolving lipid mediators such as lipoxin A4 and the resolvins RvD1 and RvD6 (REF.[82]).

T_H2 cells. Transcripts of T_H2 cell-related cytokines, such as IL-4, IL-5 and IL-13, are expressed in mouse plaques. IL-4 seems to have variable effects on atherogenesis[83,84], whereas IL-5 and IL-13 are atheroprotective[73]. IL-5 can induce the production of oxLDL-targeting IgM antibodies by B1 cells[85]. Antibody-dependent clearance of apoptotic cells dampens inflammation and prevents necrotic core formation[86]. An inverse relationship between IL-5 and atherosclerosis is evident particularly in those areas of the artery where blood flow is oscillatory and shear stress is low[87]. IL-13 deficiency, under conditions of hypercholesterolaemia, resulted in larger lesions and accelerated atherosclerosis progression[88]. Conversely, low-dose IL-13 stabilized existing plaques through increased collagen formation, inhibition of VCAM1-dependent monocyte infiltration and induction of M2-type macrophages that clear oxLDL more efficiently than pro-inflammatory M1-type macrophages[88]. Recently, group 2 innate lymphoid cells, which are an innate cell source of IL-5 and IL-13, have also been reported in atherosclerosis[8,89] (BOX 2). Further studies involving cell-type-specific deficiency of T_H2 cell-related

Trained immunity
The reprogramming of innate immune cells, mostly through epigenetic modifications, that creates a 'memory' of the initial insult and generates long-lasting innate immunity to specific triggers.

Very-low-density lipoprotein
The primary transporter of endogenous triglycerides from the liver to other tissues in the body.

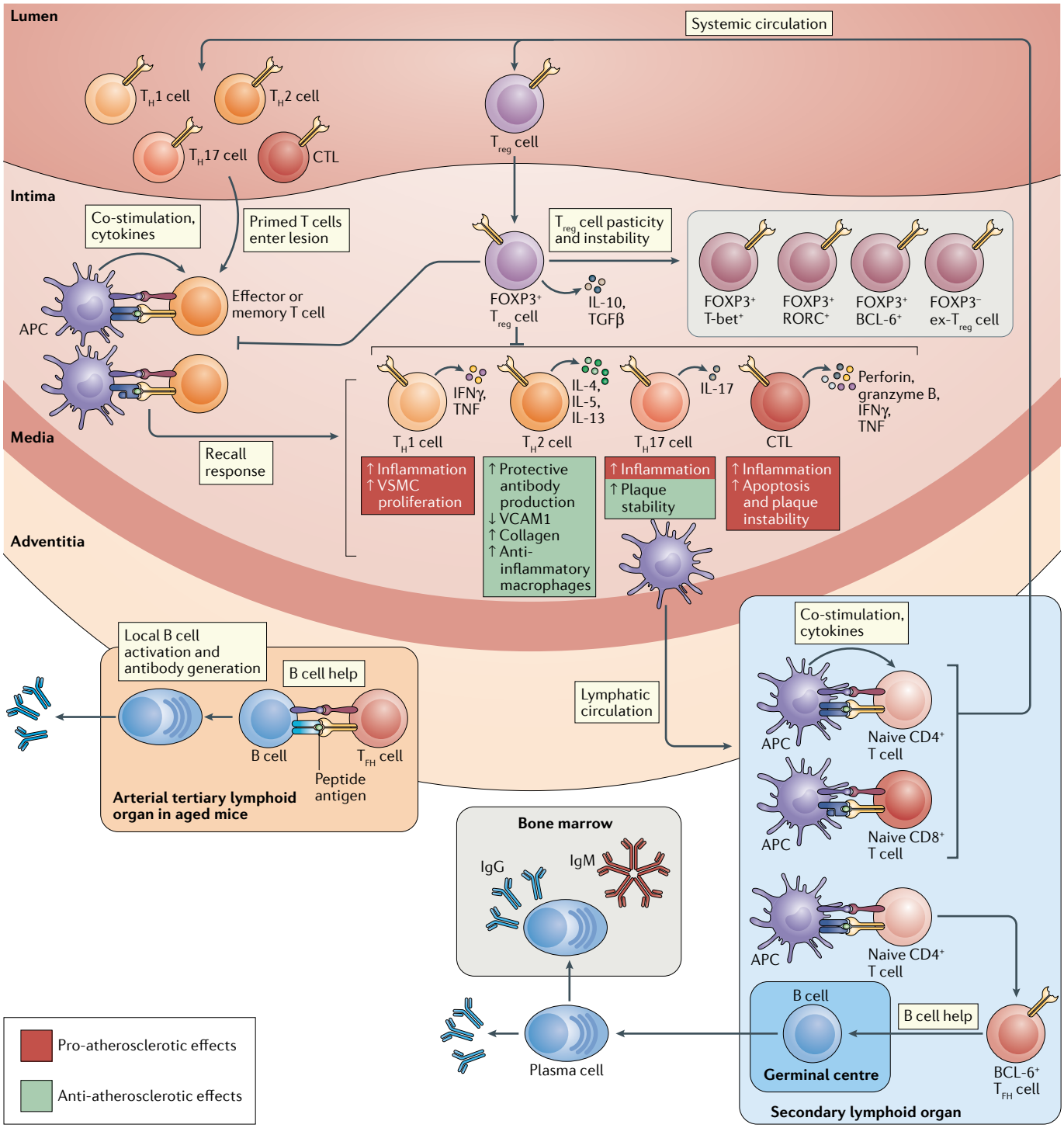


Fig. 2 | **Adaptive immune cells in atherosclerosis.** T cell subsets encounter atherosclerosis-associated antigens in secondary lymphoid organs, where antigen-presenting cells (APCs) arriving through the lymphatics present cognate peptides on MHC class I molecules to naive $CD8^+$ T cells and on MHC class II molecules to naive $CD4^+$ T cells. The nature of the co-stimulatory molecules and cytokines determines T cell polarization. $BCL\text{-}6^+$ T follicular helper (T_{FH}) cells stay in the secondary lymphoid organs and provide help to B cells in germinal centres. Other $CD4^+$ T cells — T helper 1 (T_H1) cells, T_H2 cells, T_H17 cells and regulatory T (T_{reg}) cells — as well as $CD8^+$ cytotoxic T lymphocytes (CTLs), enter the systemic circulation and home back to the plaque, where APCs initiate a recall response and trigger cytokine secretion (by $CD4^+$ T_H cells) or killing (by $CD8^+$ CTLs). The main cytokines produced are listed for each cell type, together with their pro-atherosclerotic effects (red) and anti-atherosclerotic effects (green). T_{reg} cells regulate T cell activation and responses through the production of IL-10 and transforming growth factor-β (TGFβ) and through contact-dependent mechanisms. $FOXP3^+$ T_{reg} cells are plastic (acquiring other transcription factors in addition to FOXP3) or unstable (losing FOXP3 expression and becoming ex-T_{reg} cells). B cell-derived plasma cells enter the bone marrow and produce IgM or, if class-switched, IgG antibodies to atherosclerosis antigens. In the adventitia of arteries with advanced atherosclerosis, arterial tertiary lymphoid organs form, which contain germinal centres, T_{FH} cells and B cells that mature to antibody-secreting plasma cells. IFNγ, interferon-γ; TNF, tumour necrosis factor; VCAM1, vascular cell adhesion molecule 1; VSMC, vascular smooth muscle cell.

Peripherally induced T_{reg} cells
Regulatory T (T_{reg}) cells that arise outside the thymus from conventional T cells that acquire FOXP3 expression in response to various stimuli. Unlike thymus-derived T_{reg} cells, peripherally induced T_{reg} cells do not express neuropilin 1.

Tertiary lymphoid organs
Lymphoid structures that form in peripheral tissues in response to chronic inflammation and that have functional and morphological similarities with secondary lymphoid organs.

cytokines are needed to clarify the relative contributions of innate lymphoid cells and T cells.

T_H17 cells. The role of T_H17 cells in atherosclerosis remains controversial, partly owing to the context-dependent plasticity of T_H17 cell function. The numbers of T cells expressing IL-17 or the T_H17 cell lineage-defining transcription factor RORγt (encoded by *Rorc*) were positively correlated with atherosclerosis in mice[90]. Furthermore, ablation of IL-17A–IL-17 receptor A (IL-17RA) signalling in atherosclerotic mice using neutralizing antibodies to IL-17A, bone marrow transplants from *Il17a*$^{-/-}$ mice or adenovirus-mediated in vivo expression of soluble IL-17RA inhibited pro-inflammatory cytokine and chemokine production, leukocyte infiltration and plaque formation[69,90–92]. However, a plaque-stabilizing role of IL-17A has also been documented[93,94]. Deficiency of TRIM21, a ubiquitin E3 ligase involved in inflammation, leads to an increased T_H17 cell response in hypercholesterolaemic settings[95]. These TRIM21-deficient T_H17 cells were associated with a non-pathogenic gene expression profile and led to the formation of larger but stabler plaques[95].

T follicular helper cells. A largely pro-atherogenic role of T follicular helper (T_{FH}) cells has been reported in atherosclerosis-prone mice[73]. T_{FH} cell responses regulate humoral immunity, antibody isotype switching and antibody affinity maturation. Exaggerated T_{FH} cell activity can trigger autoimmunity[96]. In models of atherosclerosis and autoimmune lupus, dyslipidaemic conditions were shown to augment lupus-related autoantibody formation and the differentiation of pro-inflammatory CXCR3$^+$ T_{FH} cells through production of IL-27, possibly released from TLR4-activated CD8α$^-$CD11b$^+$ DCs[97]. An immunomodulatory role of CD8$^+$ regulatory T cells[98] and marginal zone B cells[99] in restraining T_{FH} cell differentiation and germinal centre reactions has been reported, but the detailed roles of those cell types in atherosclerosis development remain unknown.

T cell plasticity. Atherosclerosis-related T cells are not only heterogeneous but also plastic. This means that one CD4$^+$ T cell subtype can acquire the transcriptomic and phenotypic properties of another subtype. Recent studies, using in-depth gene expression analysis and fate-mapping transgenic mouse models, have convincingly illustrated that atherogenic conditions can trigger functional plasticity of *Foxp3*-expressing T_{reg} cells such that they begin to express effector T cell-specific lineage-defining transcription factors, cytokines and chemokine receptors[70,100,101]. In some of these studies, T_{reg} cells became unstable and lost *Foxp3* expression to become ex-T_{reg} cells[101]. The overall effect is a progressive loss of T_{reg} cell-mediated immune suppression, which skews the balance towards a state of unrestrained inflammation. T_{reg} cell plasticity and instability, their possible mechanisms of action and their roles in atherosclerosis were recently reviewed[102].

Antigen specificity. B cells and T cells are activated by antigen presentation. DCs that have acquired atherosclerosis-related antigens can leave the atherosclerotic artery wall, most likely through the draining lymphatics[103], and present antigens to B cells and T cells in draining lymph nodes. Indeed, axillar and cervical lymph nodes draining the carotids and the aortic arch are enlarged and contain more cells in atherosclerotic mice[75]. In older (>1 year) mice, organized arterial tertiary lymphoid organs in the aortic adventitia contain germinal centres, in which affinity maturation and isotype switching of B cell responses can occur[104].

Although there is evidence for oligoclonal proliferation of and expression of activation markers in plaque-associated T cells[14,73], it has been technically challenging to identify their antigen specificity. Multiphoton live-cell imaging showed that antigen-specific and MHC class II-restricted productive interactions between CD4$^+$ T cells and APCs occur in the aortic walls of atherosclerotic mice[57]. This highlighted the presence of both endogenous autoantigen and antigen-experienced CD4$^+$ T cells in *Apoe*$^{-/-}$ mice. Numerous studies involving atherosclerotic mice and peptide-immunization models have established that APOB, the core protein component of LDL, is a major atherosclerosis antigen[105]. Tetramer staining and ex vivo peptide restimulation assays have detected APOB-specific T cells in mice and humans[106,107]. Analysis of lineage-defining transcription factors, cytokine secretion potential and scRNA-seq of APOB-reactive T cells revealed a prevalence of pro-inflammatory T_H1 cell and T_H17 cell phenotypes under atherosclerotic conditions[106,107].

Box 5 | Antigen presentation in atherosclerosis

Atherosclerosis is associated with loss of immune tolerance to self-antigens. CD4$^+$ T cells are the central mediators of this process because they help B cells and antibodies to mature and CD8$^+$ T cells to become fully cytotoxic. Activation of CD4$^+$ T cells requires the recognition of antigens presented by antigen-presenting cells (APCs), such as dendritic cells (DCs), on MHC class II molecules. Atherosclerotic mice that lack the invariant chain, a molecule involved in the formation and transport of peptide–MHC class II complexes, have defects in antigen-specific activation of adaptive immune responses and develop significantly smaller diet-induced aortic lesions than control mice[206]. Co-stimulatory molecules on the APCs and cytokine-mediated crosstalk instruct T cells regarding the nature of the response to be elicited[74]. Thus, MHC class II-dependent interactions between T cells and APCs are highly context dependent and can either promote pro-atherogenic inflammatory responses of effector T cells[57,64] or mediate the generation of immunosuppressive regulatory T cells[207]. DCs are the most potent activators of naive T cells in lymphoid organs and instruct the differentiation of T helper (T_H) cell subsets. Enrichment of cholesterol in membrane microdomains was shown to boost the antigen-presenting functions of DCs through enhanced MHC class II clustering[208]. Naive T cells probably encounter antigens in secondary lymphoid organs and in tertiary lymphoid organs found in the adventitia of large arteries in aged atherosclerotic mice[104].

Other APCs, including macrophages and B cells, also express MHC class II molecules and support recall (reactivation) T cell responses in lesions and germinal centres, respectively. Macrophages far outnumber DCs in atherosclerotic arteries. Abrogation of IL-27 receptor signalling resulted in the upregulation of MHC class II expression and pro-inflammatory cytokine secretion by myeloid cells, which engaged in longer and more frequent interactions with T cells in the aortic walls[209]. Unlike for myeloid cells, the role of B cell-mediated T cell activation is still controversial. Expression of MHC class II molecules on B cells was shown either to promote atherosclerosis[210] or to have no effect on lesion progression[211]. 'Innate response activator' B cells producing granulocyte–macrophage colony-stimulating factor, which accumulate in secondary lymphoid organs after the onset of atherosclerosis, indirectly modulate T_H1 cell skewing through the expansion of IL-12-secreting CD11b$^+$ DC populations[212].

Further in-depth phenotypic profiling of APOB-specific T cells in the plaque and in the blood by scRNA-seq and identification of the immunodominant epitopes that drive this pro-inflammatory T cell response will be crucial for understanding and modulating the autoimmune circuit in atherosclerosis.

Microbial and sterile triggers

Inducers of immune cell activation can be broadly classified as microbial pathogen-associated molecular patterns or endogenous danger-associated molecular patterns[108]. Recognition of these molecules by pattern recognition receptors such as TLRs and scavenger receptors results in the activation of signalling pathways and expression of pro-inflammatory cytokines such as TNF, IL-6, IFNγ and type I interferons[109–111]. Signalling through several pattern recognition receptors also culminates in the priming and activation of the NLRP3-dependent inflammasome cascade, which triggers caspase 1 activation, generation of the pro-inflammatory cytokines IL-1β and IL-18, and pyroptosis[108,112,113]. Here, we provide an overview of the microbial and host-derived inflammatory cues that have been implicated in atherogenesis.

Role of infectious agents and the host microbiota. Microorganisms can mediate chronic inflammation either through direct infection and activation of vascular cells or through indirect induction of a systemic immune response to infections at non-vascular locations[114]. Sero-epidemiological studies and sequencing of microbial DNA from atherosclerotic lesions have demonstrated an association between clinical atherosclerosis and infectious agents, including bacteria such as *Chlamydia pneumoniae*, *Helicobacter pylori*, *Porphyromonas gingivalis* and *Streptococcus sanguis* and viruses such as influenza A virus, cytomegalovirus, hepatitis C virus, HIV-1, herpes simplex viruses, Epstein–Barr virus, enteroviruses and parvovirus[114]. Although most of these studies were correlative, some microorganisms have been confirmed to promote atherosclerosis progression in mouse models[114,115]. Infection with *C. pneumoniae* was shown to breach immune tolerance through molecular mimicry, which was most evident in the case of human and bacterial forms of heat shock protein 60 (HSP60), which have significant sequence homology[116]. In a transgenic model of HIV-1 infection under atherogenic conditions, the acceleration of atherosclerosis progression by viral products was linked to inflammasome activation in monocytes and macrophages[117]. Thus, both bacterial and viral modulators may impact atherosclerosis by activating pro-atherogenic innate and adaptive immune responses. The mechanisms through which infections have been proposed to influence lesion formation, growth and rupture have been extensively described elsewhere[114,115,118].

The 'infection hypothesis' for atherogenesis led to several randomized, placebo-controlled trials that investigated whether macrolide antibiotics, which target *C. pneumoniae*, have any beneficial impact on clinical outcomes in patients with coronary artery disease. The results of these trials were unanimously negative[119], showing that macrolide antibiotics are not atheroprotective. However, antibiotic therapy does not address possible contributions to atherogenesis from viruses[114,115]. Also, it seems unlikely that a single pathogen, such as *C. pneumoniae*, would by itself affect the pathogenesis of a complex disease such as atherosclerosis. Antibiotics can also affect commensal microorganisms, which could have pro-atherogenic or anti-atherogenic effects.

In addition to infectious agents, recent studies have demonstrated a potential role of the host microbiota in influencing cardiovascular disease progression[120]. Postprandial increase in gut permeability in mice was shown to result in leakage of microbiota-derived endotoxins such as lipopolysaccharide into the circulation, which triggered a low-grade inflammatory response[121]. Furthermore, stress and dietary components can shape the composition of the gut microbiota, with some bacteria being able to mediate the production of pro-inflammatory metabolites such as trimethylamine *N*-oxide (TMAO) that have been associated with cardiovascular disease risk[120]. Trimethylamine, derived from microbial metabolism of dietary components such as choline, phosphatidylcholine and L-carnitine, can be converted to TMAO by liver flavin monooxygenases. Screening of isolates from the human gut has identified eight species of trimethylamine-producing bacteria, all of which are members of the phylum Firmicutes or proteobacteria[122]. The exact mechanism by which TMAO promotes atherosclerosis still remains speculative, possibly involving changes in cholesterol trafficking, induction of vascular inflammation[123–125] and platelet hyperactivation[126]. Multiple studies have explored the potential to use TMAO either as a prognostic biomarker[127–129] or as a target for therapy[130,131]. Some reports[132,133] in animal models have raised the question of whether TMAO has a causal role in atherogenesis. Furthermore, its potential role as a cardiovascular risk factor may be confounded by its possible involvement in kidney disease and diabetes[134]. Multi-omics association studies[135] in large cohorts and rigorous validation in animal models will be necessary to clarify the existing discrepancies.

Regular exercise and consumption of dietary fibres boost the levels of microbial metabolites such as short-chain fatty acids, particularly acetate, propionate and butyrate, which can reduce cardiovascular disease-associated risk factors, including high blood pressure, and regulate lipid homeostasis[120]. Butyrate derived from *Roseburia intestinalis* has been shown to correlate inversely with inflammation and atherosclerotic lesion size[136]. The cytokines IL-22 and IL-23 have an atheroprotective role by restraining the growth of a pro-atherogenic microbiota, which also decreases TMAO levels in the blood[137,138].

Endogenous triggers of inflammation. Immune cells in atherosclerotic lesions can also be activated by self-molecules that become immunogenic[86,139]. OxLDL, glycated proteins, lipids and nucleic acids in the vessel wall, crystalline cholesterol deposits, and cell death-associated molecules such as heat shock proteins are some of the best-characterized danger-associated

Molecular mimicry
The possible cross-reactive activation of autoreactive B cells or T cells that are specific for self-derived epitopes that have sequence similarity or structural homology with pathogen-derived foreign epitopes.

molecular patterns implicated in promoting atherosclerosis progression[108,140]. Stimulation of TLRs, particularly TLR2 and TLR4, activates distinct transcriptional programmes in plaque macrophages that trigger the release of pro-inflammatory cytokines and chemokines[108,111]. Some endogenous TLR ligands in atherosclerotic lesions, including oxLDL, are also actively recognized by scavenger receptors[86]. SR-A (also known as MSR1) and CD36 have the highest affinity for oxLDL and have a major role in modulating atherosclerosis[141]. CD36 can trigger macrophage differentiation into foam cells through the phosphorylation and activation of SRC and the MAPK pathway and can also promote the conversion of endocytosed oxLDL, amyloid-β and amyl peptides into crystals[142]. Crystalline cholesterol triggers NLRP3-dependent inflammasome activation[112] to produce IL-1β. Tissue damage and cell death can also induce inflammasome activation through the release of TLR ligands such as heat shock proteins and secondary inducers such as ATP and uric acid, which contribute to priming and activation of inflammasomes, respectively[113]. Although some studies have refuted the involvement of the NLRP3 inflammasome in atherogenesis[143], administration of NLRP3 inflammasome inhibitors significantly reduced lesion development in mouse models[144]. Furthermore, the CANTOS trial, which tested the efficacy of IL-1β blockade, provided the first clinical evidence that pro-inflammatory immune pathways can be viable targets for the prevention of cardiovascular events in humans[4,110].

Towards an atherosclerosis vaccine

Recent insights into the roles of innate and adaptive immune cells in atherosclerosis have prompted a flurry of research activities and clinical trials to test the viability of immunomodulatory therapies in cardiovascular diseases. Tolerogenic vaccination has emerged as a possible strategy to boost antigen-specific, anti-atherogenic humoral and cell-mediated immune responses in atherosclerosis[19].

Some immunization strategies in atherosclerosis aim to induce neutralizing antibodies that inhibit proteins associated with increased cardiovascular risk. This is the case for the secreted serine protease PCSK9, which interferes with LDL uptake in the liver by blocking LDL receptor recycling[145]. The approval of two fully human monoclonal antibodies to PCSK9 has initiated clinical trials to evaluate the safety and efficacy of human PCSK9 vaccines in lowering lipid levels through the induction of a long-term humoral response (ClinicalTrials.gov identifier NCT02508896). By contrast, tolerogenic vaccines aim to restrain atherosclerosis by mobilizing T_{reg} cells and restoring immune tolerance[146]. Most of the tolerogenic vaccination efforts in mouse models have focused either on LDL or on its core lipoprotein APOB. Immunization with human oxLDL induces antigen-specific $CD4^+$ T cells in transgenic mice that express human APOB[147]. Multiple studies have established that vaccinating mice with murine APOB peptides can reduce vascular inflammation, stabilize plaques and restrain disease progression in atherosclerotic mice[146]. Such atheroprotection was mediated by T_{reg} cells and Tr1 cells that suppress inflammatory cytokine production by effector T cells[19,146].

Immunization of atherosclerotic mice with the mouse APOB peptide p18, which is identical in sequence to a peptide in human APOB, reduced aortic lesion size and induced IL-10-producing $CD4^+$ T cells and $FOXP3^+$ T_{reg} cells[106]. This work provided evidence that human APOB-derived peptides could be atheroprotective and that vaccine-based therapeutic strategies could potentially be translated to the clinic. However, p18 is only one autoreactive MHC class II-restricted epitope mapped in a single atherosclerosis antigen, APOB. Identifying the autoreactive epitopes that may be driving the adaptive immune response in human atherosclerosis is technically challenging. First, autoantigen-specific T cells are rare and therefore hard to detect. Currently, only one human APOB epitope has been fully validated using tetramer technology[106]. Second, the use of tetramers is limited, because they require a priori knowledge of epitope specificity and HLA allelic restriction, are limited to one MHC allele at a time, are hard to validate and difficult to multiplex, and are expensive in both time and resources. To overcome this limitation, we recently introduced a restimulation-based assay that enabled simultaneous interrogation of APOB-specific effector T cell responses to 30 APOB-derived epitopes with heterogeneous HLA specificities[107]. Peptide stimulation of peripheral blood mononuclear cells elicited increased production of the pro-inflammatory cytokines IL-17, IFNγ and TNF by $CD4^+$ T cells from patients with coronary atherosclerosis but not from controls. Such restimulation assays promise to facilitate the discovery of epitope specificities, T cell phenotypes and interaction networks in human atherosclerosis. Whereas some epitopes can induce pro-inflammatory immune responses associated with adverse disease outcomes, other epitopes can elicit anti-atherogenic T_{reg} cell responses. To translate tolerogenic vaccine-based immunotherapy from mice to humans, a comprehensive list of immunomodulatory epitopes that can reliably induce T_{reg} cells and Tr1 cells must be identified for human APOB and other atherosclerosis antigens. Different tolerogenic adjuvants, administration routes and doses need to be tested in clinical trials. Strategies to stabilize vaccination-induced T_{reg} cells are currently being explored in clinical trials[148].

Conclusions

Atherosclerosis involves a complex interplay between metabolic imbalances and maladaptive immune responses that drive a state of chronic inflammation in the artery wall. While our understanding of the immune-mediated processes in atherosclerosis will continue to evolve, recent research has vastly expanded our knowledge of the immune cell interaction network, mainly through high-dimensional methods such as genome-wide association studies[149], mass cytometry[46,150], scRNA-seq[31,45,46,48] and combined protein and RNA sequencing[14]. Interactions between pro-atherogenic conditions and haematopoiesis (BOX 3), immune cell metabolism (BOX 4) and the microbiota are now firmly established. This new knowledge has

not only improved our mechanistic understanding of the disease pathophysiology but has also fuelled new research efforts aimed at developing strategies for immunomodulatory therapies. These have included cytokine targeting, adoptive transfer of regulatory cells, passive immunization with monoclonal antibodies and tolerogenic vaccines. However, there are major obstacles to the translation of interventions in animal models to the clinic, some of them related to the limited extent to which mouse models recapitulate human atherosclerosis (BOX 1). Recent advances in multidimensional approaches promise in-depth characterization of the cellular and molecular mediators of human atherosclerosis. Bioinformatics-based integration of datasets from plaque cells and blood cells at the genomic, transcriptomic, proteomic and metabolomic levels will help to define the human immune atlas of atherosclerosis in a more holistic manner. Large-scale longitudinal studies involving multiple cohorts will be necessary to establish which imaging, cellular and molecular biomarkers truly correlate with disease progression after confounding factors have been accounted for.

Published online: 13 August 2021

1. Michos, E. D., McEvoy, J. W. & Blumenthal, R. S. Lipid management for the prevention of atherosclerotic cardiovascular disease. *N. Engl. J. Med.* **381**, 1557–1567 (2019).
2. Sampson, U. K., Fazio, S. & Linton, M. F. Residual cardiovascular risk despite optimal LDL cholesterol reduction with statins: the evidence, etiology, and therapeutic challenges. *Curr. Atheroscler. Rep.* **14**, 1–10 (2012).
3. Ridker, P. M. Residual inflammatory risk: addressing the obverse side of the atherosclerosis prevention coin. *Eur. Heart J.* **37**, 1720–1722 (2016).
4. Ridker, P. M. et al. Antiinflammatory therapy with canakinumab for atherosclerotic disease. *N. Engl. J. Med.* **377**, 1119–1131 (2017).
5. Tardif, J. C. et al. Efficacy and safety of low-dose colchicine after myocardial infarction. *N. Engl. J. Med.* **381**, 2497–2505 (2019).
6. Basatemur, G. L., Jorgensen, H. F., Clarke, M. C. H., Bennett, M. R. & Mallat, Z. Vascular smooth muscle cells in atherosclerosis. *Nat. Rev. Cardiol.* **16**, 727–744 (2019).
7. Bennett, M. R., Sinha, S. & Owens, G. K. Vascular smooth muscle cells in atherosclerosis. *Circ. Res.* **118**, 692–702 (2016).
8. Zernecke, A. et al. Meta-analysis of leukocyte diversity in atherosclerotic mouse aortas. *Circ. Res.* **127**, 402–426 (2020).
This meta-analysis of leukocyte populations in mouse atherosclerotic aortas comprehensively summarizes the transcript profiles and gene signatures of aortic immune cell subsets obtained from nine scRNA-seq studies and two mass cytometry studies.
9. Pan, H. et al. Single-cell genomics reveals a novel cell state during smooth muscle cell phenotypic switching and potential therapeutic targets for atherosclerosis in mouse and human. *Circulation* **142**, 2060–2075 (2020).
10. Wirka, R. C. et al. Atheroprotective roles of smooth muscle cell phenotypic modulation and the TCF21 disease gene as revealed by single-cell analysis. *Nat. Med.* **25**, 1280–1289 (2019).
11. Feil, S. et al. Transdifferentiation of vascular smooth muscle cells to macrophage-like cells during atherogenesis. *Circ. Res.* **115**, 662–667 (2014).
12. Rayner, K. J. Cell death in the vessel wall: the good, the bad, the ugly. *Arterioscler. Thromb. Vasc. Biol.* **37**, e75–e81 (2017).
13. Cardoso, L. & Weinbaum, S. Microcalcifications, their genesis, growth, and biomechanical stability in fibrous cap rupture. *Adv. Exp. Med. Biol.* **1097**, 129–155 (2018).
14. Fernandez, D. M. et al. Single-cell immune landscape of human atherosclerotic plaques. *Nat. Med.* **25**, 1576–1588 (2019).
This study undertakes an integrated single-cell proteomic and transcriptomic analysis of human atherosclerotic plaques and identifies unique features of plaque T cells and macrophages that are associated with clinical cardiovascular disease.
15. Depuydt, M. A. C. et al. Microanatomy of the human atherosclerotic plaque by single-cell transcriptomics. *Circ. Res.* **127**, 1437–1455 (2020).
16. Back, M., Yurdagul, A. Jr, Tabas, I., Oorni, K. & Kovanen, P. T. Inflammation and its resolution in atherosclerosis: mediators and therapeutic opportunities. *Nat. Rev. Cardiol.* **16**, 389–406 (2019).
17. Spitz, C. et al. Regulatory T cells in atherosclerosis: critical immune regulatory function and therapeutic potential. *Cell. Mol. Life Sci.* **73**, 901–922 (2016).
18. Tse, K. et al. Atheroprotective vaccination with MHC-II restricted peptides from ApoB-100. *Front. Immunol.* **4**, 493 (2013).
19. Roy, P., Ali, A. J., Kobiyama, K., Ghosheh, Y. & Ley, K. Opportunities for an atherosclerosis vaccine: from mice to humans. *Vaccine* **38**, 4495–4506 (2020).
20. Libby, P., Lichtman, A. H. & Hansson, G. K. Immune effector mechanisms implicated in atherosclerosis: from mice to humans. *Immunity* **38**, 1092–1104 (2013).
21. Moore, K. J., Sheedy, F. J. & Fisher, E. A. Macrophages in atherosclerosis: a dynamic balance. *Nat. Rev. Immunol.* **13**, 709–721 (2013).
22. Kim, K. W., Ivanov, S. & Williams, J. W. Monocyte recruitment, specification, and function in atherosclerosis. *Cells* **10**, 15 (2020).
23. Tacke, F. et al. Monocyte subsets differentially employ CCR2, CCR5, and CX3CR1 to accumulate within atherosclerotic plaques. *J. Clin. Invest.* **117**, 185–194 (2007).
This study shows that an atherogenic diet induces systemic monocytosis in atherosclerotic mice and it defines the contributions of different chemokine receptors in mediating recruitment of distinct monocyte subsets to atherosclerotic lesions.
24. Combadiere, C. et al. Combined inhibition of CCL2, CX3CR1, and CCR5 abrogates Ly6C^{hi} and Ly6C^{lo} monocytosis and almost abolishes atherosclerosis in hypercholesterolemic mice. *Circulation* **117**, 1649–1657 (2008).
25. Tabas, I. & Lichtman, A. H. Monocyte-macrophages and T cells in atherosclerosis. *Immunity* **47**, 621–634 (2017).
26. Jakubzick, C. V., Randolph, G. J. & Henson, P. M. Monocyte differentiation and antigen-presenting functions. *Nat. Rev. Immunol.* **17**, 349–362 (2017).
27. Marcovecchio, P. M. et al. Scavenger receptor CD36 directs nonclassical monocyte patrolling along the endothelium during early atherogenesis. *Arterioscler. Thromb. Vasc. Biol.* **37**, 2043–2052 (2017).
28. Quintar, A. et al. Endothelial protective monocyte patrolling in large arteries intensified by Western diet and atherosclerosis. *Circ. Res.* **120**, 1789–1799 (2017).
29. Narasimhan, P. B., Marcovecchio, P., Hamers, A. A. J. & Hedrick, C. C. Nonclassical monocytes in health and disease. *Annu. Rev. Immunol.* **37**, 439–456 (2019).
30. Hamers, A. A. J. et al. Human monocyte heterogeneity as revealed by high-dimensional mass cytometry. *Arterioscler. Thromb. Vasc. Biol.* **39**, 25–36 (2019).
This study uses high-dimensional single-cell mass cytometry to establish the heterogeneity of human blood monocyte subsets, and it functionally characterizes non-classical monocytes expressing the carbohydrate marker 6-sulfo LacNAc, the frequency of which correlates positively with the severity of coronary artery disease.
31. Kim, K. et al. Transcriptome analysis reveals nonfoamy rather than foamy plaque macrophages are proinflammatory in atherosclerotic murine models. *Circ. Res.* **123**, 1127–1142 (2018).
32. Spann, N. J. et al. Regulated accumulation of desmosterol integrates macrophage lipid metabolism and inflammatory responses. *Cell* **151**, 138–152 (2012).
This study combines lipidomic and transcriptomic analyses to show that macrophage-derived foam cells are associated with suppression, rather than activation, of inflammatory genes and that most of the changes in gene expression and lipid metabolism are triggered by the accumulation of desmosterol, an intermediate in the cholesterol biosynthesis pathway.
33. Koelwyn, G. J., Corr, E. M., Erbay, E. & Moore, K. J. Regulation of macrophage immunometabolism in atherosclerosis. *Nat. Immunol.* **19**, 526–537 (2018).
34. Kojima, Y., Weissman, I. L. & Leeper, N. J. The role of efferocytosis in atherosclerosis. *Circulation* **135**, 476–489 (2017).
35. Moore, K. J. et al. Macrophage trafficking, inflammatory resolution, and genomics in atherosclerosis: JACC macrophage in CVD Series (part 2). *J. Am. Coll. Cardiol.* **72**, 2181–2197 (2018).
36. McArdle, S. et al. Migratory and dancing macrophage subsets in atherosclerotic lesions. *Circ. Res.* **125**, 1038–1051 (2019).
37. Robbins, C. S. et al. Local proliferation dominates lesional macrophage accumulation in atherosclerosis. *Nat. Med.* **19**, 1166–1172 (2013).
38. van Gils, J. M. et al. The neuroimmune guidance cue netrin-1 promotes atherosclerosis by inhibiting the emigration of macrophages from plaques. *Nat. Immunol.* **13**, 136–143 (2012).
This study defines the role of netrin 1, a neuroimmune guidance factor, in promoting atherosclerotic plaque progression through the retention of macrophages within inflamed blood vessels.
39. Wanschel, A. et al. Neuroimmune guidance cue semaphorin 3E is expressed in atherosclerotic plaques and regulates macrophage retention. *Arterioscler. Thromb. Vasc. Biol.* **33**, 886–893 (2013).
40. Trogan, E. et al. Gene expression changes in foam cells and the role of chemokine receptor CCR7 during atherosclerosis regression in ApoE-deficient mice. *Proc. Natl Acad. Sci. USA* **103**, 3781–3786 (2006).
41. Feig, J. E. et al. Reversal of hyperlipidemia with a genetic switch favorably affects the content and inflammatory state of macrophages in atherosclerotic plaques. *Circulation* **123**, 989–998 (2011).
42. Feig, J. E. et al. LXR promotes the maximal egress of monocyte-derived cells from mouse aortic plaques during atherosclerosis regression. *J. Clin. Invest.* **120**, 4415–4424 (2010).
43. Potteaux, S. et al. Suppressed monocyte recruitment drives macrophage removal from atherosclerotic plaques of Apoe$^{-/-}$ mice during disease regression. *J. Clin. Invest.* **121**, 2025–2036 (2011).
44. Mills, C. D., Kincaid, K., Alt, J. M., Heilman, M. J. & Hill, A. M. M-1/M-2 macrophages and the Th1/Th2 paradigm. *J. Immunol.* **164**, 6166–6173 (2000).
45. Cochain, C. et al. Single-cell RNA-seq reveals the transcriptional landscape and heterogeneity of aortic macrophages in murine atherosclerosis. *Circ. Res.* **122**, 1661–1674 (2018).
46. Winkels, H. et al. Atlas of the immune cell repertoire in mouse atherosclerosis defined by single-cell RNA-sequencing and mass cytometry. *Circ. Res.* **122**, 1675–1688 (2018).
47. King, K. R. et al. IRF3 and type I interferons fuel a fatal response to myocardial infarction. *Nat. Med.* **23**, 1481–1487 (2017).
48. Lin, J. D. et al. Single-cell analysis of fate-mapped macrophages reveals heterogeneity, including stem-like properties, during atherosclerosis progression and regression. *JCI Insight* **4**, e124574 (2019).
49. Williams, J. W. et al. Limited proliferation capacity of aortic intima resident macrophages requires monocyte recruitment for atherosclerotic plaque progression. *Nat. Immunol.* **21**, 1194–1204 (2020).
This study characterizes a monocyte-derived arterial resident macrophage population, known as Mac^{AIR}, and describes their origins and functions during early and advanced stages of atherosclerosis.

50. Ensan, S. et al. Self-renewing resident arterial macrophages arise from embryonic CX3CR1$^+$ precursors and circulating monocytes immediately after birth. *Nat. Immunol.* **17**, 159–168 (2016). **This study uses multiple fate-mapping approaches to show that tissue-resident macrophages in the arterial adventitia arise from embryonic precursor cells and undergo local proliferation**.
51. Lim, H. Y. et al. Hyaluronan receptor LYVE-1expressing macrophages maintain arterial tone through hyaluronan-mediated regulation of smooth muscle cell collagen. *Immunity* **49**, 326–341.e7 (2018).
52. Hill, C. A., Fernandez, D. M. & Giannarelli, C. Single cell analyses to understand the immune continuum in atherosclerosis. *Atherosclerosis* **330**, 85–94 (2021).
53. Zernecke, A. Dendritic cells in atherosclerosis: evidence in mice and humans. *Arterioscler. Thromb. Vasc. Biol.* **35**, 763–770 (2015).
54. Choi, J. H. et al. Flt3 signaling-dependent dendritic cells protect against atherosclerosis. *Immunity* **35**, 819–831 (2011).
55. Edelson, B. T. et al. Peripheral CD103$^+$ dendritic cells form a unified subset developmentally related to CD8α$^+$ conventional dendritic cells. *J. Exp. Med.* **207**, 823–836 (2010).
56. Weber, C. et al. CCL17-expressing dendritic cells drive atherosclerosis by restraining regulatory T cell homeostasis in mice. *J. Clin. Invest.* **121**, 2898–2910 (2011).
57. Koltsova, E. K. et al. Dynamic T cell-APC interactions sustain chronic inflammation in. *atherosclerosis. J. Clin. Invest.* **122**, 3114–3126 (2012). **This study uses live-cell imaging of explanted mouse atherosclerotic aortas to show antigen-dependent interactions between aortic APCs and CD4$^+$ T cells and that CD4$^+$ T cells from atherosclerotic mice, but not control mice, interact with APCs in the vessel walls of atherosclerotic aorta, suggesting a recall response to an endogenous antigen**.
58. Choi, J. H. et al. Identification of antigen-presenting dendritic cells in mouse aorta and cardiac valves. *J. Exp. Med.* **206**, 497–505 (2009).
59. Haddad, Y. et al. The dendritic cell receptor DNGR-1 promotes the development of atherosclerosis in mice. *Circ. Res.* **121**, 234–243 (2017).
60. Clement, M. et al. Deletion of IRF8 (interferon regulatory factor 8)-dependent dendritic cells abrogates proatherogenic adaptive immunity. *Circ. Res.* **122**, 813–820 (2018).
61. Subramanian, M., Thorp, E., Hansson, G. K. & Tabas, I. Treg-mediated suppression of atherosclerosis requires MYD88 signaling in DCs. *J. Clin. Invest.* **123**, 179–188 (2013). **This study shows that MYD88-dependent TLR signalling in CD11c$^+$ DCs, which is important for DC maturation, is necessary for the infiltration of both effector T cells and T_{reg} cells in atherosclerotic lesions but largely has an atheroprotective role through the inhibition of inflammatory monocyte recruitment by T_{reg} cells**.
62. Clement, M. et al. Impaired autophagy in CD11b$^+$ dendritic cells expands CD4$^+$ regulatory T cells and limits atherosclerosis in mice. *Circ. Res.* **125**, 1019–1034 (2019).
63. Doring, Y. et al. Auto-antigenic protein-DNA complexes stimulate plasmacytoid dendritic cells to promote atherosclerosis. *Circulation* **125**, 1673–1683 (2012).
64. Sage, A. P. et al. MHC class II-restricted antigen presentation by plasmacytoid dendritic cells drives proatherogenic T cell immunity. *Circulation* **130**, 1363–1373 (2014). **This study identifies that the pro-atherogenic functions of pDCs are mediated by MHC class II-restricted antigen presentation to CD4$^+$ T cells and that this is independent of IFNα production**.
65. Daissormont, I. T. et al. Plasmacytoid dendritic cells protect against atherosclerosis by tuning T-cell proliferation and activity. *Circ. Res.* **109**, 1387–1395 (2011).
66. Yun, T. J. et al. Indoleamine 2,3-dioxygenase-expressing aortic plasmacytoid dendritic cells protect against atherosclerosis by induction of regulatory T cells. *Cell Metab.* **23**, 852–866 (2016).
67. Sage, A. P., Tsiantoulas, D., Binder, C. J. & Mallat, Z. The role of B cells in atherosclerosis. *Nat. Rev. Cardiol.* **16**, 180–196 (2019).
68. Galkina, E. et al. Lymphocyte recruitment into the aortic wall before and during development of atherosclerosis is partially L-selectin dependent. *J. Exp. Med.* **203**, 1273–1282 (2006).
69. Butcher, M. J., Wu, C. I., Waseem, T. & Galkina, E. V. CXCR6 regulates the recruitment of pro-inflammatory IL-17A-producing T cells into atherosclerotic aortas. *Int. Immunol.* **28**, 255–261 (2016).
70. Li, J. et al. CCR5$^+$T-bet$^+$FoxP3$^+$ effector CD4 T cells drive atherosclerosis. *Circ. Res.* **118**, 1540–1552 (2016).
71. Huo, Y. et al. Circulating activated platelets exacerbate atherosclerosis in mice deficient in apolipoprotein E. *Nat. Med.* **9**, 61–67 (2003).
72. Gencer, S., Evans, B. R., van der Vorst, E. P. C., Doring, Y. & Weber, C. Inflammatory chemokines in atherosclerosis. *Cells* **10**, 226 (2021).
73. Saigusa, R., Winkels, H. & Ley, K. T cell subsets and functions in atherosclerosis. *Nat. Rev. Cardiol.* **17**, 387–401 (2020).
74. Ley, K., Gerdes, N. & Winkels, H. ATVB distinguished scientist award: how costimulatory and coinhibitory pathways shape atherosclerosis. *Arterioscler. Thromb. Vasc. Biol.* **37**, 764–777 (2017).
75. Wolf, D. & Ley, K. Immunity and inflammation in atherosclerosis. *Circ. Res.* **124**, 315–327 (2019).
76. Ait-Oufella, H., Taleb, S., Mallat, Z. & Tedgui, A. Recent advances on the role of cytokines in atherosclerosis. *Arterioscler. Thromb. Vasc. Biol.* **31**, 969–979 (2011).
77. Ley, K., Laudanna, C., Cybulsky, M. I. & Nourshargh, S. Getting to the site of inflammation: the leukocyte adhesion cascade updated. *Nat. Rev. Immunol.* **7**, 678–689 (2007).
78. Maganto-Garcia, E., Tarrio, M. L., Grabie, N., Bu, D. X. & Lichtman, A. H. Dynamic changes in regulatory T cells are linked to levels of diet-induced hypercholesterolemia. *Circulation* **124**, 185–195 (2011).
79. Klingenberg, R. et al. Depletion of FOXP3$^+$ regulatory T cells promotes hypercholesterolemia and atherosclerosis. *J. Clin. Invest.* **123**, 1323–1334 (2013).
80. Mullick, A. E. et al. Antisense oligonucleotide reduction of apoB-ameliorated atherosclerosis in LDL receptor-deficient mice. *J. Lipid Res.* **52**, 885–896 (2011).
81. Chereshnev, I. et al. Mouse model of heterotopic aortic arch transplantation. *J. Surg. Res.* **111**, 171–176 (2003).
82. Sharma, M. et al. Regulatory T cells license macrophage pro-resolving functions during atherosclerosis regression. *Circ. Res.* **127**, 335–353 (2020).
83. Engelbertsen, D. et al. T-helper 2 immunity is associated with reduced risk of myocardial infarction and stroke. *Arterioscler. Thromb. Vasc. Biol.* **33**, 637–644 (2013).
84. Engelbertsen, D. et al. Induction of T helper 2 responses against human apolipoprotein B100 does not affect atherosclerosis in ApoE$^{-/-}$ mice. *Cardiovasc. Res.* **103**, 304–312 (2014).
85. Binder, C. J. et al. IL-5 links adaptive and natural immunity specific for epitopes of oxidized LDL and protects from atherosclerosis. *J. Clin. Invest.* **114**, 427–437 (2004).
86. Binder, C. J., Papac-Milicevic, N. & Witztum, J. L. Innate sensing of oxidation-specific epitopes in health and disease. *Nat. Rev. Immunol.* **16**, 485–497 (2016).
87. Knutsson, A. et al. Associations of interleukin-5 with plaque development and cardiovascular events. *JACC Basic Transl. Sci.* **4**, 891–902 (2019).
88. Cardilo-Reis, L. et al. Interleukin-13 protects from atherosclerosis and modulates plaque composition by skewing the macrophage phenotype. *EMBO Mol. Med.* **4**, 1072–1086 (2012).
89. Newland, S. A. et al. Type-2 innate lymphoid cells control the development of atherosclerosis in mice. *Nat. Commun.* **8**, 15781 (2017). **This study uses atherosclerotic mouse models that were deficient in group 2 innate lymphoid cells to prove that IL-5 and IL-13 production by innate lymphoid cells has an atheroprotective role and to highlight the functional overlap between innate lymphoid cells and T_H cell subsets in modulating atherosclerosis**.
90. Gao, Q. et al. A critical function of Th17 proinflammatory cells in the development of atherosclerotic plaque in mice. *J. Immunol.* **185**, 5820–5827 (2010).
91. Nordlohne, J. et al. Aggravated atherosclerosis and vascular inflammation with reduced kidney function depend on interleukin-17 receptor a and are normalized by inhibition of interleukin-17A. *JACC Basic Transl. Sci.* **3**, 54–66 (2018).
92. Smith, E. et al. Blockade of interleukin-17A results in reduced atherosclerosis in apolipoprotein E-deficient mice. *Circulation* **121**, 1746–1755 (2010).
93. Danzaki, K. et al. Interleukin-17A deficiency accelerates unstable atherosclerotic plaque formation in apolipoprotein E-deficient mice. *Arterioscler. Thromb. Vasc. Biol.* **32**, 273–280 (2012).
94. Gistera, A. et al. Transforming growth factor-beta signaling in T cells promotes stabilization of atherosclerotic plaques through an interleukin-17-dependent pathway. *Sci. Transl. Med.* **5**, 196ra100 (2013).
95. Brauner, S. et al. Augmented Th17 differentiation in Trim21 deficiency promotes a stable phenotype of atherosclerotic plaques with high collagen content. *Cardiovasc. Res.* **114**, 158–167 (2018).
96. Gensous, N. et al. T follicular helper cells in autoimmune disorders. *Front. Immunol.* **9**, 1637 (2018).
97. Ryu, H. et al. Atherogenic dyslipidemia promotes autoimmune follicular helper T cell responses via IL-27. *Nat. Immunol.* **19**, 583–593 (2018).
98. Clement, M. et al. Control of the T follicular helper-germinal center B-cell axis by CD8$^+$ regulatory T cells limits atherosclerosis and tertiary lymphoid organ development. *Circulation* **131**, 560–570 (2015).
99. Nus, M. et al. Marginal zone B cells control the response of follicular helper T cells to a high-cholesterol diet. *Nat. Med.* **23**, 601–610 (2017).
100. Butcher, M. J. et al. Atherosclerosis-driven Treg plasticity results in formation of a dysfunctional subset of plastic IFNγ$^+$ Th1/Tregs. *Circ. Res.* **119**, 1190–1203 (2016). **This study shows that inflammatory conditions associated with atherosclerosis can trigger the plasticity of T_{reg} cells and result in the formation of IFNγ-expressing T cells with a mixed T_H1 cell and T_{reg} cell phenotype, which have impaired ability to suppress effector T cell proliferation**.
101. Gaddis, D. E. et al. Apolipoprotein AI prevents regulatory to follicular helper T cell switching during atherosclerosis. *Nat. Commun.* **9**, 1095 (2018).
102. Ali, A. J., Makings, J. & Ley, K. Regulatory T cell stability and plasticity in atherosclerosis. *Cells* **9**, 2665 (2020).
103. Randolph, G. J., Angeli, V. & Swartz, M. A. Dendritic-cell trafficking to lymph nodes through lymphatic vessels. *Nat. Rev. Immunol.* **5**, 617–628 (2005).
104. Srikakulapu, P. et al. Artery tertiary lymphoid organs control multilayered territorialized atherosclerosis B-cell responses in aged ApoE$^{-/-}$ mice. *Arterioscler. Thromb. Vasc. Biol.* **36**, 1174–1185 (2016).
105. Marchini, T., Hansen, S. & Wolf, D. ApoB-specific CD4$^+$ T cells in mouse and human atherosclerosis. *Cells* **10**, 446 (2021).
106. Kimura, T. et al. Regulatory CD4$^+$ T cells recognize MHC-II-restricted peptide epitopes of apolipoprotein B. *Circulation* **138**, 1130–1143 (2018). **This is the first study to detect and characterize APOB-specific autoreactive CD4$^+$ T cells in humans using peptide-specific tetramers and to show the prophylactic potential of human APOB-derived peptide vaccines in mouse models of atherosclerosis**.
107. Wolf, D. et al. Pathogenic autoimmunity in atherosclerosis evolves from initially protective apolipoprotein B100-reactive CD4$^+$ T-regulatory cells. *Circulation* **142**, 1279–1293 (2020). **This study uses a novel MHC class II tetramer to undertake detailed phenotypic and transcriptomic analyses of APOB-specific mouse CD4$^+$ T cells and establishes the feasibility of using restimulation assays to detect cytokine-expressing, MHC class II-restricted, antigen-specific T cells in human blood**.
108. Zimmer, S., Grebe, A. & Latz, E. Danger signaling in atherosclerosis. *Circ. Res.* **116**, 323–340 (2015).
109. Ley, K., Pramod, A. B., Croft, M., Ravichandran, K. S. & Ting, J. P. How mouse macrophages sense what is going on. *Front. Immunol.* **7**, 204 (2016).
110. Libby, P. Targeting inflammatory pathways in cardiovascular disease: the inflammasome, interleukin-1, interleukin-6 and beyond. *Cells* **10**, 951 (2021).
111. Miller, Y. I., Choi, S. H., Wiesner, P. & Bae, Y. S. The SYK side of TLR4: signalling mechanisms in response to LPS and minimally oxidized LDL. *Br. J. Pharmacol.* **167**, 990–999 (2012).
112. Duewell, P. et al. NLRP3 inflammasomes are required for atherogenesis and activated by cholesterol crystals. *Nature* **464**, 1357–1361 (2010).

113. Grebe, A., Hoss, F. & Latz, E. NLRP3 inflammasome and the IL-1 pathway in atherosclerosis. *Circ. Res.* **122**, 1722–1740 (2018).
114. Campbell, L. A. & Rosenfeld, M. E. Infection and atherosclerosis development. *Arch. Med. Res.* **46**, 339–350 (2015).
115. Pothineni, N. V. K. et al. Infections, atherosclerosis, and coronary heart disease. *Eur. Heart J.* **38**, 3195–3201 (2017).
116. Wick, G., Jakic, B., Buszko, M., Wick, M. C. & Grundtman, C. The role of heat shock proteins in atherosclerosis. *Nat. Rev. Cardiol.* **11**, 516–529 (2014).
117. Kearns, A. C. et al. Caspase-1 activation is related with HIV-associated atherosclerosis in an HIV transgenic mouse model and HIV patient cohort. *Arterioscler. Thromb. Vasc. Biol.* **39**, 1762–1775 (2019).
118. Epstein, S. E., Zhu, J., Najafi, A. H. & Burnett, M. S. Insights into the role of infection in atherogenesis and in plaque rupture. *Circulation* **119**, 3133–3141 (2009).
119. Andraws, R., Berger, J. S. & Brown, D. L. Effects of antibiotic therapy on outcomes of patients with coronary artery disease: a meta-analysis of randomized controlled trials. *JAMA* **293**, 2641–2647 (2005).
120. Tang, W. H. W., Backhed, F., Landmesser, U. & Hazen, S. L. Intestinal microbiota in cardiovascular health and disease: JACC state-of-the-art review. *J. Am. Coll. Cardiol.* **73**, 2089–2105 (2019).
121. Piya, M. K., Harte, A. L. & McTernan, P. G. Metabolic endotoxaemia: is it more than just a gut feeling? *Curr. Opin. Lipidol.* **24**, 78–85 (2013).
122. Romano, K. A., Vivas, E. I., Amador-Noguez, D. & Rey, F. E. Intestinal microbiota composition modulates choline bioavailability from diet and accumulation of the proatherogenic metabolite trimethylamine-*N*-oxide. *mBio* **6**, e02481 (2015).
This study uses ultra-high-pressure liquid chromatography and tandem mass spectrometry to identify choline-consuming, trimethylamine-producing bacterial species in human intestinal isolates and characterizes their relevance in gnotobiotic mice.
123. Chen, M. L. et al. Trimethylamine-*N*-oxide induces vascular inflammation by activating the NLRP3 inflammasome through the SIRT3-SOD2-mtROS signaling pathway. *J. Am. Heart Assoc.* **6**, e006347 (2017).
124. Koeth, R. A. et al. Intestinal microbiota metabolism of L-carnitine, a nutrient in red meat, promotes atherosclerosis. *Nat. Med.* **19**, 576–585 (2013).
This study establishes a mechanistic connection between the dietary component L-carnitine, gut microbial metabolism and TMAO-mediated inhibition of cholesterol reverse transport and increased cardiovascular disease risk.
125. Wang, Z. et al. Gut flora metabolism of phosphatidylcholine promotes cardiovascular disease. *Nature* **472**, 57–63 (2011).
126. Zhu, W. et al. Gut microbial metabolite TMAO enhances platelet hyperreactivity and thrombosis risk. *Cell* **165**, 111–124 (2016).
127. Li, X. S. et al. Gut microbiota-dependent trimethylamine *N*-oxide in acute coronary syndromes: a prognostic marker for incident cardiovascular events beyond traditional risk factors. *Eur. Heart J.* **38**, 814–824 (2017).
128. Schiattarella, G. G. et al. Gut microbe-generated metabolite trimethylamine-*N*-oxide as cardiovascular risk biomarker: a systematic review and dose-response meta-analysis. *Eur. Heart J.* **38**, 2948–2956 (2017).
129. Tang, W. H. et al. Prognostic value of elevated levels of intestinal microbe-generated metabolite trimethylamine-*N*-oxide in patients with heart failure: refining the gut hypothesis. *J. Am. Coll. Cardiol.* **64**, 1908–1914 (2014).
130. Roberts, A. B. et al. Development of a gut microbe-targeted nonlethal therapeutic to inhibit thrombosis potential. *Nat. Med.* **24**, 1407–1417 (2018).
131. Wang, Z. et al. Non-lethal inhibition of gut microbial trimethylamine production for the treatment of atherosclerosis. *Cell* **163**, 1585–1595 (2015).
132. Aldana-Hernandez, P. et al. Dietary choline or trimethylamine *N*-oxide supplementation does not influence atherosclerosis development in $Ldlr^{-/-}$ and $Apoe^{-/-}$ male mice. *J. Nutr.* **150**, 249–255 (2020).
133. Collins, H. L. et al. L-Carnitine intake and high trimethylamine *N*-oxide plasma levels correlate with low aortic lesions in $ApoE^{-/-}$ transgenic mice expressing CETP. *Atherosclerosis* **244**, 29–37 (2016).
134. Zeisel, S. H. & Warrier, M. Trimethylamine *N*-oxide, the microbiome, and heart and kidney disease. *Annu. Rev. Nutr.* **37**, 157–181 (2017).
135. Manor, O. et al. A multi-omic association study of trimethylamine *N*-oxide. *Cell Rep.* **24**, 935–946 (2018).
136. Kasahara, K. et al. Interactions between *Roseburia intestinalis* and diet modulate atherogenesis in a murine model. *Nat. Microbiol.* **3**, 1461–1471 (2018).
137. Fatkhullina, A. R. et al. An interleukin-23-interleukin-22 axis regulates intestinal microbial homeostasis to protect from diet-induced atherosclerosis. *Immunity* **49**, 943–957.e9 (2018).
138. Kurilenko, N., Fatkhullina, A. R., Mazitova, A. & Koltsova, E. K. Act locally, act globally-microbiota, barriers, and cytokines in atherosclerosis. *Cells* **10**, 348 (2021).
139. Chen, G. Y. & Nunez, G. Sterile inflammation: sensing and reacting to damage. *Nat. Rev. Immunol.* **10**, 826–837 (2010).
140. Que, X. et al. Oxidized phospholipids are proinflammatory and proatherogenic in hypercholesterolaemic mice. *Nature* **558**, 301–306 (2018).
This study generates a transgenic mouse model expressing antibodies to oxLDL to evaluate the pro-atherogenic role of oxidized phospholipids in vivo; it proposes antibody-mediated blockade of oxidized phospholipids as a potential therapeutic strategy for atherosclerosis.
141. Moore, K. J. & Freeman, M. W. Scavenger receptors in atherosclerosis: beyond lipid uptake. *Arterioscler. Thromb. Vasc. Biol.* **26**, 1702–1711 (2006).
142. Sheedy, F. J. et al. CD36 coordinates NLRP3 inflammasome activation by facilitating intracellular nucleation of soluble ligands into particulate ligands in sterile inflammation. *Nat. Immunol.* **14**, 812–820 (2013).
143. Menu, P. et al. Atherosclerosis in ApoE-deficient mice progresses independently of the NLRP3 inflammasome. *Cell Death Dis.* **2**, e137 (2011).
144. van der Heijden, T. et al. NLRP3 inflammasome inhibition by MCC950 reduces atherosclerotic lesion development in apolipoprotein E-deficient mice — brief report. *Arterioscler. Thromb. Vasc. Biol.* **37**, 1457–1461 (2017).
145. Weisshaar, S. & Zeitlinger, M. Vaccines targeting PCSK9: a promising alternative to passive immunization with monoclonal antibodies in the management of hyperlipidaemia? *Drugs* **78**, 799–808 (2018).
146. Nettersheim, F. S., De Vore, L. & Winkels, H. Vaccination in atherosclerosis. *Cells* **9**, 2560 (2020).
147. Gistera, A. et al. Vaccination against T-cell epitopes of native ApoB100 reduces vascular inflammation and disease in a humanized mouse model of atherosclerosis. *J. Intern. Med.* **281**, 383–397 (2017).
148. Ait-Oufella, H., Lavillegrand, J. R. & Tedgui, A. Regulatory T cell-enhancing therapies to treat atherosclerosis. *Cells* **10**, 723 (2021).
149. Mauersberger, C., Schunkert, H. & Sager, H. B. Inflammation-related risk loci in genome-wide association studies of coronary artery disease. *Cells* **10**, 440 (2021).
150. Cole, J. E. et al. Immune cell census in murine atherosclerosis: cytometry by time of flight illuminates vascular myeloid cell diversity. *Cardiovasc. Res.* **114**, 1360–1371 (2018).
151. Mulligan-Kehoe, M. J. & Simons, M. Vasa vasorum in normal and diseased arteries. *Circulation* **129**, 2557–2566 (2014).
152. Rademakers, T. et al. Plaque-associated vasa vasorum in aged apolipoprotein E-deficient mice exhibit proatherogenic functional features in vivo. *Arterioscler. Thromb. Vasc. Biol.* **33**, 249–256 (2013).
153. Kocher, O. & Krieger, M. Role of the adaptor protein PDZK1 in controlling the HDL receptor SR-BI. *Curr. Opin. Lipidol.* **20**, 236–241 (2009).
154. Pal, R. et al. Carboxy-terminal deletion of the HDL receptor reduces receptor levels in liver and steroidogenic tissues, induces hypercholesterolemia, and causes fatal heart disease. *Am. J. Physiol. Heart Circ. Physiol.* **311**, H1392–H1408 (2016).
155. Yin, W. et al. Plasma lipid profiling across species for the identification of optimal animal models of human dyslipidemia. *J. Lipid Res.* **53**, 51–65 (2012).
156. Barrett-Connor, E. Menopause, atherosclerosis, and coronary artery disease. *Curr. Opin. Pharmacol.* **13**, 186–191 (2013).
157. Maeda, N. et al. Anatomical differences and atherosclerosis in apolipoprotein E-deficient mice with 129/SvEv and C57BL/6 genetic backgrounds. *Atherosclerosis* **195**, 75–82 (2007).
158. Buscher, K. et al. Natural variation of macrophage activation as disease-relevant phenotype predictive of inflammation and cancer survival. *Nat. Commun.* **8**, 16041 (2017).
159. Bennett, B. J. et al. Genetic architecture of atherosclerosis in mice: a systems genetics analysis of common inbred strains. *PLoS Genet.* **11**, e1005711 (2015).
160. Mailer, R. K. W., Gistera, A., Polyzos, K. A., Ketelhuth, D. F. J. & Hansson, G. K. Hypercholesterolemia enhances T cell receptor signaling and increases the regulatory T cell population. *Sci. Rep.* **7**, 15655 (2017).
161. Pollock, A. H. et al. Prolonged intake of dietary lipids alters membrane structure and t cell responses in $LDLr^{-/-}$ mice. *J. Immunol.* **196**, 3993–4002 (2016).
162. Beura, L. K. et al. Normalizing the environment recapitulates adult human immune traits in laboratory mice. *Nature* **532**, 512–516 (2016).
163. Barrington, W. T. & Lusis, A. J. Atherosclerosis: association between the gut microbiome and atherosclerosis. *Nat. Rev. Cardiol.* **14**, 699–700 (2017).
164. He, S. et al. Gut intraepithelial T cells calibrate metabolism and accelerate cardiovascular disease. *Nature* **566**, 115–119 (2019).
165. Bonaccorsi, I. et al. Symptomatic carotid atherosclerotic plaques are associated with increased infiltration of natural killer (NK) cells and higher serum levels of NK activating receptor ligands. *Front. Immunol.* **10**, 1503 (2019).
166. Selathurai, A. et al. Natural killer (NK) cells augment atherosclerosis by cytotoxic-dependent mechanisms. *Cardiovasc. Res.* **102**, 128–137 (2014).
167. Nour-Eldine, W. et al. Genetic depletion or hyperresponsiveness of natural killer cells do not affect atherosclerosis development. *Circ. Res.* **122**, 47–57 (2018).
168. Kyaw, T. et al. Cytotoxic and proinflammatory $CD8^+$ T lymphocytes promote development of vulnerable atherosclerotic plaques in apoE-deficient mice. *Circulation* **127**, 1028–1039 (2013).
169. Seijkens, T. T. P. et al. Deficiency of the T cell regulator casitas B-cell lymphoma-B aggravates atherosclerosis by inducing $CD8^+$ T cell-mediated macrophage death. *Eur. Heart J.* **40**, 372–382 (2019).
170. Cochain, C. et al. $CD8^+$ T cells regulate monopoiesis and circulating Ly6C-high monocyte levels in atherosclerosis in mice. *Circ. Res.* **117**, 244–253 (2015).
171. Dimayuga, P. C. et al. Identification of apoB-100 peptide-specific $CD8^+$ T cells in atherosclerosis. *J. Am. Heart Assoc.* **6**, e005318 (2017).
172. Schafer, S. & Zernecke, A. $CD8^+$ T cells in atherosclerosis. *Cells* **10**, 37 (2020).
173. Getz, G. S. & Reardon, C. A. Natural killer T cells in atherosclerosis. *Nat. Rev. Cardiol.* **14**, 304–314 (2017).
174. Li, Y. et al. $CD4^+$ natural killer T cells potently augment aortic root atherosclerosis by perforin- and granzyme B-dependent cytotoxicity. *Circ. Res.* **116**, 245–254 (2015).
175. Vu, D. M. et al. γδT cells are prevalent in the proximal aorta and drive nascent atherosclerotic lesion progression and neutrophilia in hypercholesterolemic mice. *PLoS ONE* **9**, e109416 (2014).
176. Cheng, H. Y., Wu, R. & Hedrick, C. C. Gammadelta (γδ) T lymphocytes do not impact the development of early atherosclerosis. *Atherosclerosis* **234**, 265–269 (2014).
177. Poller, W. C., Nahrendorf, M. & Swirski, F. K. Hematopoiesis and cardiovascular disease. *Circ. Res.* **126**, 1061–1085 (2020).
178. Yvan-Charvet, L. et al. ATP-binding cassette transporters and HDL suppress hematopoietic stem cell proliferation. *Science* **328**, 1689–1693 (2010).
179. Murphy, A. J. et al. ApoE regulates hematopoietic stem cell proliferation, monocytosis, and monocyte accumulation in atherosclerotic lesions in mice. *J. Clin. Invest.* **121**, 4138–4149 (2011).
180. Seijkens, T. et al. Hypercholesterolemia-induced priming of hematopoietic stem and progenitor cells aggravates atherosclerosis. *FASEB J.* **28**, 2202–2213 (2014).
181. Robbins, C. S. et al. Extramedullary hematopoiesis generates Ly-$6C^{high}$ monocytes that infiltrate

atherosclerotic lesions. *Circulation* **125**, 364–374 (2012).
This study shows that monocyte infiltrates in atherosclerotic lesions do not arise exclusively from bone marrow haematopoietic progenitors but also are produced by haematopoiesis at extramedullary sites, particularly the spleen.
182. Westerterp, M. et al. Regulation of hematopoietic stem and progenitor cell mobilization by cholesterol efflux pathways. *Cell Stem Cell* **11**, 195–206 (2012).
183. Dutta, P. et al. Myocardial infarction accelerates atherosclerosis. *Nature* **487**, 325–329 (2012).
184. Heidt, T. et al. Chronic variable stress activates hematopoietic stem cells. *Nat. Med.* **20**, 754–758 (2014).
185. McAlpine, C. S. et al. Sleep modulates haematopoiesis and protects against atherosclerosis. *Nature* **566**, 383–387 (2019).
186. Frodermann, V. et al. Exercise reduces inflammatory cell production and cardiovascular inflammation via instruction of hematopoietic progenitor cells. *Nat. Med.* **25**, 1761–1771 (2019).
187. Nahrendorf, M. & Swirski, F. K. Lifestyle effects on hematopoiesis and atherosclerosis. *Circ. Res.* **116**, 884–894 (2015).
188. Jaiswal, S. & Libby, P. Clonal haematopoiesis: connecting ageing and inflammation in cardiovascular disease. *Nat. Rev. Cardiol.* **17**, 137–144 (2020).
189. Ketelhuth, D. F. J. et al. Immunometabolism and atherosclerosis: perspectives and clinical significance: a position paper from the Working Group on Atherosclerosis and Vascular Biology of the European Society of Cardiology. *Cardiovasc. Res.* **115**, 1385–1392 (2019).
190. Perez-Medina, C., Fayad, Z. A. & Mulder, W. J. M. Atherosclerosis immunoimaging by positron emission tomography. *Arterioscler. Thromb. Vasc. Biol.* **40**, 865–873 (2020).
191. Ouimet, M. et al. MicroRNA-33-dependent regulation of macrophage metabolism directs immune cell polarization in atherosclerosis. *J. Clin. Invest.* **125**, 4334–4348 (2015).
192. Gerriets, V. A. et al. Leptin directly promotes T-cell glycolytic metabolism to drive effector T-cell differentiation in a mouse model of autoimmunity. *Eur. J. Immunol.* **46**, 1970–1983 (2016).
193. Taleb, S. et al. Defective leptin/leptin receptor signaling improves regulatory T cell immune response and protects mice from atherosclerosis. *Arterioscler. Thromb. Vasc. Biol.* **27**, 2691–2698 (2007).
194. Shirai, T. et al. The glycolytic enzyme PKM2 bridges metabolic and inflammatory dysfunction in coronary artery disease. *J. Exp. Med.* **213**, 337–354 (2016).
This study identifies the glycolytic enzyme pyruvate kinase M2 as a crucial molecular link between altered metabolic adaptations and increased inflammation and oxidative stress in monocytes and macrophages from patients with coronary artery disease.
195. Tomas, L. et al. Altered metabolism distinguishes high-risk from stable carotid atherosclerotic plaques. *Eur. Heart J.* **39**, 2301–2310 (2018).
This study conducts targeted metabolic profiling of human carotid plaques and identifies distinct metabolic signatures that distinguish high-risk atherosclerotic plaques from low-risk atherosclerotic plaques.
196. Christ, A. et al. Western diet triggers NLRP3-dependent innate immune reprogramming. *Cell* **172**, 162–175.e14 (2018).
This study shows that a Western-style diet high in fat and cholesterol triggers NLRP3-dependent inflammasome activation and long-term epigenetic and metabolic reprogramming of myeloid progenitor cells.
197. Bekkering, S. et al. Metabolic induction of trained immunity through the mevalonate pathway. *Cell* **172**, 135–146.e9 (2018).
This study identifies mevalonate, a metabolic intermediate in the cholesterol synthesis pathway, as an inducer of trained immunity in human monocytes.
198. Flores-Gomez, D., Bekkering, S., Netea, M. G. & Riksen, N. P. Trained immunity in atherosclerotic cardiovascular disease. *Arterioscler. Thromb. Vasc. Biol.* **41**, 62–69 (2021).
199. Baardman, J. et al. A defective pentose phosphate pathway reduces inflammatory macrophage responses during hypercholesterolemia. *Cell Rep.* **25**, 2044–2052.e5 (2018).
200. Cheng, H. Y. et al. Loss of ABCG1 influences regulatory T cell differentiation and atherosclerosis. *J. Clin. Invest.* **126**, 3236–3246 (2016).
201. Westerterp, M. et al. Cholesterol accumulation in dendritic cells links the inflammasome to acquired immunity. *Cell Metab.* **25**, 1294–1304.e6 (2017).
202. Ketelhuth, D. F. J. The immunometabolic role of indoleamine 2,3-dioxygenase in atherosclerotic cardiovascular disease: immune homeostatic mechanisms in the artery wall. *Cardiovasc. Res.* **115**, 1408–1415 (2019).
203. Polyzos, K. A. et al. Inhibition of indoleamine 2,3-dioxygenase promotes vascular inflammation and increases atherosclerosis in Apoe$^{-/-}$ mice. *Cardiovasc. Res.* **106**, 295–302 (2015).
204. Cole, J. E. et al. Indoleamine 2,3-dioxygenase-1 is protective in atherosclerosis and its metabolites provide new opportunities for drug development. *Proc. Natl Acad. Sci. USA* **112**, 13033–13038 (2015).
205. Metghalchi, S. et al. Indoleamine 2,3-dioxygenase fine-tunes immune homeostasis in atherosclerosis and colitis through repression of interleukin-10 production. *Cell Metab.* **22**, 460–471 (2015).
206. Sun, J. et al. Deficiency of antigen-presenting cell invariant chain reduces atherosclerosis in mice. *Circulation* **122**, 808–820 (2010).
207. Wigren, M. et al. Lack of ability to present antigens on major histocompatibility complex class II molecules aggravates atherosclerosis in ApoE$^{-/-}$ mice. *Circulation* **139**, 2554–2566 (2019).
208. Bonacina, F. et al. Myeloid apolipoprotein E controls dendritic cell antigen presentation and T cell activation. *Nat. Commun.* **9**, 3083 (2018).
209. Peshkova, I. O., Fatkhullina, A. R., Mikulski, Z., Ley, K. & Koltsova, E. K. IL-27R signaling controls myeloid cells accumulation and antigen-presentation in atherosclerosis. *Sci. Rep.* **7**, 2255 (2017).
210. Tay, C. et al. Follicular B cells promote atherosclerosis via T cell-mediated differentiation into plasma cells and secreting pathogenic immunoglobulin G. *Arterioscler. Thromb. Vasc. Biol.* **38**, e71–e84 (2018).
211. Williams, J. W. et al. B cell-mediated antigen presentation through MHC class II Is dispensable for atherosclerosis progression. *Immunohorizons* **3**, 37–44 (2019).
212. Hilgendorf, I. et al. Innate response activator B cells aggravate atherosclerosis by stimulating T helper-1 adaptive immunity. *Circulation* **129**, 1677–1687 (2014).

Acknowledgements
The authors' work is funded by the US National Institutes of Health (HL136275, HL140976, HL145241, HL146134 and HL148094 to K.L.), a Conrad Prebys Foundation award, a postdoctoral fellowship from the American Heart Association (to M.O.) and the Tullie and Rickey Families SPARK Awards for Innovations in Immunology at La Jolla Institute for Immunology (to P.R. and M.O.).

Author contributions
All authors contributed to all aspects of the article.

Competing interests
K.L. is a co-founder of Atherovax. M.O. and K.L. are named as co-inventors on patents applied for by La Jolla Institute for Immunology relating to cardiovascular diagnostics and therapeutics, and might have the right to receive royalty payments for inventions or discoveries related to cardiovascular diagnostics or therapeutics. P.R. declares no competing interests.

Peer review information
Nature Reviews Immunology thanks K. Moore and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Publisher's note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Title: TET deficiency perturbs mature B cell homeostasis and promotes oncogenesis associated with accumulation of G-quadruplex and R-loop structures

Authors

Vipul Shukla[1#], Daniela Samaniego-Castruita[1#], Zhen Dong[1], Edahi Gonzalez-Avalos[1], Qingqing Yan[3], Kavitha Sarma[3], Anjana Rao*[1, 2, 4]

Affiliations

[1]Division of Signaling and Gene Expression, La Jolla Institute for Immunology, 9420 Athena Circle, La Jolla, California, 92037, USA.

[2]Department of Pharmacology and Moores Cancer Center, University of San Diego, California; 9500 Gilman Drive, La Jolla, California, 92093, USA.

[3]Gene Expression and Regulation Program, The Wistar Institute, Philadelphia, PA 19104, USA; Epigenetics Institute, University of Pennsylvania, Philadelphia, PA 19104, USA.

[4]Sanford Consortium for Regenerative Medicine; 2880 Torrey Pines Scenic Drive, La Jolla, California, 92037, USA.

# these authors contributed equally

* Corresponding author: Dr. Anjana Rao (arao@lji.org)

Abstract

Enzymes of the TET family are methylcytosine dioxygenases that undergo frequent mutational or functional inactivation in human cancers. Recurrent loss-of-function mutations in TET proteins are frequent in human Diffuse Large B-Cell Lymphoma (DLBCL). Here we investigate the role of TET proteins in B-cell homeostasis and development of B cell lymphomas with features of DLBCL. We show that deletion of *Tet2* and *Tet3* genes in mature B cells in mice perturbs B-cell homeostasis and results in spontaneous development of germinal center-derived B cell lymphomas with increased G-quadruplexes and R-loops. At a genome-wide level, G-quadruplexes and R-loops were associated with increased DNA double strand breaks at immunoglobulin switch regions. Deletion of the DNA methyltransferase DNMT1 in TET-deficient B cells prevented expansion of germinal center B cells, diminished the accumulation of G-quadruplexes and R-loops, and delayed B lymphoma development, consistent with the opposing functions of DNMT and TET enzymes in DNA methylation and demethylation. CRISPR-mediated depletion of nucleases and helicases that regulate G-quadruplexes and R-loops decreased the viability of TET-deficient B cells. Our studies suggest a molecular mechanism by which TET loss-of-function might predispose to development of B cell malignancies.

Introduction

The three mammalian TET enzymes (TET1, TET2 and TET3) are Fe(II)), O_2 and α-ketoglutarate dependent dioxygenases that sequentially oxidize 5-methylcytosine (5mC) to 5-hydroxymethyl- (5hmC), 5-formyl- (5fC) and 5-carboxyl- (5caC) cytosine[1-3]. TET enzymes regulate enhancer activity and DNA methylation dynamics during development (including B cell development) [4-8], cell differentiation and cell lineage specification (reviewed in[9-12]). *TET2* gene mutations and/or decreased TET activity have been observed in many hematological malignancies and solid cancers, often through impaired regulation of metabolic enzymes that affect TET activity (reviewed in[9,13-16]). For instance, *TET2* is recurrently mutated in ~10% of Diffuse Large B-cell Lymphoma (DLBCL) [17-19], a heterogeneous malignancy originating in mature B cells undergoing activation and differentiation in germinal centers (GCs). *TET2* mutations

represent an early driver event in DLBCL[6]; in mouse models, deletion of *Tet2* in hematopoietic lineages disrupted GC B cell homeostasis and promoted development of more aggressive lymphomas when the transcription factor BCL6 was constitutively overexpressed[6]. 5hmC deposition has been observed at sites of DNA double strand breaks (DSBs) in HeLa cells[20], and TET2 is associated with degradation of stalled replication forks in BRCA2-deficient mouse cells[21], suggesting that TET proteins regulate genomic integrity.

Two non-canonical DNA structures, R-loops and G-quadruplexes, can act as physical impediments to DNA and RNA polymerases during transcription and DNA replication, and have been linked to replication fork stalling and genome instability (reviewed in[22-26]). R-loops form, mostly at genomic regions with high GC content, when RNA binds to the transcribed strand of DNA, displacing the non-transcribed DNA strand[22,23]. G-quadruplexes often form on the displaced G-rich strand when four guanines, one from each of four tracts of two or more guanine bases interspersed with variable numbers of random nucleotides ($G_{>2}N_{1\text{-}n}G_{>2}N_{1\text{-}n}G_{>2}N_{1\text{-}n}G_{>2}$), form square planar structures known as G-quartets, that are stabilized by Hoogsteen hydrogen bonding[23,25,27-29]. G-quartets stack above each other and the resulting G-quadruplex structure is further stabilized by monovalent cations[24-26]. G-quadruplexes and R-loops are commonly associated with gene promoters, 5' untranslated regions, DNA replication origins, telomeres and other regulatory elements in the mammalian genome[22,24-26,30-32]. The pathological effects of these structures can be mitigated by the concerted actions of diverse nucleases and helicases, among them RNases H1 and H2 which destroy RNA in RNA:DNA hybrids[22,23], and several helicases (e.g. ATRX, FANCD2, and Werner's (WRN) and Bloom (BLM) syndrome Recq-like DNA helicases) that can bind and resolve G-quadruplexes[24-26].

Here we show that profound TET loss-of-function, induced by deletion of the *Tet2* and *Tet3* genes in mature B cells in mice, is associated with the rapid development of DLBCL-like tumors from GC B cells. Like other malignancies associated with TET loss-of-function in primary mouse cells [33-35], Tet2/3-deficient B cells exhibit increased DNA damage, based on increased staining with phosphorylated H2AX (γH2AX). 5hmC deposition has been observed at sites of DNA double strand breaks (DSBs) in HeLa cells[20], and While exploring possible mechanisms for increased γH2AX staining, we noticed a marked accumulation of G-quadruplex (G4) structures and R-loops in expanded TET-deficient B, T and myeloid cells. CRISPR-mediated depletion of RNASEH1 or ATRX, FANCD2 and BLM helicases led to a slight increase in DNA damage and apoptosis in TET-deficient compared to control B cells. Genome-wide mapping and high-throughput genome-wide translocation sequencing (HTGTS) showed a strong correlation of increased G-quadruplex and R-loop structures with increased DNA DSBs in switch regions of immunoglobulin genes in TET-deficient B cells. TET-deficient B cells also showed upregulation of the maintenance DNA methyltransferase DNMT1, as well as a slight but significant increase in DNA methylation at regions associated with G-quadruplexes and R-loops. DNMT1 was recently shown to bind G-quadruplex structures[36], and deletion of DNMT1 in TET-deficient B cells was associated with a striking reduction in GC B cells, decreased levels of G-quadruplex and R-loop structures in the surviving B cells, and a marked delay in B lymphoma development. Together, our findings suggest molecular mechanisms through which TET loss-of-function could promote oncogenesis and genome instability in multiple TET-deficient cell types.

Results

TET deficiency in mature B cells causes B cell lymphoma

Tet2 and Tet3 are the major TET paralogs expressed in mature B cells[4,5,7]. To disrupt Tet2 and Tet3 activity specifically in mature B cells, we generated compound transgenic mice bearing floxed alleles of the *Tet2* and *Tet3* genes (*Tet2*$^{fl/fl}$, *Tet3*$^{fl/fl}$, here termed *double-floxed (Dfl)*), the Cre recombinase under control of the CD19 locus (*CD19Cre*), and the yellow fluorescent protein (YFP) reporter preceded by the

loxP-stop-loxP cassette in the *Rosa26* locus (*Rosa26-YFP*^{LSL}) (**Fig. 1a**). Whereas *CD19Cre* and *Dfl* control mice, and mice with individual deletions of *Tet2* or *Tet3* (*CD19 Tet2 KO* and *CD19 Tet3 KO* mice), survived normally for more than a year of age, *Tet2*$^{fl/fl}$*, Tet3*$^{fl/fl}$*, CD19Cre* (*CD19 DKO*) mice with profound TET deficiency showed spontaneous development of lymphoma with complete penetrance and a median survival age of ~20 weeks (**Fig. 1b**). The disease was marked by lymphadenopathy and splenomegaly as early as 6 weeks of age (**Fig. 1c**), with disruption of normal splenic architecture (Extended data **Fig. 1a**), early expansion of activated CD4 and CD8 T cells (CD44^{+}CD62L^{low}) as well as T follicular helper cells (CD4^{+}PD1^{+}CXCR5^{+}) and progressive expansion of TET-deficient CD19^{+}YFP^{+} B cells (**Figs. 1d;** Extended data **Figs. 1b-1k**). Transfer of the expanded YFP^{+} B cells from 8-week-old *CD19 DKO* (CD45.2^{+}) mice into immunocompetent, CD45.1^{+} congenic host mice recapitulated the lymphoma, whereas transfer of B cells from *Dfl* control mice did not result in long-term engraftment (Extended data **Figs. 1l-1o**).

Transcriptional profiling of *CD19 DKO* and control *Dfl* B cells from 8-week-old mice revealed 2,678 differentially expressed genes (DEGs) in *CD19 DKO* B cells compared with *Dfl* control B cells. 1630 genes were upregulated and 1048 genes were downregulated (log2 fold change ≥2; FDR ≤0.05; **Fig. 1g and** Extended data **table 2**). The transcriptional profile of *CD19 DKO* B cells resembled that of early germinal center (GC) B cells (Extended data **Figs. 2a, 2b**). Compared to control *Dfl* or *CD19Cre* B cells, a subset of *CD19 DKO* B cells expressed low IgD and high Ephrin B1 (EFNB1), characteristic of GC B cells (Extended data **Fig. 2c**). The cells also expressed higher levels of BCL6, AID (activation-induced deaminase, a GC B cell-specific enzyme), IgG1, IgG2b and IgG3, indicating that they were undergoing class switch recombination, albeit without any major changes in germline transcription at the *IghM* (*μ*) and *IgG1* (*γ1*) loci (Extended data **Figs. 2d**-**2g**). Analysis of variable gene segments of immunoglobulin heavy and kappa (κ) light chains (clonotypes) from *Dfl* and *CD19 DKO* B cells showed that the *CD19 DKO* B cells had undergone oligoclonal expansion compared to *Dfl* B cells (Extended data **Figs. 2h**, **2i**). Furthermore, *CD19 DKO* mice showed expansion of GC B cells in Peyer's patches (Extended data **Figs. 2j, 2k**). These findings point to a GC origin for expanded B cells from *CD19 DKO* mice.

Consistent with previous reports in other TET-deficient cell types in mice[33-35], *CD19 DKO* B cells showed increased staining with antibodies to phosphorylated histone H2AX (γH2AX), which marks sites of DNA damage in the genome, compared to B cells isolated from *CD19Cre* and *Dfl* control mice (**Figs. 1e, 1f**). To further characterize the potential mechanisms by which TET deficiency predisposes to increased genomic instability, we scored differentially expressed genes from *CD19 DKO* B cells (**Fig. 1g**) for enrichment of annotated molecular pathways. The most downregulated pathways included signaling and scaffolding components of the Ig B cell receptor, the MHC complex and the RAS GTPase complex, which relay signals essential for B cell differentiation, but interestingly, the major upregulated pathways were linked to alterations in secondary DNA structures and DNA damage signaling (**Fig. 1h**). Compared to control B cells, TET-deficient B cells showed upregulation of mRNAs encoding several helicases and other proteins – FANCD2, BLM, WRN, PIF1, RECQL4, RNASEH2B, nucleolin (Ncl), Ataxin (Atxn1) – that recognize or regulate G-quadruplex (G4) structures and R-loops (**Fig. 1g**). At the protein level, we confirmed that four known G4-binding helicases ATRX, FANCD2, BLM and WRN[25,26], DNMT1 (which also binds G-quadruplex structures[36]), and RNASEH1 (which like the RNASEH2 complex, specifically degrades the RNA strand of RNA:DNA hybrids in R-loops[22,23]) showed increased expression in TET-deficient compared to control *Dfl* B cells (**Fig. 1i**).

TET deficiency is associated with increased G4s and R-loops

We found that TET deficiency resulted in increased levels of G-quadruplex structures and R-loops in *CD19 DKO* B cells. To detect G-quadruplex structures, we used BG4-Ig, a fusion protein of BG4, a single chain variable fragment (scFv) that recognizes G-quadruplex structures[37] fused to the mouse IgG1 constant region to improve its valency and sensitivity (**Fig. 2a**). Treatment of activated primary B cells

with a G4-stabilizing ligand, pyridostatin (PDS)[38] led to a significant increase in BG4-Ig fluorescence signal, authenticating this approach for G-quadruplex detection (Extended data **Figs. 3a, 3b**). *CD19 DKO* B cells from 8 to 10 weeks old mice showed an approximately 1.5- to 2-fold increase in staining with BG4-Ig antibody compared with *Dfl* or *CD19Cre* control B cells (**Figs. 2b, 2c**). We confirmed this result using N-methyl mesoporphyrin IX (NMM), a G4-binding compound that shows a strong increase in fluorescence when it binds G-quadruplex structures[39] (Extended data **Figs. 3c-e**). [38,40]*CD19 DKO* B cells showed a marked increase (~3 fold) in NMM fluorescence signal compared with *Dfl* and *CD19Cre* control B cells (**Figs. 2d, 2e**). Amnis imaging flow cytometry further confirmed increased staining of *CD19 DKO* B cells with both BG4-Ig and NMM, compared with *CD19Cre* control B cells (Extended data **Figs. 3h-3k**).

To detect R-loops, we used a recombinant, V5-epitope-tagged version of a catalytically inactive, mutant (D210N) RNASEH1 protein (rRNASEH1)[41-43] (**Fig. 2a**) in preference to the S9.6 antibody against RNA:DNA hybrids[44] which gave high background staining. CH12 B cells stained with rRNASEH1 displayed a strong fluorescence signal over background that was significantly diminished upon treatment with catalytically active, RNASE H enzyme (Extended data **Figs. 3f, 3g**). There was a clear increase (~2-fold) in R-loop levels in *CD19 DKO* B cells stained with rRNASEH1 compared with *Dfl* or *CD19Cre* control B cells (**Figs. 2f, 2g;** Extended data **Figs. 3l, 3m**). Together, these studies provide compelling evidence for upregulation of G-quadruplex and R-loop structures in TET-deficient B cells.

The increase in G-quadruplex structures and R-loops occurred early (within 12 days) after deletion of the *Tet2 and Tet3* genes in B cells from *Cγ1CreTet2$^{fl/fl}$Tet3$^{fl/fl}$Rosa26-YFP^{LSL}* (*Cγ1 DKO*) mice, in which Cre recombinase expression and TET deletion are induced primarily in GC B cells after immunization. Unlike *CD19 DKO* mice, unimmunized *Cγ1 DKO* mice did not show early signs of morbidity and mortality. Immunization with a model antigen, sheep red blood cells (SRBCs), led to a slight increase in spleen cellularity but no apparent change in the frequency of GC B cells in *Cγ1 DKO* compared with *Cγ1Cre* or *Dfl* control mice (Extended data **Figs. 4a-4c**), but resulted in a significant increase in G4 and R-loop levels in *Cγ1 DKO* GC B cells compared with *Cγ1Cre* or *Dfl* GC B cells (**Fig. 3a-3d**).

We used the 40LB cell culture system, in which B cells are cultured on a mouse fibroblast cell line expressing CD40 ligand and B cell activation factor (BAFF) to induce *Cγ1-Cre* recombinase activity[45], to explore the kinetic relationship between TET loss-of-function and alterations in G-quadruplex and R-loop structures. ~95 to 98% of *Cγ1 DKO* and *Cγ1Cre* B cells cultured on 40LB cells underwent genetic recombination (deletion of the *LoxP-stop-LoxP* cassette), as judged by an increase in YFP expression within 3 days (**Fig. 3e**). Compared with *Cγ1Cre* control B cells, *Cγ1 DKO* B cells displayed a significant increase in G-quadruplex and R-loop levels by 4 days (**Figs. 3f-3i**). In addition, we deleted *Tet* genes acutely by tamoxifen treatment of *ERT2-Cre Tet2$^{fl/fl}$ Tet3$^{fl/fl}$ Rosa26-YFP^{LSL} (ERT2-Cre DKO)* mice; compared to control *Dfl* mice, these mice show similar B cell development *in vivo* and similar proliferation of naïve B cells *in vitro* upon activation with lipopolysaccharide (LPS) and interleukin-4 (IL-4)[7]. Naïve *ERT2-Cre DKO* B cells, as well as *ERT2-Cre DKO* B cells activated in vitro with LPS and IL-4 for 3 days in the presence of tamoxifen (Extended data **Fig. 4d**), exhibited a significant increase in G-quadruplex levels compared to control *Dfl* B cells (Extended data **Figs. 4e, 4f**). Increased levels of G-quadruplexes were also observed in expanded TET-deficient myeloid cells (*ERT2-Cre Tet1$^{fl/fl}$ Tet2$^{fl/fl}$ Tet3$^{fl/fl}$, ERT2creTKO*) and iNKT cells (*CD4Cre Tet2$^{-/-}$ Tet3$^{fl/fl}$, CD4Cre DKO*) relative to their TET-sufficient counterparts (Extended data **Figs. 4g-4j**)[35,46]. In summary, increased levels of G-quadruplexes and R-loops are early events associated with TET deficiency in B cells and other hematopoietic cells.

TET-deficient B cells are sensitive to G4 and R-loop targeting

Since mRNAs encoding G4-resolving helicases and the R-loop resolving enzyme RNase H were upregulated in TET-deficient B cells (**Fig. 1g**), we asked whether depletion of ATRX, BLM, FANCD2 and RNASEH1 proteins would affect the growth or survival of TET-deficient B cells. Cells were nucleofected with Cas9 ribonucleoprotein complexes (Cas9 RNPs) loaded with the appropriate CRISPR guide RNAs

(Extended data **Fig. 5a**), then cultured in the presence of LPS for 48 hours. This procedure led to efficient depletion of the targeted proteins in *CD19 DKO* B cells, compared to their levels in cells nucleofected with a control Cas9 RNP directed to the *Cd4* locus (*Ctrl*) (Extended data **Fig. 5b**). Individual depletion of ATRX, BLM, FANCD2 and RNASEH1 caused a significant increase in apoptosis of *CD19 DKO* B cells but not *Dfl* control B cells (measured by flow cytometry for activated caspase 3; Extended data **Figs. 5c**, **5d**), concomitantly with increased DNA damage and increased G-quadruplex levels (assessed by flow cytometry for γH2AX and NMM respectively; Extended data **Figs. 5e, 5f**). We also used a G4 ligand, pyridostatin (PDS) (Extended data **Fig. 5g**), to stabilize G-quadruplexes in *Dfl* and *CD19 DKO* B cells cultured with LPS with or without PDS for 48 hours. PDS treatment led to an ~2.5-fold increase in apoptosis in *CD19 DKO* B cells but not in *Dfl* control B cells (Extended data **Figs. 5h**, **5k**). These studies show that depletion of enzymes that regulate G-quadruplexes and R-loops increases apoptosis in TET-deficient B cells.

Relation of G4s and R-loops to DNA modifications

We mapped the genome-wide distribution of G-quadruplexes and R-loops in control-*Dfl* and *CD19 DKO* B cells by chromatin immunoprecipitation followed by sequencing (ChIP-seq) using BG4-Ig[47] and MapR[42] (a modified CUT&RUN-based method for mapping R-loops which utilizes a recombinant catalytically inactive RNASEH1 protein fused to micrococcal nuclease) respectively. We defined a total of 9722 regions that showed enrichment for G4 and/or R-loop signals over the local background in replicate experiments (**Fig. 4a**). Consistent with the flow cytometry results, *CD19 DKO* B cells showed increased G4 and R-loop signals compared with *Dfl* control B cells (**Figs. 4b, 4c**). A majority of these (6212 regions; 64%) were present at annotated promoters and 5'UTRs (within +/- 1 kb of transcription start sites (TSS)), and showed enrichment for G-rich DNA sequences predicted to be capable of forming G-quadruplex structures compared to the control regions selected randomly from the genome (Extended data **Figs. 6a, 6b**). The remaining 3510 regions (36%) were located distant from promoters (median distance ~14 kb from TSS). The G-quadruplex signal was increased at both promoter and non-promoter regions, but only promoter regions showed a significant increase in R-loops signal in *CD19 DKO* B cells (Extended data **Figs. 6c, 6d**). The distributions of R-loops and G-quadruplexes at promoter regions were both shifted slightly 5' of the TSS in *CD19 DKO* B cells, consistent with the known propensity for G-quadruplex formation on the displaced single DNA strand of R-loops[22-26,29]. The actual shift in R-loop signal upstream of the TSS might suggest transcriptional pausing at promoter proximal regions, however, the precise reason for the change in the R-loop signal profile in *CD19 DKO* B cells remains to be determined.

To relate the distribution of TET-regulated DNA modifications to the distribution of G-quadruplexes and R-loops, we mapped 5hmC by HMCP (5hmC Pull-down), a method similar to hME-Seal[48] in which 5hmC is glucosylated and biotinylated prior to DNA precipitation using streptavidin beads. As expected, *CD19 DKO* B cells showed a strong depletion of 5hmC across the genome (Extended data **Fig. 6e**). In *Dfl* B cells, we observed a significant enrichment of 5hmC signal at and near (+/- 1kb) G-quadruplex and R-loop forming regions compared to control random regions in euchromatin (Extended data **Fig. 6f**). Whole Genome Bisulfite Sequencing (WGBS) analysis identified 6948 differentially-methylated regions (DMRs) genome-wide, of which 5934 (~85%) were hypermethylated and 1014 (~15%) were hypomethylated in *CD19 DKO* B cells compared with control cells (Extended data **Fig. 6g**). G4- and R-loop-enriched promoter and non-promoter regions displayed a slight but significant increase in DNA methylation in their flanking regions (+/- 1 to 2 kb) in *CD19 DKO* compared to *Dfl* B cells (Extended data **Figs. 6c, 6d, 6h, 7a-7c**). In contrast, random genomic regions in *CD19 DKO* B cells were hypomethylated compared with those in control B cells (Extended data **Fig. 6i**), as we have previously reported for other TET-deficient genomes[33,35,46].

In general, more highly expressed genes in both *Dfl* and *CD19 DKO* B cells showed greater enrichment for G-quadruplexes and R-loops near the TSS (Extended data **Fig. 6j**). We also asked whether the

presence of G-quadruplex structures and R-loops correlated with differential gene expression. Among all differentially expressed genes (DEGs) in *CD19 DKO* B cells, a significant proportion (~36%) harbored a G4- and R-loop-enriched region within 1 kb of their TSSs, compared to the proportion (64%) of DEGs without any G4- and R-loop-region and the proportion (22%) of non-DEGs containing a G4- and R-loop enriched region in their promoters (Extended data **Fig. 6k**, scatter plot and pie chart). However, both up- and down-regulated genes had G-quadruplexes and R-loops at their promoters (Extended data **Figs. 6k, 7a-c**), consistent with the proposed roles of G4 and R-loop structures in both up- and down-regulation of gene expression[22,25].

TET-deficient B cells show genome-wide increase in Ig-translocations

To investigate the potential link between increased G-quadruplexes/R-loop levels and increased DNA breaks in *CD19 DKO* compared to control *Dfl* B cells (**Figs. 1e, 1f**), we performed locus-directed high-throughput genome-wide translocation sequencing (HTGTS[49]). Using a biotinylated capture oligonucleotide bait located at a well-characterized G4- and R-loop-forming region present within the switch μ region in the IgH locus (**Fig. 4d**), we observed a striking increase in the absolute numbers of translocations arising from DNA breaks within or up to 500 bases 5' of this region (**Figs. 4d, 4e, 4g;** Extended data **Fig. 7d**), as well as an increase in the number of genomic loci harboring translocations in *CD19 DKO* compared with *Dfl* B cells (**Figs. 4f, 4g,** Extended data **Fig. 7d and** Extended data **table 3**). The regions captured as translocation partners of the switch μ region by HTGTS showed significant enrichment of G-quadruplex and R-loop signals compared to control regions in both *Dfl* and *CD19 DKO* B cells (**Fig. 4h**). The absolute number of DNA breaks (translocations) that directly overlap G4- and R-loop-enriched regions were also significantly increased in *CD19 DKO* B cells compared with *Dfl* B cells (**Fig. 4i**). The vast majority (~98%) of DNA breaks were located in the IgH locus associated with IgM as well as other Ig isotypes (**Fig. 4d**). Compared to control euchromatic regions, a substantial proportion (~85%) of HTGTS hits harbored canonical motifs predicted to form G-quadruplexes (Extended data **Fig. 7e**) and showed significant enrichment of DNA motifs targeted by the cytidine deaminase AID in B cells (Extended data **Fig. 7f**). Together, these studies show increased G-quadruplexes and R-loops in *CD19 DKO* B cells, correlating with an increase in the levels of DNA double-strand breaks at Ig switch regions.

DNMT1 deletion delays oncogenesis in TET-deficient mice

DNMT and TET enzymes have opposing biochemical activities in DNA methylation and demethylation respectively. A recent study showed that binding of DNMT1 to G-quadruplex structures inhibits its catalytic activity[36]. We confirmed the DNMT1-G-quadruplex interaction (Extended data **Fig. 8a, b**), then asked whether (given the strong upregulation of DNMT1 in TET-deficient B cells (**Fig. 1i**)), depleting DNMT1 in TET-deficient B cells affected the pathogenesis of B cell lymphoma. We generated *Dnmt1 fl/fl Rosa26-YFP^{LSL}* mice expressing *Cre* recombinase under control of the CD19 locus (*CD19Cre Dnmt1 knock out (KO))* and bred them with *Tet2fl/fl Tet3fl/fl* mice (Triple floxed-*Tfl*) (**Fig. 5a**). The *CD19 DKO* mice generated from these crosses succumbed to B cell lymphoma with a median survival of 20 weeks, whereas *CD19 TKO* mice (with additional deletion of *Dnmt1*) showed a substantial increase in survival (median survival 98 weeks) (**Fig. 5b**). Compared with *CD19 DKO* mice, *Tfl*, *CD19 Dnmt1 KO* and *CD19 TKO* mice displayed no apparent signs of splenomegaly at 10 weeks of age (**Figs. 5c;** Extended data Figs. **8c, 8d**); a small proportion (4 out of 14 mice) of *CD19 TKO* mice did develop splenomegaly, albeit with a very long latency (Extended data **Fig. 8e**).

Since GC B cells are the cell-of-origin of the expanded B cells in *CD19 DKO* mice, we examined the numbers of GC B cells in Peyer's patches of *CD19 DKO* versus *Dnmt1-Tet2/3 CD19 TKO* mice. Deletion of *Dnmt1* in B cells completely abrogated GC B cells (**Figs. 5d, 5e**), similar to previous reports using DNMT1 germline hypomorphs[50]. *CD19 TKO* mice displayed GC B cell frequencies considerably lower

than the expanded GC B cell population in *CD19 DKO* mice, but only slightly lower than those in *Tfl* control mice when gated on YFP+ B cell population (**Figs. 5d, 5e**). Together, these findings point to a functional interplay between TET and DNMT1 activities in GC B cell development and oncogenic transformation of TET-deficient B cells.

We assessed the levels of G-quadruplexes and R-loops in total splenic B cells and Peyer's patch GC B cells of 10-week-old mice by flow cytometry. *Tet2/3, Dnmt1*-deficient *CD19 TKO* B cells demonstrated a notable decrease in the levels of G4 (using both NMM and BG4-Ig staining) and R-loop (rRNASE H1 staining) structures compared with *Tet2/3*-deficient *CD19 DKO* B cells, almost to the levels of control *Tfl* B cells (**Figs. 5f-l;** Extended data Figs. **8f, 8g**). The decrease in the levels of G4 and R-loop structures in *CD19 TKO* B cells was also accompanied by decreased levels of γH2AX compared with *CD19 DKO* B cells (Extended data **Fig. 8h**). Similar decreases in G-quadruplex and R-loop levels were observed when *CD19 TKO* GC B cells isolated from Peyer's patches were compared to *CD19 DKO* GC B cells (Extended data **Figs. 8i-8l**). Altogether, these results show that DNMT1 deletion inhibits the development of GC B cells to delay oncogenesis, and is associated with decreased levels of G-quadruplex and R-loop structures (Extended data **Fig. 8m**).

Discussion

Our studies establish a causal relationship between TET deficiency and the development of mature B cell neoplasms. The fully penetrant progression of B cell lymphomas in *CD19 Tet2/3 DKO* mice is consistent with the frequent occurrence of *TET* gene mutations or dysregulation of TET activity in human DLBCL[17-19]. We show that profound inactivation of TET function, through deletion of both the *Tet2* and *Tet3* genes, perturbs normal B cell homeostasis, leading to spontaneous expansion of B cells with a GC phenotype. We also document a clear association between TET deficiency, increased G-quadruplexes and R-loops, and increased DNA double strand breaks, particularly at Ig-switch regions in B cells; indeed, accumulation of G-quadruplex and R-loop structures is a novel feature of TET deficiency in B cells and cells of other hematopoietic lineages. TET-deficient B cells upregulated mRNAs encoding several proteins that recognize and regulate R-loops and G-quadruplex structures, including RNase H1, *Rnaseh2b*, DNMT1, *Fancd2*/FANCD2, ATRX, BLM, WRN*, Recql4*, and *Pif1*[22-26]. Genetic deletion of one of the G4 binders, DNMT1, in *CD19 Tet2/3 DKO* B cells was associated with decreased levels of the precursor GC B cells, decreased G-quadruplex and R-loop structures and a notable increase in the survival of *Tet2/3-deficient* mice. Together with the de novo DNA methyltransferases Dnmt3a and 3b, Dnmt1 is expected to maintain the increased DNA methylation resulting from *Tet* gene deletion in cells; thus, *Dnmt1* deletion might be expected to counter, at least partially, any deleterious effects resulting from altered DNA methylation in Tet2/3-deficient B cells.

A recent study showed that TET proteins limit the activation of self-reactive B cells in the periphery and serve as critical regulators of B cell tolerance[51]. While the identity of the antigen(s) driving the expansion of *CD19 DKO* B cells in our system is not known, it is possible that TET-deficient B cells expand because they are self-reactive. This is consistent with the well-documented relationship between the expression of self-reactive B cell receptors and two mature B cell-derived malignancies, chronic lymphocytic leukemia (CLL) and the activated B cell (ABC) subtype of DLBCL, in humans[52]. The oligoclonal (rather than polyclonal) nature of the BCRs observed in expanded *CD19 DKO* B cells implies antigen-driven selection[52]. Two possibilities, not mutually exclusive, warrant further investigation. The intracellular and transcriptional signals emanating from BCRs may be regulated by TET enzymes to limit cellular expansion; alternatively, loss of TET proteins may inhibit a physiological brake on antigen-driven expansion through a mechanism that does not involve transcription.

G-quadruplex structures have been implicated in initiation of the frequent *IgH-BCL2* translocations in follicular lymphoma[53,54]. We observed an association between increased G-quadruplex and R-loop structures and DNA breaks at the switch regions (primarily switch μ) of Ig isotypes in Tet2/3-deficient B

cells. Due to the lack of robust methods for genome-wide mapping of DNA breaks, we relied on HTGTS, a locus-directed method for DNA break mapping at single base resolution[49]. Since HTGTS maps DNA breaks relative to DNA junctions formed by break ligation (with possible errors due to indels), it is likely that we have underestimated the total number of DNA breaks. Nevertheless, we observed a striking increase in the absolute numbers of DNA DSBs in *CD19 DKO* B cells when compared with control B cells. The DNA breaks could arise either from conflicts of G-quadruplex and R-loop structures with transcription or DNA replication machineries, or from specific targeting of AID cytidine deaminase, the B cell mutator, to G4 structures and R-loops. In fact, AID possesses a G4-binding activity that is important for its genome-wide targeting[55-57]. A recent study reported an association between G4 structures and AID activity at commonly mutated genes in B cell lymphoma[56]; consistent with these findings, we observed an enrichment of AID hotspot motifs within translocation sites with increased levels of G-quadruplexes and R-loops. The detailed mechanisms through which G-quadruplexes and R-loops recruit AID and/or promote genomic instability in B cells remain to be addressed.

TET-dependent methylation changes do not necessarily play a direct biochemical role in the observed increase of R-loop and G quadruplex structures in Tet2/3 deficient B cells. We based this conclusion on the fact that although we observed a slight but significant link between loss of TET activity and increased DNA methylation in the vicinity of R-loops and G-quadruplex structures, we did not observe any dramatic changes in DNA methylation at the R-loop/G-quadruplex-containing regions (median size ~750 bp) themselves. Thus, further studies are needed to fully understand how TET deficiency in multiple cell types results in increased levels of G-quadruplex and R-loop structures.

DNMT1 has been previously shown to be induced in GC B cells[50], where it may be needed to maintain the DNA methylation landscape during rapid GC B cell proliferation. Thus, the apparent upregulation of DNMT1 in TET-deficient B cells is likely due to their prominent GC phenotype compared to WT B cells. Conversely, *Dnmt1* deletion results in a dramatic decrease of GC B cells, potentially the primary reason for the delayed oncogenesis observed in triple *Tet2/3, Dnmt1*-deficient B cells. In fact, the late-onset cancers may reflect delayed outgrowth of cells that escaped deletion of one or both Dnmt1 alleles, since in one experiment (*unpublished data*), PCR analysis of YFP+ total B cells and GC B cells in Peyer's patches in much older (7-month-old) mice showed almost complete deletion of *Tet2* and *Tet3* but only ~50% deletion of Dnmt1. The decrease in R-loop and G-quadruplex structures observed upon *Dnmt1* deletion in total *Tet2/3*-deficient B cells suggests a possible role for DNMT1 activity in stabilization of these structures. DNMT1 is known to be overexpressed in several different hematological and solid cancers[58]. The potential interplay between TET and DNMT activities in regulating oncogenesis as well as R-loop and G-quadruplex levels in B cells and other cell types remains to be investigated.

Our studies suggest that G-quadruplexes and R-loops could be therapeutic vulnerabilities in cancers with TET loss-of-function. G4-stabilizing ligands were recently shown to decrease cell viability in ATRX-deficient gliomas and BRCA1/2-deficient tumor cells[59,60]. In our hands, the use of a G4-stabilizing ligand, or depletion of proteins known to regulate G4 or R-loop structures, was associated with increased DNA DSBs and a slight increase in apoptosis in TET-deficient B cells. Furthermore, deletion of DNMT1 in TET-deficient B cells prevented the accumulation of R-loop and G-quadruplex structures in splenic B cells and Peyer's patch GC B cells and rescued the survival of TET-deficient mice. Follow-up studies in pre-clinical models could test whether a combination of G4-stabilizing agents and DNA methyltransferase inhibitors might synergize to delay the onset and/or progression of B cell lymphomas and other malignancies with TET loss-of function.

Acknowledgements. We thank Drs. H. Yuita and I. Lopez-Moyado for generating the *ERT2-Cre Tet1$^{fl/fl}$ Tet2$^{fl/fl}$ Tet3$^{fl/fl}$, ERT2creTKO* mice; Dr. D. Kitamura at the Tokyo University of Science for sharing the 40LB cells; Drs. U. Basu and B. Laffleur at Columbia University for help with the HTGTS protocol; our collaborators at Cambridge Epigenetix (UK) for providing the 5hmC mapping kits; the LJI Flow Cytometry

Core team: C. Kim, D. Hinz, C. Dillingham, M Haynes, S. Ellis for help with cell sorting; and the LJI Next generation sequencing core members: J. Day, S. Alarcon, H. Dose, K. Tanaguay and A. Hernandez for help with sequencing. BD FACSAria II is supported by NIH (NIH S10OD016262, NIH S10RR027366) and our research used resources of the Advanced Light Source, which is a DOE Office of Science User Facility under contract no. DE-AC02-05CH11231. The NovaSeq 6000 and the HiSeq 2500 were acquired through the Shared Instrumentation Grant (SIG) Program (S10); NovaSeq 6000 S10OD025052 and HiSeq 2500 S10OD016262. K.S. acknowledges support from National Institutes of Health Grants DP2-NS105576. V.S. was supported by Leukemia and Lymphoma Society Postdoctoral Fellowship (grant ID: 5463-18) and currently by a K99/R00 award from National Cancer Institute (grant ID: CA248835). D.S.C. and E.G.A were supported by CONACYT/ UCMEXUS Fellowships. This work is supported by the National Institutes of Health (NIH) grants R35 CA210043, R01 AI109842 and AI128589 to A.R and, K99/R00 CA248835, research funds from LLS grant 5463-18 and the Tullie and Rickey families SPARK award from LJI to V.S.

Author Contributions Statement. V.S. and D.S.C. acquired, analyzed and interpreted the data. Z.D. performed R-loop mapping experiments and helped with the interpretation of results. E.D.A and D.S.C. performed the bioinformatics analysis. Q.Y. and K.S. provided the plasmids and suggestions for MapR. A.R. and V.S. supervised the studies, conceptualized the experiments and helped with the interpretation of data. V.S., D.S.C. and A.R. wrote the manuscript. All authors were involved in reviewing and editing the manuscript.

Competing Interests Statement. The Authors declare no competing financial interests.

Tables

Extended data table 1: Oligonucleotides and Primers

Extended data table 2: RNA-Seq analysis for differentially expressed genes

Extended data table 3: HTGTS hits

Figure Legends

Figure 1. TET deficiency in mature B cells causes B cell lymphoma. a) Breeding strategy used to generate single *Tet2* or *Tet3*-deficient mice (*CD19 Tet2KO* or *CD19 Tet3KO*) and double *Tet2, Tet3*-deficient mice (*CD19 DKO*). **b)** Kaplan-Meir curves displaying overall survival of control *Dfl* (blue) and *CD19 Cre* (green) mice, *CD19 Tet2KO* (black), *CD19 Tet3KO* (grey) and *CD19 DKO* (red) mice. Y-axis denotes percent survival and X-axis shows time in weeks. Only the *CD19 DKO* mice develop B cell lymphoma in the 52-week time period shown. **c)** Enlarged spleen and lymph nodes in 18 week-old *CD19 DKO* and control mice. **d)** Flow cytometry plots of splenocytes from 18-week-old *CD19Cre*, *Dfl and CD19 DKO* mice. Numbers show frequencies of B cells among splenocytes. **e)** Flow cytometry histograms showing γH2AX staining compared to isotype IgG controls in B cells from 8-week-old *CD19cre (YFP$^+$)*, *Dfl and CD19 DKO (YFP$^+$)* mice. **f)** Quantification of the median fluorescence intensity (MFI) of γH2AX in *CD19Cre (YFP$^+$)*, *Dfl and CD19 DKO (YFP$^+$)* B cells from 4 independent experiments. **g)** MA plot of RNA-sequencing data, displaying changes in gene expression in *CD19 DKO* B cells compared to *Dfl* B cells. The highlighted genes represent known G4 and R-loop binders. **h)** Upregulated (red) and downregulated pathways (blue) in *CD19 DKO* compared to *Dfl* B cells. X-axis, Z-score. * Asterisks highlight pathways related to DNA structures and DNA damage. **i)** Protein levels of RNASE H1 and selected G4-binding helicases and proteins in *Dfl* and *CD19 DKO* B cells, assessed by immunoblotting. The data is representative of at least 2 independent experiments. Statistical significance is calculated using the log-

rank test (**b**) or one-way ANOVA (**f**). Error bars represent mean +/- standard deviation, *** p value ≤0.001, **** p value ≤0.0001.

Extended data figure 1. TET deficiency leads to development of mature B cell lymphoma. a) H&E staining in spleens from 8 week-old *Dfl* and *CD19 DKO* mice The data is representative of 2 independent experiments. **b)** Kaplan-Meir curves displaying lymphoma-free survival of *CD19 DKO* (red) and control *Dfl* (blue) and *CD19 Cre* (green) mice. Y-axis, percent mice without lymphoma, defined by ≥ 3-fold increase in spleen weight and ≥ 2-fold increase in cellularity. **c)** – **f)** Quantification of **c)** cell numbers, **d)** spleen weights, **e)** B cell numbers, **f)** percent of B cells in spleens of 9 week-old *CD19 DKO* (red) and control *Dfl* (blue) and, *CD19 Cre* (green) mice from at least 5 biological replicates and 4 independent experiments. **g)** Percent of B cells in spleens of *CD19 DKO* (red) and control *Dfl* (blue) and *CD19 Cre* (green) mice at different ages from 3 independent experiments. **h)** Quantification of cell numbers of $CD4^+$ T, $CD8^+$ T and $CD4^+$ Tfh cells from spleens of 9 week-old *CD19 DKO* (red) and control *Dfl* (blue) mice from 3 biological replicates and 2 independent experiments. **i)** Percent of activated $CD4^+$ and, $CD8^+$ T cells, and CD4 Tfh cells from spleens of 9 week-old *CD19 DKO* (red) and control *Dfl* (blue) mice from 3 biological replicates and 2 independent experiments. **j) and k)** Flow cytometry plots of **j)** Activated $CD62L^{low}$ (X-axis) and $CD44^+$ (Y-axis) $CD4^+$ and $CD8^+$ T cells and **k)** $CD4^+$ Tfh cells $PD1^+$ (Y-axis) $CXCR5^+$ (X-axis) from spleen of 9 week-old *CD19 DKO* (red) and control *Dfl* (blue) mice. Numbers represent frequency of gated populations. **l)** Scheme of adoptive transfer experiment. B cells from *Dfl* and *CD19 DKO* CD45.2 mice were transferred retro-orbitally into sub-lethally irradiated CD45.1 immunocompetent host mice. **m)** Kaplan-Meir curves displaying overall survival of CD45.1 host mice transplanted with B cells from *Dfl* or *CD19 DKO* mice. X-axis denotes time in weeks after transplantation. **n)** Enlarged spleens in CD45.1 host mice transplanted with *CD19 DKO* compared to *Dfl* B cells 8 weeks after adoptive transfer. **o)** Representative flow cytometry data from the spleens of CD45.1 host mice 8 weeks after transplantation with *Dfl and CD19 DKO* B cells. Frequencies of the $CD45.1^+$ and $CD45.2^+$ cell populations are shown. Statistical significance is calculated using the log-rank test for **b**) and **m**), one-way ANOVA for **c**)-**f**) and two-way ANOVA for **g**)-**i**). Error bars represent mean +/- standard deviation, ** p value ≤0.01 *** p value ≤0.001, **** p value ≤0.0001.

Extended data figure 2. Expanded B cells from *CD19 DKO* mice have a germinal center (GC) origin. a), b) Gene set enrichment analysis (GSEA) plots showing enrichment for a GC B cell transcriptional signature in the transcriptional profile of *CD19 DKO* compared to *Dfl* B cells, using gene sets from **a)** GC versus follicular B cells and **b)** early GC versus late GC B cells. Y-axis denotes enrichment score. NES, Normalized enrichment score, FDR, False discovery rate. **c)** Representative flow cytometry data gated on splenic B cells from 8 week-old *Dfl*, *CD19Cre (YFP$^+$) and CD19 DKO (YFP$^+$)* mice. Numbers in the rectangles represent frequencies of GC B cells, identified as $EFNB1^+$ (Y-axis) and IgD^{low} (X-axis). **d)** Representative flow cytometry data showing Ig isotype expression, gated on splenic B cells from 8 week-old *Dfl*, *CD19Cre (YFP$^+$)* and *CD19 DKO (YFP$^+$)* mice. Top, IgG1; middle, IgG2b; bottom, IgG3 X-axis shows expression of the default IgD isotype. Numbers represent frequencies of gated cell populations. **e)** Immunoblot showing AID expression in *Dfl* and *CD19 DKO* B cells (2 replicate experiments). Actin is used as a loading control. **f)** Relative fold-change in expression of *μ* and *γ1* germline, *Tet3*, *Irf4* and *Myc* transcripts measured by qRT-PCR in *Dfl* and *CD19 DKO* B cells from 2 biological replicates. **g)** Histogram (*left panel*) and bar-graph (*right panel*) showing staining with BCL6 antibody compared to isotype IgG controls in B cell from 8 week-old *CD19Cre*, *Dfl and CD19 DKO* mice from 3 independent experiments. **h) and i)** Bar plots displaying the proportions of **(h)** IgVH and **(i)** Igκ clonotypes (rearranged variable gene segments) from *Dfl* (blue) and *CD19 DKO* (red) B cells, identified from BCR repertoire analysis of RNA-Seq data. Y-axis represents the proportion of each clonotype. Each individual IgVH and Igκ clonotype is displayed using a different color in the bar plots. Numbers at the bottom represent the number of

clonotypes identified in two independent replicates of *Dfl* (blue) and *CD19 DKO* (red) B cells. **j)** Representative flow cytometry data gated on Peyer's patch B cells from 8 week-old *Dfl and CD19 DKO* mice. Numbers represent frequency of GC B cells, identified as FAS^+ (Y-axis) and $CD38^-$ (X-axis). **k)** Quantification of GC B cell frequency in Peyer's patches of *Dfl* (blue) and *CD19 DKO* (red) mice from 3 independent experiments. Statistical significance is calculated using two-tailed student t-test in **f**), **k**) and one-way ANOVA in **g**). Error bars represent mean +/- standard deviation in **f**), **g**) and **k**). ** p value ≤0.01.

Figure 2. TET deficiency is associated with increased levels of G-quadruplexes and R-loops. **a)** Diagrammatic representation of a G-quadruplex with an associated R-loop structure, illustrating the reagents used for detection of G-quadruplexes and R-loops. All experiments were performed in 8 to 10 week-old mice. **b)** Flow cytometric detection of G-quadruplexes by staining of permeabilized cells with BG4-Ig antibody or isotype IgG controls in B cells from *CD19cre (YFP$^+$)*, *Dfl and CD19 DKO (YFP$^+$)* mice. **c)** Quantification of median fluorescence intensity (MFI) of BG4-Ig signal from *CD19cre (YFP$^+$)*, *Dfl and CD19 DKO (YFP$^+$)* B cells from 4 independent experiments. **d)** Flow cytometric detection of G-quadruplexes after incubation of cells with NMM or DMSO vehicle controls (Veh) in B cells from *CD19Cre (YFP$^+$)*, *Dfl and CD19 DKO (YFP$^+$)* mice. **e)** Quantification of median fluorescence intensity (MFI) of NMM signal from *CD19Cre*, *Dfl and CD19 DKO* B cells from 6 independent experiments. **f)** Flow cytometric detection of R-loops using V5-epitope-tagged recombinant RNASE H1 (rRNASE H1) in B cells from *CD19Cre (YFP$^+$)*, *Dfl and CD19 DKO (YFP$^+$)* mice. Samples stained with anti-V5 and anti-rabbit secondary antibodies were used as controls (IgG). **g)** Quantification of median fluorescence intensity (MFI) of R-loops (rRNASE H1) signal from *CD19cre (YFP$^+$)*, *Dfl and CD19 DKO (YFP$^+$)* B cells from 4 independent experiments. Statistical significance is calculated using one-way ANOVA in **c**), **e**) and **g**). Error bars represent mean +/- standard deviation, ** p value ≤0.01, *** p value≤0.002, **** p value ≤0.0001.

Extended data figure 3. TET deficiency is associated with increased levels of G-quadruplexes and R-loops. **a)** Flow cytometric detection of G-quadruplexes with BG4-Ig antibody or isotype IgG controls in primary B cells stimulated with 25 µg/ml LPS for 48 hours and treated with 10 µM pyridostatin (PDS, G4 ligand) for 24 hours. Numbers represent median fluorescence intensity. **b)** Quantification of median fluorescence intensity (MFI) of BG4-Ig signal from primary B cells treated with (red) or without (blue) PDS. Lines connect paired samples treated with or without PDS from 3 independent experiments. **c)** Fluorescence emission spectrum of NMM in the presence of a G4-forming oligonucleotide (oligo) from the human *c-Kit* gene promoter or a control oligo in which guanines in G4-forming regions (G-tracts) were mutated. **d)** Fluorescence enhancement over background (no oligos) for NMM at 610 nm in presence of known G4-forming oligos (Kit1, Kit2, Spb1) from the *c-Kit* gene locus or the telomeric repeat (Telo) or their respective mutated versions. **e)** G-quadruplex levels assessed by NMM or DMSO vehicle control (Veh) staining in untreated CH12 B cells or cells treated with 5 µM pyridostatin (PDS, G4 ligand) for 24 hours. Numbers represent median fluorescence intensity from 3 independent experiments. **f)** Flow cytometric detection of R-loops using V5-epitope-tagged recombinant RNASE H1 (rRNASE H1) in CH12 cells with or without RNASE H enzyme digestion during staining. Numbers represent median fluorescence intensity. **g)** Quantification of median fluorescence intensity (MFI) of R-loops (rRNASE H1) signal from CH12 cells with (red) or without (blue) RNASE H enzyme digestion. Lines connect paired samples with or without RNASE H digestion from 3 independent experiments. **h) - m)** Representative images (**h, j, l**) and quantification of mean fluorescence signal (**i, k, m**) of *CD19cre and CD19 DKO* YFP^+ B cells stained with DAPI or propidium iodide and CD19, BG4-Ig (**h, i**), NMM (**j, k**) and rRNASE H1 (**l, m**) or respective controls using the AMNIS imagestream. Data are from two independent experiments. Statistical significance is calculated using paired student t-test in **b**) and **g**), two-tailed student t-test **i**), **k**) and **m)**. Error bars represent mean +/- standard deviation, ** p value ≤0.01, **** p value ≤0.0001.

Figure 3. Acute TET deletion is associated with increased levels of G-quadruplexes and R-loops. **a)** Flow cytometric detection of G-quadruplexes with NMM versus DMSO vehicle staining controls (Veh) in GC B cells from *Cγ1Cre* or *Dfl* control and *Cγ1 DKO* mice. **b)** Quantification of median fluorescence intensity (MFI) of NMM signal from *Cγ1Cre* or *Dfl* control and *Cγ1 DKO* GC B cells from 3 independent experiments. **c)** Flow cytometric detection of R-loops (rRNASE H1) signal versus IgG staining controls (IgG) in GC B cells from *Cγ1cre* or *Dfl* control and *Cγ1 DKO* mice. **d)** Fold change in median fluorescence intensity (MFI) signal of R-loops (rRNASE H1) *Cγ1 DKO* GC B cells relative to *Cγ1Cre* or *Dfl* control GC B cells from 2 independent experiments. **e)** Percent YFP^+ B cells from *Cγ1Cre* control and *Cγ1 DKO* mice cultured on 40LB media, X-axis denotes time (in days) of culture. **f)** Flow cytometry detection of G-quadruplexes with BG4-Ig antibody or isotype IgG controls in *Cγ1Cre* and *Cγ1 DKO* B cells at day 4 of culture on 40LB feeder cells. **g)** Quantification of median fluorescence intensity (MFI) of G-quadruplexes (BG4-Ig) signal from *Cγ1cre* and *Cγ1 DKO* B cells at day 3 and 4 of culture on 40LB cells from 3 biological replicates. **h)** Flow cytometry detection of R-loops (rRNASE H1) signal versus anti-rabbit secondary antibody controls (IgG) in *Cγ1cre* and *Cγ1 DKO* B cells at day 4 of culture on 40LB cells. **i)** Quantification of median fluorescence intensity (MFI) of R-loops (rRNASE H1) signal from *Cγ1cre* and *Cγ1 DKO* B cells at day 3 and 4 of culture on 40LB cells from 3 biological replicates. Statistical significance is calculated using two-tailed student t-test **b**) and **d**) and two-way ANOVA in **g**) and **i**). Error bars represent mean +/- standard deviation, * p value ≤0.05, *** p value≤0.002.

Extended data figure 4. TET deficiency in multiple primary cell types is associated with increased DNA G-quadruplex structures. a) Representative flow cytometry data gated on splenic B cells from 8 week-old *Cγ1Cre*, *Dfl* and *Cγ1 DKO* mice 12 days after immunization with SRBCs. GC B cells are identified as FAS^+ (Y-axis) and $CD38^-$ (X-axis). Numbers represent frequency of GC B cells. **b)**, **c)** Quantification of **(b)** GC B cell frequencies and **(c)** absolute numbers of splenocytes from 8 week-old *Cγ1Cre*, *Dfl* and *Cγ1 DKO* mice 12 days after immunization with SRBCs from 3 independent experiments. **d)** Experimental design. *ERT2Cre DKO* or control *Dfl* mice were injected for 5 consecutive days with tamoxifen to induce Cre expression and *TET* deletion, then rested for 2 days. Splenic B cells were activated for 72 hours *in vitro* with LPS and IL-4 in the presence of 4-hydroxytamoxifen (4-OHT). **e)** G-quadruplex levels in naïve (*left panel*) and activated (*right panel*) B cells from tamoxifen-treated *ERT2Cre DKO (YFP^+)* or control *Dfl* mice. Numbers represent median fluorescence intensity. **f)** Quantification of median fluorescence intensity (MFI) of NMM signal from naïve and activated B cells from *ERT2Cre DKO (YFP^+)* or control *Dfl* mice from 3 independent experiments. **g)**, **i)** G-quadruplex levels assessed by NMM or DMSO vehicle staining (Con) in **(g)** transferred $CD45.2^+$ myeloid cells from *ERT2Cre TKO (YFP^+)* or control *Tfl* mice, and **(i)** transferred $CD45.2^+$ T cells from *CD4Cre DKO (YFP^+)* or control *Dfl* mice. **h)**, **j)** Quantification of median fluorescence intensity (MFI) of NMM signal in **(h)** transferred $CD45.2^+$ myeloid cells from *ERT2Cre TKO (YFP^+)* or control *Tfl* mice from 2 biological replicates and **j)** transferred $CD45.2^+$ T cells from *CD4Cre DKO (YFP+)* or control *Dfl* mice from 3 biological replicates. Statistical significance is calculated using one-way ANOVA in **c**), two-way **f**) and two-tailed student t-test in **h**) and **j**). Error bars represent mean +/- standard deviation, * p value ≤0.01, ** p value ≤0.005.

Extended data figure 5. Increased apoptosis and DNA DSBs in TET-deficient B lymphoma cells depleted of enzymes that resolve G-quadruplexes and R-loops. a) Experimental design. Primary B cells from *Dfl* and *CD19 DKO* B cells were nucleofected with Cas9 RNPs loaded with sgRNAs against *Rnase H1* or the known G4-binding helicases *Atrx, Blm* and *Fancd2*, then stimulated with 10 µg/ml LPS for 48 hours before assessing the frequency of apoptotic cells by flow cytometry for cleaved Caspase 3. **b)** Representative immunoblots showing decreased protein levels of ATRX, BLM, FANCD2 and RNASE H1 in *CD19 DKO* B cells nucleofected 48 hours earlier with the corresponding or *CD4* Cas9 RNPs (*Ctrl*).

The data is representative of at least 2 independent experiments. **c)** Representative flow cytometry plots quantifying percent apoptotic cells in *Dfl* and *CD19 DKO* B cells nucleofected with Cas9 RNPs. Y-axis, staining for cleaved Caspase 3; X-axis, forward scatter (FSC). **d)** Quantification of apoptosis, measured as percent of cells showing staining for cleaved Caspase 3, in cells nucleofected with Cas9 RNPs to *Atrx, Blm*, *Fancd2, Rnase H* from 3 biological replicates*.* **e)** Quantification of G-quadruplexes as NMM median fluorescence intensity (MFI) in *Dfl* and *CD19 DKO* B cells 48 hours after nucleofection with Cas9 RNPs. The signal is normalized to the NMM MFI of the same biological sample nucleofected with *CD4* Cas9 RNPs (*Ctrl*) from 3 biological replicates. **f)** Quantification of DNA DSBs, assessed by γH2AX median fluorescence intensity (MFI) in *Dfl* and *CD19 DKO* B cells 48 hours after nucleofection with the indicated Cas9 RNPs from 3 biological replicates. The signal is normalized to the γH2AX MFI of the same biological sample nucleofected with control Cas9 RNP loaded with sgRNA against *CD4 (Ctrl)*. **g)** Experimental design. *Dfl* and *CD19 DKO* B cells were treated for 2 days with the G-quadruplex stabilizing compound pyridostatin (PDS) prior to activation for 48 hours with LPS. **h)** Quantification of apoptosis, measured as percent of cells showing staining for cleaved caspase 3, in *Dfl* and *CD19 DKO* B cells cultured without (untreated) or with 10 μM PDS from 5 biological replicates and 3 independent experiments. **i)** Representative flow-cytometry plots quantifying percent apoptotic cells in *Dfl* and *CD19 DKO* B cells with or without PDS treatment. Y-axis, staining for cleaved caspase 3; X-axis, forward scatter (FSC). Statistical significance is calculated using two-way ANOVA. Error bars represent mean +/- standard error, * p value ≤0.05, ** p value ≤0.01, *** p value ≤0.0005 **** p value <0.0001.

Figure 4. TET-deficient B cells show a genome-wide increase in G-quadruplexes and R-loops and increased translocations to Ig switch regions. a) Heatmap showing enrichment of G-quadruplex (G4) structures (average of 2 replicates) and R-loops (average of 3 replicates) in *Dfl* and *CD19 DKO* B cells. Reads per million (RPM) values in 9722 regions with overlapping high intensity signals for both G-quadruplexes and R-loops are plotted in +/- 2 kb windows from the center of the region. **b)**, **c)** Box and whisker plots quantifying enrichment (RPKM, reads per kilobase per million) of **(b)** G-quadruplex structures (2 biological replicates) and **(c)** R-loops (3 biological replicates) in the 9722 regions from *Dfl* and *CD19 DKO* B cells. **d)** Genome browser view showing data from *Dfl* (blue tracks) and *CD19 DKO* (red tracks) B cells. Green and blue bars at the bottom show the location of Sμ and G4 regions. The red arrow indicates the bait sequence located 5' of the switch μ (Sμ) region used to capture translocations by HTGTS. The zoomed in panel on the right shows the distribution of signals around the *IghM* region. **e)- f)** Quantification of **(e)** total number of translocations (hits), and **(f)** the number of genomic loci to which translocations occur in *CD19 DKO* and *Dfl* B cells from 2 biological replicates. **g)** Circos plots to visually depict all translocations identified by HTGTS in *Dfl* and *CD19 DKO*. Colored lines connect the Sμ bait with the translocation partner regions. Color scale represents the number of translocations identified in a 10 kb window. Translocations from two *Dfl* and *CD19 DKO* replicates were concatenated for this representation. **h)** Box and whisker plots quantifying G-quadruplexes (*left panel*) and R-loops (*right panel*) RPKM signal in translocation partner regions of translocations identified in *CD19 DKO* and *Dfl* B cells. Regions were extended +/-1 kb from the center of the junctions in the translocation partner regions from 2 biological replicates. **i)** Quantification of the number of translocations (hits) overlapping G-quadruplexes and R-loops regions in *CD19 DKO* and *Dfl* B cells from 2 biological replicates. Statistical significance is calculated using Kruskal-Wallis test and the ad hoc Dunn's test in **b**), **c**) and **h)**. Boxes in box and whisker plots represent median (center) with 25th to 75 th percentile and whiskers represent maxima/minima. Error bars represent mean +/- standard deviation, ***** p value <0.000001.

Extended data figure 6. TET deficiency is associated with genome-wide accumulation of G-quadruplexes and R-loops. **a)** Genome annotations of regions enriched for G-quadruplexes (G4) and R-loops (*right bar*) compared to their representation in the mouse genome (mm10) (*left bar*). **b)** Relative

representation of different classes of motifs predicted to form G-quadruplexes (pG4) in control regions selected randomly from the genome (*left bar*) and regions enriched for G-quadruplexes and R-loops (*right bar*). **c)** Heat maps showing enrichment (RPM) for G-quadruplexes and R-loops in *CD19 DKO* and control B cells. The signal is plotted in a +/- 2 kb window from the center of the regions ordered based on decreasing intensity from top to bottom in the entire 4 kb window. R-loop signal is plotted after background subtraction of MNase-alone control. **d)** Profile histograms showing the signals for G-quadruplexes (G4) (RPM, reads per million), R-loops (RPM), WGBS (percent of 5mC+5hmC/unmodified C) and 5hmC (RPM). The 9722 regions enriched for both G-quadruplexes and R-loops are divided into two categories – 6212 regions overlapping promoters (*left panels*) and 3510 regions not at promoters (*right panels*). Dashed grey lines indicate the center of the region and the 1 kb boundaries located on either side of the center. Blue and red lines show data from *Dfl* and CD19 DKO B cells, respectively. Asterisks represent statistical significance calculated by comparing the signals between *Dfl* and *CD19 DKO* B cells, either within the G-quadruplex and R-loop forming regions, the region to +/-1kb window or +/-1kb to 2kb window for respective datasets. **e)** Profile histograms showing the 5hmC signal in *Dfl* (blue) and *CD19 DKO* (red) B cells in 23,467 regions identified as enriched for 5hmC signal. **f)** Violin plots quantifying enrichment (RPKM) of 5hmC signal in *Dfl* B cells in the +/-1 kb from G-quadruplex and R-loop forming regions at promoters, non-promoter regions and control regions randomly located in euchromatin (Hi-C A genomic compartment) from 2 biological replicates. **g)** Pie chart showing the differentially methylated regions (DMRs) in *CD19 DKO* compared to control *Dfl* B cells. Of a total of 6948 DMRs identified by WGBS, 1014 (15%) showed reduced DNA methylation (hypomethylation) and 5934 (85%) showed increased DNA methylation (hypermethylation). **h)** Box and whisker plots quantifying percent of 5mC+5hmC/unmodified C (from WGBS) at and near the G4 and R-loop forming regions overlapping promoters and regions not overlapping promoters in *Dfl* (blue) and *CD19 DKO* (red) B cells from 2 biological replicates. The signal is plotted in three windows; window 1, within the G4 and R-loop regions; window 2, from region to +/- 1kb on either side and; window 3, +/-1kb to +/-2kb on either side. **i)** Percent of 5mC+5hmC/unmodified C (from WGBS) in random genomic regions of *Dfl* (blue) and *CD19 DKO* (red) B cells. **j)** Heatmaps of enrichment (RPM) for G-quadruplexes (*left*) and R-loops (*right*) in *Dfl* and *CD19 DKO* B cells, ordered in descending order of gene expression. **k)** MA plot (*left*) showing differentially expressed genes (DEGs) in *CD19 DKO* B cells. *Red dots*, upregulated DEGs; *blue dots*, downregulated DEGs; black dots, DEGs with G-quadruplexes and R-loops at their promoters (+/-1kb of TSS); *yellow dots*, non-DEGs with G-quadruplexes and R-loops at their promoters; *grey dots*, non-DEGs without G-quadruplexes and R-loops. The pie-chart (*right*) shows the percent of DEGs with (*green*) and without (*brown*) G-quadruplexes and R-loops at their promoters. Asterisks indicate statistical significance. Statistical significance is calculated using Kruskal-Wallis test and the ad hoc Dunn's test in **d**), **f**) and **h)**, Chi-square test in **j**). Boxes in box and whisker plots represent median (center) with 25th to 75 th percentile and whiskers represent maxima/minima. **** p value ≤0.0001 and ***** p value <0.000001.

Extended data figure 7. Genome-wide analysis of TET deficient B cells. a)-c) Genome browser tracks showing the distribution of G-quadruplexes (G4), R-loops, RNA-Seq, WGBS, and 5hmC datasets for *Dfl* (blue tracks) and *CD19 DKO* (red tracks) B cells. Grey boxes indicate regions of interest. The blue arrows at the bottom show the location of the TSS and the direction of transcription. **d)** Circos plots to visually depict all translocations identified by HTGTS in *Dfl* and *CD19 DKO* replicates. Colored lines connect the Sμ bait with the translocation partner regions. Color scale represents the number of translocation partner regions identified in a 10 kb window. Translocations from two *Dfl* and *CD19 DKO* replicates are represented separately. **e)** Relative representation of different classes of motifs predicted to form G-quadruplexes (pG4) in control genomic regions selected randomly (*right bar*) and +/-300 bp from the center of translocation partner junctions identified from translocations in *Dfl* and *CD19 DKO* B cells (*right bar*). The numbers (n) of *Dfl* and *CD19* DKO hits, and control regions are included in the plots. **f)** Density of AID motifs (WRCY/RGYW) in +/-300bp from the center of translocation partner junctions

identified from translocations in *Dfl* and *CD19 DKO* B cells compared to control random regions in euchromatin (Hi-C A compartment) from 2 biological replicates. Statistical significance is calculated using the Wilcoxon signed rank test **f**), ***** p value <0.000001.

Figure 5. DNMT1 deletion delays oncogenesis in TET-deficient mice. a) Breeding strategy used to generate the triple *Dnmt1*,*Tet2*,*Tet3* deficient mice(*CD19TKO*). **b)** Kaplan-Meir curves displaying the overall survival of *Tfl* (*grey*), *CD19 Dnmt1 KO* (*purple*), *CD19 DKO* (*red*) and *CD19 TKO* (*brown*) mice. Y-axis denotes percent survival and X-axis shows time in weeks. **c)** Spleens from 10 week-old *Tfl, CD19 Dnmt1, CD19 DKO* and *CD19 TKO* mice. **d)** Representative flow cytometry data gated on Peyer's patch B cells from 10 week-old *Tfl, CD19 Dnmt1 KO, CD19 DKO and CD19 TKO* mice. Numbers represent frequency of GC B cells, identified as FAS^+ (Y-axis) and $CD38^-$ (X-axis). **e)** Quantification of GC B cell frequency in Peyer's patches of *Tfl* (YFP^-, *grey*), *CD19 Dnmt1 KO* (YFP^+, *purple*), *CD19 DKO* (YFP^+, *red*) *and CD19 TKO* (YFP^+, *brown*) mice from 3 biological replicates and 2 independent experiments. **f)** Flow cytometric detection of G-quadruplexes by staining of permeabilized cells with NMM or DMSO vehicle controls (Veh) in B cells from *Tfl* (YFP^-), *CD19 Dnmt1 KO* (YFP^+), *CD19 DKO* (YFP^+) and *CD19 TKO* (YFP^+) mice. **g)** Quantification of median fluorescence intensity (MFI) of NMM signal from *Tfl*, *CD19 Dnmt1 KO*, *CD19 DKO* and *CD19 TKO* B cells from 7 biological replicates and 5 independent experiments. **h)** Flow cytometric detection of R-loops using V5-epitope-tagged recombinant RNASE H1 (rRNASE H1) in B cells from *Tfl* (YFP^-), *CD19 Dnmt1 KO* (YFP^+), *CD19 DKO* (YFP^+) and *CD19 TKO* (YFP^+) mice. Samples stained with anti-V5 and anti-rabbit secondary antibodies were used as controls (IgG). **i)** Quantification of median fluorescence intensity (MFI) of R-loops (rRNASE H1) signal from *Tfl, CD19 Dnmt1KO, CD19 TKO* and *CD19 DKO* B cells from 3 biological replicates and 2 independent experiments. Statistical significance is calculated using one-way ANOVA. Error bars represent mean +/- standard deviation, * p value ≤0.05, ** p value ≤0.01, *** p value ≤0.0005 **** p value <0.0001.

Extended data figure 8. DNMT1 deletion delays oncogenesis in TET-deficient mice. a) Diagrammatic representation of the strategy used to confirm G-quadruplex binding. Nuclear lysates of activated B cells were incubated with biotin-conjugated single stranded G4- or non-G4-forming control oligonucleotides (Oligos) captured using streptavidin beads. **b)** Immunoblots showing flag-tagged BG4 (positive control), ATRX, BLM and DNMT1 proteins. *Left lane*, $1/10^{th}$ input from nuclear lysates ($1/10^{th}$ Input); *middle lane,* proteins pulled down with G4 forming oligonucleotides; and *right lane*, proteins pulled down with non-G4 control oligonucleotides. The data is representative of at least 2 independent experiments. **c)** - **d)** Quantification of **c)** cell numbers, **d)** spleen weights, of 10 week-old *Tfl* (grey), *CD19 Dnmt1 KO* (purple), *CD19 DKO* (red) and *CD19 TKO* (brown) mice from 5 independent experiments. **e)** Enlarged spleen of 75-week-old *CD19 TKO* mice compared with *Tfl* control mice. **f)** Flow cytometric detection of G-quadruplexes with BG4-Ig antibody or isotype IgG controls in B cells from *Tfl* (YFP^-, *grey*), *CD19 Dnmt1 KO* (YFP^+, *purple*), *CD19 DKO* (YFP^+, *red*) and *CD19 TKO* (YFP^+, *brown*) mice. **g) – h)** Quantification of median fluorescence intensity (MFI) of **g)** BG4-Ig signal, **h)** γH2AX signal from *Tfl* (YFP^-), *CD19 Dnmt1 KO* (YFP^+), *CD19 DKO* (YFP^+) and *CD19 TKO* (YFP^+) B cells from 4 independent experiments. **i), k)** Flow cytometric detection of **i**) G-quadruplexes with BG4-Ig antibody or isotype IgG controls and **k**) R-loops using V5-epitope-tagged recombinant RNASE H1 (rRNASE H1) or IgG controls in GC B cells (Fas^+) from *Tfl* (YFP^-, *grey*), *CD19 DKO* (YFP^+, *red*) and *CD19 TKO* (YFP^+, *brown*) mice. **j), l)** Quantification of median fluorescence intensity (MFI) of **j)** BG4-Ig signal and **l)** R-loops (rRNASE H1) signal in GC B cells (Fas^+) from *Tfl* (YFP^-, *grey*), *CD19 DKO* (YFP^+, *red*) and *CD19 TKO* (YFP^+, *brown*) mice from 3 independent experiments. **m)** Model proposing functional interplay between TET and DNMT activities to limit GC B cell expansion. TET deficiency in B cells leads to increased G4 and R-loop structures and is associated with altered gene expression, DNA damage and development of B cell lymphoma. Statistical significance is calculated using one-way ANOVA in **c), d), g), h), j)** and **l)**. Error

bars represent mean +/- standard deviation, * p value ≤0.05, ** p value ≤0.01, *** p value ≤0.0005 **** p value <0.0001.

Extended data figure 9. FACS gating strategy and original blots. a) Sequential gating strategy used for the flow cytometry analysis. The respective gate names are mentioned in the corresponding figures. **b**)-**e**) scanned immunoblots for extended data figures 2 (**b**), 5 (**c**) and 8 (**d**).

References

1 Tahiliani, M. *et al.* Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science* **324**, 930-935, doi:10.1126/science.1170116 (2009).

2 Ko, M. *et al.* Impaired hydroxylation of 5-methylcytosine in myeloid cancers with mutant TET2. *Nature* **468**, 839-843, doi:10.1038/nature09586 (2010).

3 Ito, S. *et al.* Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. *Science* **333**, 1300-1303, doi:10.1126/science.1210597 (2011).

4 Lio, C. W. *et al.* Tet2 and Tet3 cooperate with B-lineage transcription factors to regulate DNA modification and chromatin accessibility. *Elife* **5**, doi:10.7554/eLife.18290 (2016).

5 Orlanski, S. *et al.* Tissue-specific DNA demethylation is required for proper B-cell differentiation and function. *Proc Natl Acad Sci U S A* **113**, 5018-5023, doi:10.1073/pnas.1604365113 (2016).

6 Dominguez, P. M. *et al.* TET2 Deficiency Causes Germinal Center Hyperplasia, Impairs Plasma Cell Differentiation, and Promotes B-cell Lymphomagenesis. *Cancer Discov* **8**, 1632-1653, doi:10.1158/2159-8290.CD-18-0657 (2018).

7 Lio, C. J. *et al.* TET enzymes augment activation-induced deaminase (AID) expression via 5-hydroxymethylcytosine modifications at the Aicda superenhancer. *Sci Immunol* **4**, doi:10.1126/sciimmunol.aau7523 (2019).

8 Rosikiewicz, W. *et al.* TET2 deficiency reprograms the germinal center B cell epigenome and silences genes linked to lymphomagenesis. *Sci Adv* **6**, eaay5872, doi:10.1126/sciadv.aay5872 (2020).

9 Lio, C. J. *et al.* TET methylcytosine oxidases: new insights from a decade of research. *J Biosci* **45** (2020).

10 Wu, H. & Zhang, Y. Reversing DNA methylation: mechanisms, genomics, and biological functions. *Cell* **156**, 45-68, doi:10.1016/j.cell.2013.12.019 (2014).

11 Pastor, W. A., Aravind, L. & Rao, A. TETonic shift: biological roles of TET proteins in DNA demethylation and transcription. *Nat Rev Mol Cell Biol* **14**, 341-356, doi:10.1038/nrm3589 (2013).

12 Rasmussen, K. D. & Helin, K. Role of TET enzymes in DNA methylation, development, and cancer. *Genes Dev* **30**, 733-750, doi:10.1101/gad.276568.115 (2016).

13 Cimmino, L., Abdel-Wahab, O., Levine, R. L. & Aifantis, I. TET family proteins and their role in stem cell differentiation and transformation. *Cell Stem Cell* **9**, 193-204, doi:10.1016/j.stem.2011.08.007 (2011).

14 Huang, Y. & Rao, A. Connections between TET proteins and aberrant DNA modification in cancer. *Trends Genet* **30**, 464-474, doi:10.1016/j.tig.2014.07.005 (2014).

15 Ko, M., An, J. & Rao, A. DNA methylation and hydroxymethylation in hematologic differentiation and transformation. *Curr Opin Cell Biol* **37**, 91-101, doi:10.1016/j.ceb.2015.10.009 (2015).

16 Lio, C. J., Yuita, H. & Rao, A. Dysregulation of the TET family of epigenetic regulators in lymphoid and myeloid malignancies. *Blood* **134**, 1487-1497, doi:10.1182/blood.2019791475 (2019).

17 Reddy, A. *et al.* Genetic and Functional Drivers of Diffuse Large B Cell Lymphoma. *Cell* **171**, 481-494 e415, doi:10.1016/j.cell.2017.09.027 (2017).

18 Schmitz, R. *et al.* Genetics and Pathogenesis of Diffuse Large B-Cell Lymphoma. *N Engl J Med* **378**, 1396-1407, doi:10.1056/NEJMoa1801445 (2018).

19 Chapuy, B. *et al.* Molecular subtypes of diffuse large B cell lymphoma are associated with distinct pathogenic mechanisms and outcomes. *Nat Med* **24**, 679-690, doi:10.1038/s41591-018-0016-8 (2018).

20 Kafer, G. R. *et al.* 5-Hydroxymethylcytosine Marks Sites of DNA Damage and Promotes Genome Stability. *Cell Rep* **14**, 1283-1292, doi:10.1016/j.celrep.2016.01.035 (2016).
21 Kharat, S. S. *et al.* Degradation of 5hmC-marked stalled replication forks by APE1 causes genomic instability. *Sci Signal* **13**, doi:10.1126/scisignal.aba8091 (2020).
22 Crossley, M. P., Bocek, M. & Cimprich, K. A. R-Loops as Cellular Regulators and Genomic Threats. *Mol Cell* **73**, 398-411, doi:10.1016/j.molcel.2019.01.024 (2019).
23 Skourti-Stathaki, K. & Proudfoot, N. J. A double-edged sword: R loops as threats to genome integrity and powerful regulators of gene expression. *Genes Dev* **28**, 1384-1396, doi:10.1101/gad.242990.114 (2014).
24 Hansel-Hertsch, R., Di Antonio, M. & Balasubramanian, S. DNA G-quadruplexes in the human genome: detection, functions and therapeutic potential. *Nat Rev Mol Cell Biol* **18**, 279-284, doi:10.1038/nrm.2017.3 (2017).
25 Rhodes, D. & Lipps, H. J. G-quadruplexes and their regulatory roles in biology. *Nucleic Acids Res* **43**, 8627-8637, doi:10.1093/nar/gkv862 (2015).
26 Sauer, M. & Paeschke, K. G-quadruplex unwinding helicases and their function in vivo. *Biochem Soc Trans* **45**, 1173-1182, doi:10.1042/BST20170097 (2017).
27 De Magis, A. *et al.* DNA damage and genome instability by G-quadruplex ligands are mediated by R loops in human cancer cells. *Proc Natl Acad Sci U S A* **116**, 816-825, doi:10.1073/pnas.1810409116 (2019).
28 Gray, L. T., Vallur, A. C., Eddy, J. & Maizels, N. G quadruplexes are genomewide targets of transcriptional helicases XPB and XPD. *Nat Chem Biol* **10**, 313-318, doi:10.1038/nchembio.1475 (2014).
29 Miglietta, G., Russo, M. & Capranico, G. G-quadruplex-R-loop interactions and the mechanism of anticancer G-quadruplex binders. *Nucleic Acids Res*, doi:10.1093/nar/gkaa944 (2020).
30 Chedin, F. Nascent Connections: R-Loops and Chromatin Patterning. *Trends Genet* **32**, 828-838, doi:10.1016/j.tig.2016.10.002 (2016).
31 Niehrs, C. & Luke, B. Regulatory R-loops as facilitators of gene expression and genome stability. *Nat Rev Mol Cell Biol* **21**, 167-178, doi:10.1038/s41580-019-0206-3 (2020).
32 Sanz, L. A. *et al.* Prevalent, Dynamic, and Conserved R-Loop Structures Associate with Specific Epigenomic Signatures in Mammals. *Mol Cell* **63**, 167-178, doi:10.1016/j.molcel.2016.05.032 (2016).
33 An, J. *et al.* Acute loss of TET function results in aggressive myeloid cancer in mice. *Nat Commun* **6**, 10071, doi:10.1038/ncomms10071 (2015).
34 Cimmino, L. *et al.* TET1 is a tumor suppressor of hematopoietic malignancy. *Nat Immunol* **16**, 653-662, doi:10.1038/ni.3148 (2015).
35 Tsagaratou, A. *et al.* TET proteins regulate the lineage specification and TCR-mediated expansion of iNKT cells. *Nat Immunol* **18**, 45-53, doi:10.1038/ni.3630 (2017).
36 Mao, S. Q. *et al.* DNA G-quadruplex structures mold the DNA methylome. *Nat Struct Mol Biol* **25**, 951-957, doi:10.1038/s41594-018-0131-8 (2018).
37 Biffi, G., Tannahill, D., McCafferty, J. & Balasubramanian, S. Quantitative visualization of DNA G-quadruplex structures in human cells. *Nat Chem* **5**, 182-186, doi:10.1038/nchem.1548 (2013).
38 Muller, S. *et al.* Pyridostatin analogues promote telomere dysfunction and long-term growth inhibition in human cancer cells. *Org Biomol Chem* **10**, 6537-6546, doi:10.1039/c2ob25830g (2012).
39 Sabharwal, N. C. *et al.* N-methylmesoporphyrin IX fluorescence as a reporter of strand orientation in guanine quadruplexes. *FEBS J* **281**, 1726-1737, doi:10.1111/febs.12734 (2014).
40 Nicoludis, J. M. *et al.* Optimized end-stacking provides specificity of N-methyl mesoporphyrin IX for human telomeric G-quadruplex DNA. *J Am Chem Soc* **134**, 20446-20456, doi:10.1021/ja3088746 (2012).
41 Chen, L. *et al.* R-ChIP Using Inactive RNase H Reveals Dynamic Coupling of R-loops with Transcriptional Pausing at Gene Promoters. *Mol Cell* **68**, 745-757 e745, doi:10.1016/j.molcel.2017.10.008 (2017).
42 Yan, Q., Shields, E. J., Bonasio, R. & Sarma, K. Mapping Native R-Loops Genome-wide Using a Targeted Nuclease Approach. *Cell Rep* **29**, 1369-1380 e1365, doi:10.1016/j.celrep.2019.09.052 (2019).

43 Crossley, M. P. *et al.* Catalytically inactive, purified RNase H1: A specific and sensitive probe for RNA-DNA hybrid imaging. *J Cell Biol* **220**, doi:10.1083/jcb.202101092 (2021).

44 Boguslawski, S. J. *et al.* Characterization of monoclonal antibody to DNA.RNA and its application to immunodetection of hybrids. *J Immunol Methods* **89**, 123-130, doi:10.1016/0022-1759(86)90040-2 (1986).

45 Nojima, T. *et al.* In-vitro derived germinal centre B cells differentially generate memory B or plasma cells in vivo. *Nat Commun* **2**, 465, doi:10.1038/ncomms1475 (2011).

46 Lopez-Moyado, I. F. *et al.* Paradoxical association of TET loss of function with genome-wide DNA hypomethylation. *Proc Natl Acad Sci U S A* **116**, 16933-16942, doi:10.1073/pnas.1903059116 (2019).

47 Hansel-Hertsch, R. *et al.* G-quadruplex structures mark human regulatory chromatin. *Nat Genet* **48**, 1267-1272, doi:10.1038/ng.3662 (2016).

48 Song, C. X. *et al.* Selective chemical labeling reveals the genome-wide distribution of 5-hydroxymethylcytosine. *Nat Biotechnol* **29**, 68-72, doi:10.1038/nbt.1732 (2011).

49 Hu, J. *et al.* Detecting DNA double-stranded breaks in mammalian genomes by linear amplification-mediated high-throughput genome-wide translocation sequencing. *Nat Protoc* **11**, 853-871, doi:10.1038/nprot.2016.043 (2016).

50 Shaknovich, R. *et al.* DNA methyltransferase 1 and DNA methylation patterning contribute to germinal center B-cell differentiation. *Blood* **118**, 3559-3569, doi:10.1182/blood-2011-06-357996 (2011).

51 Tanaka, S. *et al.* Tet2 and Tet3 in B cells are required to repress CD86 and prevent autoimmunity. *Nat Immunol* **21**, 950-961, doi:10.1038/s41590-020-0700-y (2020).

52 Young, R. M. *et al.* Survival of human lymphoma cells requires B-cell receptor engagement by self-antigens. *Proc Natl Acad Sci U S A* **112**, 13447-13454, doi:10.1073/pnas.1514944112 (2015).

53 Nambiar, M. *et al.* Formation of a G-quadruplex at the BCL2 major breakpoint region of the t(14;18) translocation in follicular lymphoma. *Nucleic Acids Res* **39**, 936-948, doi:10.1093/nar/gkq824 (2011).

54 Rabkin, C. S., Hirt, C., Janz, S. & Dolken, G. t(14;18) Translocations and risk of follicular lymphoma. *J Natl Cancer Inst Monogr*, 48-51, doi:10.1093/jncimonographs/lgn002 (2008).

55 Qiao, Q. *et al.* AID Recognizes Structured DNA for Class Switch Recombination. *Mol Cell* **67**, 361-373 e364, doi:10.1016/j.molcel.2017.06.034 (2017).

56 Xu, Y. Z. *et al.* Activation-induced cytidine deaminase localizes to G-quadruplex motifs at mutation hotspots in lymphoma. *NAR Cancer* **2**, zcaa029, doi:10.1093/narcan/zcaa029 (2020).

57 Yewdell, W. T. *et al.* A Hyper-IgM Syndrome Mutation in Activation-Induced Cytidine Deaminase Disrupts G-Quadruplex Binding and Genome-wide Chromatin Localization. *Immunity*, doi:10.1016/j.immuni.2020.10.003 (2020).

58 Zhang, W. & Xu, J. DNA methyltransferases and their roles in tumorigenesis. *Biomark Res* **5**, 1, doi:10.1186/s40364-017-0081-z (2017).

59 Wang, Y. *et al.* G-quadruplex DNA drives genomic instability and represents a targetable molecular abnormality in ATRX-deficient malignant glioma. *Nat Commun* **10**, 943, doi:10.1038/s41467-019-08905-8 (2019).

60 Xu, H. *et al.* CX-5461 is a DNA G-quadruplex stabilizer with selective lethality in BRCA1/2 deficient tumours. *Nat Commun* **8**, 14432, doi:10.1038/ncomms14432 (2017).

61 Kang, J. *et al.* Simultaneous deletion of the methylcytosine oxidases Tet1 and Tet3 increases transcriptome variability in early embryogenesis. *Proc Natl Acad Sci U S A* **112**, E4236-4245, doi:10.1073/pnas.1510510112 (2015).

62 Ko, M. *et al.* Ten-Eleven-Translocation 2 (TET2) negatively regulates homeostasis and differentiation of hematopoietic stem cells in mice. *Proc Natl Acad Sci U S A* **108**, 14566-14571, doi:10.1073/pnas.1112317108 (2011).

63 Chedin, F., Hartono, S. R., Sanz, L. A. & Vanoosthuyse, V. Best practices for the visualization, mapping, and manipulation of R-loops. *EMBO J* **40**, e106394, doi:10.15252/embj.2020106394 (2021).

64 Konig, F., Schubert, T. & Langst, G. The monoclonal S9.6 antibody exhibits highly variable binding affinities towards different R-loop sequences. *PLoS One* **12**, e0178875, doi:10.1371/journal.pone.0178875 (2017).
65 Ramirez, F. *et al.* deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res* **44**, W160-165, doi:10.1093/nar/gkw257 (2016).
66 Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* **10**, R25, doi:10.1186/gb-2009-10-3-r25 (2009).
67 Zhang, Y. *et al.* Model-based analysis of ChIP-Seq (MACS). *Genome Biol* **9**, R137, doi:10.1186/gb-2008-9-9-r137 (2008).
68 Heinz, S. *et al.* Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol Cell* **38**, 576-589, doi:10.1016/j.molcel.2010.05.004 (2010).
69 Xi, Y. & Li, W. BSMAP: whole genome bisulfite sequence MAPping program. *BMC Bioinformatics* **10**, 232, doi:10.1186/1471-2105-10-232 (2009).
70 Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15-21, doi:10.1093/bioinformatics/bts635 (2013).
71 Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**, 550, doi:10.1186/s13059-014-0550-8 (2014).
72 Bolotin, D. A. *et al.* Antigen receptor repertoire profiling from RNA-seq data. *Nat Biotechnol* **35**, 908-911, doi:10.1038/nbt.3979 (2017).
73 Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841-842, doi:10.1093/bioinformatics/btq033 (2010).
74 Krzywinski, M. *et al.* Circos: an information aesthetic for comparative genomics. *Genome Res* **19**, 1639-1645, doi:10.1101/gr.092759.109 (2009).
75 Kieffer-Kwon, K. R. *et al.* Myc Regulates Chromatin Decompaction and Nuclear Architecture during B Cell Activation. *Mol Cell* **67**, 566-578 e510, doi:10.1016/j.molcel.2017.07.013 (2017).

Methods

Data Availability. All genome-wide sequencing datasets have been deposited to Gene Expression Omnibus (GEO) repository, accession number GSE161463. Any data and reagents will also be made available upon request.

Code Availability. The code used to process the NGS datasets has been deposited in GitHub repository at https://github.com/dsamanie7/Tet2-Tet3_DKO_CD19_cre

Mice. *Tet2*$^{fl/fl}$ and *Tet3*$^{fl/fl}$ mice were generated as previously described[61,62]. C57BL/6J (000664), *CD19 cre* (006785), *Rosa26-LSL-EYFP* (006148), *Cγ1Cre* (010611), *Ubc-Cre*^{ERT2} (008085; described as *ERT2cre*) and *CD45.1* mice (002014, *ptprc*^{a}) were obtained from Jackson Laboratory. To induce *ERT2cre*-mediated deletion, Cre-expressing and control mice were intra-peritoneally injected with 2 mg tamoxifen (Sigma) dissolved in 100 µL corn oil (Sigma) daily for 5 days. For transplantation studies, *CD45.1* mice were sub-lethally irradiated with 600 rads of X-rays 24 h prior to transfer of 2 million *CD19 DKO* or *Dfl* B cells through the retro-orbital sinus. Age and sex matched mice from both sexes were used in the experiments. All mice used were 8 to 16 weeks of age (unless otherwise indicated) and were on a C57BL/6 genetic background, housed in specific-pathogen free animal facility at ambient temperature and humidity with 12h light/12h dark cycle at La Jolla Institute for Immunology. Age and sex matched mice from both sexes were used in the experiments. All studies were performed according to protocols approved by the Institutional Animal Care and Use Committee.

B cell isolation and cell cultures. Primary B cells were isolated using flow sorting of YFP+ *CD19 DKO* B cells or YFP- control *Dfl* B cells for RNA-Seq, WGBS and 5hmC mapping analyses, and using EasySep Mouse Pan B cell isolation kit (#19844 Stem Cell Technology, Canada) for *in vitro* culture of B cells and CD19 positive selection (for G4 mapping, R-loop mapping and HTGTS studies) from splenocytes. Primary B cells and CH12F3 (CH12) cells were cultured at 37°C, 5% CO_2 in RPMI 1640 media supplemented with 10% FBS, 1x MEM non-essential amino acids, 10mM HEPES (pH 7.4), 2mM Glutamax, 1mM sodium pyruvate, 55µM 2-mercaptoethanol (all from Life technologies). B cells ($5x10^5$-$1x10^6$ cells/mL) were activated with 10 µg/mL LPS (for Cas9 RNP targeting studies) from *E. coli* O55:B5 (Sigma, St. Louis, MO), 25 µg/mL LPS and 10 ng/mL rmIL-4 for stimulation of *Dfl* and *ERT2cre* B cells in the presence of 1 µM 4-hydroxytamoxifen (Tocris). For pyridostatin (PDS) (#SML2690, Sigma-Aldrich) treatment, B cells were activated with 10 µg/mL LPS for 48 h in presence of 10 µM PDS before analysis. All cytokines used above were from Peprotech (Rocky Hill, NJ). CH12F3 cell line was developed in Dr. T. Honjo's laboratory at Kyoto University and was independently validated using the cell stimulation and Ig class-switching experiments.

For B cell cultures on 40LB feeder cells, 40LB cells were irradiated with 3000 rads of X-rays and plated at a density of $20X10^4$ cells per well on a 12 cm plate and cultured overnight at 37°C, 5% CO_2 in DMEM media supplemented with 10% FBS, 10mM HEPES (pH 7.4), 2mM Glutamax, 1mM sodium pyruvate, 55µM 2-mercaptoethanol (all from Life technologies). 40LB cells were obtained from Dr. D. Kitamura's laboratory (Tokyo University of Sciences) was independently validated by B cell stimulation experiments and periodically tested for mycoplasma contamination. B cells were purified from *Cγ1cre and Cγ1 DKO* mice using the Mouse B cell isolation kit (#19854 Stem Cell Technology, Canada) and isolated cells were seeded at a density of $2X10^4$ cells per well on 40LB containing 12 cm plate in RPMI 1640 media prepared as above and supplemented with 1ng/ml of rmIL-4. The expanded B cells collected from suspension were analyzed by flow cytometry.

Immunization. *Cγ1Cre, Dfl* and *Cγ1DKO* mice were immunized with sheep red blood cells-SRBCs (#31102, Colorado Serum company, CO, USA) washed two times with PBS and injected in 2 doses, first primed with $200X10^6$ SRBCs followed by a boost at day 5 with 10^9 SRBCs before analysis of splenocytes at day 12 post first immunization.

Flow cytometry. Primary cells and *in vitro* cultured cells were stained in FACS buffer (0.5% bovine serum albumin, 1mM EDTA, and 0.05% sodium azide in PBS) with indicated antibodies for 30 mins on ice. Cells were washed and then fixed with 1% paraformaldehyde (diluted from 4% with PBS; Affymetrix) for 10 min at 25°C before FACS analysis using FACS Celesta and FACS LSR II (BD Biosciences). Antibodies and dyes were from BioLegend, eBioscience, and BD Pharmingen. Data were analyzed with FlowJo (FlowJo LLC, Ashland, OR). The gating strategy for the flow cytometric analysis is displayed in extended data figure 9 and the appropriate gates used in each experiment is described in the corresponding figure legends.

Immunoblotting. Proteins isolated from cells with NP-40 lysis buffer were resolved using NuPAGE 4-12% Bis-Tris gel (ThermoFisher) and transferred from gel to PVDF membrane using Wet/Tank Blotting Systems (Bio-Rad). Membrane was blocked with 5% non-fat milk in TBSTE buffer (50mM Tris-HCl pH 7.4, 150mM NaCl, 0.05% Tween-20, 1mM EDTA), incubated with indicated primary antibodies, followed by secondary antibodies conjugated with horse-radish peroxidase (HRP) and signal was detected with enhanced chemiluminescence reagents (Invitrogen) and X-ray film. Antibodies against ATRX (1:1000, clone D5), FANCD2 (1:1000, clone FI17) and BLM (1:1000, clone B4) were purchased from Santa Cruz Biotechnology. Antibodies against RNASE H1 (1:1000, NBP2-20171) from Novus biologicals, and WRN (1:1000, clone 8H3) and β-ACTIN HRP (1:5000, clone 13E5) from Cell signaling and DNMT1 (1:1000, ab19905) from Abcam.

Cas9 RNP targeting. Alt-R crRNA and Alt-R tracrRNA (from IDT) were reconstituted at concentration of 100 µM in Nuclease-Free Duplex buffer (IDT). RNA duplexes were prepared by mixing oligonucleotides

(Alt-R crRNA and Alt-R tracrRNA) at equimolar concentrations in a sterile PCR tube (e.g. 4 µl Alt-R crRNA and 4 µl Alt-R-tracrRNA). Mixed oligonucleotides were annealed by heating at 95C for 5 min in a PCR machine, followed by incubation at 25°C for at least 1 hour. 5.9 µl of crRNA-tracrRNA duplexes (180 pmol) were then mixed with 2.1 µL IDT Cas9 v3 (80 pmol) (together Cas9 RNP) by gentle pipetting and incubated at 25°C for at least 10 min. 35 µl of cell media was aliquoted in a 24 well plate. 2 million B cells were then washed with PBS and resuspended in 15 µl resuspension buffer (4D-Nuclcofector X Kit S, #V4XP-4032; Lonza). Cells in resuspension buffer (12.5µl) were then mixed with Cas9-RNPs (7.5µl) and transferred to the nucleofection cuvette strips and electroporated using the CM-137 program with 4D nucleofector. After electroporation, cells were transferred into 35 µl media in 24 well plate, and incubated for 20 min at 37°C before transferring cells to activation media with 10 µg/ml LPS. The crRNA sequences used are listed in extended data table 1.

RNA extraction, cDNA synthesis, and quantitative RT-PCR. Total RNA was isolated with RNeasy plus kit (Qiagen, Germany) or with Trizol (ThermoFisher, Waltham, MA) following manufacturer's instructions. cDNA was synthesized using SuperScript III reverse transcriptase (ThermoFisher) and quantitative RT-PCR was performed using FastStart Universal SYBR Green Master mix (Roche, Germany) on a StepOnePlus real-time PCR system (Applied Biosytems). Gene expression was normalized to *Gapdh*. Primers are listed in extended data table 1.

G4 detection and mapping. For flow cytometry-based detection of G4s, cells were stained for cell surface markers according to the flow cytometry staining protocol described above. Following cell surface staining, cells were fixed with 4% paraformaldehyde in PBS for 12 min at 25°C. Using the intracellular transcription factor staining kit (Invitrogen), cells were then permeabilized at 4°C overnight according to manufacturer's instructions. Cells were then washed with wash buffer and treated with RNASE A (1:50, AM2269; Ambion) for 30 min at 25°C followed by washing and incubation with 1:100 dilutions of BG4-Ig antibody (Ab00174-1.1, 1mg/ml; Absolute Antibody) for 30 min at 25°C or 25°C. This was followed by washing and subsequent staining with anti-mouse IgG1 fluorophore conjugated antibody (BioLegend) to reveal the signal.

The excitation and emission spectra of NMM in presence of G4 DNA, but not in the presence of non-G4 control DNA, closely mimicked that of a fluorophore, Brilliant Violet 605 (BV605). Therefore, we also used NMM to stain for G4. Following the RNASE A treatment cells were incubated with 10 µM NMM (#NMM580; Frontier Scientific) for 30 min at 25°C or 4°C before reading the fluorescence signal in the Brilliant Violet 605 channel. Mouse CH12F3 (CH12) B cells treated with PDS, which binds G-quadruplex structures in a manner distinct from NMM[38,40], showed increased NMM fluorescence in flow cytometry (Extended data **Fig. 3e**), validating NMM as a flow cytometric probe for G-quadruplex structures in cells.

G4 mapping was performed as previously described with minor modifications[47]. Briefly, ~10 million cells were fixed with 1% formaldehyde at 25°C with nutation for 10 min mins at $1x10^6$ cell/mL in media, quenched with 125 mM glycine, washed twice with ice cold PBS. Cells were pelleted, snap-frozen in liquid nitrogen, and stored at -80°C until use. Cell pellets were then lyzed using 800 µl of ChIP hypotonic solution (Chromatrap) and 150 µl lysis buffer (Chromatrap) supplemented with protease inhibitor cocktail (invitrogen) on ice according to manufacturer's instructions. The lyzed nuclei suspension was then sonicated using the bioruptor pico sonicator (Diagenode), 16 cycles 30 seconds on / 30 seconds off and the lysates were cleared by centrifugation at >12000g for 10 min. The lysates were diluted 1:10 times in intracellular buffer (25 mM HEPES pH 7.5, 10.5 mM NaCl, 110 mM KCl, 130 nM $CaCl_2$, 1 mM $MgCl_2$) (total 200 µl) then treated with 1mg/ml RNASE A (AM2269; Ambion, 1:100 dilution) for 20 min at 37°C. Chromatin lysates were then incubated with a control IgG or BG4-Ig antibody (Ab00174-1.1, 1mg/ml; Absolute Antibody) (1:100) for 1.5 h at 16 C followed by incubation with protein G magnetic dynabeads (Invitrogen). The beads were then washed 5 times with 1 ml wash buffer (intracellular buffer with 0.05% tween 20). The beads were re-suspended in elution buffer (100 mM $NaHCO_3$, 1% SDS, 1 mg/mL RNaseA; Qiagen) treated with RNASE A (1:100, 100mg/ml) at 37 C for 30 min and proteinase K (1:40,

20mg/ml; Ambion) at 65 C for 2 h to overnight. DNA was purified with Zymo ChIP DNA Clean & Concentrator-Capped Column (Zymo Research, Irvine, CA). The Library was prepared with NEB Ultra II library prep kit (NEB) following manufacturer's instruction and was sequenced on an Illumina NovaSeq 6000 (50 bp paired-end reads). The control IgG did not pull-down sufficient amount of DNA to prepare libraries, therefore, we prepared libraries from the input DNA and used those as controls. 2 independent experiments were combined for analysis.

NMM oligonucleotide fluorescence enhancement assay. G4 forming oligonucleotides and control oligonucleotides with mutated G-stretches were ordered from Integrated DNA technologies using standard desalting purification (Extended data table 1). The oligonucleotides were reconstituted at 100µM concentration in 10mM Tris (pH 8.0) and 1mM EDTA (TE buffer) and then diluted to a final concentration of 10µM in buffer containing 20mM HEPES (pH 7.5), 250mM KCl and 1mM DTT. The diluted oligonucleotides were then heated at 95C for 10 min followed by gradual cooling to 25°C (cooled down at 0.1℃ per second ramp rate) in a PCR machine. 50 µl of 10µM was then mixed with 50 µl of 20µM NMM incubated at 25°C for 1 hour and the fluorescence was measured in the Spectramax M2 (Spectral labs) plate reader with excitation at 400nm and emission at 610nm.

Recombinant RNASE H1 purification. Coding sequence of N-terminal V5-tagged mutant human RNASE H1 (D210N)[41] was sub-cloned in frame with Glutathione-S-Transferase (GST) protein separated by a Precission protease cleavage site in the pGEX6p1 vector. BL21 (DE3) cells were transformed with the pGEX6p1 vector with GST-RNASE H1 fusion. Single colonies were picked and grown overnight in 5 ml LB at 37°C before further expanding in 500ml LB (1:100) with shaking at 37°C for 2-3 h until the OD at 600 nm reached 0.5-0.6. 1mM IPTG was then added to induce the protein expression for 3 h at 37°C with shaking. Expression of protein was confirmed by coomassie staining on SDS-PAGE gel by comparing with an uninduced control culture. Cultures were then pelleted at 4500g for 15 min and cell pellets were re-suspended in lysis buffer (50 mM Tris pH 7.5, 150 mM NaCl), 15ml for 200ml culture. The cell suspension was then lysed by sonication for a total of 4 min (5 seconds on, 30 seconds off cycle) with a probe sonicator (Missonix) followed by centrifuge at 10,000g for 20min, at 4°C. The supernatant was then mixed with 150µl washed and equilibrated glutathione agarose beads (#16100; Pierce) and rotated for 4 h to overnight. The lysate beads mixture was then washed 3 times with lysis buffer and then re-suspended in 1ml elution buffer (25mM Tris-HCl (pH 7.6), 10% Glycerol, 0.2mM EDTA, 100mM KCl, 1mM DTT) and treated overnight with precission protease (#Z02799; Genescript). The supernatant was then collected and the protein was quantified by absorbance at 280 nm using Nanodrop. The purity of protein was additionally confirmed by coomassie staining of samples on an SDS-PAGE gel.

R-loop detection and mapping (MapR). For flow cytometry-based detection of R-loops, cells were stained for cell surface marker according to the flow cytometry staining protocol described above. Following cell surface staining, cells were fixed with 4% paraformaldehyde in PBS for 12 min at 25°C. Using the intracellular transcription factor staining kit (Invitrogen), cells were then permeabilized at 4°C overnight according to manufacturer's instructions. Cells were then washed and incubated with 1:50 dilutions of V5-tagged recombinant mutant RNASE H1 for 30 min at 4°C followed by washing, subsequent staining with anti-V5 tag antibody (1:200, clone D3H8Q; Cell Signaling technologies) and with anti-rabbit secondary antibody (1:1000, #A-21245; Invitrogen) to reveal the signal. For RNASE H digestion and R-loop quantification, CH12 cells were fixed, permeabilized and treated with 20 units of RNASE H (NEB#M0297S) in 100µl of digestion buffer diluted in water for 2 h at 37°C before proceeding to R-loops staining. MapR and immunoprecipitation with S9.6 antibody identify a set of R-loops in common, as well as sets of unique R-loops that depend on differential sequence specifities and/or preferences of S9.6 and RNASE H, thus R-loops identified by MapR may represent a slightly different subset than those identified by other methods[63,64].

R-loop mapping was performed by the MapR method, as previously described[42]. The libraries were sequenced on an Illumina NovaSeq 6000 (50 bp paired-end reads). Three independent experiments were combined for analysis.

Detection of G-quadruplex binding proteins. Purified naïve B cells were activated in presence of 25 µg/ml LPS and 10ng/ml IL-4 and nuclear extracts from 20 million cells were prepared using Dignam extraction. Cells were resuspended in 10 ml dignam buffer A (10 mM KCl, 1.5 mM $MgCl_2$) supplemented with protease inhibitor cocktail (Thermofisher) and incubated on ice for at least 10 min. The cells were then homogenized and lyzed with a dounce homogenizer (10 strokes) and nuclei were collected by centrifugation at 600 g for 5 min and washed again with buffer A. Cells were then resuspended in 1ml of dignam buffer C (0.2 mM EDTA, 25% glycerol (v/v), 20 mM HEPES-KOH (pH 7.9), 0.42 M NaCl, 1.5 mM $MgCl_2$) supplemented with protease inhibitor cocktail and rotated at 4°C for 1 hour. The lysates were collected and passed through a PD-10 buffer exchange column (GE healthcare) equilibrated with intracellular buffer (25 mM HEPES pH 7.5, 10.5 mM NaCl, 110 mM KCl, 130 nM CaCl2, 1 mM MgCl2). The G4 and non-G4 oligonucleotides were folded in the presence of K^+ ions, annealed with a complementary biotinylated oligonucleotide, captured on streptavidin conjugated magnetic beads and incubated with nuclear lysates from B cells activated with LPS and IL-4 to identify G4-binding proteins (Extended data **Fig. 8a**). We observed preferential binding of the known G4-binders ATRX and BLM as well as the flag-tagged BG4 antibody to the G4 oligonucleotides (Extended data **Fig. 8b**). The DNMT1 protein also showed preferential binding to G4 forming oligonucleotides compared with non-G4 control oligonucleotides, thus independently validating previous studies (Extended data **Fig. 8b).** The nuclear lysates collected in intracellular buffer were spike-in with 1 µg flag-tagged BG4 scFv antibody (Millipore), pre-cleared with myone T1 streptavidin-conjugated dynabeads (Life technologies) for one hour and incubated with streptavidin beads conjugated with a mixture of G4 or non-G4 forming oligonucleotides for 4 h (Extended data table 1). Samples were then washed with intracellular buffer with 0.05% tween 20 for a total of 5 washes and denatured using the SDS-PAGE sample buffer for immunoblotting.

RNA-Seq. RNA from cells was isolated with Trizol (ThermoFisher, Waltham, MA) following manufacturer's instructions. Isolated RNA was further purified using RNA clean up and concentrator kit (R1013, Zymo research) according to manufacturer's instructions. 40 ng of purified RNA was then used for preparation of RNA-sequencing libraries with the Nugen ovation RNA-Seq V2 system (now Tecan Genomics) according to the manufacturer recommended protocol. Samples were sequenced on Illumina Hiseq 2500 (single-end 50 bp reads). 2 independent experiments were combined for analysis.

Whole genome bisulfite sequencing (WGBS). 1 µg of genomic DNA was spiked-in with unmethylated lambda-phage DNA (1:200) and sonicated using Bioruptor (9 cycles 30 seconds on, 30 seconds off). The sonicated DNA fragments were end repaired, A-tailed and ligated with methylated adaptors (all using NEB kits). The adaptor ligated fragments were then treated with sodium bisulfite using the Methylcode kit (Invitrogen) for a total of 4 h according to manufacturer's protocol. The bisulfite converted, adaptor ligated fragments were then amplified with Kapa Hifi Hotstart Uracil+ PCR mix with NEB universal dual indexing primers. The libraries were then sequenced on an Illumina NovaSeq 6000 (150 bp paired-end reads). 2 independent experiments were combined for analysis.

5hmC pull down (HMCP). 5hmC mapping was performed on purified genomic DNA using the HMCP method in collaboration with Cambridge Epigenetix according to the manufacturer's recommended protocol. The libraries were then sequenced on an Illumina Hiseq 2500 (single-end 50 bp reads). 2 independent experiments were combined for analysis.

Linear Amplification Mediated High Throughput Genome-wide Translocation Capture Sequencing (LAM-HTGTS or HTGTS). HTGTS protocol was adapted from previously published method[49]. Briefly, genomic DNA was sonicated with 2 pulses of 5 secs ON and 1 minute off using the Bioruptor pico sonicator (Diagenode). LAM-PCR was performed using Phusion High-Fidelity DNA Polymerase (Thermo

Fisher) and Sμ 5' biotinylated bait (TAGTAAGCGAGGCTCTAAAAAGCA). 4 independent PCRs were run with ~2.5 μg of sonicated DNA each, with 90 cycles of primer elongation. The biotinylated DNA fragments were purified using Dynabeads T1 streptavidin beads (Invitrogen) at 25°C, on a rotisserie, overnight. On bead ligation was performed with T4 DNA ligase (NEB) and annealed bridge adaptors (GCGACTATAGGGCACGCGTGGNNNNNN-NH_2 and 5'P-CCACGCGTGCCCTATAGTCGC-NH_2) for 1h at 25°C, 2h at 22°C, and overnight at 16°C. 15 cycles of nested PCR were performed on the DNA-bead complexes using an Sμ nested primer (ACACTCTTTCCCTACACGACGCTCTTCCGATCT-**GGTAAGCAAAGCTGGGCTTG**) and a reverse universal primer I7 (GACTGGAGTTCAGACGTGTGCTCTTCCGATCT-**GACTATAGGGCACGCGTGG**). The bold regions represent the complementary regions to the genomic sequence in the Sμ nested primer and the bridge adaptor for I7 primer. PCR products were cleaned with the QIAquick Gel Extraction Kit (Qiagen) and the final barcoding/indexing PCR was performed using the P5-I5 and P7-I7 universal dual indexing primers (NEB) for 15 cycles and purified using Ampure beads purification with a cutoff of 200 bp. Library quality was evaluated by Bioanalyzer (Agilent) and sequencing performed using the Illumina Miseq Reagent Kit v2 (600 cycles) following the manufacturer's instructions. DNA junctions from 2 independent experiments were combined for analyses. The complete list of primers and sequences are included in extended data table 1.

Bioinformatics analyses

The reference genome used was mm10. Heatmaps and profile plots were generated using DeepTools[65].

HMCP analysis. Single-end reads were mapped to the mouse genome mm10 GRCm38 (Dec. 2011) indexed with spiked-in phage lambda, using Bowtie (V 1.1.2) (-S -p 3)[66]. Reads that mapped to spike-ins were filtered out with samtools view; remaining mapped reads in sam files were sorted and PCR duplicates were removed with Picard (V 2.7.1). Peaks were called with MACS2 (callpeak --keep-dup all -g mm)[67]. Peaks from the replicates from same condition were merged with mergePeaks from HOMER and kept only those that intersected between the replicates. Intersected peaks from the replicates were merged between the conditions to create the total number of peaks. Tag directories were created and replicates were merged with makeTagDirectory from HOMER[68] (-genome mm10); makeMultiWigHub.pl was used to generate tracks.

WGBS analysis. WGBS reads were mapped with BSMAP (v2.90)[69] to both the mouse genome mm10 GRCm38 (Dec. 2011) from UCSC and the lambda genome. Bisulfite conversion efficiency was estimated based on cytosine methylation in all contexts. For all the samples the bisulfite conversion efficiency was higher than 0.9996. Duplicated reads caused by PCR amplification were removed by PICARD's MarkDuplicates. CpG DNA methylation at both DNA strands was called by methratio.py script, from BSMAP (v2.90) (-g - i "correct" -x CG,CHG,CHH). To identify differentially methylated cytosines and regions (DMCs and DMRs) we used RADmeth methpipe-3.4.2 (adjust -bins 1:100:1 ; merge -p 0.05).

RNA-seq and BCR analysis. Single-end (50bp) reads were mapped to the mouse genome mm10/GRCm38 using STAR[70] (v2.5.3a) (--runThreadN 8 --genomeLoad LoadAndRemove --outFilterMultimapNmax 1 --outFilterType BySJout --alignSJoverhangMin 8 --alignSJDBoverhangMin 1 --alignIntronMin 20 --alignIntronMax 1000000 --alignMatesGapMax 100000 --outFilterMismatchNmax 0). Counts were obtained with featureCounts (subread-1.4.3-p1) (-g gene_name -s 1). Differentially expressed genes were calculated with DESeq2[71], filtering out genes that did not have any count in any condition; cut off to define DEGs was an adjusted pvalue 0.05 and a log2 fold change >= +/-1

BCR sequences were retrieved from RNA-Seq data sets, and the frequency of IG chain clonotypes was determined using MiXCR[72] (mixcr-1.7–2.1) package with the default parameters "align -c IG -s mmu -p rna-seq -OallowPartialAlignments=true". Two rounds of contig assembly were performed by employing the "assemblePartial" function; extension of incomplete BCR was done with "extendAlignments" function; assembly and export of the clonotypes was performed using the "assemble" and the "exportClones" --

preset min -fraction -targets -vHits -dHits -jHits -vAlignments -dAlignments -jAlignments) functions, respectively.

G4 analysis. Paired-end (50bp) reads were mapped to the mouse genome mm10 GRCm38 (Dec. 2011) indexed with spiked-in G4 sequences, using Bowtie (V 1.1.2)[66]. Reads that mapped to spike-ins were filtered out with samtools view; remaining reads in sam format were sorted and PCR duplicates were marked and removed with Picard (V 2.7.1). Peaks were calculated with MACS2[67] (callpeak –keep-dup all -g mm -p 0.0001), using input as control. Peaks from the replicates from same condition were merged with mergePeaks from HOMER and kept only those that intersected between the replicates. Intersected peaks from the replicates were merged between the conditions to create the total number of peaks. Tag directories were created and replicates were merged with makeTagDirectory from HOMER[68] (-genome mm10); makeMultiWigHub.pl was used to generate tracks.

MapR. Paired-end (50bp) reads were mapped to the mouse genome mm10 GRCm38 (Dec. 2011, using Bowtie (V 1.1.2). Mapped reads were sorted and PCR duplicates were marked and kept with Picard (V 2.7.1). Peaks were calculated with MACS2 (callpeak –keep-dup all -g mm --broad --broad-cutoff 0.1), using input as control. Peaks from the replicates from same condition were merged with mergePeaks from HOMER and kept only those that intersected between the replicates. Intersected peaks from the replicates were merged between the conditions to create the total number of peaks. Tag directories were created and replicates were merged with makeTagDirectory from HOMER[68] (-genome mm10); makeMultiWigHub.pl was used to generate tracks. For heatmaps and profiles plots generated with Deeptools, MNase control signal was removed from the RNASEH-MNase signal.

G-quadruplexes and R-loops peaks. The 9722 G-quadruplex and R-loop regions were obtained from the union of G-quadruplexes and R-loops peaks identified by using “intersectBed” from bedtools[73].

HTGTS. Analysis was done following HTGTS[49] pipeline (https://github.com/robinmeyers/transloc_pipeline). Briefly, paired-end reads were trimmed with TranslocPreprocess.pl and full analysis was performed with TranslocWrapper.pl. Post filter analysis was done using TranslocFilter.pl with default parameters. Circos plots were generated using the Circos software[74]. HTGTS hits were identified in a 10kb window and color scale was assigned based on the number of hits per window.

Potential to form G-quadruplexes analysis. Regular expressions in ‘awk’ commands were used to categorize the potential to form G-quadruplexes and R-loops peaks from fasta sequences. The loop1-7 is defined as with the following expression: G_3+N_{1-7} G_3+N_{1-7} $G_3+N_{1-7}G_3$. The long-loop is defined as sequences with a G4 with any loop of length >7 (up to 12 for any loop and 21 for the middle loop). The simple-bulge is defined as sequences with a G4 with a bulge of 1–7 bases in one G-run or multiple 1-base bulges. And the Two Tetrad/Complex bulge is defined as sequences with G4s with two G-bases per G-run with several bulges of 1–5 bases.

AID motifs hotspots analysis. Fasta sequences were retrieved from bed files from the G-quadruplexes and R-loops peaks. Library “stringr” and function str_locate_all in R was used to identify the "[A|T][A|G]C[C|T]" or "[A|G]G[C|T][A|T]" motifs.

Random control regions. Random regions were retrieved with the function “shuffleBed” from the bedtools[15] using as the reference the whole mm10 genome. In particular cases, random regions were retrieved from euchromatin regions defined by the A compartment from Hi-C analysis in naïve and activated B cells[75].

Quantification and Statistical Analysis. Statistical analyses were performed with Graph prism 8 and R version 3.4. The statistical tests used to determine significance in each analysis are described in the figure legends of the corresponding figures. Parametric tests (t-tests, ANOVA) were used in experiments where normal distribution could be assumed, whereas in other cases, where normal

distribution could not be assumed, non-parametric tests (Wilcoxon signed-rank test, or Kruskal-Wallis test and the ad hoc Dunn's test) were used. No statistical methods were used to pre-determine sample sizes but our sample sizes are similar to those reported in previous publications[7,35]. No data points were excluded from the analysis and appropriate animals/samples for each experiment were chosen randomly. Since our analysis required genotyping of the experimental mouse groups, we were not able to blind ourselves during data collection.

Figure 1

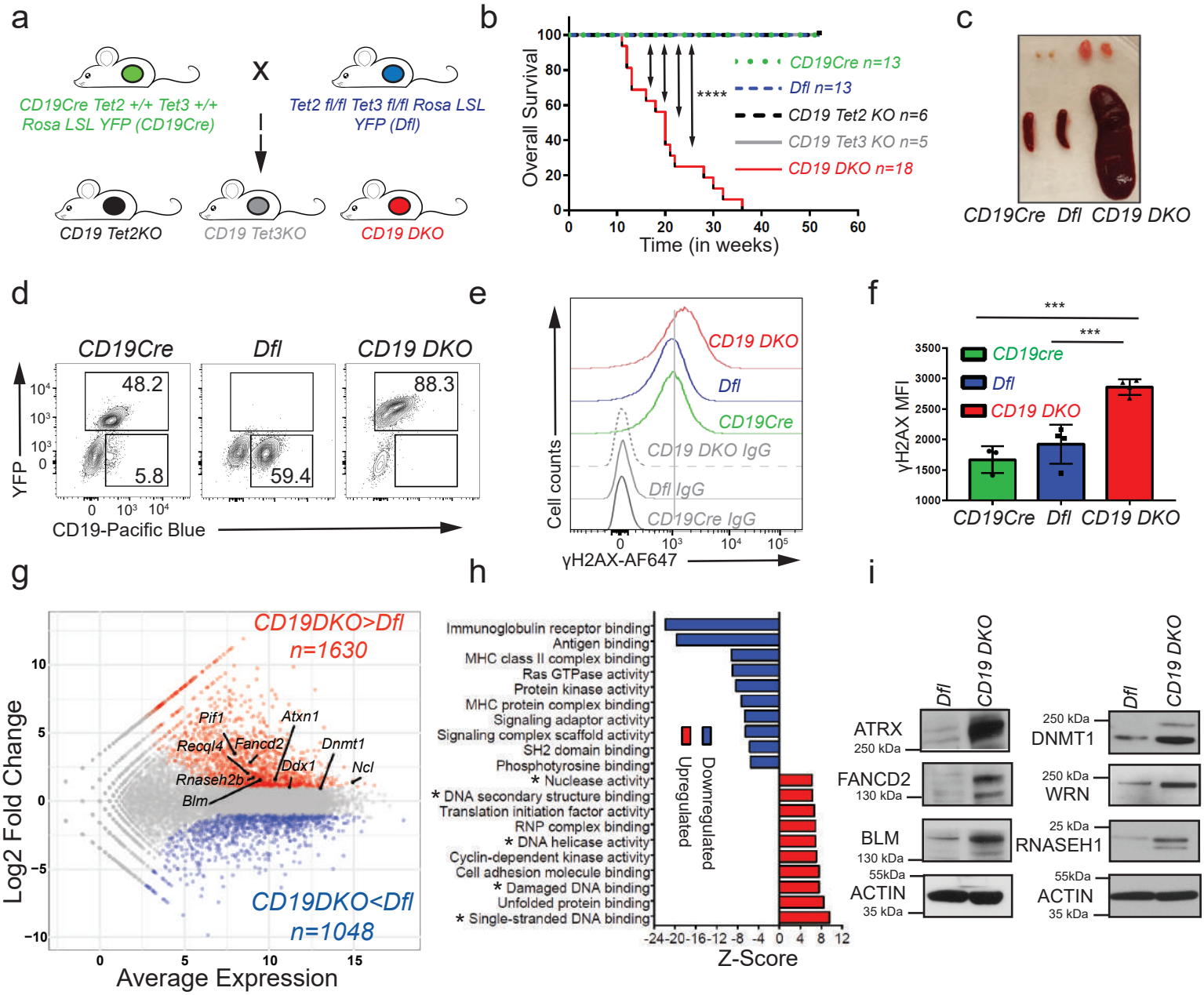

a
CD19Cre Tet2 +/+ Tet3 +/+ Rosa LSL YFP (CD19Cre)
x
Tet2 fl/fl Tet3 fl/fl Rosa LSL YFP (Dfl)
CD19 Tet2KO
CD19 Tet3KO
CD19 DKO
b
Overall Survival
Time (in weeks)
CD19Cre n=13
Dfl n=13
CD19 Tet2 KO n=6
CD19 Tet3 KO n=5
CD19 DKO n=18
****
c
CD19Cre Dfl CD19 DKO
d
CD19Cre
Dfl
CD19 DKO
48.2
5.8
59.4
88.3
YFP
CD19-Pacific Blue
e
CD19 DKO
Dfl
CD19Cre
CD19 DKO IgG
Dfl IgG
CD19Cre IgG
Cell counts
γH2AX-AF647
f
***
***
CD19cre
Dfl
CD19 DKO
γH2AX MFI
CD19Cre Dfl CD19 DKO
g
CD19DKO>Dfl
n=1630
CD19DKO<Dfl
n=1048
Pif1
Atxn1
Recql4
Fancd2
Dnmt1
Ddx1
Ncl
Rnaseh2b
Blm
Log2 Fold Change
Average Expression
h
Immunoglobulin receptor binding
Antigen binding
MHC class II complex binding
Ras GTPase activity
Protein kinase activity
MHC protein complex binding
Signaling adaptor activity
Signaling complex scaffold activity
SH2 domain binding
Phosphotyrosine binding
* Nuclease activity
* DNA secondary structure binding
Translation initiation factor activity
RNP complex binding
* DNA helicase activity
Cyclin-dependent kinase activity
Cell adhesion molecule binding
* Damaged DNA binding
Unfolded protein binding
* Single-stranded DNA binding
Upregulated
Downregulated
Z-Score
i
Dfl
CD19 DKO
ATRX
250 kDa
FANCD2
130 kDa
BLM
130 kDa
55kDa
ACTIN
35 kDa
DNMT1
250 kDa
WRN
250 kDa
RNASEH1
25 kDa
55kDa
ACTIN
35 kDa

Figure 2

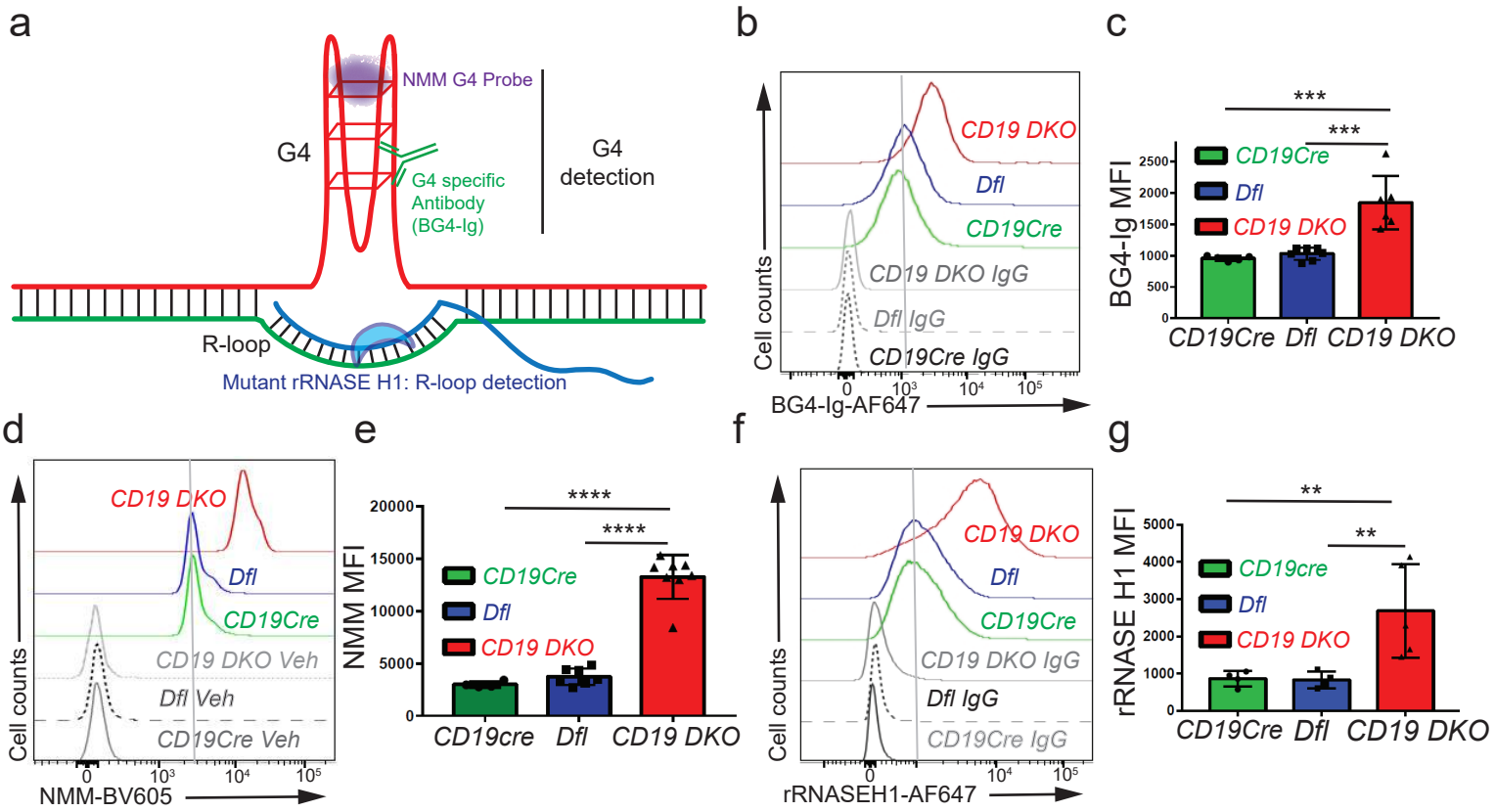

a
NMM G4 Probe
G4
G4 specific
Antibody
(BG4-Ig)
G4
detection
R-loop
Mutant rRNASE H1: R-loop detection
b
CD19 DKO
Dfl
CD19Cre
CD19 DKO IgG
Dfl IgG
CD19Cre IgG
Cell counts
BG4-Ig-AF647
c
***
***
CD19Cre
Dfl
CD19 DKO
BG4-Ig MFI
CD19Cre Dfl CD19 DKO
d
CD19 DKO
Dfl
CD19Cre
CD19 DKO Veh
Dfl Veh
CD19Cre Veh
Cell counts
NMM-BV605
e
****
****
CD19Cre
Dfl
CD19 DKO
NMM MFI
CD19cre Dfl CD19 DKO
f
CD19 DKO
Dfl
CD19Cre
CD19 DKO IgG
Dfl IgG
CD19Cre IgG
Cell counts
rRNASEH1-AF647
g
**
**
CD19cre
Dfl
CD19 DKO
rRNASE H1 MFI
CD19Cre Dfl CD19 DKO

Figure 3

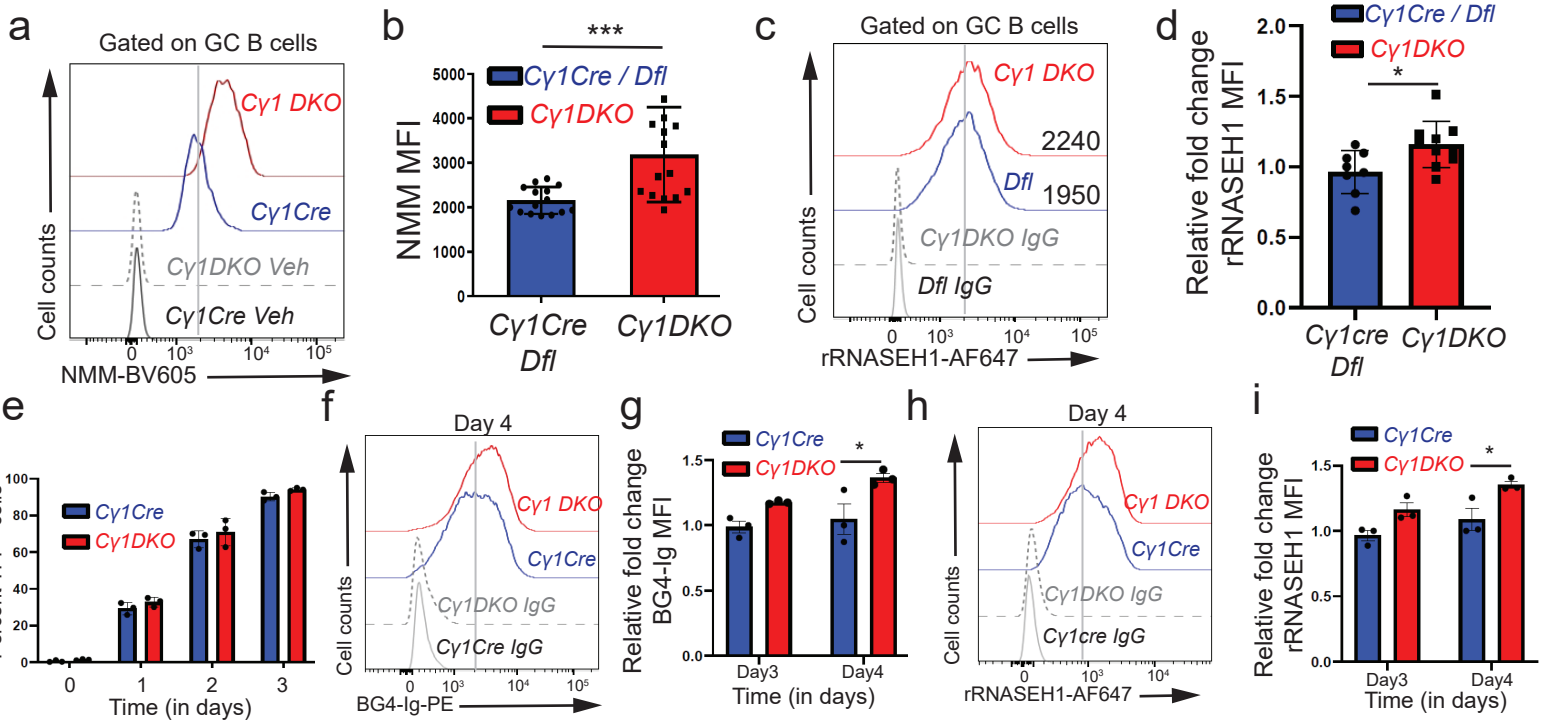

a
Gated on GC B cells
Cγ1 DKO
Cγ1Cre
Cγ1DKO Veh
Cγ1Cre Veh
Cell counts
NMM-BV605
b
***
Cγ1Cre / Dfl
Cγ1DKO
NMM MFI
Cγ1Cre Dfl
Cγ1DKO
c
Gated on GC B cells
Cγ1 DKO
2240
Dfl
1950
Cγ1DKO IgG
Dfl IgG
Cell counts
rRNASEH1-AF647
d
Cγ1Cre / Dfl
Cγ1DKO
*
Relative fold change rRNASEH1 MFI
Cγ1cre Dfl
Cγ1DKO
e
Cγ1Cre
Cγ1DKO
Percent YFP+ cells
Time (in days)
f
Day 4
Cγ1 DKO
Cγ1Cre
Cγ1DKO IgG
Cγ1Cre IgG
Cell counts
BG4-Ig-PE
g
Cγ1Cre
Cγ1DKO
*
Relative fold change BG4-Ig MFI
Day3
Day4
Time (in days)
h
Day 4
Cγ1 DKO
Cγ1Cre
Cγ1DKO IgG
Cγ1cre IgG
Cell counts
rRNASEH1-AF647
i
Cγ1Cre
Cγ1DKO
*
Relative fold change rRNASEH1 MFI
Day3
Day4
Time (in days)

Figure 4

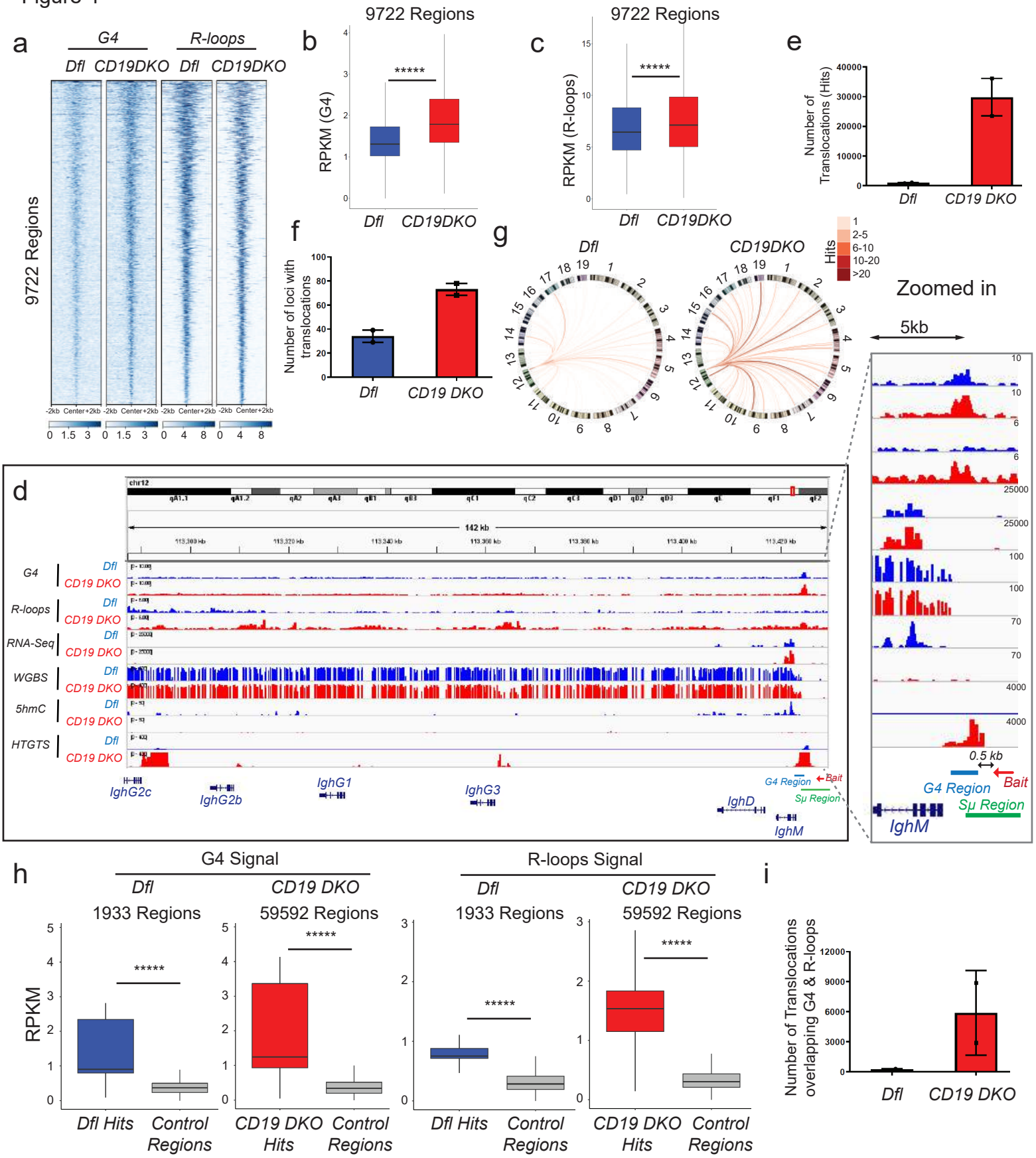
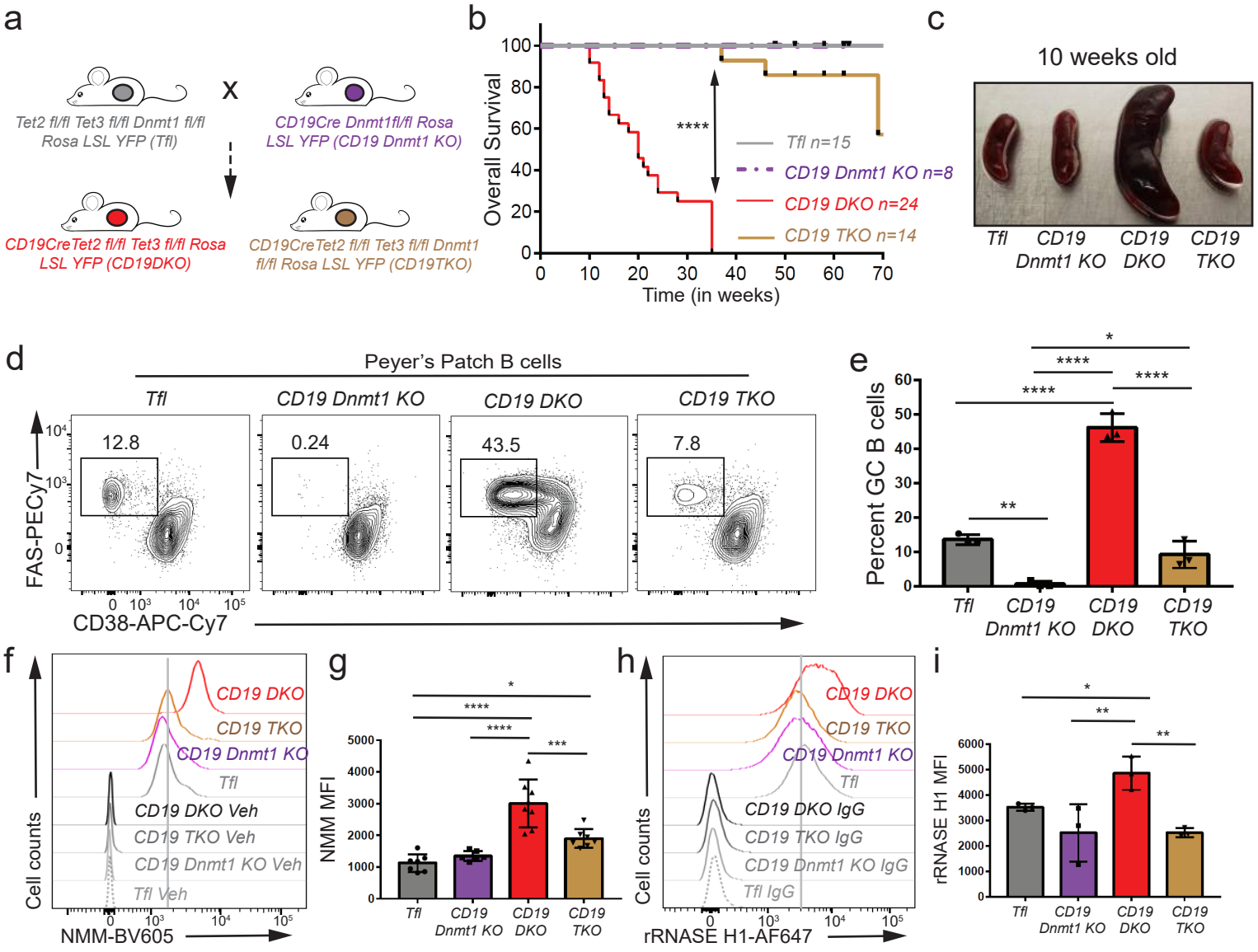
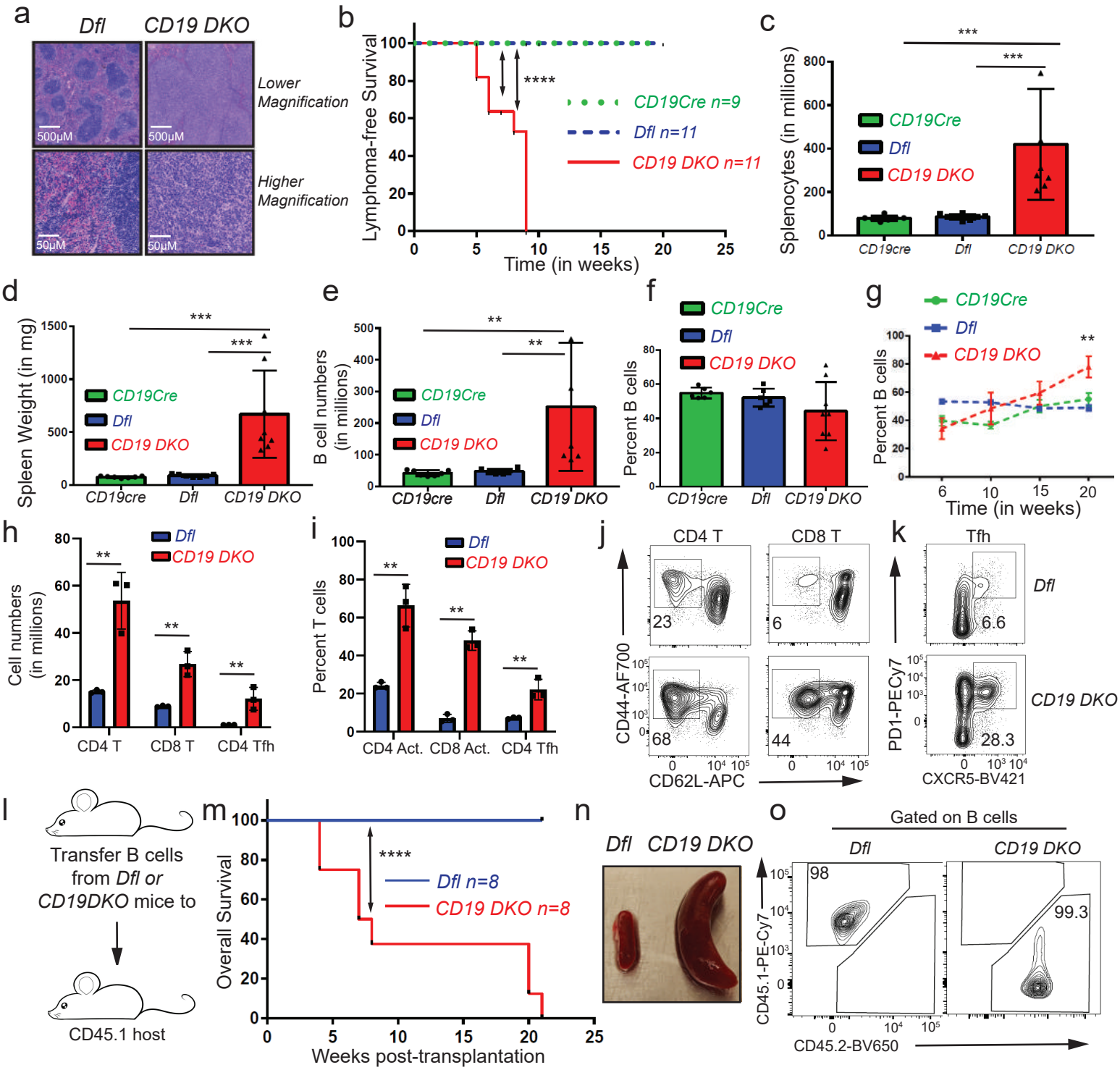

a
G4
R-loops
Dfl CD19DKO Dfl CD19DKO
9722 Regions
-2kb Center+2kb
b
9722 Regions
RPKM (G4)
*****
Dfl CD19DKO
c
9722 Regions
RPKM (R-loops)
*****
Dfl CD19DKO
e
Number of Translocations (Hits)
Dfl CD19 DKO
f
Number of loci with translocations
Dfl CD19 DKO
g
Dfl
CD19DKO
Hits
1
2-5
6-10
10-20
>20
Zoomed in
5kb
d
chr12
142 kb
G4
R-loops
RNA-Seq
WGBS
5hmC
HTGTS
Dfl
CD19 DKO
IghG2c
IghG2b
IghG1
IghG3
IghD
IghM
G4 Region
Bait
Sμ Region
0.5 kb
h
G4 Signal
R-loops Signal
Dfl
CD19 DKO
1933 Regions
59592 Regions
RPKM
*****
Dfl Hits
Control Regions
CD19 DKO Hits
i
Number of Translocations overlapping G4 & R-loops
Dfl
CD19 DKO

Figure 5

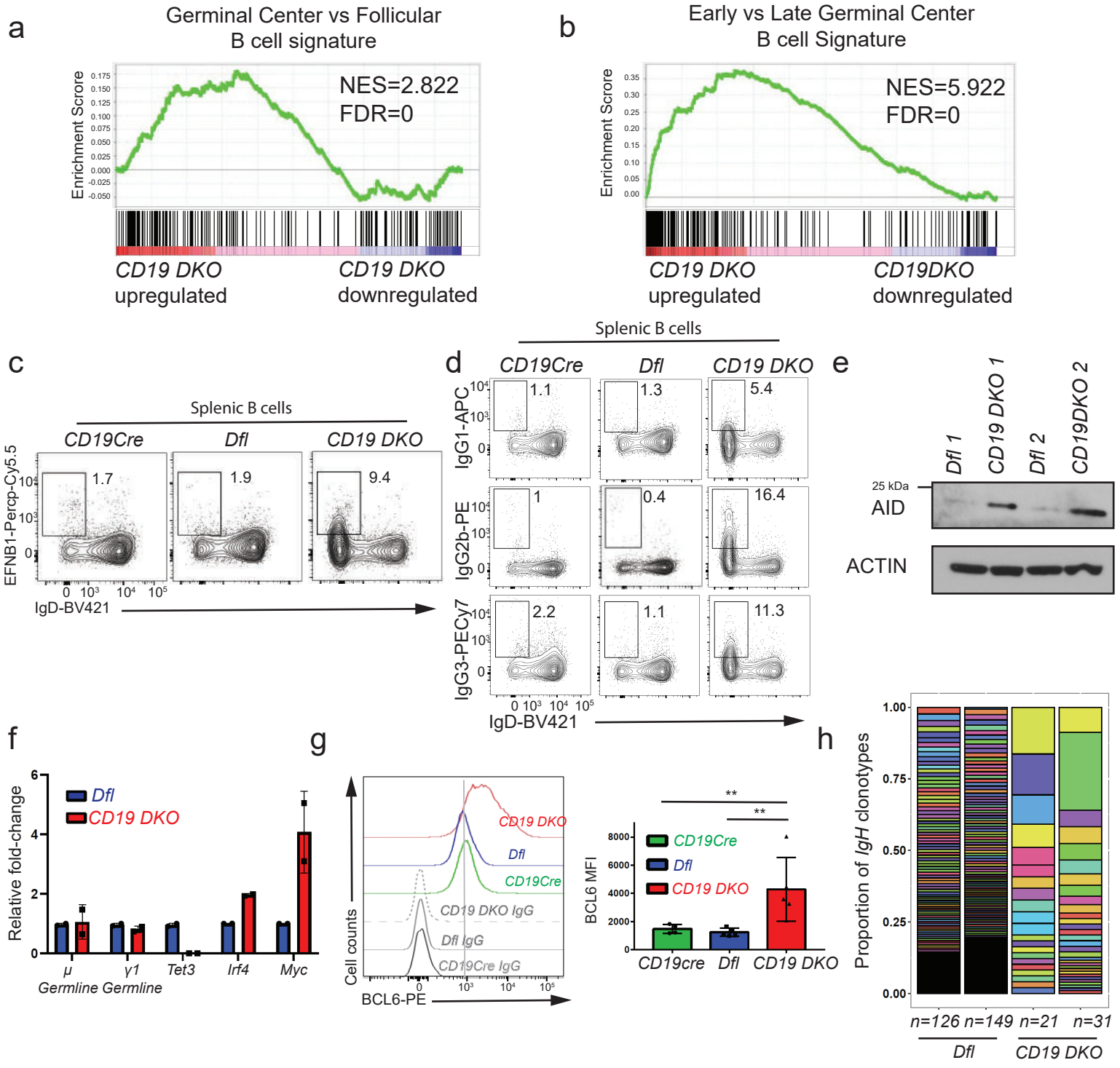

a
Tet2 fl/fl Tet3 fl/fl Dnmt1 fl/fl Rosa LSL YFP (Tfl)
x
CD19Cre Dnmt1fl/fl Rosa LSL YFP (CD19 Dnmt1 KO)
CD19CreTet2 fl/fl Tet3 fl/fl Rosa LSL YFP (CD19DKO)
CD19CreTet2 fl/fl Tet3 fl/fl Dnmt1 fl/fl Rosa LSL YFP (CD19TKO)
b
Overall Survival
Time (in weeks)
****
Tfl n=15
CD19 Dnmt1 KO n=8
CD19 DKO n=24
CD19 TKO n=14
c
10 weeks old
Tfl
CD19 Dnmt1 KO
CD19 DKO
CD19 TKO
d
Peyer's Patch B cells
Tfl
CD19 Dnmt1 KO
CD19 DKO
CD19 TKO
12.8
0.24
43.5
7.8
FAS-PECy7
CD38-APC-Cy7
e
Percent GC B cells
f
CD19 DKO
CD19 TKO
CD19 Dnmt1 KO
Tfl
CD19 DKO Veh
CD19 TKO Veh
CD19 Dnmt1 KO Veh
Tfl Veh
Cell counts
NMM-BV605
g
NMM MFI
h
CD19 DKO IgG
CD19 TKO IgG
CD19 Dnmt1 KO IgG
Tfl IgG
rRNASE H1-AF647
i
rRNASE H1 MFI

Extended data figure 1



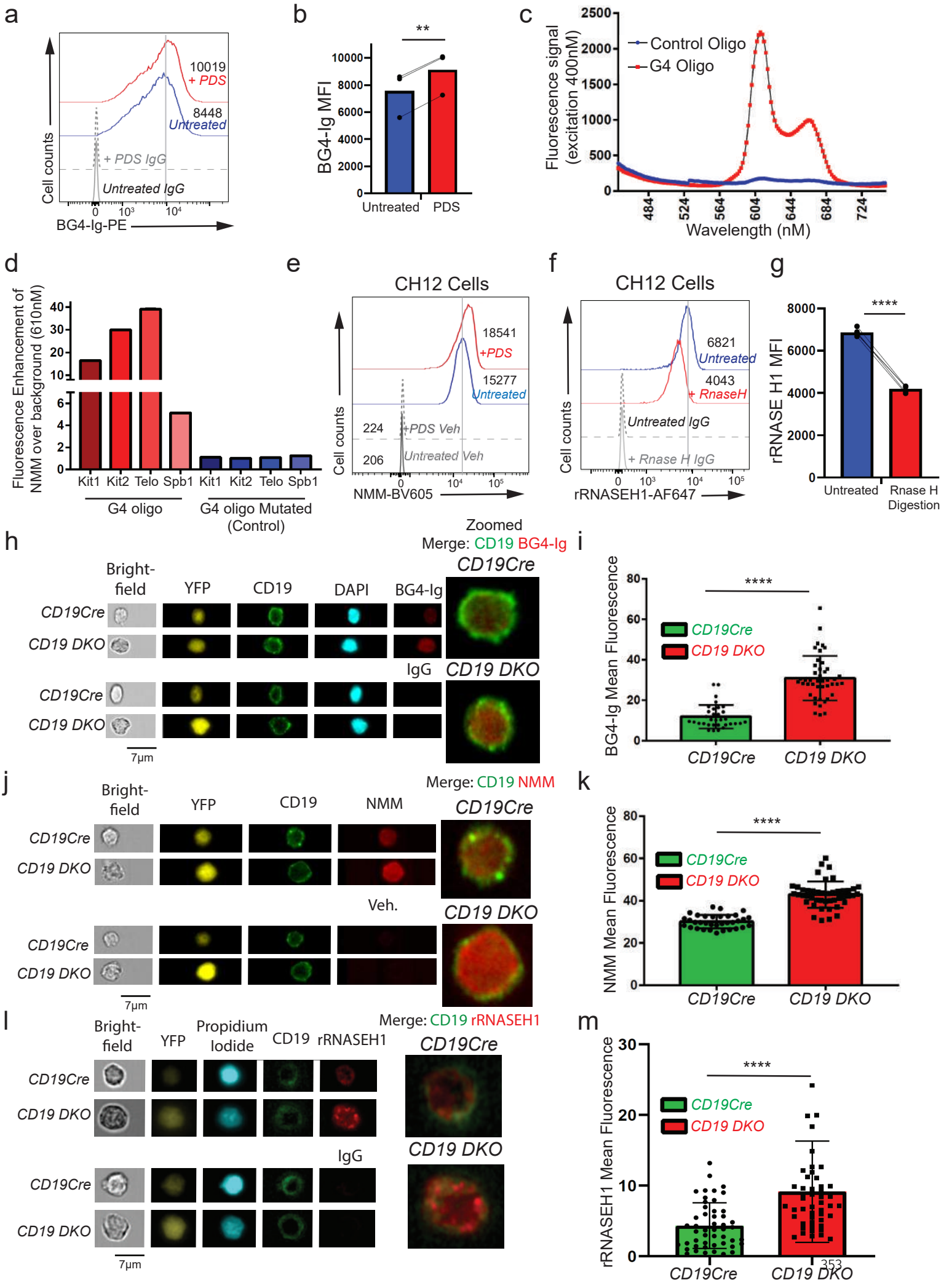
Extended data figure 1. TET deficiency leads to development of mature B cell lymphoma. a) H&E staining in spleens from 8 week-old *Dfl* and *CD19 DKO* mice The data is representative of 2 independent experiments. **b)** Kaplan-Meir curves displaying lymphoma-free survival of *CD19 DKO* (red) and control *Dfl* (blue) and *CD19 Cre* (green) mice. Y-axis, percent mice without lymphoma, defined by $\geq$ 3-fold increase in spleen weight and $\geq$ 2-fold increase in cellularity. **c)** – **f)** Quantification of **c)** cell numbers, **d)** spleen weights, **e)** B cell numbers, **f)** percent of B cells in spleens of 9 week-old *CD19 DKO* (red) and control *Dfl* (blue) and, *CD19 Cre* (green) mice from at least 5 biological replicates and 4 independent experiments. **g)** Percent of B cells in spleens of *CD19 DKO* (red) and control *Dfl* (blue) and *CD19 Cre* (green) mice at different ages from 3 independent experiments. **h)** Quantification of cell numbers of $CD4^+$ T, $CD8^+$ T and $CD4^+$ Tfh cells from spleens of 9 week-old *CD19 DKO* (red) and control *Dfl* (blue) mice from 3 biological replicates and 2 independent experiments. **i)** Percent of activated $CD4^+$ and, $CD8^+$ T cells, and CD4 Tfh cells from spleens of 9 week-old *CD19 DKO* (red) and control *Dfl* (blue) mice from 3 biological replicates and 2 independent experiments. **j) and k)** Flow cytometry plots of **j)** Activated $CD62L^{low}$ (X-axis) and $CD44^+$ (Y-axis) $CD4^+$ and $CD8^+$ T cells and **k)** $CD4^+$ Tfh cells $PD1^+$ (Y-axis) $CXCR5^+$ (X-axis) from spleen of 9 week-old *CD19 DKO* (red) and control *Dfl* (blue) mice. Numbers represent frequency of gated populations. **l)** Scheme of adoptive transfer experiment. B cells from *Dfl* and *CD19 DKO* CD45.2 mice were transferred retro-orbitally into sub-lethally irradiated CD45.1 immunocompetent host mice. **m)** Kaplan-Meir curves displaying overall survival of CD45.1 host mice transplanted with B cells from *Dfl* or *CD19 DKO* mice. X-axis denotes time in weeks after transplantation. **n)** Enlarged spleens in CD45.1 host mice transplanted with *CD19 DKO* compared to *Dfl* B cells 8 weeks after adoptive transfer. **o)** Representative flow cytometry data from the spleens of CD45.1 host mice 8 weeks after transplantation with *Dfl and CD19 DKO* B cells. Frequencies of the $CD45.1^+$ and $CD45.2^+$ cell populations are shown. Statistical significance is calculated using the log-rank test for **b**) and **m**), one-way ANOVA for **c**)-**f**) and two-way ANOVA for **g**)-**i**). Error bars represent mean +/- standard deviation, ** p value ≤0.01 *** p value ≤0.001, **** p value ≤0.0001.

Extended data figure 2



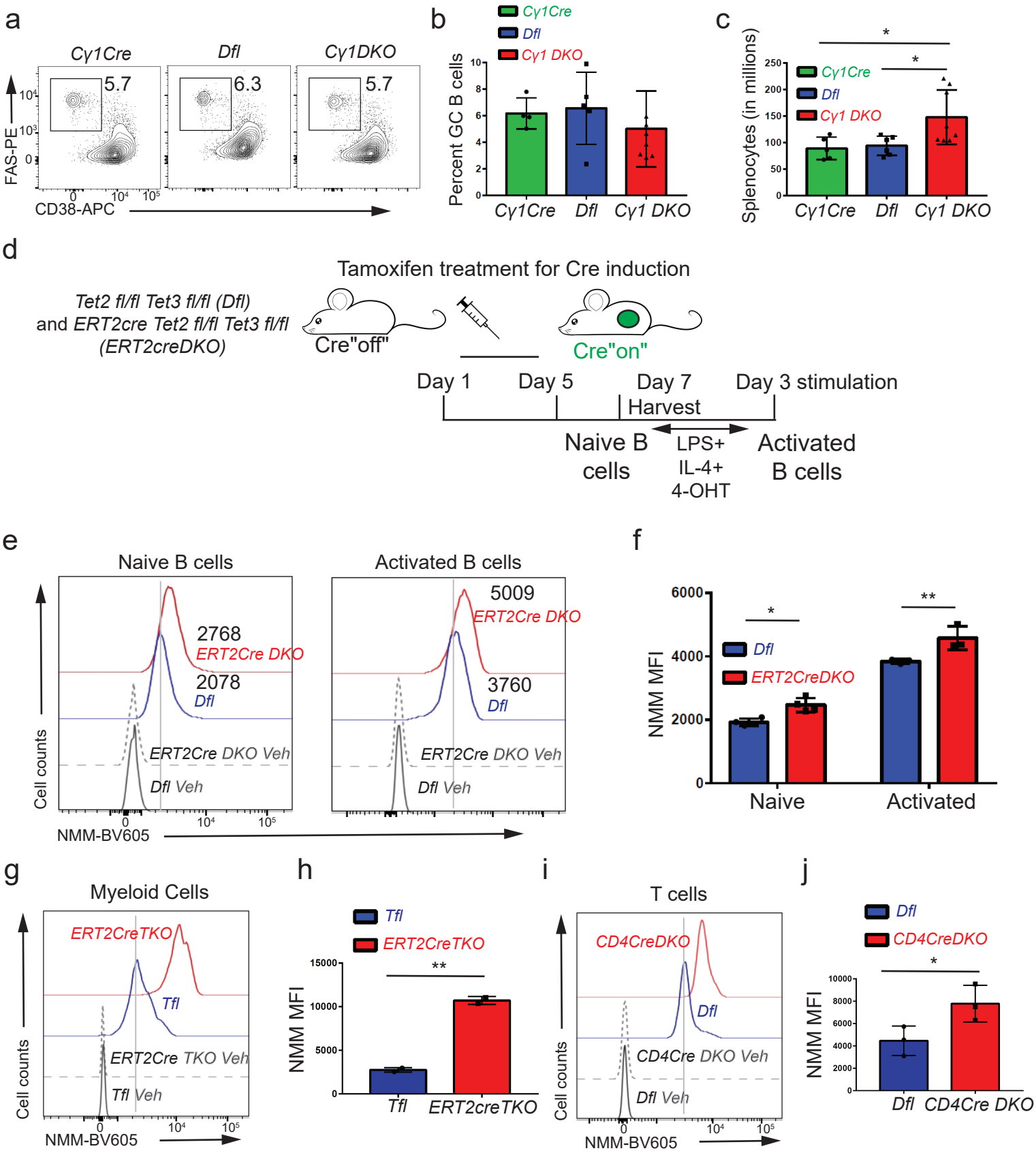
Extended data figure 2. Expanded B cells from *CD19 DKO* mice have a germinal center (GC) origin. **a), b)** Gene set enrichment analysis (GSEA) plots showing enrichment for a GC B cell transcriptional signature in the transcriptional profile of *CD19 DKO* compared to *Dfl* B cells, using gene sets from **a)** GC versus follicular B cells and **b)** early GC versus late GC B cells. Y-axis denotes enrichment score. NES, Normalized enrichment score, FDR, False discovery rate. **c)** Representative flow cytometry data gated on splenic B cells from 8 week-old *Dfl*, *CD19Cre (YFP$^+$) and CD19 DKO (YFP$^+$)* mice. Numbers in the rectangles represent frequencies of GC B cells, identified as EFNB1$^+$ (Y-axis) and IgD^{low} (X-axis). **d)** Representative flow cytometry data showing Ig isotype expression, gated on splenic B cells from 8 week-old *Dfl*, *CD19Cre (YFP$^+$)* and *CD19 DKO (YFP$^+$)* mice. Top, IgG1; middle, IgG2b; bottom, IgG3 X-axis shows expression of the default IgD isotype. Numbers represent frequencies of gated cell populations. **e)** Immunoblot showing AID expression in *Dfl* and *CD19 DKO* B cells (2 replicate experiments). Actin is used as a loading control. **f)** Relative fold-change in expression of *μ* and *γ1* germline, *Tet3*, *Irf4* and *Myc* transcripts measured by qRT-PCR in *Dfl* and *CD19 DKO* B cells from 2 biological replicates. **g)** Histogram (*left panel*) and bar-graph (*right panel*) showing staining with BCL6 antibody compared to isotype IgG controls in B cell from 8 week-old *CD19Cre*, *Dfl and CD19 DKO* mice from 3 independent experiments. **h) and i)** Bar plots displaying the proportions of **(h)** IgVH and **(i)** Igκ clonotypes (rearranged variable gene segments) from *Dfl* (blue) and *CD19 DKO* (red) B cells, identified from BCR repertoire analysis of RNA-Seq data. Y-axis represents the proportion of each clonotype. Each individual IgVH and Igκ clonotype is displayed using a different color in the bar plots. Numbers at the bottom represent the number of clonotypes identified in two independent replicates of *Dfl* (blue) and *CD19 DKO* (red) B cells. **j)** Representative flow cytometry data gated on Peyer's patch B cells from 8 week-old *Dfl and CD19 DKO* mice. Numbers represent frequency of GC B cells, identified as FAS$^+$ (Y-axis) and CD38$^-$ (X-axis). **k)** Quantification of GC B cell frequency in Peyer's patches of *Dfl* (blue) and *CD19 DKO* (red) mice from 3 independent experiments. Statistical significance is calculated using two-tailed student t-test in **f**), **k**) and one-way ANOVA in **g**). Error bars represent mean +/- standard deviation in **f**), **g**) and **k**). ** p value ≤0.01.

Extended data figure 3



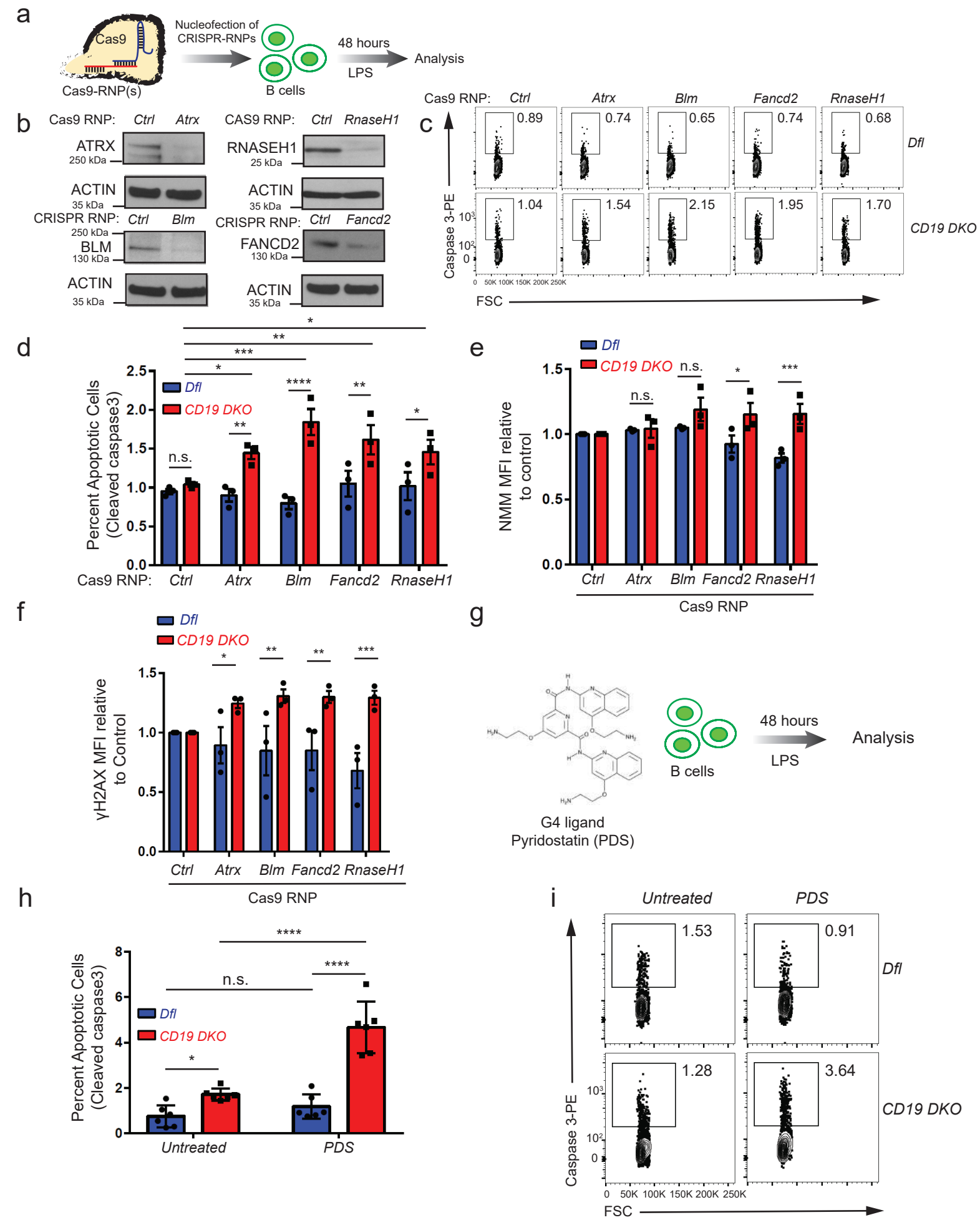
Extended data figure 3. TET deficiency is associated with increased levels of G-quadruplexes and R-loops. **a)** Flow cytometric detection of G-quadruplexes with BG4-Ig antibody or isotype IgG controls in primary B cells stimulated with 25 µg/ml LPS for 48 hours and treated with 10 µM pyridostatin (PDS, G4 ligand) for 24 hours. Numbers represent median fluorescence intensity. **b)** Quantification of median fluorescence intensity (MFI) of BG4-Ig signal from primary B cells treated with (red) or without (blue) PDS. Lines connect paired samples treated with or without PDS from 3 independent experiments. **c)** Fluorescence emission spectrum of NMM in the presence of a G4-forming oligonucleotide (oligo) from the human *c-Kit* gene promoter or a control oligo in which guanines in G4-forming regions (G-tracts) were mutated. **d)** Fluorescence enhancement over background (no oligos) for NMM at 610 nm in presence of known G4-forming oligos (Kit1, Kit2, Spb1) from the *c-Kit* gene locus or the telomeric repeat (Telo) or their respective mutated versions. **e)** G-quadruplex levels assessed by NMM or DMSO vehicle control (Veh) staining in untreated CH12 B cells or cells treated with 5 µM pyridostatin (PDS, G4 ligand) for 24 hours. Numbers represent median fluorescence intensity from 3 independent experiments. **f)** Flow cytometric detection of R-loops using V5-epitope-tagged recombinant RNASE H1 (rRNASE H1) in CH12 cells with or without RNASE H enzyme digestion during staining. Numbers represent median fluorescence intensity. **g)** Quantification of median fluorescence intensity (MFI) of R-loops (rRNASE H1) signal from CH12 cells with (red) or without (blue) RNASE H enzyme digestion. Lines connect paired samples with or without RNASE H digestion from 3 independent experiments. **h) - m)** Representative images (**h, j, l**) and quantification of mean fluorescence signal (**i, k, m**) of *CD19cre and CD19 DKO* YFP^+ B cells stained with DAPI or propidium iodide and CD19, BG4-Ig (**h, i**), NMM (**j, k**) and rRNASE H1 (**l, m**) or respective controls using the AMNIS imagestream. Data are from two independent experiments. Statistical significance is calculated using paired student t-test in **b**) and **g**), two-tailed student t-test **i**), **k**) and **m)**. Error bars represent mean +/- standard deviation, ** p value ≤0.01, **** p value ≤0.0001.

Extended data figure 4



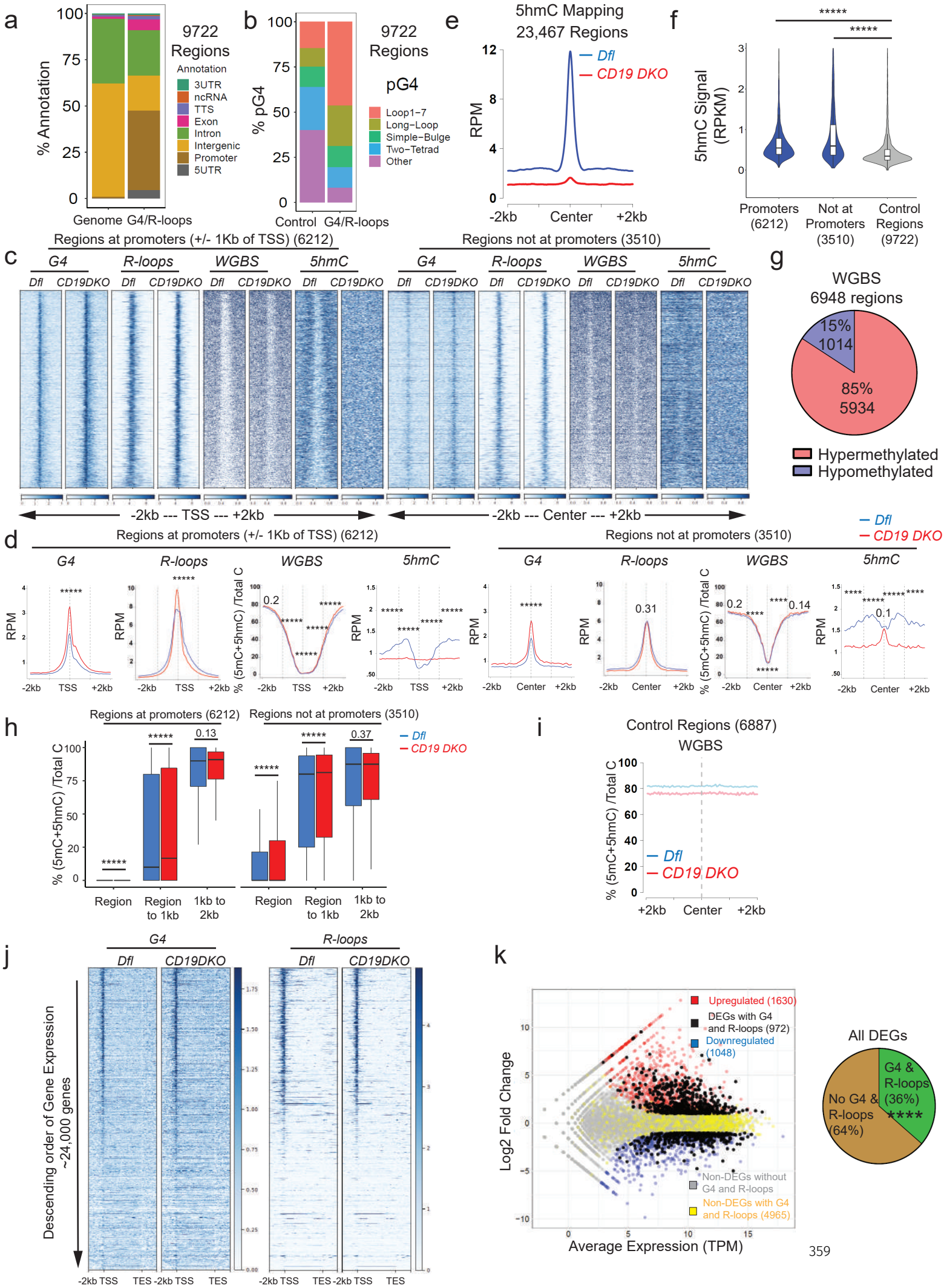
Extended data figure 4. TET deficiency in multiple primary cell types is associated with increased DNA G-quadruplex structures. a) Representative flow cytometry data gated on splenic B cells from 8 week-old *Cγ1Cre*, *Dfl* and *Cγ1 DKO* mice 12 days after immunization with SRBCs. GC B cells are identified as FAS^+ (Y-axis) and $CD38^-$ (X-axis). Numbers represent frequency of GC B cells. **b)**, **c)** Quantification of **(b)** GC B cell frequencies and **(c)** absolute numbers of splenocytes from 8 week-old *Cγ1Cre*, *Dfl* and *Cγ1 DKO* mice 12 days after immunization with SRBCs from 3 independent experiments. **d)** Experimental design. *ERT2Cre DKO* or control *Dfl* mice were injected for 5 consecutive days with tamoxifen to induce Cre expression and *TET* deletion, then rested for 2 days. Splenic B cells were activated for 72 hours *in vitro* with LPS and IL-4 in the presence of 4-hydroxytamoxifen (4-OHT). **e)** G-quadruplex levels in naïve (*left panel*) and activated (*right panel*) B cells from tamoxifen-treated *ERT2Cre DKO (YFP^+)* or control *Dfl* mice. Numbers represent median fluorescence intensity. **f)** Quantification of median fluorescence intensity (MFI) of NMM signal from naïve and activated B cells from *ERT2Cre DKO (YFP^+)* or control *Dfl* mice from 3 independent experiments. **g)**, **i)** G-quadruplex levels assessed by NMM or DMSO vehicle staining (Con) in **(g)** transferred $CD45.2^+$ myeloid cells from *ERT2Cre TKO (YFP^+)* or control *Tfl* mice, and **(i)** transferred $CD45.2^+$ T cells from *CD4Cre DKO (YFP^+)* or control *Dfl* mice. **h)**, **j)** Quantification of median fluorescence intensity (MFI) of NMM signal in **(h)** transferred $CD45.2^+$ myeloid cells from *ERT2Cre TKO (YFP^+)* or control *Tfl* mice from 2 biological replicates and **j)** transferred $CD45.2^+$ T cells from *CD4Cre DKO (YFP+)* or control *Dfl* mice from 3 biological replicates. Statistical significance is calculated using one-way ANOVA in **c**), two-way **f**) and two-tailed student t-test in **h**) and **j**). Error bars represent mean +/- standard deviation, * p value ≤0.01, ** p value ≤0.005.

Extended data figure 5



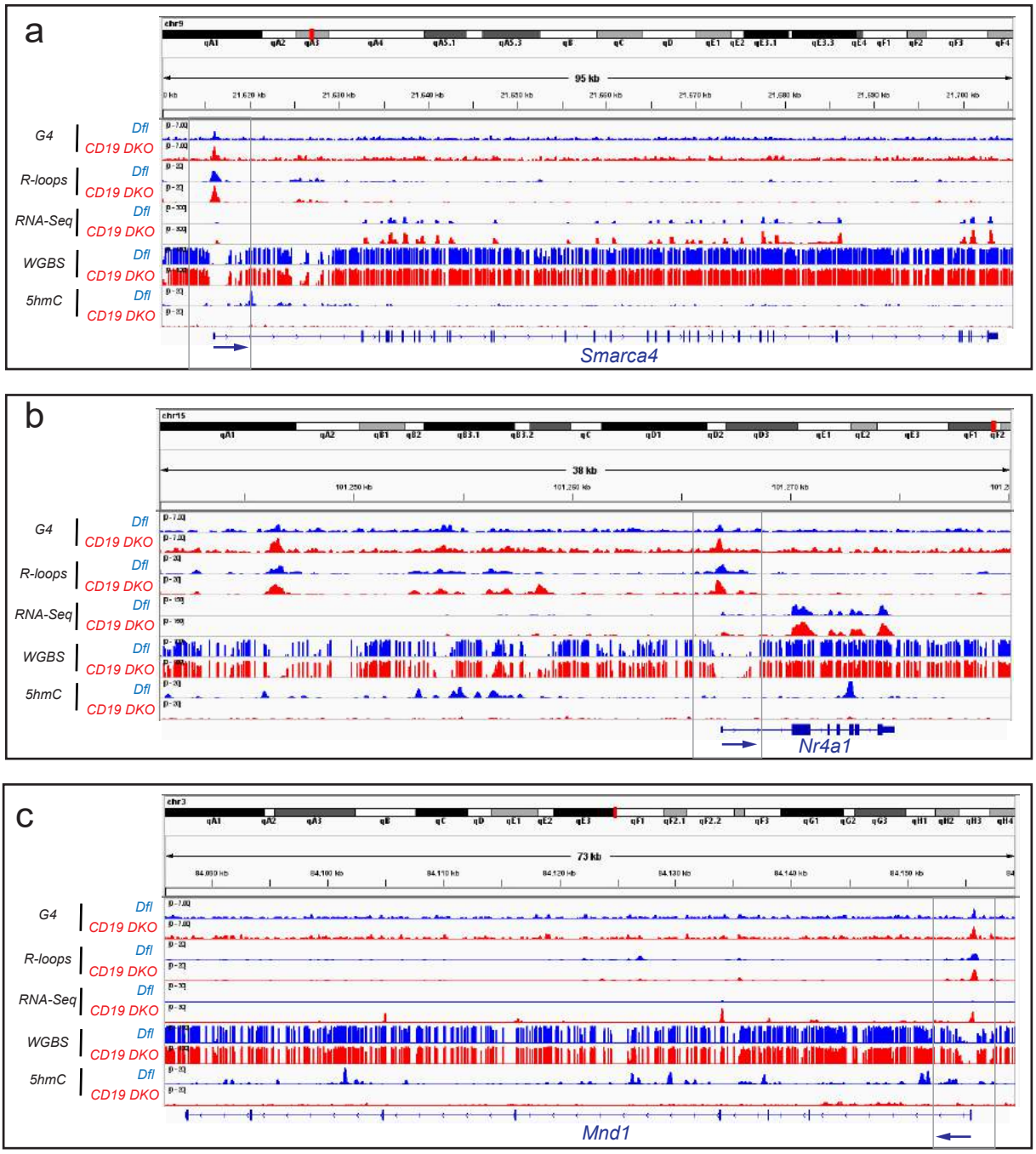
Extended data figure 5. Increased apoptosis and DNA DSBs in TET-deficient B lymphoma cells depleted of enzymes that resolve G-quadruplexes and R-loops. a) Experimental design. Primary B cells from *Dfl* and *CD19 DKO* B cells were nucleofected with Cas9 RNPs loaded with sgRNAs against *Rnase H1* or the known G4-binding helicases *Atrx, Blm* and *Fancd2*, then stimulated with 10 µg/ml LPS for 48 hours before assessing the frequency of apoptotic cells by flow cytometry for cleaved Caspase 3. **b)** Representative immunoblots showing decreased protein levels of ATRX, BLM, FANCD2 and RNASE H1 in *CD19 DKO* B cells nucleofected 48 hours earlier with the corresponding or *CD4* Cas9 RNPs (*Ctrl*). The data is representative of at least 2 independent experiments. **c)** Representative flow cytometry plots quantifying percent apoptotic cells in *Dfl* and *CD19 DKO* B cells nucleofected with Cas9 RNPs. Y-axis, staining for cleaved Caspase 3; X-axis, forward scatter (FSC). **d)** Quantification of apoptosis, measured as percent of cells showing staining for cleaved Caspase 3, in cells nucleofected with Cas9 RNPs to *Atrx, Blm*, *Fancd2, Rnase H* from 3 biological replicates. **e)** Quantification of G-quadruplexes as NMM median fluorescence intensity (MFI) in *Dfl* and *CD19 DKO* B cells 48 hours after nucleofection with Cas9 RNPs. The signal is normalized to the NMM MFI of the same biological sample nucleofected with *CD4* Cas9 RNPs (*Ctrl*) from 3 biological replicates. **f)** Quantification of DNA DSBs, assessed by γH2AX median fluorescence intensity (MFI) in *Dfl* and *CD19 DKO* B cells 48 hours after nucleofection with the indicated Cas9 RNPs from 3 biological replicates. The signal is normalized to the γH2AX MFI of the same biological sample nucleofected with control Cas9 RNP loaded with sgRNA against *CD4 (Ctrl)*. **g)** Experimental design. *Dfl* and *CD19 DKO* B cells were treated for 2 days with the G-quadruplex stabilizing compound pyridostatin (PDS) prior to activation for 48 hours with LPS. **h)** Quantification of apoptosis, measured as percent of cells showing staining for cleaved caspase 3, in *Dfl* and *CD19 DKO* B cells cultured without (untreated) or with 10 µM PDS from 5 biological replicates and 3 independent experiments. **i)** Representative flow-cytometry plots quantifying percent apoptotic cells in *Dfl* and *CD19 DKO* B cells with or without PDS treatment. Y-axis, staining for cleaved caspase 3; X-axis, forward scatter (FSC). Statistical significance is calculated using two-way ANOVA. Error bars represent mean +/- standard error, * p value ≤0.05, ** p value ≤0.01, *** p value ≤0.0005 **** p value <0.0001.

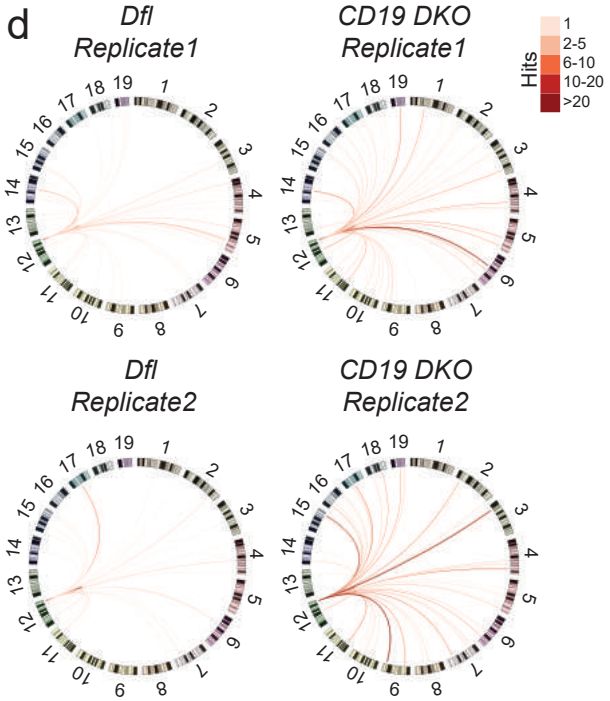
Extended data figure 6

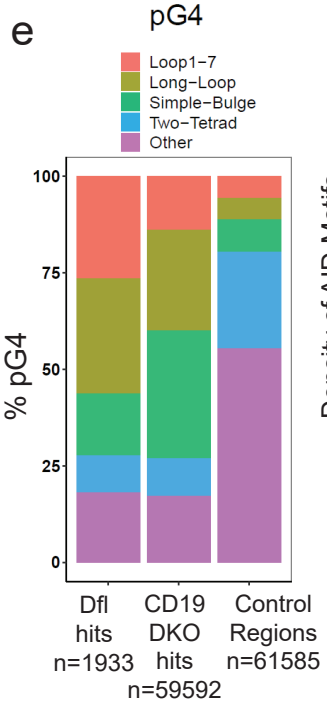


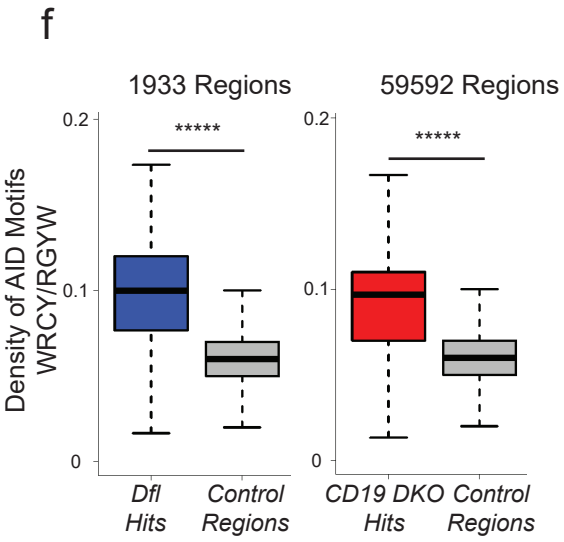
Extended data figure 6. TET deficiency is associated with genome-wide accumulation of G-quadruplexes and R-loops. **a)** Genome annotations of regions enriched for G-quadruplexes (G4) and R-loops (*right bar*) compared to their representation in the mouse genome (mm10) (*left bar*). **b)** Relative representation of different classes of motifs predicted to form G-quadruplexes (pG4) in control regions selected randomly from the genome (*left bar*) and regions enriched for G-quadruplexes and R-loops (*right bar*). **c)** Heat maps showing enrichment (RPM) for G-quadruplexes and R-loops in *CD19 DKO* and control B cells. The signal is plotted in a +/- 2 kb window from the center of the regions ordered based on decreasing intensity from top to bottom in the entire 4 kb window. R-loop signal is plotted after background subtraction of MNase-alone control. **d)** Profile histograms showing the signals for G-quadruplexes (G4) (RPM, reads per million), R-loops (RPM), WGBS (percent of 5mC+5hmC/unmodified C) and 5hmC (RPM). The 9722 regions enriched for both G-quadruplexes and R-loops are divided into two categories – 6212 regions overlapping promoters (*left panels*) and 3510 regions not at promoters (*right panels*). Dashed grey lines indicate the center of the region and the 1 kb boundaries located on either side of the center. Blue and red lines show data from *Dfl* and CD19 DKO B cells, respectively. Asterisks represent statistical significance calculated by comparing the signals between *Dfl* and *CD19 DKO* B cells, either within the G-quadruplex and R-loop forming regions, the region to +/-1kb window or +/-1kb to 2kb window for respective datasets. **e)** Profile histograms showing the 5hmC signal in *Dfl* (blue) and *CD19 DKO* (red) B cells in 23,467 regions identified as enriched for 5hmC signal. **f)** Violin plots quantifying enrichment (RPKM) of 5hmC signal in *Dfl* B cells in the +/-1 kb from G-quadruplex and R-loop forming regions at promoters, non-promoter regions and control regions randomly located in euchromatin (Hi-C A genomic compartment) from 2 biological replicates. **g)** Pie chart showing the differentially methylated regions (DMRs) in *CD19 DKO* compared to control *Dfl* B cells. Of a total of 6948 DMRs identified by WGBS, 1014 (15%) showed reduced DNA methylation (hypomethylation) and 5934 (85%) showed increased DNA methylation (hypermethylation). **h)** Box and whisker plots quantifying percent of 5mC+5hmC/unmodified C (from WGBS) at and near the G4 and R-loop forming regions overlapping promoters and regions not overlapping promoters in *Dfl* (blue) and *CD19 DKO* (red) B cells from 2 biological replicates. The signal is plotted in three windows; window 1, within the G4 and R-loop regions; window 2, from region to +/- 1kb on either side and; window 3, +/-1kb to +/-2kb on either side. **i)** Percent of 5mC+5hmC/unmodified C (from WGBS) in random genomic regions of *Dfl* (blue) and *CD19 DKO* (red) B cells. **j)** Heatmaps of enrichment (RPM) for G-quadruplexes (*left*) and R-loops (*right*) in *Dfl* and *CD19 DKO* B cells, ordered in descending order of gene expression. **k)** MA plot (*left*) showing differentially expressed genes (DEGs) in *CD19 DKO* B cells. *Red dots*, upregulated DEGs; *blue dots*, downregulated DEGs; black dots, DEGs with G-quadruplexes and R-loops at their promoters (+/-1kb of TSS); *yellow dots*, non-DEGs with G-quadruplexes and R-loops at their promoters; *grey dots*, non-DEGs without G-quadruplexes and R-loops. The pie-chart (*right*) shows the percent of DEGs with (*green*) and without (*brown*) G-quadruplexes and R-loops at their promoters. Asterisks indicate statistical significance. Statistical significance is calculated using Kruskal-Wallis test and the ad hoc Dunn's test in **d**), **f**) and **h)**, Chi-square test in **j**). Boxes in box and whisker plots represent median (center) with 25th to 75 th percentile and whiskers represent maxima/minima. **** p value ≤0.0001 and ***** p value <0.000001.

Extended data figure 7



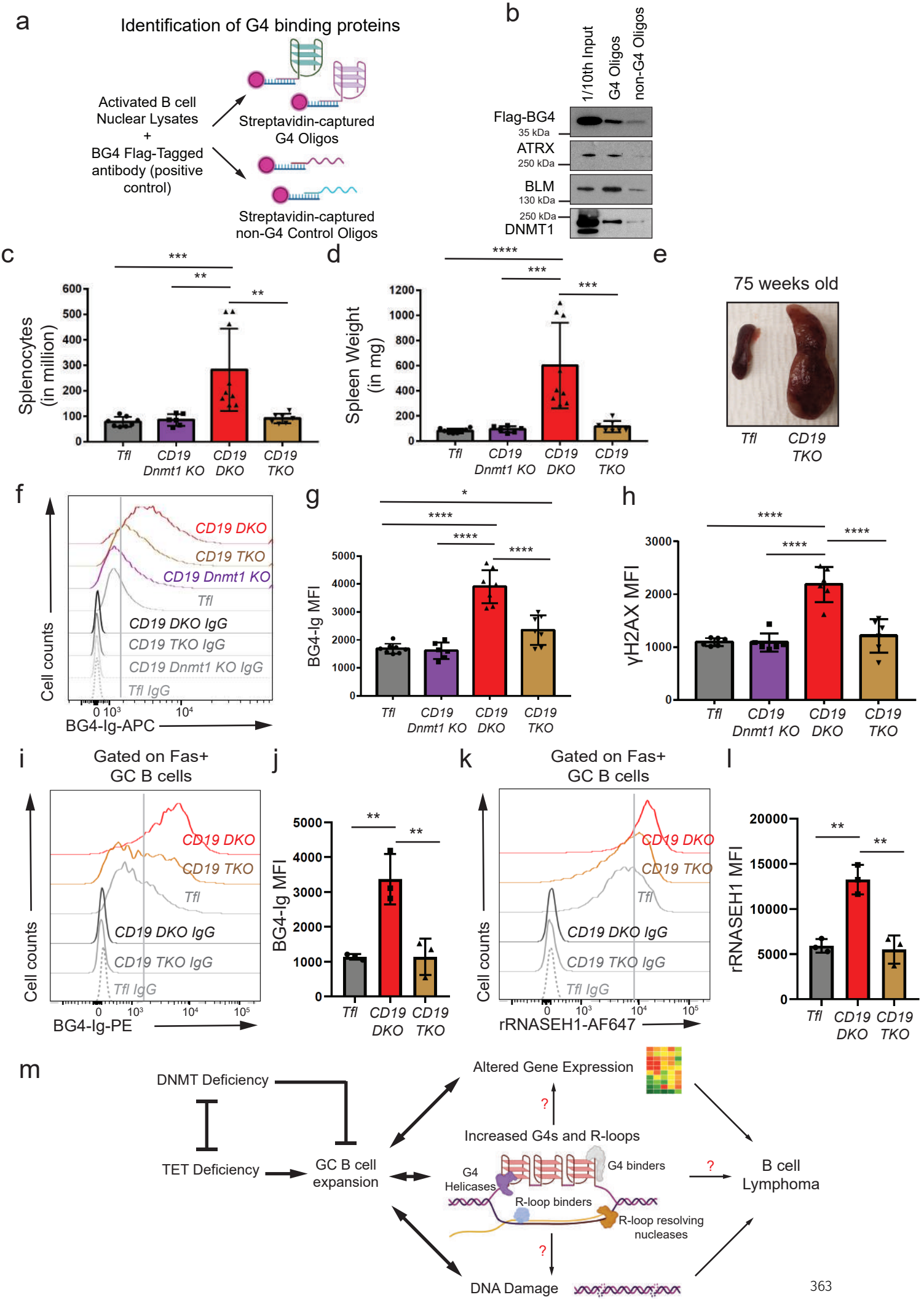






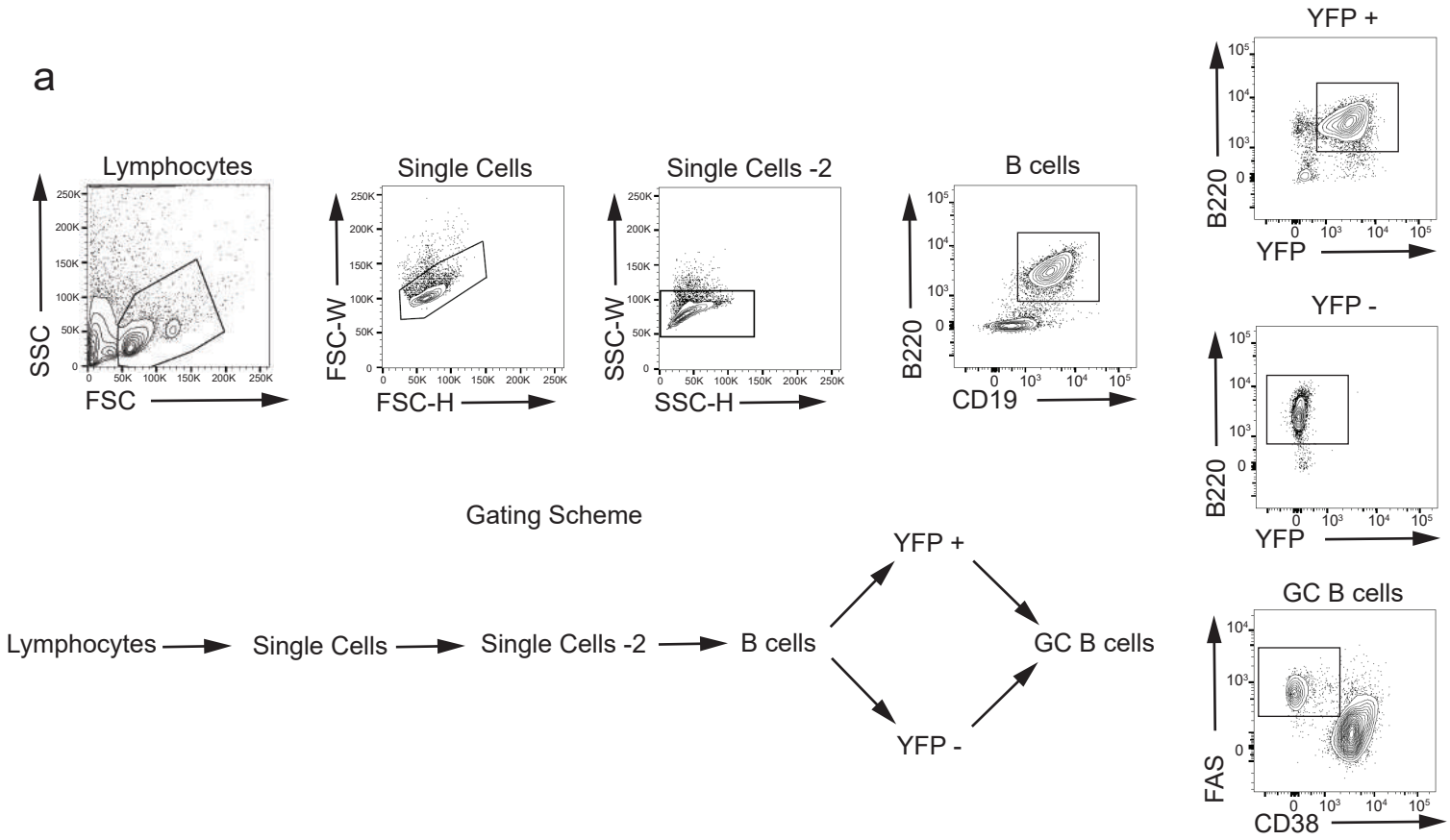
Extended data figure 7. Genome-wide analysis of TET deficient B cells. a)-c) Genome browser tracks showing the distribution of G-quadruplexes (G4), R-loops, RNA-Seq, WGBS, and 5hmC datasets for *Dfl* (blue tracks) and *CD19 DKO* (red tracks) B cells. Grey boxes indicate regions of interest. The blue arrows at the bottom show the location of the TSS and the direction of transcription. **d)** Circos plots to visually depict all translocations identified by HTGTS in *Dfl* and *CD19 DKO* replicates. Colored lines connect the Sμ bait with the translocation partner regions. Color scale represents the number of translocation partner regions identified in a 10 kb window. Translocations from two *Dfl* and *CD19 DKO* replicates are represented separately. **e)** Relative representation of different classes of motifs predicted to form G-quadruplexes (pG4) in control genomic regions selected randomly (*right bar*) and +/-300 bp from the center of translocation partner junctions identified from translocations in *Dfl* and *CD19 DKO* B cells (*right bar*). The numbers (n) of *Dfl* and *CD19* DKO hits, and control regions are included in the plots. **f)** Density of AID motifs (WRCY/RGYW) in +/-300bp from the center of translocation partner junctions identified from translocations in *Dfl* and *CD19 DKO* B cells compared to control random regions in euchromatin (Hi-C A compartment) from 2 biological replicates. Statistical significance is calculated using the Wilcoxon signed rank test **f**), ***** p value <0.000001.

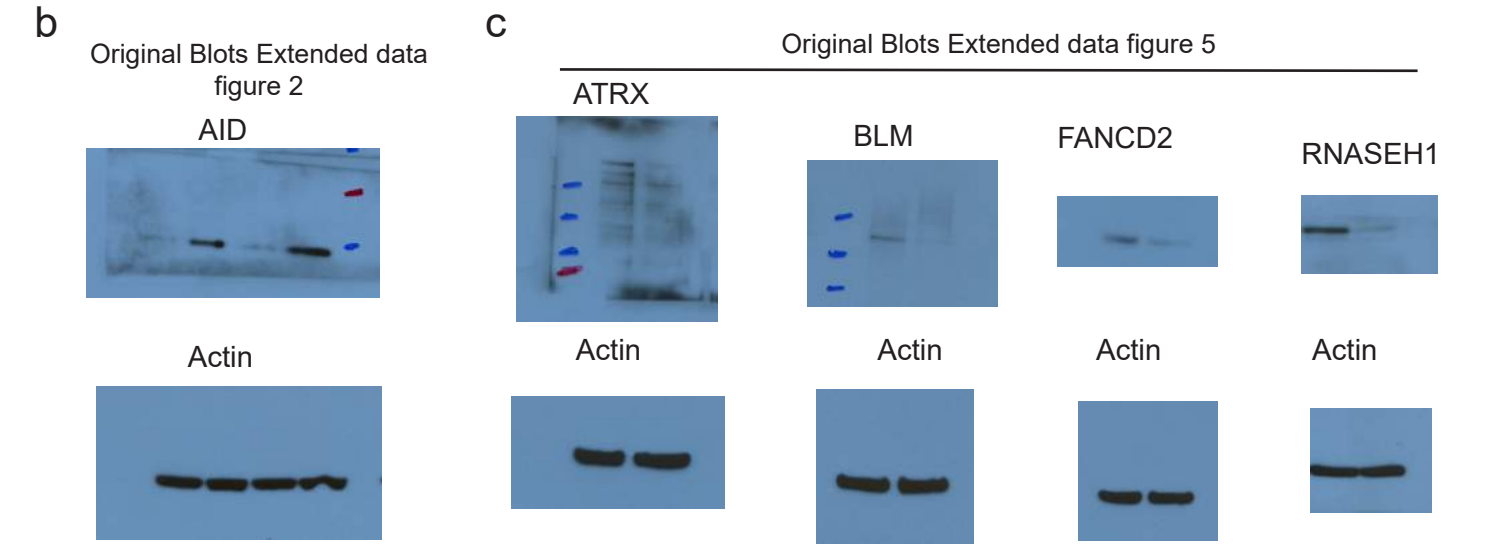
Extended data figure 8

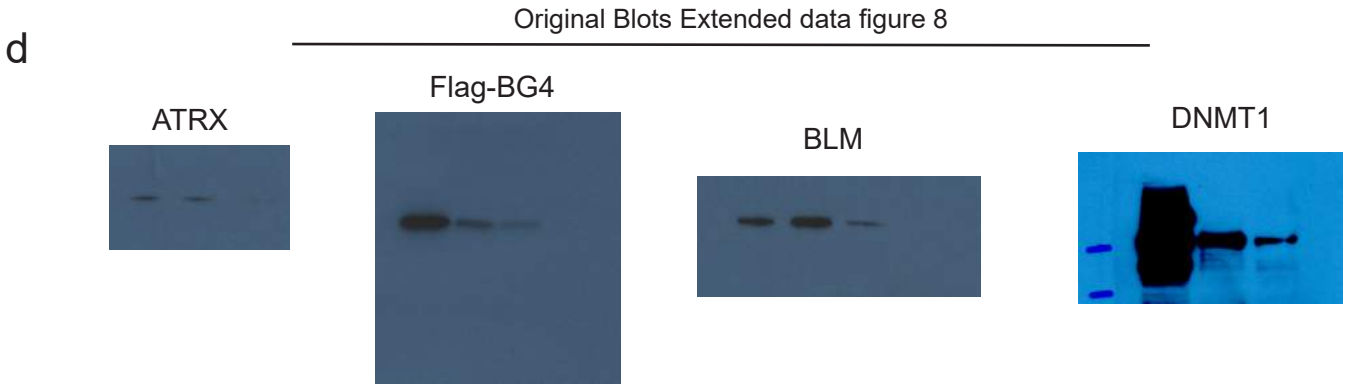


Extended data figure 8. DNMT1 deletion delays oncogenesis in TET-deficient mice. a) Diagrammatic representation of the strategy used to confirm G-quadruplex binding. Nuclear lysates of activated B cells were incubated with biotin-conjugated single stranded G4- or non-G4-forming control oligonucleotides (Oligos) captured using streptavidin beads. **b)** Immunoblots showing flag-tagged BG4 (positive control), ATRX, BLM and DNMT1 proteins. *Left lane*, 1/10th input from nuclear lysates (1/10th Input); *middle lane,* proteins pulled down with G4 forming oligonucleotides; and *right lane*, proteins pulled down with non-G4 control oligonucleotides. The data is representative of at least 2 independent experiments. **c)** - **d)** Quantification of **c)** cell numbers, **d)** spleen weights, of 10 week-old *Tfl* (grey), *CD19 Dnmt1 KO* (purple), *CD19 DKO* (red) and *CD19 TKO* (brown) mice from 5 independent experiments. **e)** Enlarged spleen of 75-week-old *CD19 TKO* mice compared with *Tfl* control mice. **f)** Flow cytometric detection of G-quadruplexes with BG4-Ig antibody or isotype IgG controls in B cells from *Tfl* (YFP^{-}, *grey*), *CD19 Dnmt1 KO* (YFP^{+}, *purple*), *CD19 DKO* (YFP^{+}, *red*) and *CD19 TKO* (YFP^{+}, *brown*) mice. **g) – h)** Quantification of median fluorescence intensity (MFI) of **g)** BG4-Ig signal, **h)** γH2AX signal from *Tfl* (YFP^{-}), *CD19 Dnmt1 KO* (YFP^{+}), *CD19 DKO* (YFP^{+}) and *CD19 TKO* (YFP^{+}) B cells from 4 independent experiments. **i), k)** Flow cytometric detection of **i**) G-quadruplexes with BG4-Ig antibody or isotype IgG controls and **k**) R-loops using V5-epitope-tagged recombinant RNASE H1 (rRNASE H1) or IgG controls in GC B cells (Fas^{+}) from *Tfl* (YFP^{-}, *grey*), *CD19 DKO* (YFP^{+}, *red*) and *CD19 TKO* (YFP^{+}, *brown*) mice. **j), l)** Quantification of median fluorescence intensity (MFI) of **j)** BG4-Ig signal and **l)** R-loops (rRNASE H1) signal in GC B cells (Fas^{+}) from *Tfl* (YFP^{-}, *grey*), *CD19 DKO* (YFP^{+}, *red*) and *CD19 TKO* (YFP^{+}, *brown*) mice from 3 independent experiments. **m)** Model proposing functional interplay between TET and DNMT activities to limit GC B cell expansion. TET deficiency in B cells leads to increased G4 and R-loop structures and is associated with altered gene expression, DNA damage and development of B cell lymphoma. Statistical significance is calculated using one-way ANOVA in **c), d), g), h), j)** and **l)**. Error bars represent mean +/- standard deviation, * p value ≤0.05, ** p value ≤0.01, *** p value ≤0.0005 **** p value <0.0001.

Extended data figure 9







Extended data figure 9. FACS gating strategy and original blots. a) Sequential gating strategy used for the flow cytometry analysis. The respective gate names are mentioned in the corresponding figures. **b**)-**e**) scanned immunoblots for extended data figures 2 (**b**), 5 (**c**) and 8 (**d**).

CORONAVIRUS

SARS-CoV-2 infection generates tissue-localized immunological memory in humans

Maya M.L. Poon[1,2]†, Ksenia Rybkina[1]†, Yu Kato[3]†, Masaru Kubota[4], Rei Matsumoto[4], Nathaniel I. Bloom[3], Zeli Zhang[3], Kathryn M. Hastie[3], Alba Grifoni[3], Daniela Weiskopf[3], Steven B. Wells[5], Basak B. Ural[1], Nora Lam[1,6], Peter A. Szabo[1], Pranay Dogra[1], Yoon S. Lee[1], Joshua I. Gray[1], Marissa C. Bradley[7], Maigan A. Brusko[8], Todd M. Brusko[8], Erica O. Saphire[3], Thomas J. Connors[7], Alessandro Sette[3,9], Shane Crotty[3,9]*, Donna L. Farber[1,4]*

Adaptive immune responses to SARS-CoV-2 infection have been extensively characterized in blood; however, most functions of protective immunity must be accomplished in tissues. Here, we report from examination of SARS-CoV-2 seropositive organ donors (ages 10 to 74) that CD4$^+$ T, CD8$^+$ T, and B cell memory generated in response to infection is present in the bone marrow, spleen, lung, and multiple lymph nodes (LNs) for up to 6 months after infection. Lungs and lung-associated LNs were the most prevalent sites for SARS-CoV-2–specific memory T and B cells with significant correlations between circulating and tissue-resident memory T and B cells in all sites. We further identified SARS-CoV-2–specific germinal centers in the lung-associated LNs up to 6 months after infection. SARS-CoV-2–specific follicular helper T cells were also abundant in lung-associated LNs and lungs. Together, the results indicate local tissue coordination of cellular and humoral immune memory against SARS-CoV-2 for site-specific protection against future infectious challenges.

INTRODUCTION

Ending the global coronavirus disease 2019 (COVID-19) pandemic caused by the novel coronavirus SARS-CoV-2 (severe acute respiratory syndrome coronavirus 2) depends on the establishment of immunological memory. SARS-CoV-2 infects the respiratory tract and induces adaptive immune responses, resulting in virus-specific T and B lymphocytes mediating viral clearance at the infection site and inhibiting viral dissemination through T cell effector functions and antibodies. It is now well documented that mild and severe infection generates circulating virus-specific T cells and antibodies detectable in peripheral blood for up to a year or more (*1*–*8*). Moreover, the presence of neutralizing antibodies specific for the viral spike (S) protein correlates with protection for SARS-CoV-2 vaccines (*9*, *10*). However, the emergence of viral variants with potential for immune evasion (*11*–*15*) and variability in vaccination rates among populations enable ongoing SARS-CoV-2 spread. An understanding of the breadth and functional potential of virus-specific T and B cell memory is needed for developing improved strategies to protect against continually evolving strains.

A major limitation in studying human immune responses is that sampling is largely confined to peripheral blood, whereas adaptive immune responses are generated and carry out their protective functions in a range of tissues. Memory cells are also maintained in diverse tissues—including infection sites and lymphoid organs [reviewed in (*16*, *17*)]. Virus-specific memory CD4$^+$ and CD8$^+$ T cells comprise heterogeneous subsets of circulating subsets and noncirculating tissue-resident memory (T_{RM}) cells in various sites (*18*, *19*). In mouse models, T_{RM} cells in the lung mediate optimal protective responses to respiratory infection (*20*–*22*), and this localized protection also involves responses in lung-associated lymph nodes (LNs) (*23*, *24*). In humans, the majority of T cells in adults are memory cells. The subset composition of human T cell memory is specific to the tissue site; in mucosal, exocrine, and barrier sites, T_{RM} cells predominate, whereas lymphoid sites contain circulating effector and central memory (T_{EM} and T_{CM}) along with T_{RM} subsets (*19*, *25*, *26*). All types of memory T cells in tissues express gene expression signatures distinct from memory T cells in peripheral blood (*27*), suggesting that they are differentially maintained compared with circulation. In SARS-CoV-2 infection, activated T_{EM} and T_{RM} were identified in the airways of severe COVID-19 (*28*, *29*), although the distribution and functional capacity of SARS-CoV-2–specific memory T cells across tissues remain uncharacterized.

The generation of memory B cells following an infection occurs in secondary lymphoid organs (LNs and spleen) and requires virus-specific follicular helper T (T_{FH}) cells, which promote B cell differentiation, survival, and somatic hypermutation in germinal centers (GCs) (*30*). Memory B cells can persist in multiple sites and exhibit tissue-resident phenotypes (*31*). In mouse respiratory infection models, resident memory B (B_{RM}) cells in lung and lung-associated LNs can be important for protection (*32*). In human lymphoid and mucosal sites, memory B cells are the predominant subset, whereas naïve B cells prevail in circulation (*33*). SARS-CoV-2 infection elicits generation of S-, receptor binding domain (RBD)–, and nucleocapsid (N)-specific memory B cells detectable in peripheral blood (*1*, *3*, *7*); however, the distribution and maintenance of SARS-CoV-2–specific memory B cells and GC B cells in tissues following natural infection

[1]Department of Microbiology and Immunology, Columbia University Irving Medical Center, New York, NY 10032, USA. [2]Medical Scientist Training Program, Columbia University Irving Medical Center, New York, NY 10032, USA. [3]Center of Infectious Disease and Vaccine Research, La Jolla Institute for Immunology, La Jolla, CA 92037, USA. [4]Department of Surgery, Columbia University Irving Medical Center, New York, NY 10032, USA. [5]Department of Systems Biology, Columbia University Irving Medical Center, New York, NY 10032, USA. [6]Department of Pathology and Cell Biology, Columbia University Irving Medical Center, New York, NY 10032, USA. [7]Department of Pediatrics, Columbia University Irving Medical Center, New York, NY 10032, USA. [8]Department of Pathology, Immunology, and Laboratory Medicine, University of Florida, Gainesville, FL 32611, USA. [9]Division of Infectious Diseases and Global Public Health, Department of Medicine, University of California, San Diego, La Jolla, CA 92037, USA.
*Corresponding author. Email: df2396@cumc.columbia.edu (D.L.F.); shane@lji.org (S.C.)
†These authors contributed equally to this work.

have not been reported. Moreover, the relationship between human B and T cell memory in tissues is largely unexplored.

The use of physiologically healthy tissues from organ donors has enabled study of human immune cells across multiple sites (*16*, *34*, *35*). Investigating tissue immunity to SARS-CoV-2 is particularly challenging, as previously infected, but unvaccinated, donors are required. Here, we present an investigation of SARS-CoV-2–specific memory T and B cell populations in lymphoid and mucosal sites of previously infected, seropositive organ donors, which we identified in the bone marrow (BM), spleen, lung, and LNs up to 6 months after infection. Lung and lung-associated LNs were the most prevalent sites for SARS-CoV-2–specific memory T and B cells, with a proportion exhibiting tissue-resident profiles. We also detected virus-specific GC B cells in lung-associated LNs along with T_{FH}, suggesting ongoing generation of humoral immunity. Together, our results reveal local coordination of cellular and humoral memory immune responses for site-specific protective immunity.

RESULTS

Organ donor cohorts for analysis of SARS-CoV-2–specific immune responses in tissues

We have established a human tissue resource for obtaining multiple tissues from organ donors through collaborations with organ procurement organizations (OPOs) (*16*, *34*, *35*). Use of organ donor tissue allows for rapid isolation of live immune cells for functional analysis, thus enabling assessment of immune responses in multiple sites within an individual. We identified four organ donors ages 10 to 74 years with previous SARS-CoV-2 infection (Fig. 1A) who died of noninfectious-related causes and were SARS-CoV-2 polymerase chain reaction (PCR) negative at the time of organ procurement. Previous SARS-CoV-2 infection history was based on post-procurement testing for antibodies to N protein (see Materials and Methods) and/or a confirmed history of COVID-19 2 to 6 months previously (table S1). Controls were prepandemic seronegative organ donors procured before November 2019 who also died of noninfectious causes (table S1).

All seropositive donors had detectable serum immunoglobulin G (IgG) to N, S, and RBD, along with SARS-CoV-2–specific neutralizing antibodies, consistent with antibody responses generated from acute infection (*36*, *37*), whereas serum from prepandemic organ donors lacked antibodies to SARS-CoV-2 (Fig. 1, B and C). Using this cohort, we examined SARS-CoV-2–specific T and B cell responses across the blood, BM, spleen, lung, lung-associated LNs, as well as gut-associated LNs of seropositive and seronegative donors.

SARS-CoV-2–specific T cells in lung and lymphoid tissues

SARS-CoV-2–specific T cells in different sites from seropositive and seronegative organ donors were measured based on expression of T cell receptor–dependent activation-induced markers (AIMs) after stimulation with SARS-CoV-2–specific peptide megapools (MPs), which enable simultaneous presentation of a large number of virus-specific epitopes (see fig. S1 and Materials and Methods) (*2*, *38*, *39*). Mononuclear cells from each site were stimulated for 24 hours in vitro with four different SARS-CoV-2–specific MPs: MP_S (containing overlapping peptides representing the entirety of S), MP_CD4_R [containing predicted human leukocyte antigen (HLA) class II viral epitopes minus S], and MP_CD8_A and MP_CD8_B (each containing predicted HLA class I epitopes from all viral proteins) (*2*). SARS-CoV-2–specific $CD4^+$T cells were identified on the basis of coexpression of two or more of the following three AIMs [OX40, 4-1BB, and/or CD40 ligand (CD40L)] (Fig. 1D and fig. S2A), whereas virus-reactive $CD8^+$ T cells were identified on the basis of coexpression of CD25 and 4-1BB (Fig. 1E and fig. S2B). Quantification of SARS-CoV-2–specific $CD4^+$ and $CD8^+$ T cell responses was based on dimethyl sulfoxide (DMSO) background–subtracted frequencies of AIM^+ $CD4^+$ and $CD8^+$ T cells (Fig. 1, F and H).

For $CD4^+$ T cells, significant responses to S protein were found in all sites examined (blood, BM, spleen, lung, lung-associated LNs, and gut-associated LNs) relative to prepandemic control samples (Fig. 1F, left). For non–S SARS-CoV-2 epitopes, there were significant $CD4^+$ T cell frequencies in the BM, lung-associated LNs, and gut-associated LNs (Fig. 1F, right). Total SARS-CoV-2–specific $CD4^+$ T cell responses largely reflected the pattern observed with S-specific responses (Fig. 1, F and G). SARS-CoV-2–specific $CD8^+$ T cell frequencies were generally lower in magnitude than for $CD4^+$ T cells and more variable between donors (Fig. 1, G to I). Significant SARS-CoV-2–specific $CD8^+$ T cell frequencies above controls were observed for lung- and gut-associated LNs for class I epitopes (Fig. 1, H and I). Comparing across all donors and sites, $CD4^+$ T cells responding to S protein epitopes were the predominant SARS-CoV-2–specific T cells across tissue sites for all individuals (Fig. 1, F and G, and fig. S2C). Although the distribution patterns of SARS-CoV-2–specific T cell responses across tissue sites varied between donors (Fig. 1 and fig. S3), the lung and lung-associated LNs were most consistently the dominant sites for virus-specific $CD4^+$ and $CD8^+$ T cells (fig. S2D). These results indicate that SARS-CoV-2 infection generates virus-specific T cell responses across blood, multiple lymphoid sites, and lungs, with higher frequencies localized in lung tissue and lung-associated LNs.

SARS-CoV-2–specific T cells are maintained in tissues as circulating and resident memory subsets

We analyzed the subset distribution of SARS-CoV-2–specific T cells based on coordinate expression of CD45RA and CCR7, defining T_{CM} cells ($CD45RA^-CCR7^+$), T_{EM} cells ($CD45RA^-CCR7^-$), terminally differentiated effector T (T_{EMRA}) cells ($CD45RA^+CCR7^-$), and naïve or stem-like memory cells ($CD45RA^+CCR7^+$) (*40*, *41*). Each tissue had a distinct T cell subset composition that is conserved between individuals as we previously determined (*26*, *35*, *42*); T cell phenotypes for each site and for SARS-CoV-2–specific T cells are shown in representative flow cytometry plots (Fig. 2A). The majority of SARS-CoV-2–specific $CD4^+$ T cells were maintained as T_{EM} ($\geq$75%) in the blood and lung and as T_{EM} or T_{CM} ($\geq$80%) in lymphoid sites (Fig. 2B). For SARS-CoV-2–specific $CD8^+$ T cells, the majority were maintained as T_{EM} and T_{EMRA} cells ($\geq$50%) for all sites; the proportion of T_{EMRA} was higher than T_{EM} for the BM, spleen, and lung, whereas T_{EM} were more prevalent than T_{EMRA} for LNs (Fig. 2C). Donors D495 and D498 had a particularly high proportion of SARS-CoV-2–specific $CD8^+$ T_{EMRA} cells in the lung, whereas the pediatric donor (HDL113) harbored more SARS-CoV-2–specific $CD45RA^+CCR7^+$ $CD8^+$ T cells in the lung and lung-associated LN (Fig. 2C).

We also analyzed coexpression of residency markers CD69 and CD103 by SARS-CoV-2–specific T cells (see fig. S1 for gating), as assessment of CD69 alone as a T_{RM} marker was confounded because of CD69 up-regulation by T cell receptor–stimulated T cells. Virus-specific $CD69^+CD103^+$ memory $CD4^+$ and $CD8^+$ T cells (T_{RM}) were mostly confined to the lung, whereas lower frequencies of SARS-CoV-2–specific $CD8^+$ T_{RM} were also detected in LNs (Fig. 2, D and E).

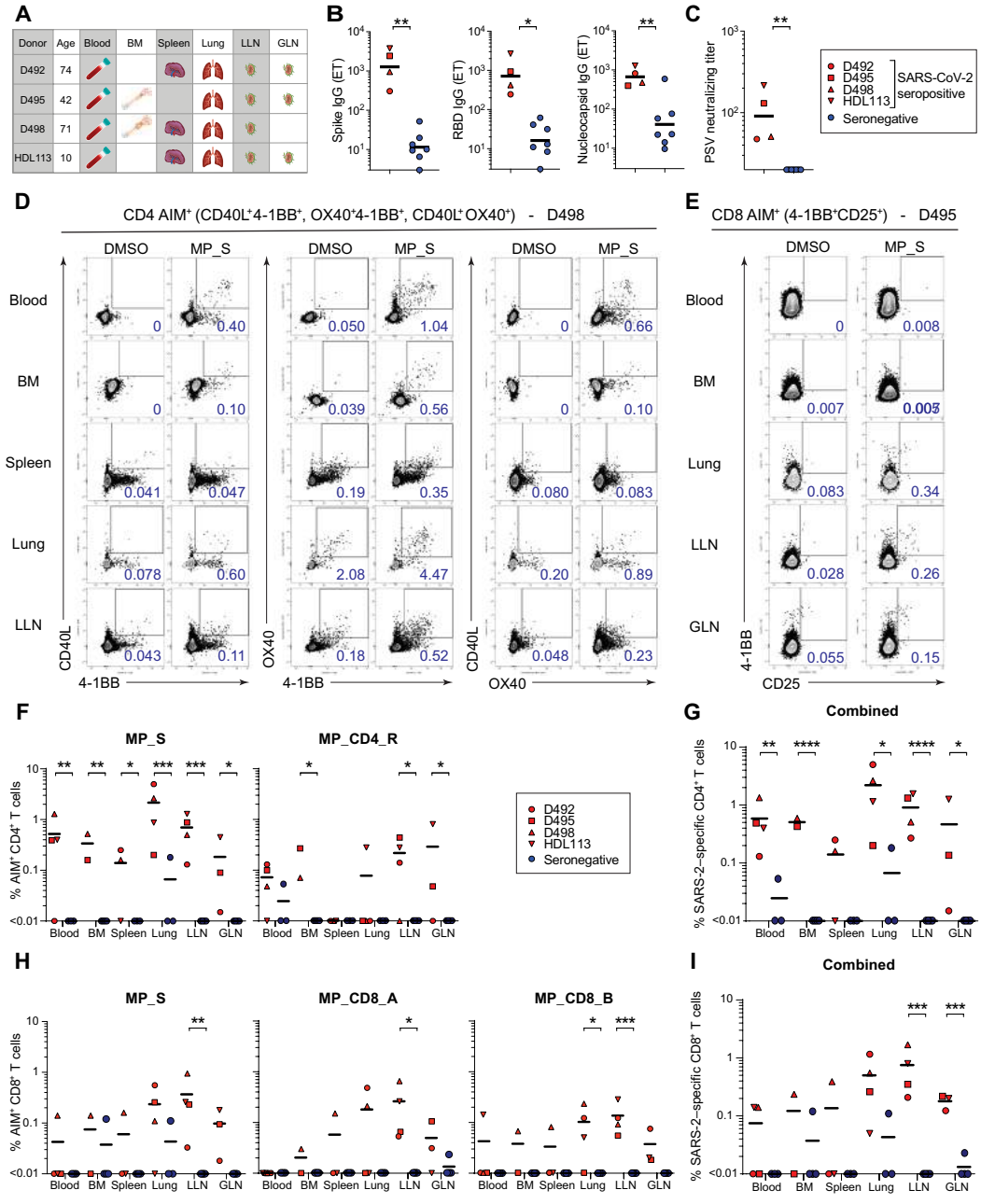


Fig. 1. SARS-CoV-2–specific CD4$^+$ and CD8$^+$ T cells in blood and tissues of previously infected organ donors. (**A**) SARS-CoV-2 seropositive donors and tissues used from each donor for this study. (**B**) Anti–SARS-CoV-2 antibody reactivities for seropositive and seronegative donors. Graphs show ET of IgG specific for SARS-CoV-2 Spike, RBD, and nucleocapsid. (**C**) SARS-CoV-2 Spike PSV neutralizing titers for seropositive and seronegative donors. Serology statistical analyses were performed using Mann-Whitney *U* test. *$P \leq 0.05$; **$P \leq 0.01$; ***$P \leq 0.001$; ****$P \leq 0.0001$. (**D**) Identification of SARS-CoV-2–specific CD4$^+$ T cells using the AIM assay. Mononuclear cells isolated from blood, BM, spleen, lung, and lung-associated lymph node (LLN) were stimulated with SARS-CoV-2 peptide pools (see Materials and Methods) and responding CD4$^+$ T cells were identified based on induction of OX40, 4-1BB, and CD40L as shown in representative flow cytometry plots from D498 reactive to MP_S. SARS-CoV-2–specific CD4$^+$ T cells were defined based on combined gates CD40L$^+$4-1BB$^+$ (left), OX40$^+$4-1BB$^+$ (middle), or CD40L$^+$OX40$^+$ (right) of total CD4$^+$ T cells for each stimulation condition and tissue site (see fig. S1 for gating strategy). (**E**) AIM assay for detection of SARS-CoV-2–specific CD8$^+$ T cells in blood, BM, lung, LLN, and gut-associated lymph node (GLN) showing induction of 4-1BB and CD25 in representative flow cytometry plots from D495 reactive to MP_S. AIM$^+$ CD8 T cells were defined based on frequency 4-1BB$^+$CD25$^+$ from total CD8$^+$ T cells for each stimulation condition and tissue site (see fig. S1 for gating strategy). (**F**) SARS-CoV-2 epitope-specific CD4$^+$T cells identified following stimulation with MP_S (left) and MP_CD4_R (right) peptide MPs from indicated tissues sites of seropositive and seronegative donors. (**G**) Total SARS-CoV-2–specific CD4$^+$ T cells in each site from individual donors based on responses to all epitopes. (**H**) SARS-CoV-2 epitope-specific CD8$^+$ T cells identified following stimulation with MP_S (left), MP_CD8_A (middle), and MP_CD8_B (right) peptide MPs from indicated sites of seropositive and seronegative donors. (**I**) Total SARS-CoV-2–specific CD8$^+$ T cells in each site from individual donors based on compiled responses to all epitopes. $n = 4$ SARS-CoV-2 seropositive donors ($n = 4$ for blood, lung, LLN; $n = 3$ for spleen and GLN; $n = 2$ for BM). $n = 10$ for seronegative donors ($n = 4$ for BM, LLN, and GLN; $n = 3$ for blood, spleen, and lung). Statistical analysis was performed using one-way analysis of variance (ANOVA), corrected for multiple comparisons by false discovery rate (FDR) using two-stage linear step-up procedure of Benjamini, Krieger, and Yekutieli. *$q \leq 0.05$, **$q \leq 0.01$, ***$q \leq 0.001$, and ****$q \leq 0.0001$. Datasets were log-transformed before statistical analysis.

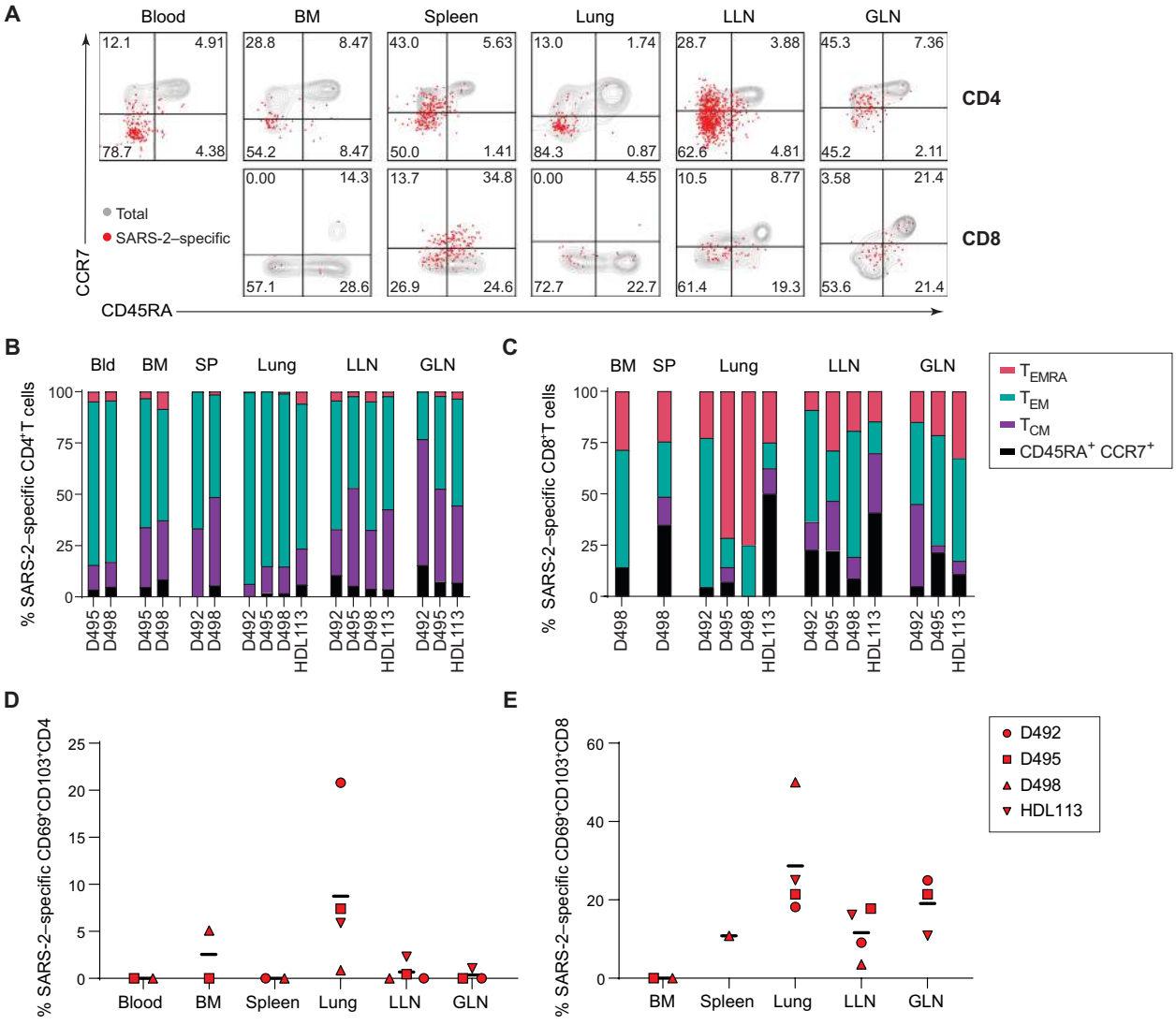


Fig. 2. SARS-CoV-2–specific T cells are maintained as memory subsets in diverse tissues of seropositive donors. (**A**) Subset phenotypes of total (gray contour) and SARS-CoV-2–specific (red dots) $CD4^+$ (top row) and $CD8^+$ (bottom row) T cells based on CD45RA and CCR7 expression ($CD45RA^+CCR7^+$; $CD45RA^-CCR7^+$, T_{CM}; $CD45RA^-CCR7^-$, T_{EM}; $CD45RA^+CCR7^-$, T_{EMRA}) shown in representative flow cytometry plots. (**B**) T cell subset delineation of SARS-CoV-2–specific $CD4^+$ T cells in blood and indicated tissues of seropositive organ donors. (**C**) T cell memory subset delineation of SARS-CoV-2–specific $CD8^+$ T cells in blood and indicated tissues of seropositive organ donors. (**D**) Expression of tissue residency markers CD69 and CD103 by SARS-CoV-2–specific $CD4^+$ T cells in indicated sites of seropositive donors. (**E**) Expression of tissue residency markers CD69 and CD103 by SARS-CoV-2–specific $CD8^+$ T cells in indicated sites of seropositive donors. Memory subset differentiation and residency marker analysis were conducted on tissue sites for which number of SARS-CoV-2–specific T cells was ≥5 based on AIM assays. SARS-CoV-2–specific $CD8^+$ T cells in the blood were not detected above this threshold. $n = 4$ SARS-CoV-2 seropositive donors ($n = 4$ for blood, lung, and LLN; $n = 3$ for spleen and GLN; $n = 2$ for BM). $n = 10$ for seronegative donors ($n = 4$ for BM, LLN, and GLN; $n = 3$ for blood, spleen, and lung). Statistical analysis was performed using one-way ANOVA, corrected for multiple comparisons by FDR using two-stage linear step-up procedure of Benjamini, Krieger, and Yekutieli. No comparisons were statistically significant. Datasets were log-transformed before statistical analysis. SARS-2, SARS-CoV-2; T_{EMRA}, terminally differentiated effector T cell; Bld, blood; SP, spleen.

Together, these results show that SARS-CoV-2–specific T cells are maintained across diverse tissue sites as memory T cells, with a portion of T cells persisting as tissue-resident populations, particularly within the lung.

Tissue specificity and heterogeneity of functional responses to SARS-CoV-2

The functional responses of SARS-CoV-2–responding cells from different sites were assessed by multiplex quantification of 50 immune mediators from culture supernatants of peptide-stimulated mononuclear cells as in Fig. 1. Blood and tissues exhibited distinct functional profiles, and the magnitude of responses varied between donors (Fig. 3A and fig. S4). There was heterogeneity between individuals in the distribution of functional responses across sites. In particular, SARS-CoV-2–specific functional responses were primarily located in the lung for donor D492 and in the lung-associated LN and blood for donor D495, whereas D498 and HDL113 exhibited a broad functional response across multiple sites (Fig. 3A). For both

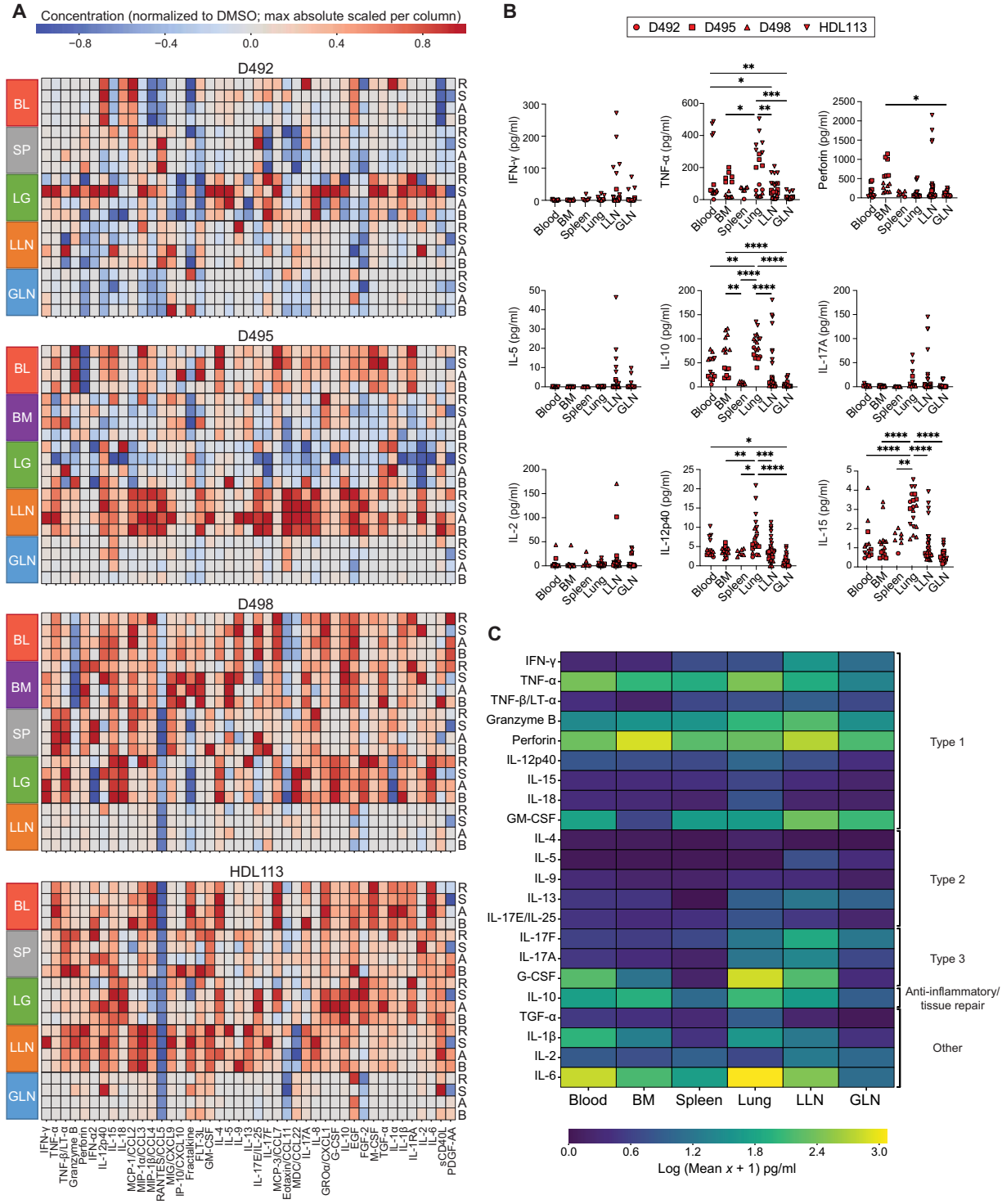


Fig. 3. Heterogeneity and tissue specificity of functional responses to SARS-CoV-2 epitopes. (**A**) Profiles of immune mediators produced for multiple tissue sites within SARS-CoV-2 seropositive donors following stimulation with peptide MPs MP_S (S), MP_CD4_R (R), MP_CD8_A (A), and MP_CD8_B (B), shown as a heatmap. The color intensity of each cell represents DMSO background–subtracted analyte concentration (max absolute scaled per column) within each donor (see Materials and Methods). (**B**) Concentration of indicated immune mediators measured within supernatants from in vitro stimulations of blood, BM, spleen, lung, LLN, and GLN mononuclear cells with SARS-CoV-2 MPs for which SARS-CoV-2–specific T cells were identified based on DMSO background–subtracted frequencies of AIM^+ $CD4^+$ and $CD8^+$ T cells. Statistical analysis was performed using one-way ANOVA, corrected for multiple comparisons by Tukey's multiple comparisons test. $^*P \leq 0.05$, $^{**}P \leq 0.01$, $^{***}P \leq 0.001$, and $^{****}P \leq 0.0001$. (**C**) Immune mediator milieu for each site. Heatmap showing log (Mean $x + 1$) pg/ml levels of immune mediators averaged across donors and stimulation conditions for each tissue site derived from samples for which significant frequencies of SARS-CoV-2–specific T cells were identified above background in Fig. 1. BL, blood; LG, lung.

elderly donors (D492 and D498), lung-associated LN responses were much weaker compared with other tissue sites, even when the frequency AIM^+ $CD4^+$ and $CD8^+$ T cells in the lung-associated LN were high relative to other tissue sites (Fig. 3A and fig. S3). Overall, these data show that virus-specific functional responses are differentially maintained across sites and individuals.

Tissue-specific functional profiles were apparent in the profile of soluble mediators produced in response to stimulation with SARS-CoV-2 peptide pools (Fig. 3, B and C). The functional responses in the LN were the most diverse and included type 1 proinflammatory cytokines and cytolytic mediators [interferon-γ (IFN-γ), tumor necrosis factor–α (TNF-α), granzyme B, perforin, and granulocyte-macrophage colony-stimulating factor (GM-CSF)], type 2 cytokines [interleukin-5 (IL-5), IL-9, and IL-13], and type 3 cytokines [IL-17A, IL-17F, and granulocyte colony-stimulating factor (G-CSF)]—most at increased levels compared with other sites (Fig. 3, B and C). The functional response in the lung was distinct from other sites and included proinflammatory profiles (TNF-α, perforin, granzyme, IL-12, IL-17A, and G-CSF), IL-10 associated with regulation of inflammation during respiratory infections (*43*), and higher levels of IL-6 and the homeostatic cytokine IL-15 compared with other sites (Fig. 3, B and C). BM responses were mostly TNF-α, perforin, granzyme B, and IL-10 and blood responses were similar to BM (Fig. 3, B and C). Comparing the immune mediator milieu between sites revealed that certain cytokines are produced across sites, whereas others are distinct to specific sites (Fig. 3C). Together, these results indicate that SARS-CoV-2–specific memory T cells in different sites exhibit distinct functional responses to viruses that are likely adapted to the site and contribute to a multifaceted protective response.

SARS-CoV-2–specific memory B cells and resident phenotypes in tissues

To characterize the nature of memory B cell responses to SARS-CoV-2 in tissues, we used fluorescently labeled, biotinylated, and multimerized probes of full-length S and RBD proteins to detect antigen-binding B cells among IgM^+, IgG^+, or IgA^+ memory B cells (fig. S5A), as previously described (*1*). SARS-CoV-2–specific memory B cells were detected at frequencies substantially higher in seropositive donors than seronegative donors in all tissues examined, including the lungs, BM, spleen, lung-associated LNs, and gut-associated LNs (Fig. 4, A and B). IgG^+ was the dominant isotype of SARS-CoV-2–specific memory B cells in almost all samples, although IgM^+ and IgA^+ memory B cells were present (Fig. 4C). For D498, whereas few SARS-CoV-2–specific memory B cells were IgM^+ in the LNs, >40% were IgM^+ in the BM and spleen. For D495, whereas almost no IgA^+ SARS-CoV-2–specific memory B cells were found in the lung, ~25% were IgA^+ in gut-associated LNs (Fig. 4C). SARS-CoV-2–specific memory B cells were present at significantly higher frequencies in lung and lung-associated LNs than in the spleen or gut-associated LNs (Fig. 4, B and D). The highest proportion of SARS-CoV-2–specific IgG^+ memory B cells was found in the lung and lung-associated LNs (Fig. 4D).

Memory B cells can persist as tissue-resident cells (B_{RM}) in lymphoid or nonlymphoid tissues and are identified by CD69 expression (*32*, *33*, *44*). In tissues of seropositive and seronegative donors, we detected significant populations (>50 to 75%) of $CD69^+$ B cells comprising >50% of total B cells in the lungs and LNs, whereas the BM and spleen contained much lower frequencies (0 to 10%) of CD69-expressing B cells (fig. S5, B and C). Substantial frequencies of SARS-CoV-2–specific memory B cells exhibited CD69 expression indicative of tissue-resident profiles in lungs (50 to 80%) and LNs (20 to 40%) (Fig. 4, E and F). By contrast, negligible frequencies (<3%) of $CD69^+$ SARS-CoV-2 S/RBD–specific memory B cells were detected in the BM of organ donors (Fig. 4, E and F), and in previously obtained peripheral blood samples from convalescent individuals with COVID-19 (fig. S5, D to I), consistent with previous analysis of polyclonal B cells in these sites (*33*). Together, these results provide direct evidence for human antigen-specific B_{RM} in lungs and LNs that are distinct from corresponding populations in the blood. Thus, SARS-CoV-2 infection leads to the preferential formation and/or retention of antigen-specific B cell memory within lungs and lung-associated LNs, with $CD69^+$ tissue-resident cells representing the majority of the SARS-CoV-2 S–specific memory B cells in lungs.

SARS-CoV-2–specific GC B cells and T_{FH} cells in LNs

GCs within lymphoid organs are important microanatomical sites in which activated B cells receive cognate help from T_{FH} to undergo somatic hypermutation to evolve higher affinity antibody recognition of pathogens (*45*). While previous studies have demonstrated affinity-matured SARS-CoV-2–specific memory B cells in the blood (*3*), BM plasma cells (*46*), and circulating virus-specific T_{FH} cells (*47*), which all indicate GC responses in COVID-19 (*48*), direct evidence of SARS-CoV-2 antigen-specific GCs (and S-specific GC B cells, in particular) is lacking. We identified GC B cells by assessing coexpression of Bcl6 [a transcription factor required for GC B cell differentiation (*49*)] and Ki67 (a marker of active cellular proliferation) among total $CD19^+$ B cells (fig. S5A). The frequency of GC B cells ($Bcl6^+Ki67^+CD19^+$ B cells) trended higher in lung-associated LNs of seropositive donors than seronegative donors ($q = 0.053$; Fig. 5, A and B). SARS-CoV-2–specific GC B cells identified based on binding to S and RBD proteins were identified in lung-associated LNs of three of four seropositive donors and in gut-associated LNs of one donor (Fig. 5, C and D, and fig. S5J). Virus-specific GC B cells were not detected in the BM, spleen, or lung of seropositive donors or in any tissues of seronegative donors (Fig. 5, C and D). These results provide direct evidence that SARS-CoV-2–specific GC responses are induced by SARS-CoV-2 infection and are maintained in lung-associated LNs after resolution of infection. In addition, long-lasting GCs can even be generated in human gut-associated LNs.

T_{FH} cells can be identified by coexpression of CXCR5 and PD-1 (*50*). T_{FH}-phenotype cells ($CXCR5^+PD\text{-}1^+CD4^+$ T cells) were found at low frequencies (5 to 10%) in lymphoid sites (spleen, lung-associated LNs, and gut-associated LNs) and in even lower frequencies in BM and most lungs of seronegative and seropositive donors (fig. S6), demonstrating the relative rarity of this population among total T cells. However, SARS-CoV-2–specific T_{FH} cells (identified by AIMs as in Fig. 1) were found in multiple sites of seropositive donors comprising 20 to 50% of SARS-CoV-2–reactive $CD4^+$ T cells in LNs and lower but significant frequencies in the lung, spleen, and BM (Fig. 5, E to G). In one seropositive donor, more than 80% of virus-specific $CD4^+$ T cells in the lung were T_{FH}, representing ~4% of non-naïve (NN) SARS-CoV-2–specific $CD4^+$ T cells, consistent with the higher overall frequency of T_{FH} cells in the lungs of that donor (Fig. 5, F and G, and fig. S6). Together, the SARS-CoV-2–specific GC B cell and T_{FH} data indicated robust GC responses to SARS-CoV-2 infection, distributed among lymphoid tissues and sites of viral infection, with some GCs being active for months after infection.

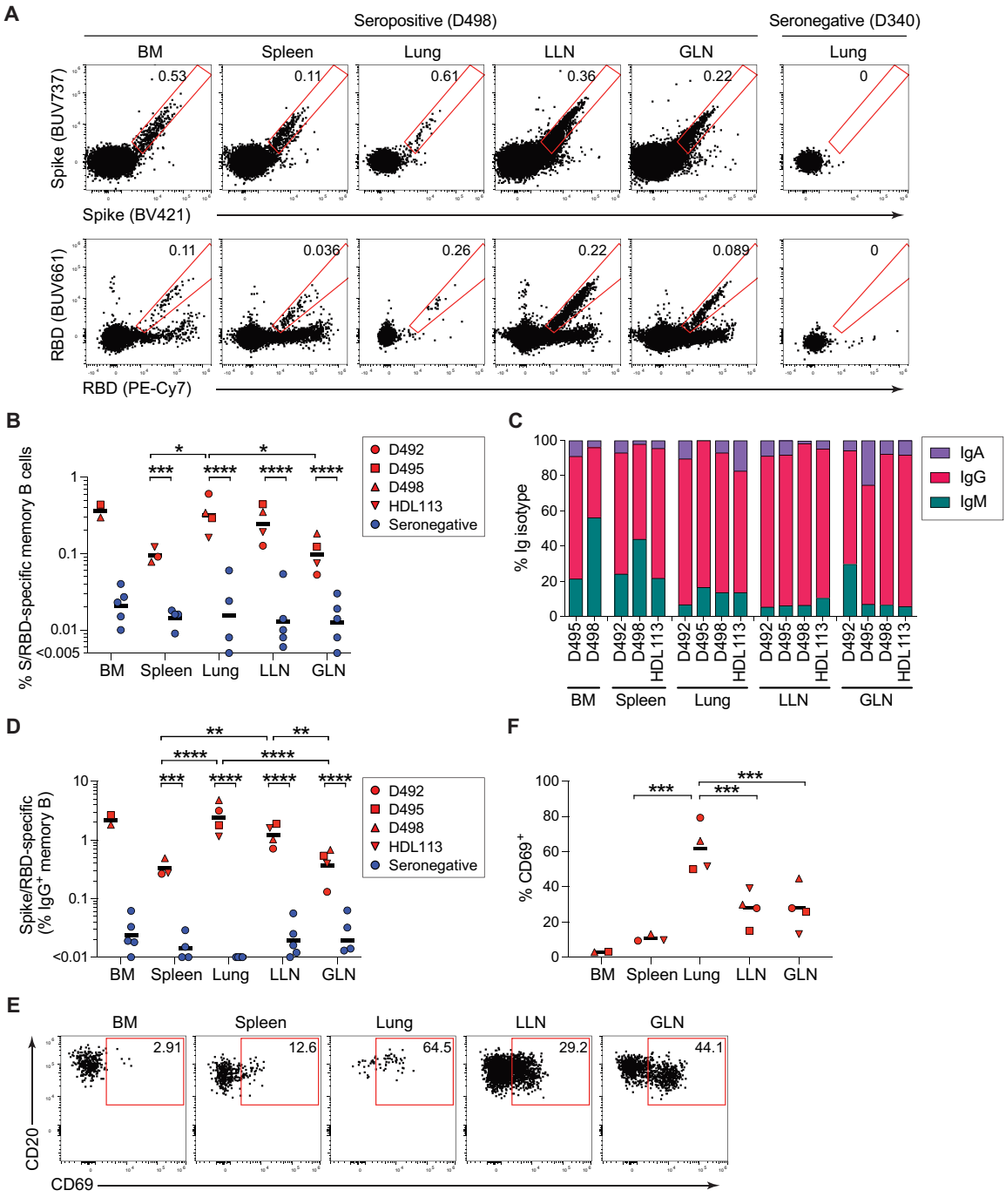


Fig. 4. SARS-CoV-2–specific memory B cells in tissues. (**A**) Representative flow cytometry plots showing staining patterns of probes for SARS-CoV-2 Spike (top) and RBD (bottom) on memory B cells, defined here as $CD19^+CD20^+IgD^-$ non-GC B cells (see fig. S5A for gating). Memory B cells in BM, spleen, lung, LLN, and GLN from a SARS-CoV-2 Spike seropositive donor (D498) and memory B cells in lung from a seronegative donor (D340). Percentages are indicated. (**B**) SARS-CoV-2–specific memory B cells in tissues. Graph shows frequency of memory B cells specific to Spike and/or RBD (S/RBD) in indicated sites expressed as a percentage of $CD19^+CD20^+$ total B cells. (**C**) Fraction of SARS-CoV-2 S/RBD–specific memory B cells that belong to indicated Ig isotypes. (**D**) Frequency (percentage) of IgG^+ memory B cells that are specific to S/RBD. (**E**) Representative flow cytometry plots showing CD69 expression on memory B cells specific to S/RBD. Percentages are indicated. (**F**) Frequency (percentage) of SARS-CoV-2 S/RBD–specific memory B cells that are $CD69^+$. $n = 4$ Seropositive donors ($n = 4$ for lung, LLN, and GLN; $n = 3$ for spleen; $n = 2$ for BM). $n = 7$ Seronegative donors ($n = 4$ for lung and spleen; $n = 5$ for LLN, GLN, and BM). Statistical analysis was performed using one-way ANOVA corrected for multiple comparisons by FDR using two-stage linear step-up procedure of Benjamini, Krieger, and Yekutieli. *$q \leq 0.05$, **$q \leq 0.01$, ***$q \leq 0.001$, and ****$q \leq 0.0001$. For (B) and (D), datasets were log-transformed before statistical analysis.

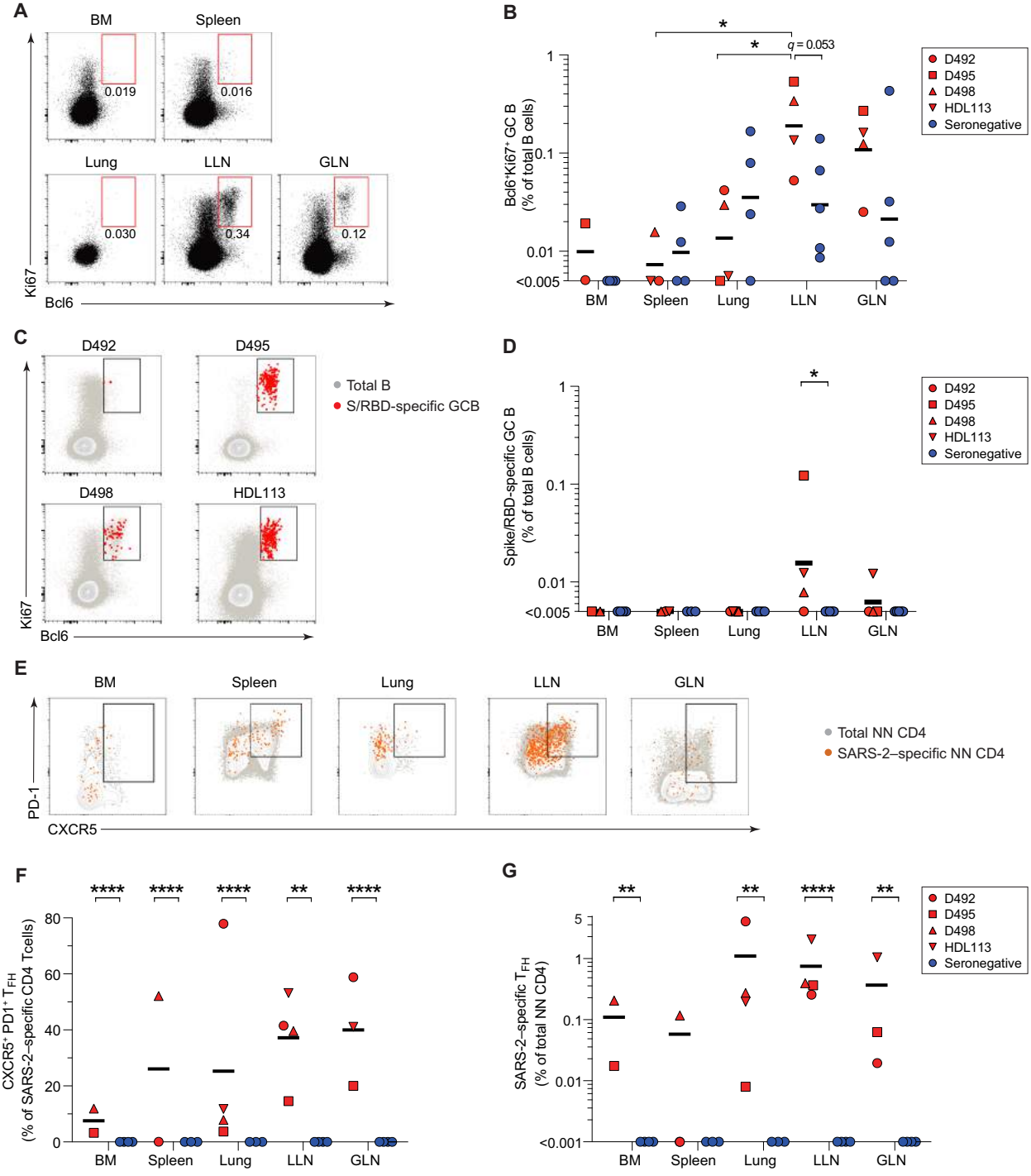


Fig. 5. SARS-CoV-2–specific GC B cells and follicular helper T cells in seropositive donors. (**A**) $Bcl6^+Ki67^+$ GC B cells in different tissues, shown in representative flow cytometry plots from seropositive donor D498 (see fig. S5 for gating). (**B**) Frequency of $Bcl6^+Ki67^+$ GC B cells in tissue sites of seropositive and seronegative donors expressed as a percentage of total $CD19^+CD20^+$ total B cells. (**C**) GC B cells in the LLN for each donor shown in representative flow cytometry plots. GC B cells specific to S/RBD are depicted in red. (**D**) Frequency of S/RBD-specific GC B cells as a percentage of total B cells. (**E**) T_{FH} phenotype and frequency among NN $CD4^+$ T cells per tissue as depicted by the rectangle gate shown in representative flow cytometry plots from seropositive donor D498 (lung, LLN, spleen, and BM) and D495 (GLN); see Materials and Methods for gating strategy. SARS-CoV-2–specific NN $CD4^+$ T cells are highlighted in orange. (**F**) Frequency of $CXCR5^+PD\text{-}1^+$ T_{FH} cells per tissue as percentage of SARS-CoV-2–specific $CD4^+$ T cells. (**G**) Frequency of SARS-CoV-2–specific T_{FH} cells per tissue as percentage of total NN $CD4^+$ T cells. Statistical analysis was performed on log-transformed datasets using one-way ANOVA, corrected for multiple comparisons by FDR using two-stage linear step-up procedure of Benjamini, Krieger, and Yekutieli. *$q \leq 0.05$, **$q \leq 0.01$; ****$q \leq 0.0001$. S/RBD, SARS-CoV-2 Spike and/or RBD; T_{FH}, Follicular helper T cell.

Coordinated adaptive immunity across tissues

The identification of SARS-CoV-2–specific memory T and B cells in multiple tissues, along with T_{FH} and GC B cells in LNs, suggested site-directed coordination of cellular and humoral immunity. To identify potential associations between SARS-CoV-2–specific lymphocyte populations across sites, we performed an exploratory correlation analysis (Fig. 6 and fig. S7). The frequencies of S-specific $CD4^+$ and $CD8^+$ T cells (but not total SARS-CoV-2–specific T cells; fig. S7A) were positively correlated across tissue sites ($P = 0.0116$; Fig. 6A). Significant associations also emerged between SARS-CoV-2–specific B and T cells (Fig. 6, B to F, and fig. S7, B to K). SARS-CoV-2–specific $CD4^+$ T cell frequencies correlated positively with S/RBD-specific memory B cells across all tissues ($P = 0.0009$), as well as IgG^+ ($P = 0.0009$) and IgA^+ ($P = 0.0203$) subpopulations (Fig. 6, B and C, and fig. S7, D and E). For tissue-resident SARS-CoV-2–specific lymphocytes, $CD69^+CD103^+$ $CD8^+$ T_{RM} cells correlated with $CD69^+$ B_{RM}—both as frequencies of total lymphocytes ($P = 0.0069$) and as fractions within their respective antigen-specific populations ($P = 0.01$) (Fig. 6E and fig. S7J). Positive correlations were also observed between $CD4^+$ T_{RM} and B_{RM} across all tissues ($P = 0.0198$; fig. S7I). For follicular responses, S/RBD-specific memory B cells correlated with SARS-CoV-2–specific T_{FH} cells across tissues ($P = 0.0236$; Fig. 6F and fig. S7K).

Given that lung and lung-associated LNs contained the highest frequencies of SARS-CoV-2–specific lymphocytes, we performed a targeted correlation matrix analysis to identify potential associations within and between these sites (Fig. 6G). SARS-CoV-2–specific T_{FH} cells in the lung-associated LNs were significantly associated with lung memory B cells, whereas SARS-CoV-2–specific $CD4^+$ T cell frequencies positively correlated with SARS-CoV-2–specific GC B cells lung-associated LNs. Conversely, certain LN and lung populations were inversely correlated. In particular, the amount of S/RBD-specific GC B cells in the lung-associated LNs negatively correlated with S-specific $CD4^+$ T cells in the lung. Similarly, S/RBD-specific memory B cell frequencies in the lung-associated LNs negatively correlated with SARS-CoV-2–specific and S-specific $CD8^+$ T cell frequencies in the lung. Similar correlations were observed when including blood and plasma antigen-specific lymphocyte populations in correlation matrix analyses (fig. S8). Together, these results suggest opposing or compensatory effects of humoral and cellular immune responses in lung-associated LNs and lungs.

DISCUSSION

Immunological memory is maintained by heterogeneous subsets of virus-specific T and B cells in nonlymphoid tissue sites of infection and multiple lymphoid organs. A comprehensive assessment of memory responses is therefore difficult to accomplish in humans. Here, we reveal the cellular complexity and functional diversity of SARS-CoV-2–specific memory T and B cells in lymphoid and mucosal tissues of previously infected organ donors up to 6 months after infection (see fig. S9 for summary schematic). SARS-CoV-2–specific $CD4^+$ T, $CD8^+$ T, and B cells predominantly localized in the lung and lung-associated LNs and were maintained as memory cell populations. Tissue-resident T and B cells, known to participate in protection against secondary viral infections, were found most abundantly in the lung and were correlated across multiple sites. Moreover, SARS-CoV-2–specific GC B cells and T_{FH} cells were found in lung-associated LNs, suggesting persisting GC responses months after resolution of infection. Together, these results indicate that the maintenance of SARS-CoV-2–specific immune memory is characterized by localized, ongoing coordination of cellular and humoral immunity within tissues.

SARS-CoV-2–specific memory T and B cells were found throughout the body and localized preferentially to lung and lung-associated LNs, providing direct evidence that those sites are key locations for establishment of immune memory after SARS-CoV-2 infection. Gut-draining LNs in some donors were also significant sites for SARS-CoV-2–specific memory T and B cells (particularly T_{RM} and B_{RM}), which could be because of the gut being a major site for SARS-CoV-2 replication in some cases (*3*, *51*). The low frequency of SARS-CoV-2–specific memory T or B cells in the spleen further suggests that virus infection is generally limited to mucosal sites of entry. Our results show that a proportion of SARS-CoV-2–specific memory T cells in the lung were T_{RM}, consistent with findings in airways of severe COVID-19 (*28*) and surgical lung samples from previously infected patients (*52*). Additionally, we show that the majority of the SARS-CoV-2–specific memory B cells were resident. In mouse models of influenza infection, localization and tissue residence of T and B cells to the lung and lung-associated LNs are correlated with optimal protective responses (*20*, *23*, *32*). Therefore, tissue-localized and resident memory T and B cells in the lung are likely important for site-specific protection and could be targets for site-specific boosting in vaccination.

SARS-CoV-2–specific memory $CD4^+$ T cells were identified at significantly higher frequencies than $CD8^+$ T cells across tissue sites, reflecting previous studies of peripheral blood showing that $CD4^+$ responses are more robust than $CD8^+$ responses months after resolution of infection (*2*, *53*). In addition, SARS-CoV-2–specific T cells exhibited tissue-specific functional profiles with cytotoxic proinflammatory, regulatory, and tissue repair functions variably manifested across different sites. In the lung-associated LNs, memory T cells exhibited broad proinflammatory, helper, and regulatory functional profiles. SARS-CoV-2–specific lung T cells produced higher levels of IL-10 compared with other sites consistent with a role for T cell–derived IL-10 in regulating lung inflammation in mice (*43*). We previously showed in paired airway and blood samples of patients with severe COVID-19 that the cytokine and chemokine profile in airway washes was distinct from that in plasma (*28*). Here, we further demonstrate that the functional responses of virus-specific T cells are tissue-specific—not only at the site of infection but also across numerous lymphoid tissues. Together, these results suggest that T cells in tissues mediate responses that are functionally adapted to the tissue site, resulting in heterogeneity of immune memory stored throughout the body.

SARS-CoV-2–specific memory B cells were distributed across multiple sites. While frequencies were highest in lung and LNs, there were also significant frequencies in BM. In all sites, virus-specific memory B cells exhibited a predominantly IgG^+ memory phenotype. The finding of S/RBD-specific B_{RM} in lung and lymphoid sites was notable, as was the low frequency of IgA^+ SARS-CoV-2–specific memory B cells in mucosal tissue and associate LNs.

Our results directly demonstrate ongoing, persistent GC responses in LNs following resolution of SARS-CoV-2 infection—including at least one example 6 months after infection. Despite a report of potentially impaired GC responses in fatal COVID-19 (*54*), our data show coordinated T_{FH} and GC B cells in the lung-associated LNs to nonfatal SARS-CoV-2 infection. These results provide evidence of

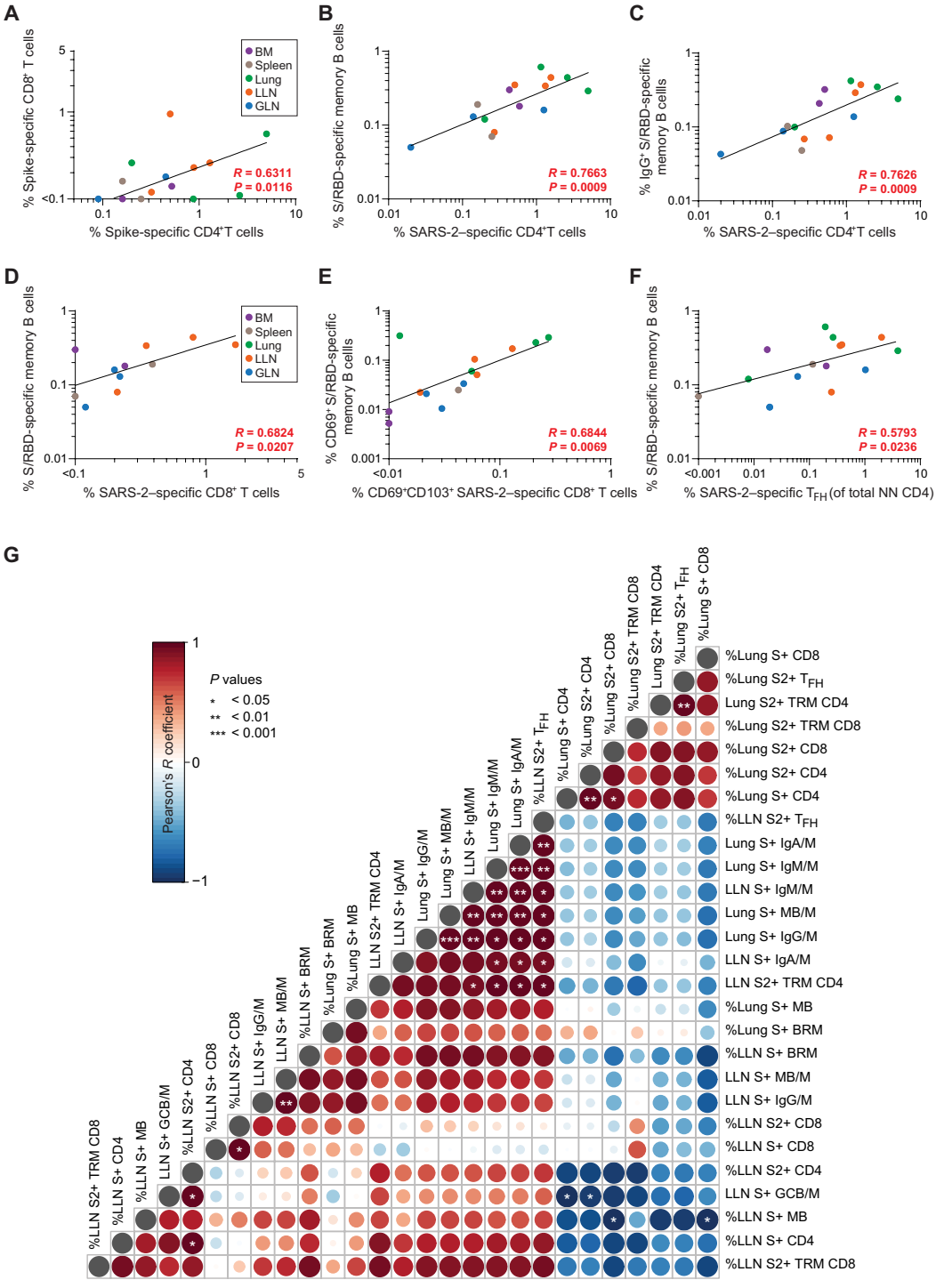


Fig. 6. SARS-CoV-2–specific immune memory relationships across organs. (**A**) Correlation between SARS-CoV-2 Spike–specific $CD4^+$ T cells as frequency of total $CD4^+$ T cells and Spike-specific $CD8^+$ T cells as frequency of total $CD8^+$ T cells. (**B**) Correlation between the frequency of SARS-CoV-2–specific $CD4^+$ T cells and SARS-CoV-2 S/RBD–specific memory B cells. (**C**) Correlation between SARS-CoV-2–specific $CD4^+$ T cells and IgG^+ SARS-CoV-2 S/RBD–specific memory B cells. (**D**) Correlation between SARS-CoV-2–specific $CD8^+$ T cells and SARS-CoV-2 S/RBD–specific memory B cells within lymphoid tissues. (**E**) Correlation between SARS-CoV-2–specific $CD69^+CD103^+$ $CD8^+$ T_{RM} cells and $CD69^+$ SARS-CoV-2 S/RBD–specific BRM. (**F**) Correlation between SARS-CoV-2–specific $CXCR5^+PD\text{-}1^+$ T_{FH} cells as frequency of NN $CD4^+$ T cells and SARS-CoV-2 S/RBD–specific memory B cells as frequency of total memory B cells. (**G**) Correlogram of SARS-CoV-2–specific lung and LLN lymphocyte populations. Pearson's *R* coefficients are shown from blue (−1.0) to red (1.0); *R* values are indicated by color and circle size. SARS-CoV-2–specific lymphocyte frequencies are depicted as a percentage of the parent population (%) or as counts per million peripheral blood mononuclear cells (/M); SARS-CoV-2–specific T_{FH} cells are a percentage of total NN $CD4^+$ T cells. Statistical analysis was performed on datasets using Pearson correlation. $^*P \le 0.05$, $^{**}P \le 0.01$, and $^{***}P \le 0.001$. GCB, GC B cell; MB, memory B cell; T_{RM}, $CD69^+CD103^+$-resident memory T cell; BRM, $CD69^+$-resident memory B cell; S+, Spike protein–specific; S2+, SARS-CoV-2–specific.

ongoing GC reactions after resolution of infection, consistent with reports of prolonged evolution of humoral responses in peripheral blood up to 6 months after SARS-CoV-2 infection (*3*, *55*). GC B cells were detected in donors spanning a broad age range—from 10 to 74 years, providing compelling evidence that the ability to establish robust GC responses to novel pathogens can be maintained with age.

These results also indicate ongoing interaction and coordination between T and B cells within LNs, which we also found related to memory populations in the lung. Significant correlations were also found between SARS-CoV-2–specific memory B and T cell populations across tissue sites, consistent with correlations between virus-specific T and B cell responses in peripheral blood of previously infected individuals (*1*, *7*). We also identified potential inverse correlations between frequencies of virus-specific $CD8^+$T cells in the lung and memory B cells in the lung-associated LN, suggesting that lung responses in situ can affect the magnitude or requirement for humoral responses in the associated LN. Together, these findings suggest that dynamic coordination of adaptive immune responses across the body is a feature of antiviral immunity to SARS-CoV-2.

This work has certain limitations. Namely, our study focuses on four seropositive donors across seven decades of life to provide a representative profile of tissue-specific antiviral immune responses. In addition to the challenges of obtaining live cells for immunological studies from organ donors, the findings here also depended on SARS-CoV-2 seropositive donors who had not been vaccinated, thus limiting the size of the donor pool and the timeframe of collection (before December 2020). The consistency in cell type and site-specific trends and correlations across all profiled donors, as well as corroboration of larger scale blood studies, demonstrates how this project provides new insights into tissue-specific immune memory maintenance and persistence of humoral and cellular responses after SARS-CoV-2 infection.

In conclusion, we reveal here that immunological memory from SARS-CoV-2 infection is maintained as heterogeneous subsets across multiple sites, with active and preferential maintenance in lung and associated LNs, as well as site-specific functional adaptations. These findings support the development of site-specific strategies for monitoring immune memory to infections and vaccines and for fortifying immune responses at the infection sites.

MATERIALS AND METHODS

Study design

The objective of this study was to measure adaptive immune responses to SARS-CoV-2 in blood and tissues of seropositive individuals after resolution of infection. We measured the frequency of SARS-CoV-2–specific $CD4^+$ T, $CD8^+$ T, and B cells in seropositive organ donors compared with prepandemic seronegative donors to understand the maintenance of immunological memory to SARS-CoV-2 as T and B cell subsets across the body, the functional immune response in tissues, and the immune memory relationships across circulating and tissue-resident SARS-CoV-2–specific T and B cell populations.

Human samples

Human tissues were obtained from deceased organ donors at the time of organ acquisition for clinical transplantation through an approved protocol and material transfer agreement with LiveOnNY, the OPO for the New York metropolitan area, as previously described (*34*, *35*, *56*–*60*). Human tissues from the pediatric donor (HDL113) were obtained through arrangements with multiple OPOs across the United States through the Human Atlas for Neonatal Development-Immunity program, which is an extension of the coordinating center set up for nPOD (network for Pancreatic Organ Donors with Diabetes) (*61*). A list of donors from which tissues were used in this study is presented in table S1. Organ donors were tested for SARS-CoV-2 infection (and confirmed as SARS-CoV-2 negative) by PCR testing of nasal swabs, tracheal aspirates, and/or bronchoalveolar lavage. A history of previous COVID-19 (D492 and D498) and positive serology (D495) was provided in the donor summary, and SARS-CoV-2 serology for all donors was measured in the Center for Advanced Laboratory Medicine at Columbia University Irving Medical Center. Tissues from all seropositive donors were obtained before December 2020, and all donors were free of cancer and seronegative for hepatitis B, hepatitis C, and HIV. Because tissues were obtained from brain-dead organ donors, this study does not qualify as "human subjects" research, as confirmed by the Columbia University Institutional Review Board.

Isolation of single-cell suspensions from tissue samples

Tissue samples were maintained in cold saline or media and transported to the laboratory within 2 to 4 hours of organ procurement for adult organs and shipped to the laboratory on ice within 24 hours of procurement for pediatric donors. Tissue processing protocols were adapted from protocols previously described (*34*, *56*–*60*), with some recent optimizations. Briefly, mononuclear cells were isolated from the blood and BM samples by density centrifugation using Ficoll-Paque PLUS (GE Healthcare, catalog no. 17-1440-03). Spleen was processed using mechanical dissociation, followed by pushing through 100-μm filters (Thermo Fisher Scientific, catalog no. 50-146-1428), and Ficoll-Paque density centrifugation as above. Lung and LN samples were first incubated with collagenase D (1 mg/ml) (Sigma-Aldrich, catalog no. 11088882001) and deoxyribonuclease (0.1 mg/ml) (Thermo Fisher Scientific, catalog no. NC9709009) in Iscove's modified Dulbecco's medium (Thermo Fisher Scientific, catalog no. 12-440-053) for 30 min at 37°C on a shaker followed by addition of 0.5 M EDTA (pH 8.0) (Thermo Fisher Scientific, catalog no. 150575-020), filtration, and density centrifugation as above, resulting in high yields of live leukocytes.

SARS-CoV-2 serology testing

Blood from deceased organ donors (D495, D498, and HDL113) was collected, and serum was obtained after centrifugation using serum separating clot activator tubes (Thermo Fisher Scientific, catalog no. 22040546). SARS-CoV-2 serology testing for N protein was then performed by the Center for Advanced Laboratory Medicine at Columbia University Irving Medical Center to determine previous exposure to SARS-CoV-2 for inclusion in the study.

SARS-CoV-2 enzyme-linked immunosorbent assay titers were determined as previously described (*1*). Briefly, Corning 96-well half-area plates (Thermo Fisher Scientific, catalog no.3690) were coated with SARS-CoV-2 S protein (1 μg/ml), RBD protein, or N protein (Sino Biological, catalog no. 40588-V07E) overnight at 4°C. The next day, plates were blocked with 3% milk (skim milk powder, Thermo Fisher Scientific, catalog no. LP0031) in phosphate-buffered saline (PBS) containing 0.05% Tween 20 (ThermoScientific, catalog no. J260605-AP) for 2 hours at room temperature. Heat-inactivated serum (30 min at 56°C) was then added to the plates and incubated for 1.5 hours at room temperature. Plates were washed five times with 0.05% PBS/Tween

20. Secondary antibodies were diluted in 1% milk containing 0.05% Tween 20 in PBS. IgG titers were determined using anti-human IgG peroxidase antibody (Hybridoma Reagent Laboratory, catalog no. HP6123-HRP) at 1:1000 dilution. End-point titers were plotted for each sample using background subtracted data. The limit of detection was defined as 1:3 for IgG.

Pseudovirus neutralization assay

The pseudovirus (PSV) neutralization assays were performed as previously described (*1*). Briefly, 2.5×10^4 Vero cells [American Type Culture Collection (ATCC), catalog no. CCL-81] were seeded in clear flat-bottom 96-well plates (Thermo Fisher Scientific, catalog no. 165305) to produce a monolayer at the time of infection. Recombinant SARS-CoV-2 S–D614G–pseudotyped VSV-ΔG-GFP were generated by transfecting human embryonic kidney 293T cells (ATCC, catalog no. CRL-321) with plasmid phCMV3–SARS-CoV-2 S and then infecting with VSV-ΔG-GFP. Pretitrated rVSV-SARS-CoV-2-S-D614G was incubated with serially diluted human heat-inactivated serum at 37°C for 1 to 1.5 hours before addition to confluent Vero cell monolayers. Cells were incubated for 16 hours at 37°C in 5% CO_2, fixed in 4% paraformaldehyde in PBS (pH 7.4) (Santa Cruz, catalog no. sc-281692) with Hoechst (10 μg/ml) (Thermo Fisher Scientific, catalog no. 62249), and imaged using a CellInsight CX5 imager to quantify the total number of cells and infected green fluorescent protein (GFP)–expressing cells to determine the percentage of infection. Neutralization titers or inhibition dose 50 (ID_{50}) were calculated using the One-Site Fit Log IC_{50} model in Prism 8.0 (GraphPad). As internal quality control to define the interassay variation, three samples were included across the PSV neutralization assays. Samples that did not reach 50% inhibition at the lowest serum dilution of 1:20 were considered as non-neutralizing.

In vitro T cell stimulations with SARS-CoV-2 peptide MPs

Mononuclear cells from the blood, BM, spleen, lung, lung-associated LNs, and gut-associated LNs of SARS-CoV-2 seropositive donors were thawed, and dead cells were removed using the EasySep Dead Cell Removal (annexin V) Kit (STEMCELL Technologies, catalog no. 17899) containing 10% heat-inactivated human AB serum (Gemini, catalog no. 507533010) and penicillin-streptomycin-glutamine (Thermo Fisher Scientific, catalog no. 10378016) and incubated overnight at 37°C in 5% CO_2. Cells were stimulated for 6 or 24 hours by the addition of SARS-CoV-2–specific CD4 and CD8 MPs (1 μg/ml) (MP_S, MP_CD4_R, MP_CD8_A, and MP_CD8_B) designed and synthesized as previously described. Briefly, MP_S consists of 253 15-mer peptides overlapping by 10 residues and covering the entire S protein. MP_CD4_R consists of 221-predicted HLA class II $CD4^+$ T cell epitopes covering all proteins apart from S protein. For CD8 epitopes, MPs were synthesized based on epitope predictions for 12 most common HLA class I A and B alleles; these resulted in 628 predicted $CD8^+$ T cell epitopes, which were separated into MP_CD8_A and MP_CD8_B (*1*, *2*, *62*). Equimolar amount of DMSO was used as negative control. Before the addition of peptide MPs, cells were blocked for 15 min with anti-CD40 monoclonal antibody (0.5 μg/ml) (Miltenyi Biotec, catalog no. 130-094-133), as previously described (*63*). After either 6 or 24 hours, the supernatant was collected for multiplex detection of cytokines, and cells were stained for AIMs and analyzed via flow cytometry.

AIM^+ antigen-specific $CD4^+$ T cells were identified as positive following Boolean OR gating of the $CD40L^+OX40^+$, $4\text{-}1BB^+OX40^+$, $4\text{-}1BB^+CD40L^+$ subsets (see fig. S1 for gating strategy). The resultant gate was used to quantify AIM^+ $CD4^+$ T cell frequency. AIM^+ antigen-specific $CD8^+$ T cells were identified as $4\text{-}1BB^+CD25^+$. Antigen-specific T_{FH} calculated as a frequency of NN $CD4^+$ T cells were gated on total $CD4^+$ T cells excluding $CD45RA^+CCR7^+$ $CD4^+$ T cells. Antigen-specific $CD4^+$ and $CD8^+$ T cells were measured as DMSO background–subtracted data. For quantification of the frequency of total SARS-CoV-2–specific $CD4^+$ T cells, a weighted average was taken for percentage of AIM^+ $CD4^+$ T cells identified for samples stimulated with MP_S or MP_CD4_R MPs. For quantification of the frequency of total SARS-CoV-2–specific $CD8^+$ T cells, a weighted average was taken for percentage of AIM^+ $CD8^+$ T cells identified for samples stimulated with MP_S, MP_CD8_A, or MP_CD8_B MPs.

Flow cytometry

For flow cytometry analysis of SARS-CoV-2 antigen-reactive T cells, cells were stained in 96-well U-bottom plates protected from light using fluorochrome-conjugated antibodies (see table S2 for antibodies in the T cell flow cytometry panel). Briefly, cells were washed with fluorescence-activated cell sorting (FACS) buffer (PBS with 2% heat-inactivated fetal bovine serum) and then resuspended with surface staining antibody cocktail for 20 min at room temperature. Surface-stained cells were fixed for 30 min at room temperature in fixation buffer (Tonbo, catalog no. TNB-0607-KIT), washed with permeabilization buffer (Tonbo, catalog no. TNB-0607-KIT), and washed again with FACS buffer. Flow cytometry data were collected using the five-laser Cytek Aurora flow cytometer (Cytek Bio) and analyzed using FlowJo V 10.7 software and Prism 9.0.1. software.

For flow cytometry analysis of SARS-CoV-2–specific B cells, biotinylated protein antigens multimerized on fluorescently labeled streptavidin were used as probes to detect antigen-specific B cells (see table S3 for antibodies used in the B cell flow cytometry panel). Avitagged full-length SARS-CoV-2 S (2P-stabilized, double streptavidin-tagged) and RBD proteins were generated in-house. Biotinylation was performed using biotin protein ligase standard reaction kit (Avidity, catalog no. Bir500A) following the manufacturer's protocol and dialyzed against PBS. Biotinylated S was mixed with streptavidin BV421 (BioLegend, catalog no. 405225) and streptavidin BUV737 (BD Bioscience, catalog no. 612775) at 20:1 ratio (~6:1 molar ratio). Biotinylated RBD was mixed with streptavidin phycoerythrin (PE)–Cy7 (BioLegend, catalog no. 405206) and streptavidin BUV661 (BD Bioscience, catalog no. 612979) at 2.2:1 ratio (~4:1 molar ratio). Streptavidin PE-Cy5.5 (Thermo Fisher Scientific, catalog no. SA1018) was used as a decoy probe for nonspecific streptavidin-binding B cells. The probes were then mixed in Brilliant Stain Buffer (BD Bioscience, catalog no. 566349) containing 5 μM free D-biotin (Avidity, catalog no. Bir500A). Cells (~10^7) were prepared in U-bottom 96-well plates and stained with 50 μl of antigen cocktail containing 400 ng of S (200 ng per probe), 100-ng of RBD (50 ng per probe), and 20 ng of streptavidin PE-Cy5.5 at 4°C for 1 hour followed by staining for surface markers in Brilliant Stain Buffer at 4°C for 30 min. Dead cells were stained using the LIVE/DEAD Fixable Blue Stain Kit (Thermo Fisher Scientific, catalog no. L34962) in PBS at 4°C for 30 min. Cells were then fixed and permeabilized using eBioscience Intracellular Fixation and Permeabilization Buffer Set (Thermo Fisher Scientific, catalog no. 88-8824-00) before staining with antibodies against transcription factors in eBioscience Permeabilization Buffer (Thermo Fisher Scientific, catalog no. 00-8333-56). Samples were acquired on Cytek Aurora and analyzed using FlowJo10.7.1 (BD Biosciences). In each

experiment, peripheral blood mononuclear cells from a known convalescent individual with COVID-19 and an unexposed individual were included to ensure consistent sensitivity and specificity of the assay.

Multiplex detection of cytokines

Cryopreserved supernatant from in vitro T cell stimulation experiments was sent to Eve Technologies Corp. (Calgary, Alberta) for quantification of 50 total human cytokines, chemokines, and growth factors. Luminex xMAP technology was used for multiplexed quantification of two human cytokines in one array (perforin and granzyme B) and 48 human cytokines, chemokines, and growth factors in a separate array [sCD40L, EGF (endothelial growth factor), eotaxin, FGF-2 (fibroblast growth factor 2), Flt-3 ligand, fractalkine, G-CSF, GM-CSF, GROα, IFN-α2, IFN-γ, IL-1α, IL-1β, IL-1ra, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, IL-10, IL-12 (p40), IL-12 (p70), IL-13, IL-15, IL-17A, IL-17E/IL-25, IL-17F, IL-18, IL-22, IL-27, IP-10, MCP-1 (monocyte chemoattractant protein 1), MCP-3, M-CSF, MDC (macrophage-derived chemokine) (CCL22), MIG (monokine induced by gamma interferon), MIP-1α (macrophage inflammatory protein 1α), MIP-1β, PDGF-AA (platelet-derived growth factor AA), PDGF-AB/BB, RANTES, TGF-α (transforming growth factor–α), TNF-α, TNF-β, and VEGF-A (vascular endothelial growth factor A)]. The multiplexing analysis was performed using the Luminex 200 system with assay kits sourced by Millipore MILLIPEX (MilliporeSigma, Burlington, Massachusetts, USA) according to the manufacturer's protocol.

Observed concentrations were calculated with the standard curve based on the fluorescence intensity of the bead population for a specific analyte. For analysis and visualization of cytokine/chemokine production by antigen-responding cells in multiple tissues sites within each individual donor, observed concentrations for each analyte were first subtracted from DMSO-negative control and then scaled across samples for each individual donor on a maximum absolute scale, with values ranging from −1 to 1 across all analytes using the MaxAbsScaler features of sklearn.preprocessing function of the Python *scikit-learn* library (*36*, *63*). Heatmap visualizations were generated using the Python data visualization library *seaborn* (*64*). For analysis comparing, the production of analytes across donors and tissue sites either absolute observed concentrations were used or observed concentrations for each analyte were normalized to DMSO-negative control.

Correlogram plot and visualization

Correlograms were analyzed and plotted using the Pearson product moment correlation coefficient (*r*) between all parameter pairs from blood, lung, and lung-associated LN lymphocyte populations (see data file 2 for raw data, *R* and *P* values). Correlograms were created with the corrplot package (v0.88) (*65*) running under R (v4.0.2) in RStudio (1.4.1103). Visual clustering of parameters was performed using the "hclust" option of corrMatOrder. Two-sided *P* values were calculated using corr.test (stats v4.0.2) and graphed (corrplot v0.88) based on $*P \leq 0.05$, $**P \leq 0.01$, and $***P \leq 0.001$.

Statistical analysis

Descriptive statistics of compiled flow cytometry data and statistical testing were performed using Prism (GraphPad). Graphs were generated using Prism (GraphPad), Python *matplotlib* and *seaborn* libraries (*65*, *67*), and RStudio corrplot package (*66*). Differences in means between two sample groups were compared using nonparametric test of null hypothesis Mann-Whitney *U* test. Pearson correlations were used to evaluate immune memory relationships. Multiple group comparisons were done using one-way analysis of variance (ANOVA, corrected for multiple comparisons by false discovery rate using two-stage linear step-up procedure of Benjamini, Krieger, and Yekutieli when comparing seropositive and seronegative donors. For comparing immune mediator profiles across tissue sites, statistical analyses were performed via one-way ANOVA corrected for multiple comparisons by Tukey's multiple comparison test. $P < 0.05$ was considered as statistically significant. $****P \leq 0.0001$, $***P \leq 0.001$, $**P \leq 0.01$, and $*P \leq 0.05$.

SUPPLEMENTARY MATERIALS

www.science.org/doi/10.1126/sciimmunol.abl9105
Figs. S1 to S9
Tables S1 to S3
Data files S1 and S2

View/request a protocol for this paper from *Bio-protocol*.

REFERENCES AND NOTES

1. J. M. Dan, J. Mateus, Y. Kato, K. M. Hastie, E. D. Yu, C. E. Faliti, A. Grifoni, S. I. Ramirez, S. Haupt, A. Frazier, C. Nakao, V. Rayaprolu, S. A. Rawlings, B. Peters, F. Krammer, V. Simon, E. O. Saphire, D. M. Smith, D. Weiskopf, A. Sette, S. Crotty, Immunological memory to SARS-CoV-2 assessed for up to 8 months after infection. *Science* **371**, eabf4063 (2021).
2. A. Grifoni, D. Weiskopf, S. I. Ramirez, J. Mateus, J. M. Dan, C. R. Moderbacher, S. A. Rawlings, A. Sutherland, L. Premkumar, R. S. Jadi, D. Marrama, A. M. de Silva, A. Frazier, A. F. Carlin, J. A. Greenbaum, B. Peters, F. Krammer, D. M. Smith, S. Crotty, A. Sette, Targets of T cell responses to SARS-CoV-2 coronavirus in humans with COVID-19 disease and unexposed individuals. *Cell* **181**, 1489–1501.e15 (2020).
3. C. Gaebler, Z. Wang, J. C. C. Lorenzi, F. Muecksch, S. Finkin, M. Tokuyama, A. Cho, M. Jankovic, D. Schaefer-Babajew, T. Y. Oliveira, M. Cipolla, C. Viant, C. O. Barnes, Y. Bram, G. Breton, T. Hagglof, P. Mendoza, A. Hurley, M. Turroja, K. Gordon, K. G. Millard, V. Ramos, F. Schmidt, Y. Weisblum, D. Jha, M. Tankelevich, G. Martinez-Delgado, J. Yee, R. Patel, J. Dizon, C. Unson-O'Brien, I. Shimeliovich, D. F. Robbiani, Z. Zhao, A. Gazumyan, R. E. Schwartz, T. Hatziioannou, P. J. Bjorkman, S. Mehandru, P. D. Bieniasz, M. Caskey, M. C. Nussenzweig, Evolution of antibody immunity to SARS-CoV-2. *Nature* **591**, 639–644 (2021).
4. Z. Wang, F. Muecksch, D. Schaefer-Babajew, S. Finkin, C. Viant, C. Gaebler, H. H. Hoffmann, C. O. Barnes, M. Cipolla, V. Ramos, T. Y. Oliveira, A. Cho, F. Schmidt, J. Da Silva, E. Bednarski, L. Aguado, J. Yee, M. Daga, M. Turroja, K. G. Millard, M. Jankovic, A. Gazumyan, Z. Zhao, C. M. Rice, P. D. Bieniasz, M. Caskey, T. Hatziioannou, M. C. Nussenzweig, Naturally enhanced neutralizing breadth against SARS-CoV-2 one year after infection. *Nature* **595**, 426–431 (2021).
5. L. B. Rodda, J. Netland, L. Shehata, K. B. Pruner, P. A. Morawski, C. D. Thouvenel, K. K. Takehara, J. Eggenberger, E. A. Hemann, H. R. Waterman, M. L. Fahning, Y. Chen, M. Hale, J. Rathe, C. Stokes, S. Wrenn, B. Fiala, L. Carter, J. A. Hamerman, N. P. King, M. Gale Jr., D. J. Campbell, D. J. Rawlings, M. Pepper, Functional SARS-CoV-2-specific immune memory persists after mild COVID-19. *Cell* **184**, 169–183.e17 (2021).
6. T. Bilich, A. Nelde, J. S. Heitmann, Y. Maringer, M. Roerden, J. Bauer, J. Rieth, M. Wacker, A. Peter, S. Horber, D. Rachfalski, M. Marklin, S. Stevanovic, H. G. Rammensee, H. R. Salih, J. S. Walz, T cell and antibody kinetics delineate SARS-CoV-2 peptides mediating long-term immune responses in COVID-19 convalescent individuals. *Sci. Transl. Med.* **13**, eabf7517 (2021).
7. K. W. Cohen, S. L. Linderman, Z. Moodie, J. Czartoski, L. Lai, G. Mantus, C. Norwood, L. E. Nyhoff, V. V. Edara, K. Floyd, S. C. De Rosa, H. Ahmed, R. Whaley, S. N. Patel, B. Prigmore, M. P. Lemos, C. W. Davis, S. Furth, J. B. O'Keefe, M. P. Gharpure, S. Gunisetty, K. Stephens, R. Antia, V. I. Zarnitsyna, D. S. Stephens, S. Edupuganti, N. Rouphael, E. J. Anderson, A. K. Mehta, J. Wrammert, M. S. Suthar, R. Ahmed, M. J. McElrath, Longitudinal analysis shows durable and broad immune memory after SARS-CoV-2 infection with persisting antibody responses and memory B and T cells. *Cell Rep. Med.* **2**, 100354 (2021).
8. J. Zuo, A. C. Dowell, H. Pearce, K. Verma, H. M. Long, J. Begum, F. Aiano, Z. Amin-Chowdhury, K. Hoschler, T. Brooks, S. Taylor, J. Hewson, B. Hallis, L. Stapley, R. Borrow, E. Linley, S. Ahmad, B. Parker, A. Horsley, G. Amirthalingam, K. Brown, M. E. Ramsay, S. Ladhani, P. Moss, Robust SARS-CoV-2-specific T cell immunity is maintained at 6 months following primary infection. *Nat. Immunol.* **22**, 620–626 (2021).
9. D. S. Khoury, D. Cromer, A. Reynaldi, T. E. Schlub, A. K. Wheatley, J. A. Juno, K. Subbarao, S. J. Kent, J. A. Triccas, M. P. Davenport, Neutralizing antibody levels are highly predictive of immune protection from symptomatic SARS-CoV-2 infection. *Nat. Med.* **27**, 1205–1211 (2021).

10. K. A. Earle, D. M. Ambrosino, A. Fiore-Gartland, D. Goldblatt, P. B. Gilbert, G. R. Siber, P. Dull, S. A. Plotkin, Evidence for antibody as a protective correlate for COVID-19 vaccines. *Vaccine* **39**, 4423–4428 (2021).
11. S. S. Abdool Karim, T. de Oliveira, New SARS-CoV-2 Variants—Clinical, public health, and vaccine implications. *N. Eng. J. Med.* **384**, 1866–1868 (2021).
12. L. J. Abu-Raddad, H. Chemaitelly, A. A. Butt; National Study Group for COVID-19 Vaccination, Effectiveness of the BNT162b2 Covid-19 Vaccine against the B.1.1.7 and B.1.351 Variants. *N. Engl. J. Med.* **385**, 187–189 (2021).
13. R. E. Chen, X. Zhang, J. B. Case, E. S. Winkler, Y. Liu, L. A. VanBlargan, J. Liu, J. M. Errico, X. Xie, N. Suryadevara, P. Gilchuk, S. J. Zost, S. Tahan, L. Droit, J. S. Turner, W. Kim, A. J. Schmitz, M. Thapa, D. Wang, A. C. M. Boon, R. M. Presti, J. A. O'Halloran, A. H. J. Kim, P. Deepak, D. Pinto, D. H. Fremont, J. E. Crowe Jr., D. Corti, H. W. Virgin, A. H. Ellebedy, P.-Y. Shi, M. S. Diamond, Resistance of SARS-CoV-2 variants to neutralization by monoclonal and serum-derived polyclonal antibodies. *Nat. Med.* **27**, 717–726 (2021).
14. N. R. Faria, T. A. Mellan, C. Whittaker, I. M. Claro, D. D. S. Candido, S. Mishra, M. A. E. Crispim, F. C. S. Sales, I. Hawryluk, J. T. McCrone, R. J. G. Hulswit, L. A. M. Franco, M. S. Ramundo, J. G. de Jesus, P. S. Andrade, T. M. Coletti, G. M. Ferreira, C. A. M. Silva, E. R. Manuli, R. H. M. Pereira, P. S. Peixoto, M. U. G. Kraemer, N. Gaburo Jr., C. D. C. Camilo, H. Hoeltgebaum, W. M. Souza, E. C. Rocha, L. M. de Souza, M. C. de Pinho, L. J. T. Araujo, F. S. V. Malta, A. B. de Lima, J. D. P. Silva, D. A. G. Zauli, A. C. S. Ferreira, R. P. Schnekenberg, D. J. Laydon, P. G. T. Walker, H. M. Schluter, A. L. P. D. Santos, M. S. Vidal, V. S. Del Caro, R. M. F. Filho, H. M. D. Santos, R. S. Aguiar, J. L. Proenca-Modena, B. Nelson, J. A. Hay, M. Monod, X. Miscouridou, H. Coupland, R. Sonabend, M. Vollmer, A. Gandy, C. A. Prete Jr., V. H. Nascimento, M. A. Suchard, T. A. Bowden, S. L. K. Pond, C. H. Wu, O. Ratmann, N. M. Ferguson, C. Dye, N. J. Loman, P. Lemey, A. Rambaut, N. A. Fraiji, M. Carvalho, O. G. Pybus, S. Flaxman, S. Bhatt, E. C. Sabino, Genomics and epidemiology of the P.1 SARS-CoV-2 lineage in Manaus, Brazil. *Science* **372**, 815–821 (2021).
15. M. McCallum, A. De Marco, F. A. Lempp, M. A. Tortorici, D. Pinto, A. C. Walls, M. Beltramello, A. Chen, Z. Liu, F. Zatta, S. Zepeda, J. di Iulio, J. E. Bowen, M. Montiel-Ruiz, J. Zhou, L. E. Rosen, S. Bianchi, B. Guarino, C. S. Fregni, R. Abdelnabi, S. C. Foo, P. W. Rothlauf, L. M. Bloyet, F. Benigni, E. Cameroni, J. Neyts, A. Riva, G. Snell, A. Telenti, S. P. J. Whelan, H. W. Virgin, D. Corti, M. S. Pizzuto, D. Veesler, N-terminal domain antigenic mapping reveals a site of vulnerability for SARS-CoV-2. *Cell* **184**, 2332–2347.e16 (2021).
16. D. L. Farber, Tissues, not blood, are where immune cells function. *Nature* **593**, 506–509 (2021).
17. S. P. Weisberg, B. B. Ural, D. L. Farber, Tissue-specific immunity for a changing world. *Cell* **184**, 1517–1529 (2021).
18. D. Masopust, A. G. Soerens, Tissue-resident T cells and other resident leukocytes. *Annu. Rev. Immunol.* **37**, 521–546 (2019).
19. P. A. Szabo, M. Miron, D. L. Farber, Location, location, location: Tissue resident memory T cells in mice and humans. *Sci. Immunol.* **4**, eaas9673 (2019).
20. J. R. Teijaro, D. Turner, Q. Pham, E. J. Wherry, L. Lefrancois, D. L. Farber, Cutting edge: Tissue-retentive lung memory CD4 T cells mediate optimal protection to respiratory virus infection. *J. Immunol.* **187**, 5510–5514 (2011).
21. T. Wu, Y. Hu, Y. T. Lee, K. R. Bouchard, A. Benechet, K. Khanna, L. S. Cauley, Lung-resident memory CD8 T cells (TRM) are indispensable for optimal cross-protection against pulmonary virus infection. *J. Leukoc. Biol.* **95**, 215–224 (2014).
22. J. Zhao, J. Zhao, A. K. Mangalam, R. Channappanavar, C. Fett, D. K. Meyerholz, S. Agnihothram, R. S. Baric, C. S. David, S. Perlman, Airway Memory CD4(+) T cells mediate protective immunity against emerging respiratory Coronaviruses. *Immunity* **44**, 1379–1391 (2016).
23. D. H. Paik, D. L. Farber, Anti-viral protective capacity of tissue resident memory T cells. *Curr. Opin. Virol.* **46**, 20–26 (2021).
24. D. H. Paik, D. L. Farber, Influenza infection fortifies local lymph nodes to promote lung-resident heterosubtypic immunity. *J. Exp. Med.* **218**, e20200218 (2021).
25. J. Kubler-Kielb, F. Majadly, Z. Biesova, C. P. Mocca, C. Guo, R. Nussenzweig, V. Nussenzweig, S. Mishra, Y. Wu, L. H. Miller, J. M. Keith, T. Y. Liu, J. B. Robbins, R. Schneerson, A bicomponent Plasmodium falciparum investigational vaccine composed of protein-peptide conjugates. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 1172–1177 (2010).
26. B. V. Kumar, T. J. Connors, D. L. Farber, Human T cell development, localization, and function throughout life. *Immunity* **48**, 202–213 (2018).
27. P. A. Szabo, H. M. Levitin, M. Miron, M. E. Snyder, T. Senda, J. Yuan, Y. L. Cheng, E. C. Bush, P. Dogra, P. Thapa, D. L. Farber, P. A. Sims, Single-cell transcriptomics of human T cells reveals tissue and activation signatures in health and disease. *Nat. Commun.* **10**, 4706 (2019).
28. P. A. Szabo, P. Dogra, J. I. Gray, S. B. Wells, T. J. Connors, S. P. Weisberg, I. Krupska, R. Matsumoto, M. M. L. Poon, E. Idzikowski, S. E. Morris, C. Pasin, A. J. Yates, A. Ku, M. Chait, J. Davis-Porada, X. V. Guo, J. Zhou, M. Steinle, S. Mackay, A. Saqi, M. R. Baldwin, P. A. Sims, D. L. Farber, Longitudinal profiling of respiratory and systemic immune responses reveals myeloid cell-driven lung inflammation in severe COVID-19. *Immunity* **54**, 797–814.e6 (2021).
29. M. Liao, Y. Liu, J. Yuan, Y. Wen, G. Xu, J. Zhao, L. Cheng, J. Li, X. Wang, F. Wang, L. Liu, I. Amit, S. Zhang, Z. Zhang, Single-cell landscape of bronchoalveolar immune cells in patients with COVID-19. *Nat. Med.* **26**, 842–844 (2020).
30. S. Crotty, T follicular helper cell biology: A decade of discovery and diseases. *Immunity* **50**, 1132–1148 (2019).
31. F. Weisel, M. Shlomchik, Memory B cells of mice and humans. *Annu. Rev. Immunol.* **35**, 255–284 (2017).
32. S. R. Allie, J. E. Bradley, U. Mudunuru, M. D. Schultz, B. A. Graf, F. E. Lund, T. D. Randall, The establishment of resident memory B cells in the lung requires local antigen encounter. *Nat. Immunol.* **20**, 97–108 (2019).
33. N. M. Weisel, F. J. Weisel, D. L. Farber, L. A. Borghesi, Y. Shen, W. Ma, E. T. Luning Prak, M. J. Shlomchik, Comprehensive analyses of B-cell compartments across the human body reveal novel subsets and a gut-resident memory phenotype. *Blood* **136**, 2774–2785 (2020).
34. D. J. Carpenter, T. Granot, N. Matsuoka, T. Senda, B. V. Kumar, J. J. C. Thome, C. L. Gordon, M. Miron, J. Weiner, T. Connors, H. Lerner, A. Friedman, T. Kato, A. D. Griesemer, D. L. Farber, Human immunology studies using organ donors: Impact of clinical variations on immune parameters in tissues and circulation. *Am. J. Transplant.* **18**, 74–88 (2018).
35. J. J. Thome, N. Yudanin, Y. Ohmura, M. Kubota, B. Grinshpun, T. Sathaliyawala, T. Kato, H. Lerner, Y. Shen, D. L. Farber, Spatial map of human T cell compartmentalization and maintenance over decades of life. *Cell* **159**, 814–828 (2014).
36. C. Rydyznski Moderbacher, S. I. Ramirez, J. M. Dan, A. Grifoni, K. M. Hastie, D. Weiskopf, S. Belanger, R. K. Abbott, C. Kim, J. Choi, Y. Kato, E. G. Crotty, C. Kim, S. A. Rawlings, J. Mateus, L. P. V. Tse, A. Frazier, R. Baric, B. Peters, J. Greenbaum, E. Ollmann Saphire, D. M. Smith, A. Sette, S. Crotty, Antigen-specific adaptive immunity to SARS-CoV-2 in acute COVID-19 and associations with age and disease severity. *Cell* **183**, 996–1012.e19 (2020).
37. D. F. Gudbjartsson, G. L. Norddahl, P. Melsted, K. Gunnarsdottir, H. Holm, E. Eythorsson, A. O. Arnthorsson, D. Helgason, K. Bjarnadottir, R. F. Ingvarsson, B. Thorsteinsdottir, S. Kristjansdottir, K. Birgisdottir, A. M. Kristinsdottir, M. I. Sigurdsson, G. A. Arnadottir, E. V. Ivarsdottir, M. Andresdottir, F. Jonsson, A. B. Agustsdottir, J. Berglund, B. Eiriksdottir, R. Fridriksdottir, E. E. Gardarsdottir, M. Gottfredsson, O. S. Gretarsdottir, S. Gudmundsdottir, K. R. Gudmundsson, T. R. Gunnarsdottir, A. Gylfason, A. Helgason, B. O. Jensson, A. Jonasdottir, H. Jonsson, T. Kristjansson, K. G. Kristinsson, D. N. Magnusdottir, O. T. Magnusson, L. B. Olafsdottir, S. Rognvaldsson, L. le Roux, G. Sigmundsdottir, A. Sigurdsson, G. Sveinbjornsson, K. E. Sveinsdottir, M. Sveinsdottir, E. A. Thorarensen, B. Thorbjornsson, M. Thordardottir, J. Saemundsdottir, S. H. Kristjansson, K. S. Josefsdottir, G. Masson, G. Georgsson, M. Kristjansson, A. Moller, R. Palsson, T. Gudnason, U. Thorsteinsdottir, I. Jonsdottir, P. Sulem, K. Stefansson, Humoral immune response to SARS-CoV-2 in Iceland. *N. Engl. J. Med.* **383**, 1724–1734 (2020).
38. J. M. Dan, C. S. Lindestam Arlehamn, D. Weiskopf, R. da Silva Antunes, C. Havenar-Daughton, S. M. Reiss, M. Brigger, M. Bothwell, A. Sette, S. Crotty, A cytokine-independent approach to identify antigen-specific human germinal center T follicular helper cells and rare antigen-specific CD4+ T cells in blood. *J. Immunol.* **197**, 983–993 (2016).
39. C. Havenar-Daughton, S. M. Reiss, D. G. Carnathan, J. E. Wu, K. Kendric, A. Torrents de la Peña, S. P. Kasturi, J. M. Dan, M. Bothwell, R. W. Sanders, B. Pulendran, G. Silvestri, S. Crotty, Cytokine-independent detection of antigen-specific germinal center T follicular helper cells in immunized nonhuman primates using a live cell activation-induced marker technique. *J. Immunol.* **197**, 994–1002 (2016).
40. D. L. Farber, N. A. Yudanin, N. P. Restifo, Human memory T cells: Generation, compartmentalization and homeostasis. *Nat. Rev. Immunol.* **14**, 24–35 (2014).
41. F. Sallusto, J. Geginat, A. Lanzavecchia, Central memory and effector memory T cell subsets: Function, generation, and maintenance. *Annu. Rev. Immunol.* **22**, 745–763 (2004).
42. T. Sathaliyawala, M. Kubota, N. Yudanin, D. Turner, P. Camp, J. J. Thome, K. L. Bickham, H. Lerner, M. Goldstein, M. Sykes, T. Kato, D. L. Farber, Distribution and compartmentalization of human circulating and tissue-resident memory T cell subsets. *Immunity* **38**, 187–197 (2013).
43. J. Sun, R. Madan, C. L. Karp, T. J. Braciale, Effector T cells control lung inflammation during acute influenza virus infection by producing IL-10. *Nat. Med.* **15**, 277–284 (2009).
44. N. R. Mathew, J. K. Jayanthan, I. V. Smirnov, J. L. Robinson, H. Axelsson, S. S. Nakka, A. Emmanouilidi, P. Czarnewski, W. T. Yewdell, K. Schön, C. Lebrero-Fernández, V. Bernasconi, W. Rodin, A. M. Harandi, N. Lycke, N. Borcherding, J. W. Yewdell, V. Greiff, M. Bemark, D. Angeletti, Single-cell BCR and transcriptome analysis after influenza infection reveals spatiotemporal dynamics of antigen-specific B cells. *Cell Rep.* **35**, 109286 (2021).
45. G. D. Victora, M. C. Nussenzweig, Germinal centers. *Annu. Rev. Immunol.* **30**, 429–457 (2012).
46. J. S. Turner, W. Kim, E. Kalaidina, C. W. Goss, A. M. Rauseo, A. J. Schmitz, L. Hansen, A. Haile, M. K. Klebert, I. Pusic, J. A. O'Halloran, R. M. Presti, A. H. Ellebedy, SARS-CoV-2 infection induces long-lived bone marrow plasma cells in humans. *Nature* **595**, 421–425 (2021).

47. S. Boppana, K. Qin, J. K. Files, R. M. Russell, R. Stoltz, F. Bibollet-Ruche, A. Bansal, N. Erdmann, B. H. Hahn, P. A. Goepfert, SARS-CoV-2-specific circulating T follicular helper cells correlate with neutralizing antibodies and increase during early convalescence. *PLOS Pathog.* **17**, e1009761 (2021).
48. A. Sette, S. Crotty, Adaptive immunity to SARS-CoV-2 and COVID-19. *Cell* **184**, 861–880 (2021).
49. K. Basso, R. Dalla-Favera, Roles of BCL6 in normal and transformed germinal center B cells. *Immunol. Rev.* **247**, 172–183 (2012).
50. S. Crotty, Follicular helper CD4 T cells (TFH). *Annu. Rev. Immunol.* **29**, 621–663 (2011).
51. Y. Xu, X. Li, B. Zhu, H. Liang, C. Fang, Y. Gong, Q. Guo, X. Sun, D. Zhao, J. Shen, H. Zhang, H. Liu, H. Xia, J. Tang, K. Zhang, S. Gong, Characteristics of pediatric SARS-CoV-2 infection and potential evidence for persistent fecal viral shedding. *Nat. Med.* **26**, 502–505 (2020).
52. J. Grau-Expósito, N. Sánchez-Gaona, N. Massana, M. Suppi, A. Astorga-Gamaza, D. Perea, J. Rosado, A. Falcó, C. Kirkegaard, A. Torrella, B. Planas, J. Navarro, P. Suanzes, D. Álvarez-Sierra, A. Ayora, I. Sansano, J. Esperalba, C. Andrés, A. Antón, S. Ramón Y Cajal, B. Almirante, R. Pujol-Borrell, V. Falcó, J. Burgos, María J Buzón, M. Genescà, Peripheral and lung resident memory T cell responses against SARS-CoV-2. *Nat. Commun.* **12**, 3010 (2021).
53. G. Breton, P. Mendoza, T. Hagglof, T. Y. Oliveira, D. Schaefer-Babajew, C. Gaebler, M. Turroja, A. Hurley, M. Caskey, M. C. Nussenzweig, Persistent cellular immunity to SARS-CoV-2 infection. *J. Exp. Med.* **218**, e20202515 (2021).
54. N. Kaneko, H. H. Kuo, J. Boucau, J. R. Farmer, H. Allard-Chamard, V. S. Mahajan, A. Piechocka-Trocha, K. Lefteri, M. Osborn, J. Bals, Y. C. Bartsch, N. Bonheur, T. M. Caradonna, J. Chevalier, F. Chowdhury, T. J. Diefenbach, K. Einkauf, J. Fallon, J. Feldman, K. K. Finn, P. Garcia-Broncano, C. A. Hartana, B. M. Hauser, C. Jiang, P. Kaplonek, M. Karpell, E. C. Koscher, X. Lian, H. Liu, J. Liu, N. L. Ly, A. R. Michell, Y. Rassadkina, K. Seiger, L. Sessa, S. Shin, N. Singh, W. Sun, X. Sun, H. J. Ticheli, M. T. Waring, A. L. Zhu, G. Alter, J. Z. Li, D. Lingwood, A. G. Schmidt, M. Lichterfeld, B. D. Walker, X. G. Yu, R. F. Padera, S. Pillai; Massachusetts Consortium on Pathogen Readiness Specimen Working Group, Loss of Bcl-6-expressing T follicular helper cells and germinal centers in COVID-19. *Cell* **183**, 143–157.e13 (2020).
55. M. Sakharkar, C. G. Rappazzo, W. F. Wieland-Alter, C. L. Hsieh, D. Wrapp, E. S. Esterman, C. I. Kaku, A. Z. Wec, J. C. Geoghegan, J. S. McLellan, R. I. Connor, P. F. Wright, L. M. Walker, Prolonged evolution of the human B cell response to SARS-CoV-2 infection. *Sci. Immunol.* **6**, eabg6916 (2021).
56. T. Granot, T. Senda, D. J. Carpenter, N. Matsuoka, J. Weiner, C. L. Gordon, M. Miron, B. V. Kumar, A. Griesemer, S. H. Ho, H. Lerner, J. J. C. Thome, T. Connors, B. Reizis, D. L. Farber, Dendritic cells display subset and tissue-specific maturation dynamics over human life. *Immunity* **46**, 504–515 (2017).
57. B. V. Kumar, R. Kratchmarov, M. Miron, D. J. Carpenter, T. Senda, H. Lerner, A. Friedman, S. L. Reiner, D. L. Farber, Functional heterogeneity of human tissue-resident memory T cells based on dye efflux capacities. *JCI Insight* **3**, e123568 (2018).
58. M. Miron, B. V. Kumar, W. Meng, T. Granot, D. J. Carpenter, T. Senda, D. Chen, A. M. Rosenfeld, B. Zhang, H. Lerner, A. L. Friedman, U. Hershberg, Y. Shen, A. Rahman, E. T. Luning Prak, D. L. Farber, Human lymph nodes maintain TCF-1^{hi} Memory T cells with high functional potential and clonal diversity throughout life. *J. Immunol.* **201**, 2132–2140 (2018).
59. T. Senda, P. Dogra, T. Granot, K. Furuhashi, M. E. Snyder, D. J. Carpenter, P. A. Szabo, P. Thapa, M. Miron, D. L. Farber, Microanatomical dissection of human intestinal T-cell immunity reveals site-specific changes in gut-associated lymphoid tissues over life. *Mucosal Immunol.* **12**, 378–389 (2019).
60. P. Dogra, C. Rancan, W. Ma, M. Toth, T. Senda, D. J. Carpenter, M. Kubota, R. Matsumoto, P. Thapa, P. A. Szabo, M. M. Li Poon, J. Li, J. Arakawa-Hoyt, Y. Shen, L. Fong, L. L. Lanier, D. L. Farber, Tissue determinants of human NK cell development, function, and residence. *Cell* **180**, 749–763.e13 (2020).
61. A. Pugliese, M. Yang, I. Kusmarteva, T. Heiple, F. Vendrame, C. Wasserfall, P. Rowe, J. M. Moraski, S. Ball, L. Jebson, D. A. Schatz, R. Gianani, G. W. Burke, C. Nierras, T. Staeva, J. S. Kaddis, M. Campbell-Thompson, M. A. Atkinson, The Juvenile Diabetes Research Foundation Network for Pancreatic Organ Donors with Diabetes (nPOD) Program: Goals, operational model and emerging findings. *Pediatr. Diabetes* **15**, 1–9 (2014).
62. D. Weiskopf, K. S. Schmitz, M. P. Raadsen, A. Grifoni, N. M. A. Okba, H. Endeman, J. P. C. van den Akker, R. Molenkamp, M. P. G. Koopmans, E. C. M. van Gorp, B. L. Haagmans, R. L. de Swart, A. Sette, R. D. de Vries, Phenotype and kinetics of SARS-CoV-2-specific T cells in COVID-19 patients with acute respiratory distress syndrome. *Sci. Immunol.* **5**, eabd2071 (2020).
63. S. Reiss, A. E. Baxter, K. M. Cirelli, J. M. Dan, A. Morou, A. Daigneault, N. Brassard, G. Silvestri, J. P. Routy, C. Havenar-Daughton, S. Crotty, D. E. Kaufmann, Comparative analysis of activation induced marker (AIM) assays for sensitive identification of antigen-specific CD4 T cells. *PLOS ONE* **12**, e0186998 (2017).
64. F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, E. Duchesnay, Scikit-learn: Machine learning in Python. *JMLR* **12**, 2825–2830 (2011).
65. W. L. Waskom, seaborn: Statistical data visualization. *J. Open Source Soft.* **6**, 3021 (2021).
66. T. Wei, V. Simko, *R package "corrplot": Visualization of a correlation matrix (Version 0.89)* (2021).
67. J. D. Hunter, Matplotlib: A 2D graphics environment. *Comput. Sci. Eng.* **9**, 90–95 (2007).

Funding: This work was supported by NIH grants AI100119, AI128949 (Human Immunology Project Consortium) and AI106697 (to D.L.F.), AI142742 (Cooperative Centers for Human Immunology) (to A.S. and S.C.), NIH grant contract no. 75N9301900065 (to D.W. and A.S.), grant K23AI141686 (to T.J.C.), and a Helmsley Charitable Trust grant (to T.M.B. and D.L.F.). The CCTI Flow Cytometry Core was supported by NIH S10RR027050 and S10OD020056. P.D. was supported by a Cancer Research Institute (CRI) Irvington Postdoctoral Fellowship. N.L., was supported by a National Science Foundation Graduate Research Fellowship Program (NSF-GRFP). P.S. was supported by the Canadian Institutes of Health Research (CIHR) Fellowship. **Author contributions:** Conceptualization: M.M.L.P., K.R., Y.K., D.L.F., and S.C. Methodology: M.M.L.P., K.R., Y.K., N.I.B., and Z.Z. Investigation: M.M.L.P., K.R., Y.K., N.I.B., and Z.Z. Visualization: M.M.L.P., K.R., and Y.K. Resources: A.S., A.G., D.W., M.K., R.M., K.M.H., E.O.S., M.A.B., T.M.B., T.J.C., B.B.U., N.L., P.A.S., P.D., Y.S.L., J.I.G., and M.C.B. Supervision: D.L.F. and S.C. Writing (original draft): M.M.L.P., K.R., Y.K., D.L.F., and S.C. Writing (review and editing): M.M.L.P., K.R., Y.K., D.L.F., and S.C. Project administration: D.L.F., S.C., and S.B.W. Funding acquisition: D.L.F., S.C., and A.S. **Competing interests:** A.S. is a consultant for Gritstone, Flow Pharma, Merck, Epitogenesis, Gilead, and Avalia. S.C. has consulted for Avalia, Roche, and GSK. L.J.I. has filed for patent protection for various aspects of T cell epitope and vaccine design work. The other authors declare that they have no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions of the paper are available in the paper or the Supplementary Materials.

Submitted 12 August 2021
Accepted 4 October 2021
Published First Release 7 October 2021
Published 19 November 2021
10.1126/sciimmunol.abl9105

Science Immunology

SARS-CoV-2 infection generates tissue-localized immunological memory in humans

Maya M.L. PoonKsenia RybkinaYu KatoMasaru KubotaRei MatsumotoNathaniel I. BloomZeli ZhangKathryn M. HastieAlba GrifoniDaniela WeiskopfSteven B. WellsBasak B. UralNora LamPeter A. SzaboPranay DograYoon S. LeeJoshua I. GrayMarissa C. BradleyMaigan A. BruskoTodd M. BruskoErica O. SaphireThomas J. ConnorsAlessandro SetteShane CrottyDonna L. Farber

Sci. Immunol., 6 (65), eabl9105. • DOI: 10.1126/sciimmunol.abl9105

Science Immunology (ISSN) is published by the American Association for the Advancement of Science. 1200 New York Avenue NW, Washington, DC 20005. The title *Science Immunology* is a registered trademark of AAAS.

RESEARCH ARTICLE SUMMARY

CORONAVIRUS

Low-dose mRNA-1273 COVID-19 vaccine generates durable memory enhanced by cross-reactive T cells

Jose Mateus, Jennifer M. Dan†, Zeli Zhang†, Carolyn Rydyznski Moderbacher, Marshall Lammers, Benjamin Goodwin, Alessandro Sette*, Shane Crotty*, Daniela Weiskopf*

INTRODUCTION: Understanding human immune responses to severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) RNA vaccines is of interest for a panoply of reasons. mRNA vaccines have demonstrated impressive protection against COVID-19, but the durability of immunity has been a major unknown. Moreover, a better understanding of age-associated differences and mRNA vaccine dose response curves for dose sparing considerations is needed. Additionally, the impact of preexisting cross-reactive memory on immune responses to SARS-CoV-2 proteins remains an open question. Cross-reactive memory $CD4^+$ T cells recognizing SARS-CoV-2 have been found in ~50% of individuals. A vaccine trial is a controlled context for testing the relevance of such cross-reactive T cells. Each of these topics was addressed in this study using blood samples from a National Institutes of Health clinical trial of 25-µg mRNA-1273 COVID-19 vaccinees as well as from 100-µg mRNA-1273 COVID-19 vaccinees and SARS-CoV-2–infected individuals.

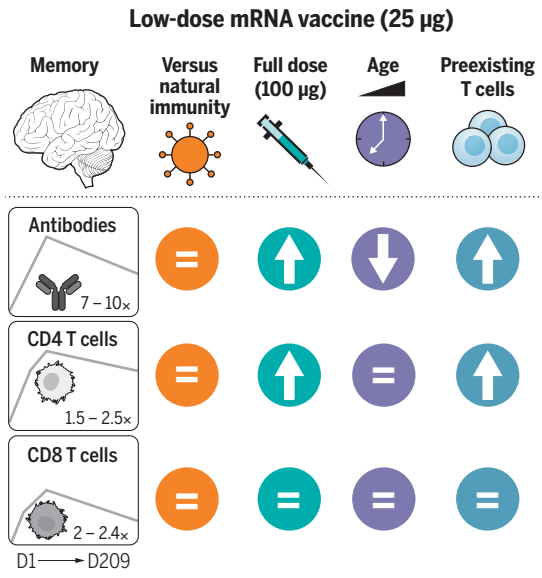
RATIONALE: Vaccination and infection are two different paths to immunity. Comparison of vaccine-generated and infection-generated immune memory is of value. Given evidence that antibodies, $CD4^+$ T cells, and $CD8^+$ T cells can each participate in protective immunity against COVID-19, we measured acute and memory SARS-CoV-2 spike–specific antibodies, $CD4^+$ T cells, and $CD8^+$ T cells in the blood of subjects who received a low-dose (25 µg) or standard-dose (100 µg) mRNA-1273 COVID-19 vaccine. Immunological measurements were used to address the four issues described above: namely, the durability of immune memory over 7 months after vaccination, mRNA vaccine dose responses, age differences, and the impact of preexisting cross-reactive T cells.

RESULTS: Longitudinal samples from 35 volunteers immunized with 25 µg of mRNA-1273 on days 1 and 29 were used to measure SARS-CoV-2 spike–binding antibodies, receptor binding domain (RBD)–binding antibodies, SARS-CoV-2 pseudovirus (PSV) neutralizing antibodies, spike-specific $CD4^+$ T cells, and spike-specific $CD8^+$ T cells. Overall, substantial anti-spike, anti-RBD, and PSV neutralizing antibodies were induced in response to two 25-µg mRNA-1273 vaccinations, were maintained in 88 to 100% of vaccinees for at least 6 months after the second immunization, and were comparable in magnitude and quality to those observed 6 to 7 months after infection with SARS-CoV-2.

Spike-specific $CD4^+$ T cells were generated by low-dose mRNA-1273 and were maintained as memory $CD4^+$ T cells. We observed strong T follicular helper (T_{FH}) and type 1 T helper cell polarization of these cells, which is advantageous for antiviral immunity. Spike-specific $CD8^+$ T cells were detectable in 88% of vaccinees and maintained for at least 6 months in 67% of vaccinees. Spike-specific $CD4^+$ or $CD8^+$ T cell frequencies were not lower in older vaccinee groups than in 18- to 55-year-olds, either in the acute or memory phase. Thus, 25-µg mRNA-1273 vaccination induced spike antibody levels and memory T cell frequencies at 7 months after vaccination similar to those observed for COVID-19 cases 7 months after symptom onset.

Next, to assess the impact of mRNA dosing, we compared immune responses between 25-µg and 100-µg doses of mRNA-1273 vaccine. Peak anti-spike, anti-RBD, and PSV neutralizing antibody levels were about twofold higher in 100-µg vaccinees than in 25-µg vaccinees. Spike-specific $CD4^+$ T cells responses were ~1.4-to-2.0-fold higher in 100-µg vaccinees, whereas peak $CD8^+$ T cell responses were comparable between 25-µg and 100-µg dose regimens.

Finally, to address potential positive or negative effects of preexisting cross-reactive memory T cells, we compared 25-µg mRNA-1273 COVID-19 vaccine responses between subjects with or without measurable preexisting SARS-CoV-2 spike–reactive memory $CD4^+$ T cells. Preexisting immunity enhanced vaccine antibody responses after a single vaccine dose, which was associated with higher spike-specific T_{FH} cells and total spike-specific $CD4^+$ T cell responses. Individuals with preexisting cross-reactive memory T cells also sustained higher SARS-CoV-2–neutralizing antibodies 6 months after vaccination.



Response to low-dose mRNA-1273 vaccination over 7 months. Immunological memory of antibodies, $CD4^+$ T cells, and $CD8^+$ T cells was examined after low-dose mRNA vaccination. Levels of spike-specific immune memory were then compared to immune memory observed after natural infection with SARS-CoV-2 or after full-dose vaccination. Robust immune memory comparable to natural infection but lower than after full-dose vaccination was observed. Increased vaccinee age correlated with reduced antibody levels but had no effect on cellular immune memory. Immune memory was enhanced by preexisting cross-reactive T cells. D1, day 1.

CONCLUSION: The 25-µg dose of mRNA-1273 vaccine induces durable and functional T cell and antibody memory at comparable magnitude to natural infection. This work expands our understanding of immune memory to mRNA vaccine in humans, vaccine dose sparing, and possible timing of boosters. Finally, these data provide evidence that cross-reactive memory $CD4^+$ T cells are biologically relevant and can exert a considerable positive influence on immunity generated by vaccination, with potential implications for vaccines and SARS-CoV-2 infections. ■

The list of author affiliations is available in the full article online.
*Corresponding author. Email: alex@lji.org (A.S.); shane@lji.org (S.C.); daniela@lji.org (D.W.)
†These authors contributed equally to this work.
Cite this article as J. Mateus *et al.*, *Science* **374**, eabj9853 (2021). DOI: 10.1126/science.abj9853

READ THE FULL ARTICLE AT
https://doi.org/10.1126/science.abj9853

RESEARCH ARTICLE

CORONAVIRUS

Low-dose mRNA-1273 COVID-19 vaccine generates durable memory enhanced by cross-reactive T cells

Jose Mateus[1], Jennifer M. Dan[1,2]†, Zeli Zhang[1]†, Carolyn Rydyznski Moderbacher[1], Marshall Lammers[1], Benjamin Goodwin[1], Alessandro Sette[1,2]*, Shane Crotty[1,2]*, Daniela Weiskopf[1]*

Vaccine-specific CD4$^+$ T cell, CD8$^+$ T cell, binding antibody, and neutralizing antibody responses to the 25-μg Moderna messenger RNA (mRNA)–1273 vaccine were examined over the course of 7 months after immunization, including in multiple age groups, with a particular interest in assessing whether preexisting cross-reactive T cell memory affects vaccine-generated immunity. Vaccine-generated spike-specific memory CD4$^+$ T cells 6 months after the second dose of the vaccine were comparable in quantity and quality to COVID-19 cases, including the presence of T follicular helper cells and interferon-γ–expressing cells. Spike-specific CD8$^+$ T cells were generated in 88% of subjects, with equivalent memory at 6 months post-boost compared with COVID-19 cases. Lastly, subjects with preexisting cross-reactive CD4$^+$ T cell memory exhibited stronger CD4$^+$ T cell and antibody responses to the vaccine, demonstrating the biological relevance of severe acute respiratory syndrome coronavirus 2–cross-reactive CD4$^+$ T cells.

Understanding human immune responses to severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) RNA vaccines is of interest for a panoply of reasons. mRNA vaccines have demonstrated impressive protection against COVID-19 (*1*–*7*). The COVID-19 vaccine mRNA-1273, developed as a collaboration between Moderna and the National Institutes of Health Vaccine Research Center, encodes a stabilized SARS-CoV-2 full-length spike (*8*, *9*). Durability of immunity has been, and remains, a major unknown for mRNA vaccines in humans. Encouraging reports from both Pfizer-BioNTech and Moderna indicate protective immunity of 91 and 93%, respectively, over the 6-month period after the second immunization (7 months after the first immunization) (*10*, *11*), down modestly from the 95% maximal protection observed for each of those two vaccines within 1 to 2 months after the second immunization (*1*, *2*). Although neutralizing antibodies are a clear correlate of immunity after two immunizations (*12*), the underlying mechanisms remain unclear. Moreover, those mechanisms of immunity may change as the immune response develops (e.g., after a single immunization) or as immune memory changes composition (*13*–*15*). Direct measurements of immune memory compartments in humans are necessary to provide insights into these important topics.

Infection and vaccination are two different paths to immunity. Comparison of vaccine-generated immune memory with immune memory of persons infected with SARS-CoV-2 is of value, as studies have indicated that natural immunity is 93 to 100% protective against symptomatic reinfection for 7 to 8 months (*16*–*19*), although natural immunity protection against certain variants of concern (VOCs) is likely to be lower (*20*). After SARS-CoV-2 infection, immunological memory has been observed for ≥8 months for CD4$^+$ T cells, CD8$^+$ T cells, memory B cells, and antibodies (*21*, *22*). The immune memory in response to SARS-CoV-2 infection exhibits a relatively gradual decline that partially stabilizes within 1 year (*23*–*26*). The 100-μg mRNA-1273 vaccination has been shown to induce durable antibody responses (*27*), but it is unknown whether immune memory to the mRNA-1273 vaccine months after immunization is similar to or different than memory generated by SARS-CoV-2 infection. Additionally, both 25-μg- and 100-μg-dose mRNA-1273 vaccinations have been tested in clinical trials (*9*, *28*), with 100-μg mRNA-1273 proceeding toward licensure (*2*, *29*).

Preexisting cross-reactive memory CD4$^+$ T cells that recognize SARS-CoV-2 have been found in ~50% of individuals pre-pandemic (*30*–*37*). There has been intense interest in understanding whether these preexisting cross-reactive memory CD4$^+$ T cells, identified in vitro, are biologically relevant in vivo (*33*, *38*, *39*). One approach to test the relevance of such T cells in a controlled fashion is in the context of a vaccine trial, as individuals in a clinical trial are all exposed to a well-defined dose of antigen at a specific time. Additionally, exposure to a low antigen dose may be more sensitive to influence by cross-reactive memory. Thus, we examined immune responses to the 25-μg dose of the mRNA-1273 COVID-19 vaccine.

[1]Center for Infectious Disease and Vaccine Research, La Jolla Institute for Immunology (LJI), La Jolla, CA 92037, USA.
[2]Department of Medicine, Division of Infectious Diseases and Global Public Health, University of California, San Diego (UCSD), La Jolla, CA 92037, USA.
*Corresponding author. Email: alex@lji.org (A.S.); shane@lji.org (S.C.); daniela@lji.org (D.W.)
†These authors contributed equally to this work.

Spike-specific antibody elicited by the 25-μg mRNA-1273 vaccine dose over time

An open-label, age de-escalation phase 1 trial used the mRNA-1273 vaccine with 25-μg immunizations on days 1 and 29 (*9*, *28*), with blood samples collected on study days 1, 15, 43, and 209. SARS-CoV-2 spike–binding antibodies, receptor binding domain (RBD)–binding antibodies, and SARS-CoV-2 pseudovirus (PSV) neutralization titers were determined (Fig. 1). Anti-spike and anti-RBD immunoglobulin G (IgG) were maintained at detectable levels for at least 7 months after the first vaccination in 100% (33/33) of subjects (Fig. 1, A and B). RBD IgG was induced by one immunization in 94% (33/35) of subjects. This response rate increased to 100% (33/33) after the second immunization and was maintained for at least 6 months after the second vaccination. SARS-CoV-2 PSV neutralization titers were detected in 29% (10/35) of subjects after one vaccination and 100% of subjects after two vaccinations (33/33), and 88% (29/33) of subjects maintained detectable neutralizing antibodies for at least 6 months after the second vaccination (Fig. 1C). All three antibody measurements demonstrated similar kinetics (Fig. 1, A to C) and were highly correlated (correlation coefficient, r = 0.89 to 0.90, fig. S1). Peak spike IgG, RBD IgG, and PSV titers were 6.8-, 9.5-, and 9.5-fold higher, respectively, than at 7 months (study day 209; 181 days after the second immunization). Similar fold changes were reported for 100-μg mRNA-1273 vaccination, indicating similar memory quality and durability (*40*). The 25-μg mRNA-1273 vaccine–generated antibodies were comparable to antibodies from SARS-CoV-2–infected subjects collected at a similar time after exposure [7 months post–symptom onset (PSO), 170 to 195 days] (Fig. 1D). Thus, increased anti-spike IgG, anti-RBD IgG, and PSV neutralizing antibodies were induced in response to two 25-μg mRNA-1273 vaccinations. These levels were maintained in 88 to 100% of vaccinees for at least 6 months after the second immunization and were comparable to those observed after infection with SARS-CoV-2.

Spike-specific CD4$^+$ T cells elicited by the 25-μg mRNA-1273 vaccine dose over time

SARS-CoV-2 spike–specific CD4$^+$ T cell responses were first measured using a flow cytometry activation-induced marker (AIM) assay (Fig. 2A and fig. S2). On day 1, before vaccination, spike-specific CD4$^+$ T cells with a predominantly memory phenotype were detected in 49% of clinical trial subjects (17/35), demonstrating the presence of preexisting SARS-CoV-2 spike–cross-reactive memory CD4$^+$ T cells, as discussed in the latter part of this Report. Spike-specific CD4$^+$ T cell responses were observed after the first vaccination in 97% of subjects (34/35) (Fig. 2A). Cytomegalovirus (CMV)–specific CD4$^+$ T cells were unchanged, as

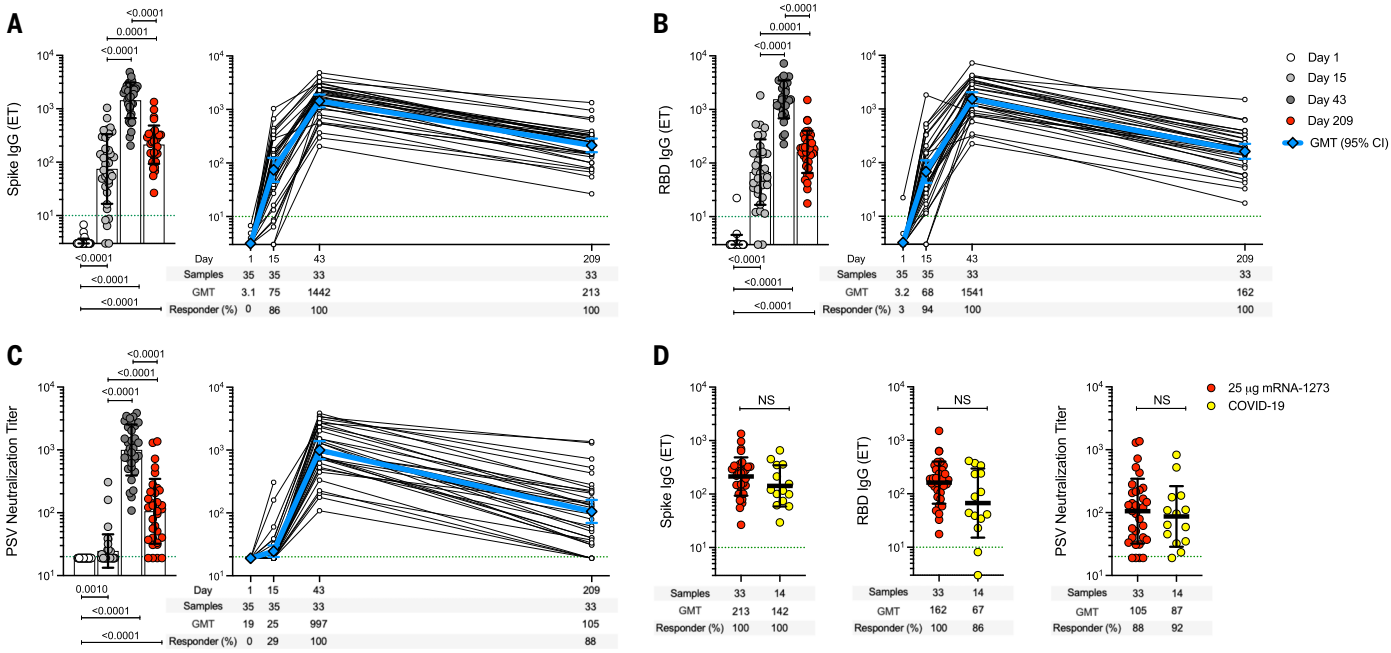


Fig. 1. Spike antibodies induced by 25-µg mRNA-1273 vaccination. Participants received two injections of the 25-µg mRNA-1273 vaccine, 28 days apart. PBMC samples were collected on day 1, day 15 ± 2 (2 weeks after first dose), day 43 ± 2 (2 weeks after second dose), and day 209 ± 7 days (6 months after second dose). (**A**) Longitudinal anti-SARS-CoV-2 spike IgG binding titers, (**B**) longitudinal anti-SARS-CoV-2 RBD IgG binding titers, and (**C**) longitudinal SARS-CoV-2 spike pseudovirus neutralizing titers (PSV). (**D**) Comparison of anti-spike IgG, anti-RBD IgG, and PSV neutralizing titers induced by two doses of 25-µg mRNA-1273 vaccine at day 209 ± 7 (n = 33) and COVID-19 convalescent donors at 170 to 195 days PSO (n = 14). Dotted green lines indicate the limit of quantification (LOQ). The bars in (A), (B), (C), and (D) indicate the geometric mean titers (GMTs) and geometric SD for anti-spike IgG (endpoint titer, ET), anti-RBD IgG (ET), and PSV neutralizing titers, respectively. Data were analyzed for statistical significance using Wilcoxon signed-rank test [(A), (B), and (C)] and Mann-Whitney U test (D). NS, nonsignificant. Background-subtracted and log data analyzed in all cases.

expected, indicating no bystander influence of the mRNA-1273 vaccination (fig. S3). The SARS-CoV-2 spike–specific $CD4^+$ T cell response rate increased to 100% (32/32) after the second vaccination and was maintained for at least 6 more months. Spike-specific memory $CD4^+$ T cell frequencies at 7 months were similar to those observed for COVID-19 cases (COVID-19 samples collected 170 to 195 days PSO) (Fig. 2B). Median mRNA-1273 vaccine–generated spike-specific $CD4^+$ T cell frequencies at all time points after vaccination also exceeded CMV-specific $CD4^+$ T cell frequencies (Fig. 2A and fig. S3A).

T follicular helper (T_{FH}) cell differentiation and cytokine production by vaccine-generated spike-specific $CD4^+$ T cells were then assessed (Fig. 2, C to F). T_{FH} cells are the specialized subset of $CD4^+$ T cells required for B cell help and are critical for the generation of neutralizing antibodies in most conditions (*41*). Spike-specific circulating T_{FH} (cT_{FH}) cells were detected in 71% (25/35) and 75% (24/32) of subjects after the first and second vaccination, respectively (Fig. 2C, right panel). Spike-specific cT_{FH} cells were detectable in 94% of subjects overall (33/35). Different response kinetics were observed at the level of individual subjects (Fig. 2C, right panel). Spike-specific memory cT_{FH} cells were still detected in 63% of vaccinees 6 months after the second vaccination (20/32) (Fig. 2C). Vaccine-specific $CD4^+$ T cell cytokine profiles were determined by intracellular cytokine staining (ICS) (Fig. 2, D to F, and fig. S4). Interferon-γ positive ($IFN\gamma^+$) spike-specific $CD4^+$ T cells were detected in 85% (28/33), tumor necrosis factor–α positive ($TNF\alpha^+$) in 97% (32/33), interleukin-2 positive ($IL\text{-}2^+$) in 100% (33/33), and granzyme B positive (GzB^+) in 76% (25/33) of subjects at day 43 (Fig. 2D). Little-to-no IL-4, IL-17A, or IL-10 was detected (fig. S5). Cytokine-producing spike-specific $CD4^+$ T cells ($CD40L^+$ cells producing IFNγ, TNFα, IL-2, and/or GzB) (Fig. 2E and fig. S6) were observed in 94% (33/35) and 100% (33/33) of subjects after the first and second vaccination, respectively, and were maintained for at least 6 months after the second vaccination [97% (32/33)] (Fig. 2E). Spike-specific $CD4^+$ T cells generated by the 25-µg mRNA-1273 vaccine exhibited multifunctionality comparable with that of CMV-specific cells (Fig. 2F, fig. S3C, and table S1). Thus, robust spike-specific $CD4^+$ T cells and T cell memory were generated by the low-dose mRNA-1723 vaccine, with strong T_{FH} and T helper 1 cell polarization advantageous for antiviral immunity.

Spike-specific $CD8^+$ T cells elicited by the 25-µg mRNA-1273 vaccine dose over time

SARS-CoV-2 spike–specific $CD8^+$ T cells were measured by AIM ($CD69^+$ and $CD137^+$) (fig. S2) and were observed in 34% (12/35) and 53% (17/32) of subjects after the first and second 25-µg mRNA-1273 vaccination, respectively (Fig. 3A). Spike-specific $CD8^+$ T cells were detectable for >6 months after the second vaccination, with a response rate comparable to that observed for COVID-19 cases (COVID-19 samples collected 170 to 195 days PSO) (Fig. 3B). Next, SARS-CoV-2 spike–specific $CD8^+$ T cells were measured by ICS (IFNγ, TNFα, IL-2, or GzB) (Fig. 3C and fig. S4). The first immunization induced significant spike-specific $CD8^+$ T cell cytokine responses in 51% (18/35) of subjects (Fig. 3D and fig. S7), increasing to 70% (23/33) of subjects after the second vaccination (Fig. 3D). $IFN\gamma^+$ spike-specific $CD8^+$ T cells were detected in 70% (23/33), $TNF\alpha$ in 39% (13/33), and $IL\text{-}2^+$ in 12% (4/33) of vaccinees at day 43 (Fig. 3C). Multiple positive- and negative-control samples and experimental conditions were used to demonstrate the specificity of the spike $CD8^+$ T cells (Fig. 3 and figs. S2 and S4, F and G). Correlation between AIM and ICS methods was highly significant for both spike-specific $CD4^+$ and $CD8^+$ T cells ($P < 0.0001$) (fig. S8). The fraction of multifunctional spike-specific $CD8^+$ T cells increased between first and second vaccination (three or more effector molecules expressed) (Fig. 3E and table S2). The most prevalent profile of $CD8^+$ T cells with

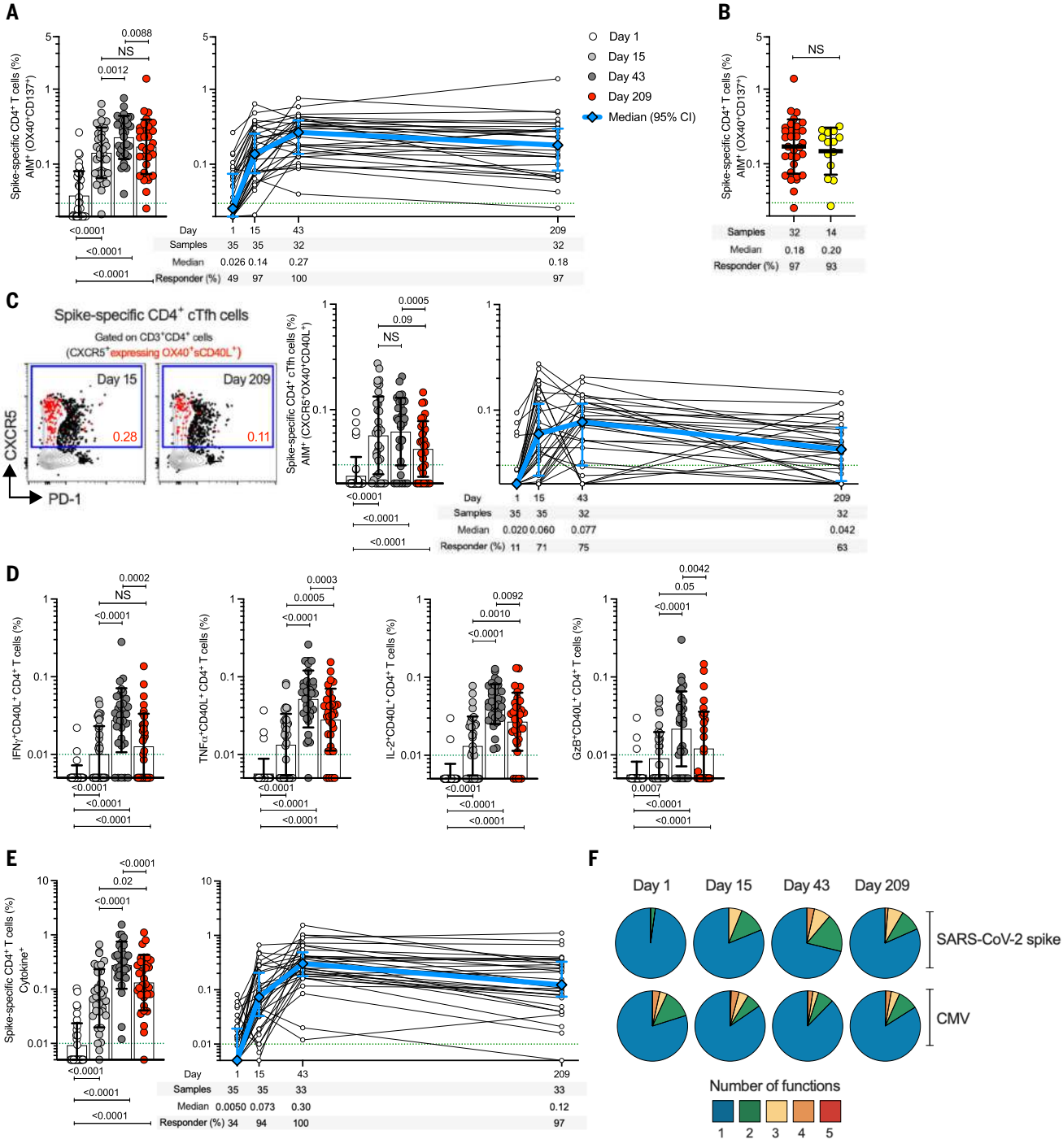


Fig. 2. mRNA-1273 vaccination induces durable and multifunctional spike-specific CD4$^+$ T cell responses. (**A**) Longitudinal spike-specific CD4$^+$ T cells in mRNA-1273 vaccinees measured by AIM. Spike-specific CD4$^+$ T cells quantified by AIM (surface OX40$^+$CD137$^+$) after stimulation with spike megapool (MP) in mRNA-1273 vaccinees (see fig. S2 for gating strategy). (**B**) Comparison of spike-specific AIM$^+$ CD4$^+$ T cell frequencies between 25-μg mRNA-1273 vaccine at day 209 ± 7 (red circles, n = 32) and COVID-19 convalescent donors at 170 to 195 days PSO (yellow circles, n = 14). (**C**) Quantitation of spike-specific circulating T follicular helper (cT$_{FH}$) cells (CXCR5$^+$OX40$^+$surface CD40L$^+$, as percentage of CD4$^+$ T cells) after stimulation with spike MP. Representative examples of spike-specific cT$_{FH}$ cells (red), overlaid on total CD4$^+$ T cells, at days 15 ± 2 and 209 ± 7. (**D**) Spike-specific CD4$^+$ T cells expressing intracellular CD40L (iCD40L) and producing IFNγ, TNFα, IL-2, or granzyme B (GzB) in mRNA-1273 vaccinees. (**E**) Longitudinal spike-specific CD4$^+$ cytokine$^+$ T cells expressing iCD40L or producing IFNγ, TNFα, IL-2, or GzB in 25-μg mRNA-1273 vaccinees (see fig. S4 for gating strategy). Dotted green lines indicate the limit of quantification (LOQ). White, day 1; light gray, day 15 ± 2; dark gray, day 43 ± 2; red, day 209 ± 7. The bars in (A) to (E) indicate the geometric mean and geometric SD in the analysis of the spike-specific CD4$^+$ T cell frequencies. (**F**) Longitudinal multifunctional spike-specific CD4$^+$ T cells in mRNA-1273 vaccinees. Proportions of multifunctional activity profiles of the spike-specific CD4$^+$ T cells from mRNA-1273 vaccinees evaluated on days 1, 15 ± 2, 43 ± 2, and 209 ± 7. The blue, green, yellow, orange, and red colors in the pie charts depict the production of one, two, three, four, or five functions, respectively (see figs. S4 and S6 for details). Data were analyzed for statistical significance using Wilcoxon signed-rank test [(A), (C), (D), and (E)] and Mann-Whitney U test (B). Background-subtracted and log data analyzed in all cases.

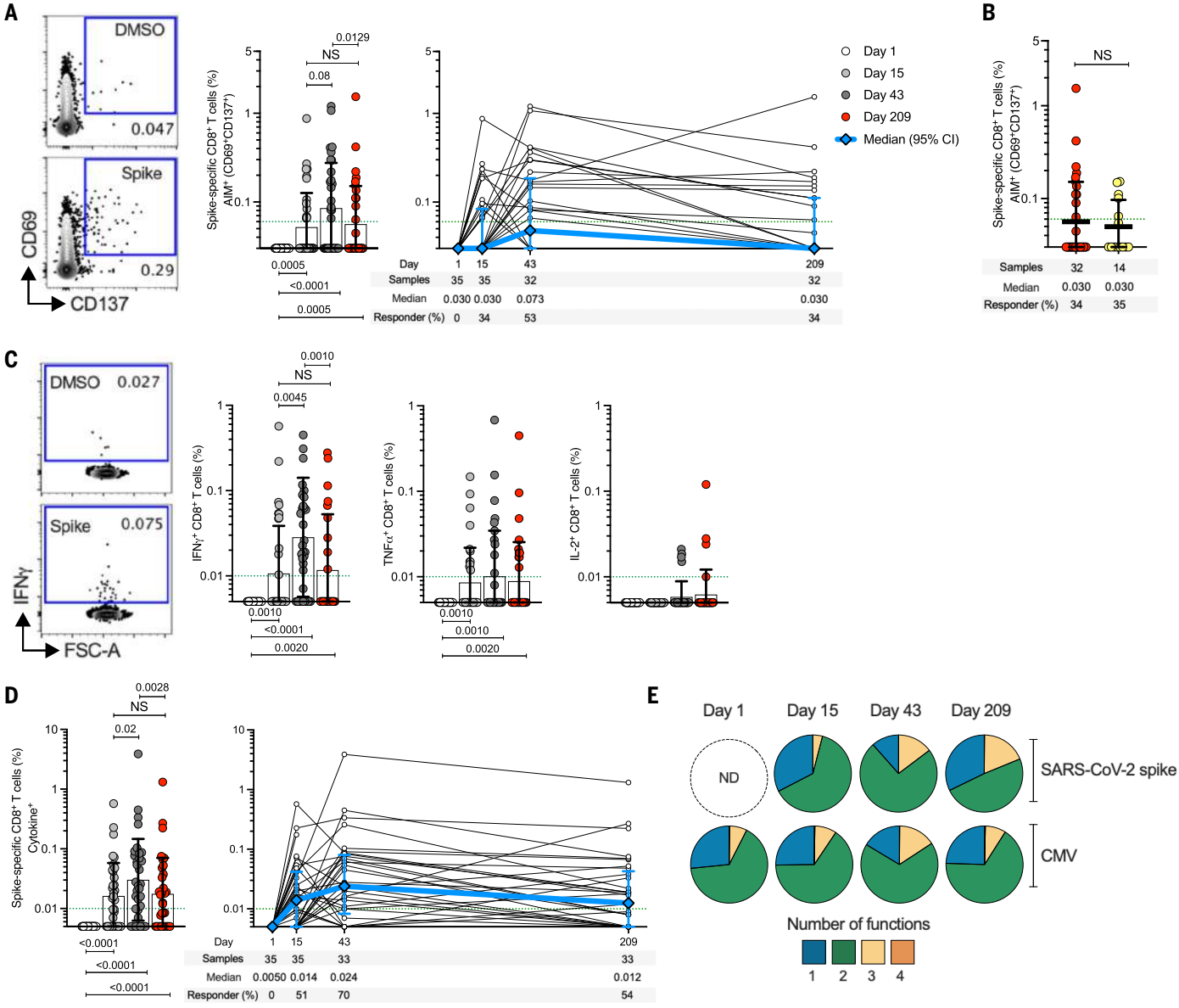


Fig. 3. mRNA-1273 vaccination induces multifunctional spike-specific CD8⁺ T cells. (**A**) Longitudinal spike-specific CD8⁺ T cells in mRNA-1273 vaccinees measured by AIM (surface CD69⁺CD137⁺). (Left panel) Representative examples of flow cytometry plots of spike-specific CD8⁺ T cells compared with DMSO control (see fig. S2 for gating strategy). (Right panel) Spike-specific CD8⁺ T cells quantified. (**B**) Comparison of spike-specific AIM⁺ CD8⁺ T cell frequencies between 25-μg mRNA-1273 vaccinees at day 209 ± 7 (n = 32) and COVID-19 convalescent donors at 170 to 195 days PSO (n = 14). (**C**) Spike-specific CD8⁺ T cells producing IFNγ, TNFα, or IL-2 by intracellular cytokine staining (ICS) in 25-μg mRNA-1273 vaccinees. (**D**) Longitudinal spike-specific CD8⁺ cytokine⁺ T cells producing IFNγ, TNFα, IL-2, or GzB in 25-μg mRNA-1273 vaccinees (see fig. S4 for gating strategy). Dotted green lines indicate the LOQ. The bars in (A) to (D) indicate the geometric mean and geometric SD. White, day 1; light gray, day 15 ± 2; dark gray, day 43 ± 2; red, day 209 ± 7. (**E**) Multifunctional activity profiles of spike-specific CD8⁺ T cells from 25-μg mRNA-1273 vaccinees, evaluated for IFNγ, TNFα, IL-2, or GzB (see figs. S4 and S8 for details). The blue, green, yellow, and orange colors in the pie charts depict the production of one, two, three, or four functions, respectively. ND, nondetectable. Data were analyzed for statistical significance using Wilcoxon signed-rank test [(A), (C), and (D)] and Mann-Whitney U test (B). Background-subtracted and log data analyzed in all cases.

three functions was GzB⁺IFNγ⁺TNFα⁺ (fig. S7), similar to the profile seen in CMV-specific CD8⁺ T cells (Fig. 3D and fig. S9). Thus, 25-μg mRNA-1273 vaccination induces multifunctional spike-specific memory CD8⁺ T cells.

Anti-spike antibody and CD4⁺ and CD8⁺ T cell responses generated by 25-μg mRNA-1273 vaccination were multifunctional, durable, and comparable in magnitude to those induced by natural infection (Table 1) (*22*). A concern has been raised that vaccination may not induce adequate immune memory in the elderly (*42*). Our vaccination cohort consisted of volunteers from three different age groups (*9*, *28*). Spike IgG and RBD IgG in the older groups (56–70 and >70 years) were reduced about twofold on day 209 (Fig. 4, A and B), similar to what was reported for 100-μg mRNA-1273 vaccination (*40*). Spike-specific CD4⁺ or CD8⁺ T cells were not reduced in the older vaccinee groups compared with the 18- to 55-year-old age group. Memory CD4⁺ and CD8⁺ T cell frequencies at day 209 were at least as strong in the older age groups as in younger adults (Fig. 4, D and H). Thus, although the study size is underpowered for in-depth examination of the three age groups, a small reduction in antibody but not T cell

Table 1. Spike-specific immune responses detected in 25-μg mRNA-1273 vaccinees. ELISA, enzyme-linked immunosorbent assay; AIM, activation-induced markers; ICS, intracellular cytokine staining.

Component	Assay	Days after vaccination			
		1	15 ± 2	43 ± 2	209 ± 7
Antibodies					
Anti-spike IgG	ELISA	0	86%	100%	100%
Anti-RBD IgG	ELISA	3%	94%	100%	100%
Neutralizing	Neutralization	0	29%	100%	88%
T cells					
Spike-specific $CD4^+$ T cells	AIM*	49%	97%	100%	97%
	ICS†	34%	94%	100%	97%
	Total‡	49%	97%	100%	97%
Spike-specific $CD8^+$ T cells	AIM*	0	34%	53%	34%
	ICS†	0	51%	70%	54%
	Total‡	0	69%	88%	67%

*Antigen-specific T cells using AIM were measured as a percentage $CD4^+$ T expressing $OX40^+CD137^+$ and $CD8^+$ T cells expressing $CD69^+CD137^+$ after stimulation of PBMC with spike overlapping peptides spanning the entire protein. †Antigen-specific T cells using ICS were measured as a percentage $CD4^+$ T expressing CD40L or producing IFNγ, TNFα, IL-2, or GzB; and $CD8^+$ T cells producing IFNγ, TNFα, IL-2, or GzB after stimulation of PBMC with spike overlapping peptides spanning the entire protein. ‡The overall spike-specific T cell response was calculated on the basis of the AIM and ICS results per donor and time point.

memory was observed in older adults compared with younger adults.

Spike-specific immune responses elicited by the 100-μg mRNA-1273 vaccine dose

More than 100 million doses of the 100-μg mRNA-1273 vaccine have been administered in the United States to date. We compared immune responses to the 25-μg and 100-μg doses of mRNA-1273 (Fig. 4, I to K). Anti-spike IgG, anti-RBD IgG, and PSV neutralizing titers were about twofold higher in 100-μg vaccinees than in those who received the 25-μg dose (Fig. 4I), which is consistent with earlier reports (*9*, *28*, *40*). Spike-specific $CD4^+$ T cells responses were ~1.4- to 2.0-fold higher in 100-μg vaccinees than in 25-μg vaccinees (Fig. 4J). Furthermore, the spike-specific $CD8^+$ T cell responses were comparable between the doses (Fig. 4K).

Preexisting cross-reactive memory

Preexisting cross-reactive memory T cells recognizing SARS-CoV-2 in vitro are found in many individuals (*30*–*37*). It was hypothesized that the existence of preexisting spike-recognizing immune memory may modulate immune responses to infection or vaccination (*43*). To address this question, we separated our cohort as a function of whether each subject had preexisting cross-reactive memory $CD4^+$ T cells reactive against SARS-CoV-2 spike (Fig. 5, A and B). As noted above (Fig. 2A), preexisting SARS-CoV-2 spike–specific $CD4^+$ T cells were present in 49% (17/35) of the 25-μg mRNA-1273 vaccinees. After vaccination, spike-specific $CD4^+$ T cells were significantly higher at day 15 in subjects with cross-reactive memory than in subjects with no cross-reactive memory (2.3-fold, $P < 0.0001$) (Fig. 5C). Spike-specific memory $CD4^+$ T cell frequencies were also higher after the second vaccination in subjects with cross-reactive memory than in those without ($P = 0.02$) (Fig. 5) and remained higher for ≥6 months ($P = 0.01$) (Fig. 5C). The impact of preexisting cross-reactive spike-specific $CD4^+$ T cell memory was more than additive (day 15, $P = 0.018$) (fig. S13), demonstrating that the spike-specific $CD4^+$ T cell response to the vaccine was enhanced by cross-reactive memory. Higher frequencies of cytokine-positive spike-specific $CD4^+$ T cells ($P = 0.0051$) (Fig. 5D) and multifunctional cells ($P = 0.02$) (Fig. 5E and fig. S10) were also observed after the first vaccination in individuals with cross-reactive memory. Preexisting cross-reactive $CD4^+$ T cells were observed in all three age groups (Fig. 4D). We did not detect preexisting cross-reactive spike-specific $CD8^+$ T cell memory (Figs. 3, A and D, and 4, G and H), and we observed no modulation of vaccine $CD8^+$ T cell responses by preexisting $CD4^+$ T cell memory (fig. S11). Cross-reactive memory $CD4^+$ T cells recognizing SARS-CoV-2 nonspike epitopes were present (fig. S12), as expected (*31*, *34*). The nonspike cross-reactive memory $CD4^+$ T cell frequencies remained unchanged over 7 months and were not modulated by mRNA-1273 vaccination (fig. S12), consistent with the vaccine containing only spike antigen and not causing bystander activation. Thus, preexisting cross-reactive $CD4^+$ T cell memory can influence mRNA-1723 vaccine–generated $CD4^+$ T cell responses.

T_{FH} cells in subjects with and without cross-reactive memory were of particular interest, because of their relevance in antibody responses. Frequencies of spike-specific cT_{FH} cells 2.6-fold higher were observed on day 15 in vaccinees with preexisting cross-reactive memory ($P = 0.0023$) (Fig. 5F). Likewise, significantly higher levels of anti-spike IgG ($P = 0.02$) (Fig. 5G) and anti-RBD IgG ($P = 0.047$) (Fig. 5H) were detected on day 15 in vaccinees with preexisting cross-reactive memory. The group with preexisting cross-reactive $CD4^+$ T cell memory demonstrated higher SARS-CoV-2 neutralizing titers 7 months after vaccination than did the group without preexisting cross-reactive $CD4^+$ T cell memory ($P = 0.04$) (Fig. 5I). Thus, a coordinated increase in spike-specific cT_{FH} responses, anti-spike IgG, and anti-RBD IgG are detected after a single vaccination in subjects with preexisting cross-reactive memory. Furthermore, 6 months after the second vaccination, higher levels of vaccine-specific memory $CD4^+$ T cells and higher titers of SARS-CoV-2 PSV neutralizing antibodies are present in individuals who have preexisting cross-reactive $CD4^+$ T cell memory.

Concluding remarks

The SARS-CoV-2 mRNA vaccines have been extraordinary successes. It is important to better understand the immunology of these vaccines in order to better appreciate (i) the mechanisms of protective immunity provided by the vaccines, (ii) the durability of immune memory generated by the vaccines (and thus infer trajectories of protective immunity), and (iii) the immunological features of these vaccines that may be relevant for vaccine design against other pathogens. Here, studying 35 vaccinated subjects 7 months out from the initial immunization, we found that two-dose 25-μg mRNA-1273 vaccination generated immune memory against spike comparable to that of SARS-CoV-2 infection for antibodies, $CD4^+$ T cells, and $CD8^+$ T cells. Furthermore, immune responses were considerably enhanced by the presence of preexisting cross-reactive $CD4^+$ T cell memory.

We consistently found spike-specific memory $CD4^+$ T cells in vaccinated subjects 6 months after second-dose 25-μg mRNA-1273 immunization. Less than a twofold difference in spike-specific $CD4^+$ T cell frequencies was observed between peak and 6 months after the second vaccination, indicative of durable vaccine T cell memory. Spike-specific memory $CD4^+$ T cell frequencies were also similar between low-dose mRNA-1273 vaccinated persons and COVID-19 cases. The vaccine-generated cells also exhibited an antiviral functional profile, including substantial T_{FH} cell counts and IFNγ expression, and the presence of multi-cytokine-expressing cells in proportions similar to CMV-specific memory $CD4^+$ T cells.

Uncertainty has surrounded whether the mRNA-1273 vaccine elicits effector and memory $CD8^+$ T cells in humans (*9*, *44*, *45*). Here

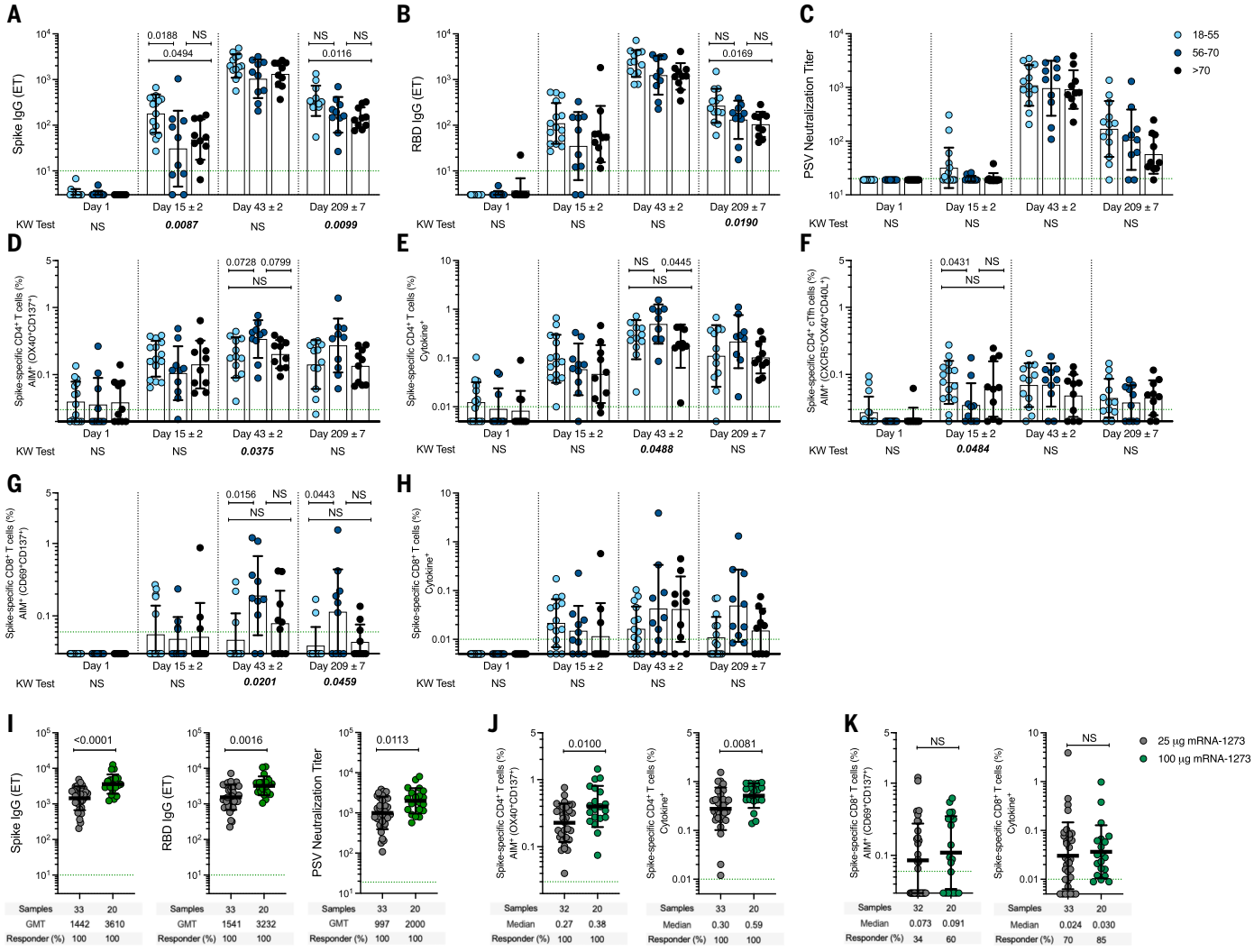


Fig. 4. Spike-specific antibody and T cell responses induced by mRNA-1273 vaccination. (**A** to **H**) Immune responses to 25-µg mRNA-1273 vaccination in three adult age groups: 18–55 (light blue symbols), 56–70 (dark blue), and over 70 years of age (black). [(A) to (C)] Anti-spike IgG, anti-RBD IgG, and PSV neutralizing titers. [(D) and (E)] Spike-specific $CD4^+$ T cells by AIM (D) or ICS (E). (F) Spike-specific cT_{FH} cells. [(G) and (H)] Spike-specific $CD8^+$ T cells by AIM (G) or ICS (H). Data were analyzed for statistical significance using nonparametric analysis of variance Kruskal-Wallis (KW) test and Dunn's post-test for multiple comparisons. The *P* values plotted on the bottom show the KW test results, and the *P* values plotted on the top show the post-test analysis comparing age groups. (**I**) Anti-SARS-CoV-2 spike IgG, RBD IgG, and PSV neutralizing titers in 25-µg and 100-µg mRNA-1273 vaccinees at day 43 (2 weeks after second dose). (**J**) Spike-specific $CD4^+$ AIM^+ (left) and $cytokine^+$ (right) T cells in mRNA-1273 vaccinees. (**K**) Spike-specific $CD8^+$ AIM^+ (left) and $cytokine^+$ (right) T cells in mRNA-1273 vaccinees. Dotted green lines indicate the limit of quantification (LOQ). The bars indicate geometric mean and geomean SD. Data in (I) to (K) were analyzed for statistical significance using Mann-Whitney *U* test. Background-subtracted and log data analyzed in all cases.

we report, at the peak of the immune response, spike-specific $CD8^+$ T cells by AIM or ICS assays were detected in 88% of subjects receiving low-dose mRNA-1273, which is a $CD8^+$ T cell response rate equivalent to that of COVID-19 cases (*21*, *22*, *26*, *31*). We speculate that absence of detection of spike-specific $CD8^+$ T cells in some studies reflects the stringency of the experimental conditions used. Here, allowance for 24 hours of antigen stimulation revealed vaccine-generated $CD8^+$ T cells in most individuals. $CD8^+$ T cells have been observed in response to the Pfizer-BioNTech BNT162b2 vaccine using peptide–major histocompatibility complex multimers (*44*). Moreover, $CD8^+$ T cells have been found 2 months after second immunization with BNT162b2 (*44*). In this study, spike-specific memory $CD8^+$ T cells were detected 6 months after the second immunization with 25-µg mRNA-1273. These mRNA-1273 vaccine–generated memory $CD8^+$ T cells were detected in 67% of subjects and were not dissimilar in magnitude to spike-specific memory $CD8^+$ T cells in COVID-19 cases. Limitations of this study include the relatively small sample size and limited cell availability. Overall, the data show that $CD4^+$ and $CD8^+$ T cell memory are generated by both low-dose and 100-µg dose COVID-19 mRNA-1273 vaccine.

Low-dose RNA vaccines have potential advantages for future needs and applications such as dose sparing. Low-dose immunization is also less reactogenic (*9*, *44*), which may also be appealing in contexts of multidose regimens. It is of interest to consider different vaccine doses across age groups, or high- versus low-risk groups, but a better understanding of immune memory to different doses is key for such considerations. Data reported here are encouraging demonstrations of the potential of RNA vaccines to generate durable T cell and antibody immune memory, including at lower vaccine doses.

Preexisting immunity in the form of cross-reactive memory $CD4^+$ T cells affected immune

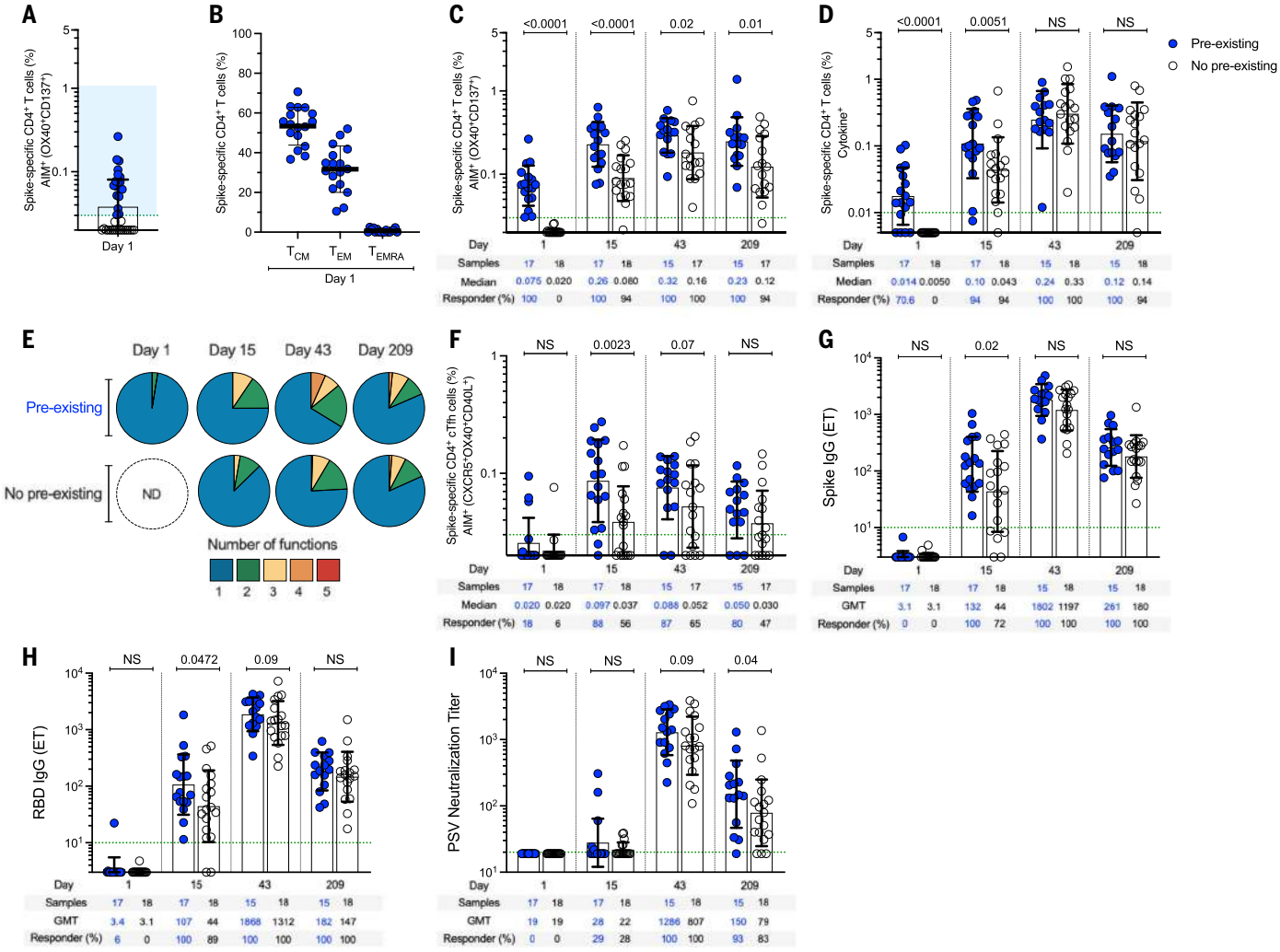


Fig. 5. Preexisting anti-spike immunity modulates T cell and antibody responses. (**A**) Preexisting spike-specific CD4$^+$ AIM$^+$ T cells at day 1 (see Fig. 2 for details). (**B**) Memory phenotype of preexisting spike-specific CD4$^+$ AIM$^+$ T cells from (A). T_{CM}, central memory cells; T_{EM}, effector memory cells; T_{EMRA}, terminally differentiated effector memory cells. (**C**) Spike-specific CD4$^+$ AIM$^+$ T cells in mRNA-1273 vaccinees with ("preexisting," blue) and without ("no preexisting," white) preexisting cross-reactive spike-reactive memory CD4$^+$ T cells evaluated on days 1, 15 ± 2, 43 ± 2, and 209 ± 7 after immunization (see fig. S13 for details). (**D**) Spike-specific CD4$^+$ cytokine$^+$ T cells in mRNA-1273 vaccinees with and without preexisting cross-reactive memory CD4$^+$ T cells. (**E**) Proportions of multifunctional spike-specific CD4$^+$ T cells in mRNA-1273 vaccinees with ("pre-existing") and without ("no preexisting") preexisting cross-reactive spike-reactive memory CD4$^+$ T cells evaluated on days 1, 15 ± 2, 43 ± 2, and 209 ± 7 after immunization (see fig. S10 for details). (**F**) Spike-specific cT_{FH} cells (as percentage of CD4$^+$ T cells), (**G**) anti-spike IgG, (**H**) anti-RBD IgG, and (**I**) SARS-CoV-2 PSV neutralizing titers in mRNA-1273 vaccinees without and with preexisting cross-reactive spike-reactive memory CD4$^+$ T cells. Dotted green lines indicate the LOQ. The bars in (A) and (C) to (I) indicate the geometric mean and geomean SD in the analysis of the antibody levels or spike-specific CD4$^+$ and CD8$^+$ T cells in mRNA-1273 vaccinees with ("preexisting") and without ("no preexisting") pre-existing cross-reactive spike-reactive memory CD4$^+$ T cells evaluated on days 1, 15 ± 2, 43 ± 2, and 209 ± 7 after immunization. The bars in (B) indicate the mean and SD in the analysis of the memory phenotype of spike-specific CD4$^+$ T cells. Data were analyzed for statistical significance using Mann-Whitney *U* test. Background-subtracted and log data analyzed in all cases.

responses to the mRNA COVID-19 vaccine in this cohort. This indicates that cross-reactive memory T_{FH} cells may both accelerate B cell priming and antibody responses to a new antigen and increase robustness of long-term humoral immunity, as evidenced by the higher neutralizing antibody titers. Both total spike-specific CD4$^+$ T cells and spike-specific T_{FH} cells were enhanced after one immunization in persons with cross-reactive memory, suggesting that the spike cross-reactive memory CD4$^+$ T cells are recalled upon vaccination and affect the CD4$^+$ T cell repertoire. By contrast, CD8$^+$ T cell responses were unchanged. Our findings of preexisting immunity enhancing spike-specific CD4$^+$ T cell, T_{FH}, and antibody responses after immunization with an RNA vaccine do not represent the exact scenario of SARS-CoV-2 infection. It therefore remains unresolved whether preexisting T cells have a biological function during human SARS-CoV-2 infection (*33*). Nevertheless, these data provide evidence that the cross-reactive CD4$^+$ T cells are biologically relevant in the context of vaccination. Thus, it is plausible that the presence and magnitude of cross-reactive memory T cells could accelerate the speed and magnitude of CD4$^+$ T cell and antibody responses to SARS-CoV-2 infection, as compared with persons who have undetectable levels of cross-reactive memory T cells. Moreover, early T cell responses have been linked to less-severe COVID-19 clinical outcomes (*46*, *47*). These findings show substantial immune responses and immune memory to a low-dose mRNA vaccine and indicate

biological relevance of cross-reactive memory T cells.

Materials and methods

Human subjects

Samples from the phase 1 25-μg mRNA-1273 SARS-CoV-2 trial

A total of 140 peripheral blood samples were obtained from 35 participants who received the 25-μg dose mRNA-1273 SARS-CoV-2 vaccine in a dose-escalation, open-label phase 1 trial (mRN-1273 ClinicalTrials.gov number NCT04283461) (*9*). Participants received two injections of the trial vaccine 28 days apart between March and April of 2020 at one of the two study sites, Kaiser Permanente Washington Health Research Institute in Seattle or at the Emory University School of Medicine in Atlanta. Peripheral blood mononuclear cell (PBMC) samples used in this study were collected at the start of the trial (day 1), 14 days after each vaccination (day 15 ± 2 and 43 ± 2) and 7 months after the first dose (day 209 ± 7).

Participants were divided into three age groups: 18 to 55 (n = 15), 56 to 70 (n = 10), and >70 (n = 10) years of age (table S3). Two participants from the 18–55 group received only one dose of the 25-μg mRNA-1273 vaccine. The day 40 and day 209 samples (both collected after the second dose) were excluded from analysis for these two vaccinees. Both sexes were represented [20:15, male:female (M:F)], and the average age of the participants was 58 years. Participants identified as white (n = 34), Hispanic (n = 1), or mixed race (Asian/Hispanic, n = 1.) An overview of samples analyzed in this study is provided in table S3.

The trial was reviewed and approved by the Advarra institutional review board, as previously published (*9*). All experiments performed at the La Jolla Institute (LJI) were approved by the institutional review board (IRB) of the La Jolla Institute (IRB#: VD-214). Collection and processing of vaccinee PBMC samples was performed at one of the two study location sites, as previously described (*9*).

Samples from convalescent COVID-19 donors

To compare levels of immune memory responses induced by 25-μg mRNA-1273 SARS-CoV-2 vaccination with immune memory responses induced by natural infection with SARS-CoV-2, we collected blood from individuals that experienced natural infection with SARS-CoV-2. We matched the 7-month (209-day) postvaccination samples with samples from convalescent donors collected on average 181 days (range: 170 to 195 days) post–symptom onset (PSO). Assuming an average incubation period of 12 days until symptom onset, the time point after exposure or vaccination is comparable between the cohorts (*48*). To further match ethnicities between the cohorts, we selected 13 samples from Caucasian donors and one sample from a Hispanic donor. Convalescent donors were California residents who were either referred to the study by a health care provider or self-referred. In the overall cohort, both sexes were represented (10:4, M:F), and the average age of the donors was 35 years (±16.53). Details of this COVID-19 convalescent cohort are listed in table S4. All of the convalescent donors experienced mild illness, defined as patients with a SARS-CoV-2 positive test who have never been hospitalized (*46*). Seropositivity against SARS-CoV-2 was confirmed by enzyme-linked immunosorbent assay (ELISA), as describe below. At the time of enrollment, all COVID-19 convalescent donors provided informed consent to participate in the present and future studies.

100-μg mRNA-1273 vaccinees

We also included a cohort of individuals vaccinated locally in San Diego, California, who received the emergency use authorization–approved 100-μg mRNA-1273 SARS-CoV-2 vaccine. Samples from 20 vaccinees collected 42 ± 6 days after first immunization (15 days after second immunization) were compared with samples from the 25-μg mRNA-1273 cohort on day 43 ± 2. To match ethnicities between the cohorts, we selected 12 samples from Caucasian donors, 4 samples from Asian donors, and 4 samples from Hispanic donors. In the overall cohort, both sexes were represented (6:14, M:F) and the average age of the donors was 48 years (±14.45). Details of this cohort are listed in table S5. At the time of enrollment, all 100-μg mRNA-1273 SARS-CoV-2 vaccinees provided informed consent to participate in the present and future studies.

Peripheral blood mononuclear cells (PBMCs) and plasma isolation

Whole blood samples from convalescent COVID-19 donors and 100-μg mRNA-1273 vaccinees were collected at La Jolla Institute in heparin-coated blood bags and centrifuged for 15 min at 803*g* to separate the cellular fraction and plasma. The plasma was then carefully removed from the cell pellet and stored at −20°C. PBMCs were isolated by density-gradient sedimentation using Ficoll-Paque (Lymphoprep, Nycomed Pharma, Oslo, Norway), as previously described (*22*, *31*). Isolated PBMCs were cryopreserved in cell recovery media containing 10% dimethyl sulfoxide (DMSO) (Gibco), supplemented with 10% heat-inactivated fetal bovine serum (FBS; Hyclone Laboratories, Logan UT), and stored in liquid nitrogen until used in the assays.

Antibody measurements

SARS-CoV-2 ELISAs

SARS-CoV-2 ELISA titers in vaccinated and convalescent samples were determined as previously described (*22*, *31*, *46*, *49*). Endpoint titers (ETs) were plotted for each sample, using background subtracted data, with ETs calculated at the dilution giving a reading above the optical density cutoff of 0.1. Negative and positive controls were used to standardize each assay and normalize across experiments. A positive control standard was created by pooling plasma from six convalescent COVID-19 donors to normalize between experiments. The limit of detection was defined as 1:3 for IgG. The limit of quantification (LOQ) for vaccinated individuals was established on the basis of uninfected subjects, using plasma from healthy donors never exposed to or vaccinated against SARS-CoV-2.

Pseudovirus (PSV) neutralization assay

The PSV neutralization assays of vaccinated and convalescent samples were performed as previously described (*22*). Neutralization titers or inhibition dose 50 (ID_{50}) were calculated using the One-Site Fit Log IC50 model in Prism 9.2 (GraphPad). As internal quality control to define the interassay variation, three samples were included across the PSV neutralization assays. Samples that did not reach 50% inhibition at the lowest serum dilution of 1:20 were considered to be non-neutralizing, and the values were set to 19. PSV neutralization titers were performed as two independent experiments on different days with two replicates per experiment. Results were comparable between experiments and results from the first experiment are graphed. We included the World Health Organization (WHO) International Reference Panel for anti-SARS-Cov2 immunoglobulin (20/268) to calibrate our PSV neutralization titers. The ID_{50} of WHO-High and WHO-Mid were measured by four independent experiments with two replicates per experiment. The geometric means (GMTs) of ID_{50} of WHO-High and WHO-Mid were 2658 and 364, respectively, by our PSV neutralization assay. The WHO assigned neutralization activity unitage of 1473 and 210 IU/ml for the WHO-High and WHO-Mid standards. The calibration factor was thus calculated as ((2658/1473) + (364/ 210))/ 2 = 1.766. The GMTs of PSV neutralization ID_{50} in 25-μg and 100-μg mRNA-1273 vaccinees at day 43 were 997 and 2000, respectively, in our figures. The WHO IU calibrated neutralization ID_{50} (cID_{50}) GMTs in 25-μg mRNA-1273 vaccinees at day 43 would be 565 IU (997/1.766), and for 100-μg mRNA-1273 vaccinees at day 43 would be 1133 IU (2000/1.766). The limit of detection was calculated as 10.7 IU (19/1.766).

Peptide megapools (MPs)

We have previously developed the MP approach to allow simultaneous testing of a large number of epitopes (*22*, *31*, *46*, *50*). Here, three MPs to evaluate the antigen-specific T cell

response against SARS-CoV-2 were used, as described below. A cytomegalovirus (CMV) MP was used as a control against a ubiquitous pathogen in the experiments.

SARS-CoV-2 MPs

To characterize SARS-CoV-2–specific T cell response, we used three MPs previously described (*31*, *43*). First, we used a spike MP of 253 overlapping peptides spanning the entire sequence of the spike protein. As this peptide pool consists of peptides with a length of 15 amino acids, both $CD4^+$ and $CD8^+$ T cells have the capacity to recognize this MP (*51*). In addition, to confirm that the spike-specific $CD8^+$ T cell response observed in the mRNA-1273 vaccinees is also induced in the absence of the spike-specific $CD4^+$ T cell response, we performed some experiments with an optimal MP of HLA class I epitopes (CD8-S MP). This CD8-S MP consists of 197 9- and 10-mers derived spike peptides that have previously been described to be recognized by $CD8^+$ T cells in SARS-CoV-2 exposed donors (*22*, *31*, *46*, *52*). Lastly, we used a predicted SARS-CoV-2–specific MP (CD4-R) to evaluate the nonspike response or the remainder of the SARS-CoV-2 genome of 246 HLA class II $CD4^+$ epitopes previously described (*31*). We have previously shown that these MPs are suitable for stimulating T cell responses from either COVID-19–exposed or SARS-CoV-2–unexposed individuals (*31*, *34*).

Cytomegalovirus (CMV) MP

As a control, we used a MP of 313 experimentally defined epitopes. This CMV MP consists of HLA class I and class II epitopes and $CD4^+$ and $CD8^+$ T cells have the capacity to recognize this MP, as has been previously published (*50*).

Flow cytometry assays

Activation-induced markers (AIM) assay

Antigen-specific $CD4^+$ T cells were measured as a percentage of AIM^+ ($OX40^+CD137^+$) $CD4^+$ and ($CD69^+CD137^+$) $CD8^+$ T cells after stimulation of PBMCs from mRNA-1273 vaccinees and COVID-19 convalescent donors with peptide MPs. Antigen-specific circulating T follicular helper (cT_{FH}) cells ($CXCR5^+OX40^+CD40L^+$, as percentage of $CD4^+$ T cells) were defined by the AIM assay.

Before addition of MPs, cells were blocked at 37°C for 15 min with 0.5 μg/ml of anti-CD40 monoclonal antibody (mAb) (Miltenyi Biotec). Then, cells were incubated at 37°C for 24 hours in the presence of fluorescently labeled anti-chemokine receptor antibodies (anti-CCR4, -CCR6, -CCR7, -CXCR3, and -CXCR5) and SARS-CoV-2 MPs (1 μg/ml) or CMV MP (1 μg/ml) in 96-well U-bottom plates, as previously described (*31*, *46*). In addition, PBMCs were incubated with an equimolar amount of DMSO as negative control and with phytohemagglutinin (5 μg/ml) (PHA, Roche) as a positive control.

For the surface stain, 1×10^6 PBMCs were resuspended in phosphate-buffered saline (PBS), incubated with BD human FC block (BD Biosciences, San Diego, CA) and the LIVE/DEAD marker in the dark for 15 min and washed with PBS. Then, an antibody mix containing the rest of the surface antibodies was added directly to cells and incubated for 60 min at 4°C in the dark. After surface staining, cells were washed twice with PBS containing 3% FBS (FACS buffer). All samples were acquired on a Cytek Aurora (Cytek Biosciences, Fremont, CA). A list of antibodies used in this panel can be found in table S6, and a representative gating strategy of spike-specific $CD4^+$ and $CD8^+$ T cells using the AIM assay is shown in fig. S2.

Antigen-specific $CD4^+$ and $CD8^+$ T cells were measured as background (DMSO)–subtracted data, with a minimal DMSO level set to 0.005%. A response of >0.02% and a stimulation index (SI) of >2 for $CD4^+$, and a response of >0.03% and SI of >3 for $CD8^+$ T cells, were considered positive. The LOQ for antigen-specific $CD4^+$ T cell responses (0.03%) and antigen-specific $CD8^+$ T cell responses (0.06%) was calculated using the median twofold standard deviation of all negative controls. As an internal quality control to define interassay variation, the CMV-specific $CD4^+$ and $CD8^+$ T cell responses were evaluated for a SARS-CoV-2–unexposed donor included in each independent experiment. The antigen-specific response against CMV and the response to the positive control was compared across experiments and revealed a coefficient of variation (CV) between 10 and 13% for the antigen-specific stimulation with the CMV-MP and a CV between 6 and 12% for mitogenic stimulation with PHA (fig. S2).

Intracellular cytokine staining (ICS) assay

To optimize spike-specific detection of cytokine-producing $CD4^+$ and $CD8^+$ T cells, we experimented with different incubation times (5, 6, and 24 hours) in the presence of GolgiPlug containing brefeldin A and GolgiStop containing monensin (BD Biosciences, San Diego, CA) for an additional 1, 3, or 4 hours, respectively. To establish optimal conditions for the ICS assay, we evaluated the IFNγ-producing $CD4^+$ and $CD8^+$ T cells in 100-μg mRNA-1273 vaccinees (n = 4) and COVID-19 convalescent donors (n = 4) (table S7). The highest signal of IFNγ-producing $CD4^+$ and $CD8^+$ T cells was detected after 24+4 hours incubation in both vaccinees and convalescent donors. Thus, we chose 24 hours as the best condition to identify spike-specific $CD4^+$ and $CD8^+$ T cells producing intracellular cytokines.

Before the addition of MPs, cells were blocked at 37°C for 15 min with 0.5 μg/ml of anti-CD40 mAb, as previously described (*46*). PBMCs were cultured in the presence of SARS-CoV-2 MPs (1 μg/ml) for 24 hours at 37°C. In addition, PBMCs were incubated with an equimolar amount of DMSO as a negative control and also CMV MP (1 μg/ml) as a positive control. After 24 hours, Golgi-Plug and Golgi-Stop were added to the culture for 4 hours, as described above. Cells were then washed and surface-stained for 30 min at 4°C in the dark and fixed with 1% of paraformaldehyde (Sigma-Aldrich, St. Louis, MO). Antibodies used in the ICS assay are listed in table S8, and a representative gating strategy of spike-specific $CD4^+$ and $CD8^+$ T cells using the ICS assay is shown in fig. S4.

Antigen-specific $CD4^+$ and $CD8^+$ T cells were measured as background (DMSO) subtracted data, with a minimal DMSO level set to 0.001%. Responses of >0.005% and a SI of >2 for $CD4^+$ and $CD8^+$ T cells were considered positive. The limit of quantification for antigen-specific $CD4^+$ and $CD8^+$ T cell responses (0.01%) was calculated using the median twofold standard deviation of all negative controls. As internal quality control to define interassay variation, the CMV-specific $CD4^+$ and $CD8^+$ T cell responses were evaluated for a SARS-CoV-2–unexposed donor included in each independent experiment. The antigen-specific response of $CD4^+$ and $CD8^+$ T cells against CMV were compared across experiments and revealed a CV of 14% for the antigen-specific stimulation with the CMV MP (fig. S4E).

The gates applied for the identification of $CD4^+$ and $CD8^+$ T cells producing cytokines were defined according to the cells cultured with DMSO for each individual as is shown in fig. S4. A Boolean analysis was performed to define the multifunctional profiles on FlowJo 10.7.1. The analysis included CD40L, GzB, IFNγ, IL-2, and TNFα gated on $CD3^+CD4^+$ cells and GzB, IFNγ, IL-2, and TNFα gated on $CD3^+CD8^+$ cells. The overall response to spike was defined as the sum of the background subtracted responses to each combination of individual cytokines. To define the multifunctional profiles of antigen-specific T cells, all positive background-subtracted data (>0.005% and a SI > 2 for $CD4^+$ T cells and >0.002% and a SI > 2 for $CD8^+$ T cells) was aggregated into a combined sum of antigen-specific $CD4^+$ or $CD8^+$ T cells on the basis of the number of functions. Values higher than the LOQ (0.01%) were considered for the analysis of the multifunctional antigen-specific T cell responses. The average of the relative $CD4^+$ and $CD8^+$ T cell response was calculated per group to define the proportion of multifunctional antigen-specific T cell responses (figs. S6 and S7 and tables S1 and S2).

Statistical analysis

Data were analyzed using FlowJo 10.7.1. Statistical analyses were performed in GraphPad Prism 9.2, unless otherwise stated. The statistical details of the experiments are provided in the respective figure legends. Data plotted in linear scale were expressed as means ± standard deviations (SD). Data plotted in logarithmic scales were expressed as geometric means ± geometric standard

deviations (SD). Mann-Whitney *U* or Wilcoxon tests were applied for unpaired or paired comparisons, respectively. Differences among age groups were evaluated using Kruskal-Wallis and Dunn's post-test for multiple comparisons. Details pertaining to significance are also noted in the respective legends.

REFERENCES AND NOTES

1. F. P. Polack *et al*., Safety and efficacy of the BNT162b2 mRNA Covid-19 vaccine. *N. Engl. J. Med.* **383**, 2603–2615 (2020). doi: 10.1056/NEJMoa2034577; pmid: 33301246
2. L. R. Baden *et al*., Efficacy and safety of the mRNA-1273 SARS-CoV-2 vaccine. *N. Engl. J. Med.* **384**, 403–416 (2021). doi: 10.1056/NEJMoa2035389; pmid: 33378609
3. R. L. Soiza, C. Scicluna, E. C. Thomson, Efficacy and safety of COVID-19 vaccines in older people. *Age Ageing* **50**, 279–283 (2021). doi: 10.1093/ageing/afaa274; pmid: 33320183
4. C. Pawlowski *et al*., FDA-authorized mRNA COVID-19 vaccines are effective per real-world evidence synthesized across a multi-state health system. *Med* **2**, 979–992.e8 (2021). doi: 10.1016/j.medj.2021.06.007; pmid: 34223401
5. S. Amit, G. Regev-Yochay, A. Afek, Y. Kreiss, E. Leshem, Early rate reductions of SARS-CoV-2 infection and COVID-19 in BNT162b2 vaccine recipients. *Lancet* **397**, 875–877 (2021). doi: 10.1016/S0140-6736(21)00448-7; pmid: 33610193
6. R. W. Frenck Jr. *et al*., Safety, immunogenicity, and efficacy of the BNT162b2 Covid-19 vaccine in adolescents. *N. Engl. J. Med.* **385**, 239–250 (2021). doi: 10.1056/NEJMoa2107456; pmid: 34043894
7. V. J. Hall *et al*., COVID-19 vaccine coverage in health-care workers in England and effectiveness of BNT162b2 mRNA vaccine against infection (SIREN): A prospective, multicentre, cohort study. *Lancet* **397**, 1725–1735 (2021). doi: 10.1016/S0140-6736(21)00790-X; pmid: 33901423
8. D. Wrapp *et al*., Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science* **367**, 1260–1263 (2020). doi: 10.1126/science.abb2507; pmid: 32075877
9. L. A. Jackson *et al*., An mRNA vaccine against SARS-CoV-2 — Preliminary report. *N. Engl. J. Med.* **383**, 1920–1931 (2020). doi: 10.1056/NEJMoa2022483; pmid: 32663912
10. S. J. Thomas *et al*., Six month safety and efficacy of the BNT162b2 mRNA COVID-19 vaccine. *medRxiv* 2021.07.28.21261159 [Preprint] (2021). https://doi.org/10.1101/2021.07.28.21261159.
11. Moderna, "Moderna Reports Second Quarter Fiscal Year 2021 Financial Results and Provides Business Updates," press release (5 August 2021); https://investors.modernatx.com/news-releases/news-release-details/moderna-reports-second-quarter-fiscal-year-2021-financial.
12. P. B. Gilbert *et al*., Immune correlates analysis of the mRNA-1273 COVID-19 vaccine efficacy trial. *medRxiv* 2021.08.09.21261290 [Preprint] (2021). https://doi.org/10.1101/2021.08.09.21261290.
13. D. S. Khoury *et al*., Neutralizing antibody levels are highly predictive of immune protection from symptomatic SARS-CoV-2 infection. *Nat. Med.* **27**, 1205–1211 (2021). doi: 10.1038/s41591-021-01377-8; pmid: 34002089
14. A. Tauzin *et al*., A single dose of the SARS-CoV-2 vaccine BNT162b2 elicits Fc-mediated antibody effector functions and T cell responses. *Cell Host Microbe* **29**, 1137–1150.e6 (2021). doi: 10.1016/j.chom.2021.06.001; pmid: 34133950
15. A. Sette, S. Crotty, Adaptive immunity to SARS-CoV-2 and COVID-19. *Cell* **184**, 861–880 (2021). doi: 10.1016/j.cell.2021.01.007; pmid: 33497610
16. L. J. Abu-Raddad *et al*., SARS-CoV-2 antibody-positivity protects against reinfection for at least seven months with 95% efficacy. *EClinicalMedicine* **35**, 100861 (2021). doi: 10.1016/j.eclinm.2021.100861; pmid: 33937733
17. S. F. Lumley *et al*., Antibody status and incidence of SARS-CoV-2 infection in health care workers. *N. Engl. J. Med.* **384**, 533–540 (2021). doi: 10.1056/NEJMoa2034545; pmid: 33369366
18. V. J. Hall *et al*., SARS-CoV-2 infection rates of antibody-positive compared with antibody-negative health-care workers in England: A large, multicentre, prospective cohort study (SIREN). *Lancet* **397**, 1459–1469 (2021). doi: 10.1016/S0140-6736(21)00675-9; pmid: 33844963
19. A. Leidi *et al*., Risk of reinfection after seroconversion to SARS-CoV-2: A population-based propensity-score matched cohort study. *Clin. Infect. Dis.* ciab495 (2021). doi: 10.1093/cid/ciab495; pmid: 34043763
20. N. R. Faria *et al*., Genomics and epidemiology of the P.1 SARS-CoV-2 lineage in Manaus, Brazil. *Science* **372**, 815–821 (2021). doi: 10.1126/science.abh2644; pmid: 33853970
21. J. Zuo *et al*., Robust SARS-CoV-2-specific T cell immunity is maintained at 6 months following primary infection. *Nat. Immunol.* **22**, 620–626 (2021). doi: 10.1038/s41590-021-00902-8; pmid: 33674800
22. J. M. Dan *et al*., Immunological memory to SARS-CoV-2 assessed for up to 8 months after infection. *Science* **371**, eabf4063 (2021). doi: 10.1126/science.abf4063; pmid: 33408181
23. R. J. Boyton, D. M. Altmann, Risk of SARS-CoV-2 reinfection after natural infection. *Lancet* **397**, 1161–1163 (2021). doi: 10.1016/S0140-6736(21)00662-0; pmid: 33743219
24. C. Li *et al*., Twelve-month specific IgG response to SARS-CoV-2 receptor-binding domain among COVID-19 convalescent plasma donors in Wuhan. *Nat. Commun.* **12**, 4144 (2021). doi: 10.1038/s41467-021-24230-5; pmid: 34230476
25. Z. Wang *et al*., Naturally enhanced neutralizing breadth against SARS-CoV-2 one year after infection. *Nature* **595**, 426–431 (2021). doi: 10.1038/s41586-021-03696-9; pmid: 34126625
26. K. W. Cohen *et al*., Longitudinal analysis shows durable and broad immune memory after SARS-CoV-2 infection with persisting antibody responses and memory B and T cells. *Cell Rep. Med.* **2**, 100354 (2021). doi: 10.1016/j.xcrm.2021.100354; pmid: 34250512
27. A. Pegu *et al*., Durability of mRNA-1273 vaccine-induced antibodies against SARS-CoV-2 variants. *Science* eabj4176 (2021). doi: 10.1126/science.abj4176; pmid: 34385356
28. E. J. Anderson *et al*., Safety and immunogenicity of SARS-CoV-2 mRNA-1273 vaccine in older adults. *N. Engl. J. Med.* **383**, 2427–2438 (2020). doi: 10.1056/NEJMoa2028436; pmid: 32991794
29. US Food and Drug Administration, "Vaccines and Related Biological Products Advisory Committee Meeting Presentation," 22 October 2020.
30. J. Braun *et al*., SARS-CoV-2-reactive T cells in healthy donors and patients with COVID-19. *Nature* **587**, 270–274 (2020). doi: 10.1038/s41586-020-2598-9; pmid: 32726801
31. A. Grifoni *et al*., Targets of T cell responses to SARS-CoV-2 coronavirus in humans with COVID-19 disease and unexposed individuals. *Cell* **181**, 1489–1501.e15 (2020). doi: 10.1016/j.cell.2020.05.015; pmid: 32473127
32. N. Le Bert *et al*., SARS-CoV-2-specific T cell immunity in cases of COVID-19 and SARS, and uninfected controls. *Nature* **584**, 457–462 (2020). doi: 10.1038/s41586-020-2550-z; pmid: 32668444
33. M. Lipsitch, Y. H. Grad, A. Sette, S. Crotty, Cross-reactive memory T cells and herd immunity to SARS-CoV-2. *Nat. Rev. Immunol.* **20**, 709–713 (2020). doi: 10.1038/s41577-020-00460-4; pmid: 33024281
34. J. Mateus *et al*., Selective and cross-reactive SARS-CoV-2 T cell epitopes in unexposed humans. *Science* **370**, 89–94 (2020). doi: 10.1126/science.abd3871; pmid: 32753554
35. B. J. Meckiff *et al*., Imbalance of regulatory and cytotoxic SARS-CoV-2-reactive $CD4^+$ T cells in COVID-19. *Cell* **183**, 1340–1353.e16 (2020). doi: 10.1016/j.cell.2020.10.001; pmid: 33096020
36. A. Sette, S. Crotty, Pre-existing immunity to SARS-CoV-2: The knowns and unknowns. *Nat. Rev. Immunol.* **20**, 457–458 (2020). doi: 10.1038/s41577-020-0389-z; pmid: 32636479
37. D. Weiskopf *et al*., Phenotype and kinetics of SARS-CoV-2-specific T cells in COVID-19 patients with acute respiratory distress syndrome. *Sci. Immunol.* **5**, eabd2071 (2020). doi: 10.1126/sciimmunol.abd2071; pmid: 32591408
38. E. M. Anderson *et al*., Seasonal human coronavirus antibodies are boosted upon SARS-CoV-2 infection but not associated with protection. *Cell* **184**, 1858–1864.e10 (2021). doi: 10.1016/j.cell.2021.02.010; pmid: 33631096
39. M. Sagar *et al*., Recent endemic coronavirus infection is associated with less-severe COVID-19. *J. Clin. Invest.* **131**, e143380 (2021). doi: 10.1172/JCI143380; pmid: 32997649
40. N. Doria-Rose *et al*., Antibody persistence through 6 months after the second dose of mRNA-1273 vaccine for Covid-19. *N. Engl. J. Med.* **384**, 2259–2261 (2021). doi: 10.1056/NEJMc2103916; pmid: 33822494
41. S. Crotty, T follicular helper cell biology: A decade of discovery and diseases. *Immunity* **50**, 1132–1148 (2019). doi: 10.1016/j.immuni.2019.04.011; pmid: 31117010
42. C. E. Gustafson, C. Kim, C. M. Weyand, J. J. Goronzy, Influence of immune aging on vaccine responses. *J. Allergy Clin. Immunol.* **145**, 1309–1321 (2020). doi: 10.1016/j.jaci.2020.03.017; pmid: 32386655
43. A. Grifoni *et al*., A sequence homology and bioinformatic approach can predict candidate targets for immune responses to SARS-CoV-2. *Cell Host Microbe* **27**, 671–680.e2 (2020). doi: 10.1016/j.chom.2020.03.002; pmid: 32183941
44. U. Sahin *et al*., BNT162b2 vaccine induces neutralizing antibodies and poly-specific T cells in humans. *Nature* **595**, 572–577 (2021). doi: 10.1038/s41586-021-03653-6; pmid: 34044428
45. K. S. Corbett *et al*., Evaluation of the mRNA-1273 vaccine against SARS-CoV-2 in nonhuman primates. *N. Engl. J. Med.* **383**, 1544–1555 (2020). doi: 10.1056/NEJMoa2024671; pmid: 32722908
46. C. Rydyznski Moderbacher *et al*., Antigen-specific adaptive immunity to SARS-CoV-2 in acute COVID-19 and associations with age and disease severity. *Cell* **183**, 996–1012.e19 (2020). doi: 10.1016/j.cell.2020.09.038; pmid: 33010815
47. A. T. Tan *et al*., Early induction of functional SARS-CoV-2-specific T cells associates with rapid viral clearance and mild disease in COVID-19 patients. *Cell Rep.* **34**, 108728 (2021). doi: 10.1016/j.celrep.2021.108728; pmid: 33516277
48. S. A. Lauer *et al*., The incubation period of coronavirus disease 2019 (COVID-19) from publicly reported confirmed cases: Estimation and application. *Ann. Intern. Med.* **172**, 577–582 (2020). doi: 10.7326/M20-0504; pmid: 32150748
49. F. Amanat *et al*., A serological assay to detect SARS-CoV-2 seroconversion in humans. *Nat. Med.* **26**, 1033–1036 (2020). doi: 10.1038/s41591-020-0913-5; pmid: 32398876
50. S. Carrasco Pro *et al*., Automatic generation of validated specific epitope sets. *J. Immunol. Res.* **2015**, 763461 (2015). doi: 10.1155/2015/763461; pmid: 26568965
51. M. F. Kotturi *et al*., The $CD8^+$ T-cell response to lymphocytic choriomeningitis virus involves the L antigen: Uncovering new tricks for an old virus. *J. Virol.* **81**, 4928–4940 (2007). doi: 10.1128/JVI.02632-06; pmid: 17329346
52. A. Grifoni *et al*., SARS-CoV-2 human T cell epitopes: Adaptive immune response against COVID-19. *Cell Host Microbe* **29**, 1076–1092 (2021). doi: 10.1016/j.chom.2021.05.010; pmid: 34237248

ACKNOWLEDGMENTS

This work used samples from the phase 1 mRNA-1273 study (NCT04283461) (*9*, *28*). The study was conducted in collaboration with ModernaTX, and funding for the manufacture of mRNA-1273 phase 1 material was provided by the Coalition for Epidemic Preparedness Innovation. We thank E. O. Saphire for providing the spike plasmids and the LJI Clinical Core for healthy donor enrollment and blood sample procurement. **Funding:** The mRNA-1273 phase 1 study was sponsored and primarily funded by the National Institute of Allergy and Infectious Diseases (NIAID), National Institutes of Health (NIH), Bethesda, MD; in part with federal funds from the NIAID under grant awards UM1AI148373 to Kaiser Washington; UM1AI148576, UM1AI148684, and NIH P51 OD011132 to Emory University; NIH AID AI149644 and contract award HHSN272201500002C to Emmes. This work was funded by the NIH NIAID under awards AI142742 (Cooperative Centers for Human Immunology) (A.S., S.C.) and NIH contract 75N93019900065 (D.W., A.S.). This work was also supported in part by LJI Institutional Funds and the NIAID under K08 award AI135078 (J.M.D.) **Author contributions:** Conceptualization: A.S., S.C., and D.W. Methodology: J.M., J.M.D., Z.Z., C.R.M., M.L., B.G., and D.W. Formal analysis: J.M., J.M.D., Z.Z., A.S., S.C., and D.W. Investigation: J.M., A.S., S.C., and D.W. Funding acquisition: A.S., S.C., and D.W. Writing: J.M., J.M.D., Z.Z., A.S., S.C., and D.W. Supervision: A.S., S.C., and D.W. **Competing interests:** A.S. is a consultant for Gritstone Bio, Flow Pharma, Arcturus Therapeutics, ImmunoScape, CellCarta, Oxford Immunotec, and Avalia Immunotherapies. S.C. has consulted for Avalia Immunotherapies, Roche, and GSK. LJI has filed for patent protection for various aspects of T cell epitope and vaccine design work. All other authors declare no conflicts of interest. **Data and materials availability:** Epitope pools used in this paper are available to the scientific community upon request and execution of a material transfer agreement (MTA). All other data are available in the main text or the supplementary materials.

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abj9853
Figs. S1 to S13
Tables S1 to S8
References
MDAR Reproducibility Checklist
Data S1

View/request a protocol for this paper from *Bio-protocol*.

16 June 2021; accepted 7 September 2021
Published online 14 September 2021
10.1126/science.abj9853

Science

Low-dose mRNA-1273 COVID-19 vaccine generates durable memory enhanced by cross-reactive T cells

Jose MateusJennifer M. DanZeli ZhangCarolyn Rydyznski ModerbacherMarshall LammersBenjamin GoodwinAlessandro SetteShane CrottyDaniela Weiskopf

Science, 374 (6566), eabj9853. • DOI: 10.1126/science.abj9853

A smaller-dose jab does the job

Low-dose messenger RNA (mRNA) vaccines potentially allow health providers to administer more doses from a limited vaccine supply and can be less reactogenic. Whether low-dose COVID-19 mRNA vaccines generate immune responses comparable to currently approved doses remains an open question, however. Mateus *et al*. report the results of a clinical trial comparing patients who received a 25-μg mRNA-1273 (Moderna) COVID-19 vaccine to 100-μg mRNA-1273 COVID-19 vaccinees and severe acute respiratory syndrome coronavirus 2–infected individuals. The low-dose Moderna vaccine generated long-lived T cell immunity that was equivalent between younger and older patients and that could be enhanced by the presence of cross-reactive T cells. Moreover, antibody and T cell responses induced by the low-dose vaccine were comparable to natural infection and about half as strong as those seen with high-dose vaccination. —STS

Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

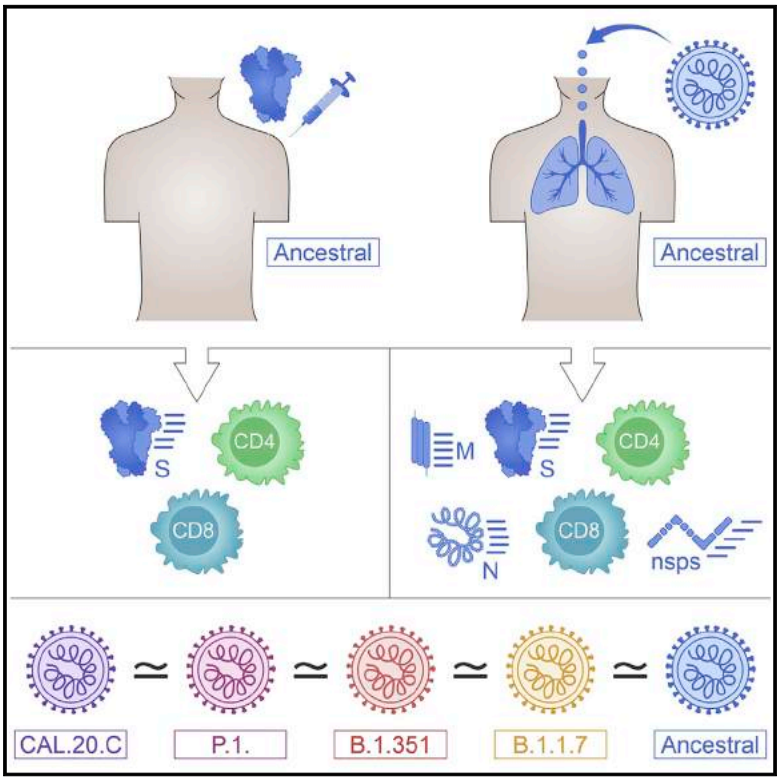
Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.

Cell Reports
Medicine

Article

Impact of SARS-CoV-2 variants on the total CD4$^+$ and CD8$^+$ T cell reactivity in infected or vaccinated individuals

Graphical abstract



Highlights

- T cells of exposed donors or vaccinees effectively recognize SARS-CoV-2 variants
- Effective recognition in AIM and FluoroSPOT assays, for spike and other proteins
- 93% and 97% of CD4 and CD8 epitopes are 100% conserved across variants

Authors

Alison Tarke, John Sidney, Nils Methot, ..., Shane Crotty, Alba Grifoni, Alessandro Sette

Correspondence

agrifoni@lji.org (A.G.),
alex@lji.org (A.S.)

In brief

Tarke et al. show that SARS-CoV-2-specific memory CD4 and CD8 T cells exposed to the ancestral strain by infection or vaccination effectively recognize the variants B.1.1.7, B.1.351, P.1, and CAL.20C. The majority of T cell epitopes are unaffected by mutations in these variant strains.

Tarke et al., 2021, Cell Reports Medicine 2, 100355
July 20, 2021
https://doi.org/10.1016/j.xcrm.2021.100355

Cell Reports Medicine

CellPress
OPEN ACCESS

Article

Impact of SARS-CoV-2 variants on the total CD4+ and CD8+ T cell reactivity in infected or vaccinated individuals

Alison Tarke,[1,2] John Sidney,[1] Nils Methot,[1] Esther Dawen Yu,[1] Yun Zhang,[4] Jennifer M. Dan,[1,3] Benjamin Goodwin,[1] Paul Rubiro,[1] Aaron Sutherland,[1] Eric Wang,[1] April Frazier,[1] Sydney I. Ramirez,[1,3] Stephen A. Rawlings,[3] Davey M. Smith,[3] Ricardo da Silva Antunes,[1] Bjoern Peters,[1,3] Richard H. Scheuermann,[1,4,5] Daniela Weiskopf,[1] Shane Crotty,[1,3] Alba Grifoni,[1,6,*] and Alessandro Sette[1,3,6,7,*]

[1]Center for Infectious Disease and Vaccine Research, La Jolla Institute for Immunology (LJI), La Jolla, CA 92037, USA
[2]Department of Internal Medicine and Center of Excellence for Biomedical Research (CEBR), University of Genoa, Genoa 16132, Italy
[3]Department of Medicine, Division of Infectious Diseases and Global Public Health, University of California, San Diego (UCSD), La Jolla, CA 92037, USA
[4]J. Craig Venter Institute, La Jolla, CA 92037, USA
[5]Department of Pathology, University of California, San Diego, San Diego, CA 92093, USA
[6]These authors contributed equally
[7]Lead contact
*Correspondence: agrifoni@lji.org (A.G.), alex@lji.org (A.S.)
https://doi.org/10.1016/j.xcrm.2021.100355

SUMMARY

The emergence of SARS-CoV-2 variants with evidence of antibody escape highlight the importance of addressing whether the total CD4+ and CD8+ T cell recognition is also affected. Here, we compare SARS-CoV-2-specific CD4+ and CD8+ T cells against the B.1.1.7, B.1.351, P.1, and CAL.20C lineages in COVID-19 convalescents and in recipients of the Moderna (mRNA-1273) or Pfizer/BioNTech (BNT162b2) COVID-19 vaccines. The total reactivity against SARS-CoV-2 variants is similar in terms of magnitude and frequency of response, with decreases in the 10%–22% range observed in some assay/VOC combinations. A total of 7% and 3% of previously identified CD4+ and CD8+ T cell epitopes, respectively, are affected by mutations in the various VOCs. Thus, the SARS-CoV-2 variants analyzed here do not significantly disrupt the total SARS-CoV-2 T cell reactivity; however, the decreases observed highlight the importance for active monitoring of T cell reactivity in the context of SARS-CoV-2 evolution.

INTRODUCTION

The emergence of several severe acute respiratory syndrome-coronavirus-2 (SARS-CoV-2) variants of concern (VOC) with multiple amino acid replacements has implications for the future control of the coronavirus disease 2019 (COVID-19) pandemic.[1–4] VOCs include the United Kingdom (UK) variant 501Y.V1 lineage B.1.1.7,[3] the South Africa (SA) variant 501Y.V2 lineage B.1.351,[4] the BR (Brazilian) variant 501Y.V3 lineage P.1,[5,6] and the California (CA) variant CAL.20C lineages B.1.427–429.[7,8] The B.1.1.7 variant is associated with increased transmissibility,[9,10] and similar epidemiological observations have been reported for the SA and BR variants.[4,5]

The mutations of greatest concern are present in the viral spike (S) protein and include notable mutations in the receptor-binding domain (RBD), the N-terminal domain (NTD), and the furin cleavage site region. Several of the RBD mutations directly affect angiotensin-converting enzyme 2 (ACE2) receptor-binding affinity, which may affect infectivity, viral load, or transmissibility.[11–14] Multiple mutations were also noted in regions bound by neutralizing antibodies, so it is crucial to address the extent to which mutations associated with the variants may affect immunity induced by either SARS-CoV-2 infection or vaccination.

Numerous reports address the effect of variant spike (S) mutations on antibody binding and function.[11,14–23] In general, the impact of B.1.1.7 mutations on neutralizing antibody titers is moderate.[15–18,20,24] In contrast, mutations in B.1.351 and P.1 variants are associated with the more pronounced loss of neutralizing capacity.[15,16,22,23,25] Concerning vaccination responses, the AstraZeneca ChAdOx1 vaccine has been associated with a partial loss of neutralizing antibody activity against B.1.1.7[15] and a large loss of neutralizing activity against B.1.351.[26] ChAdOx1 maintains efficacy against B.1.1.7[24,27] but has a major loss in efficacy against mild COVID-19 with the B.1.351 variant.[26] Current epidemiological evidence is that the BNT162b2 Pfizer/BioNTech COVID-19 vaccine retains its efficacy against B.1.1.7 in the UK[24] and in reports from Israel.[28,29] The Novavax COVID-19 vaccine (NVX-CoV2373) has differential protective immunity against the parental strain, B.1.1.7, and B.1.351 in clinical trials (96%, 86%, and 60%, respectively),[30] whereas the Janssen Ad26.COV2.S 1-dose COVID-19 vaccine has relatively similar protection for moderate COVID-19 against

Please cite this article in press as: Tarke et al., Impact of SARS-CoV-2 variants on the total CD4+ and CD8+ T cell reactivity in infected or vaccinated individuals, Cell Reports Medicine (2021), https://doi.org/10.1016/j.xcrm.2021.100355

both the ancestral strain and B.1.351 (72% and 64%, respectively).[31,32]

Several lines of evidence suggest that CD4+ and CD8+ T cell responses play important roles in the resolution of SARS-CoV-2 infection and COVID-19,[33] including modulating disease severity in humans[34,35] and reducing viral loads in non-human primates.[36] Furthermore, individuals with agammaglobulinemia or pharmaceutical depletion of B cells generally experience an uncomplicated COVID-19 disease course.[33,37,38] Robust CD4+ and CD8+ T cell memory is induced after COVID-19,[22,39–41] and multiple COVID-19 vaccines elicit CD4+ and CD8+ T cell responses.[26,42–45] It is therefore key to address the potential impact of mutations associated with SARS-CoV-2 variants on T cell reactivity; however, few data are available on this topic.[46]

Here, we take a combined experimental and bioinformatics approach to address T cell reactivity to SARS-CoV-2 VOCs. We directly assess T cell responses from individuals recovered from COVID-19 and T cell responses from recent Moderna mRNA-1273 or Pfizer/BioNTech BNT162b2 vaccinees for their capacity to recognize peptides derived from the ancestral reference sequence and the B.1.1.7, B1.351, P.1, and CAL.20C variants. As a complementary approach, bioinformatics analyses were used to predict the impact of mutations in the VOCs with sets of previously reported CD4+ and CD8+ T cell epitopes derived from the ancestral reference sequence.

RESULTS

Sequence analysis and peptide pool generation

As a first step, we mapped the specific mutations (amino acid replacements and deletions) of the main current VOCs, including B.1.1.7, B1.351, P.1, and CAL.20C, as compared to the SARS-CoV-2 Wuhan ancestral sequence (NCBI: NC_045512.2) (Table S1). Then, we synthesized the corresponding peptides associated with the different variants and generated new peptide pools (megapools [MPs]) spanning the full genome sequences of the ancestral Wuhan strain and the respective B.1.1.7, B1.351, P.1, and CAL.20C variants (Table S2). As described below, the resulting peptide pools were assessed for their capacity to be recognized by memory T cell responses derived from natural infection in convalescents and vaccinees, and responses to the variant and ancestral genome antigen-specific pools were compared.

Cohorts of COVID-19 convalescent, vaccinees, and unexposed controls

We selected three donor cohorts to investigate T cell reactivity against VOCs. Our convalescent donors were adults ranging from 20 to 67 years of age (median 38); 43% were male and 57% were female (Table 1). SARS-CoV-2 infection in these donors was determined by PCR-based testing during the acute phase of their infection, if available (46% of the cases), and/or seropositivity determined by plasma SARS-CoV-2 S protein RBD immunoglobulin G (IgG) ELISA (Figure S1A). From these donors, peripheral blood mononuclear cell (PBMC) samples were collected between May and October 2020, a period when none of the VOCs analyzed were prevalent in the San Diego, California, area, where the donations from convalescent donors were obtained. The convalescent donor samples reflect the local ethnic demographics (81% non-Hispanic white). The peak of COVID-19 severity was representative of the distribution observed in the general population, with a prevalence of mild cases (86% of the cohort analyzed).

From the vaccinated donors, we collected PBMCs after recent vaccination with the Moderna mRNA-1273 or the Pfizer/BioNTech BNT162b2 vaccines, ~14 days following second dose administration (Table 1). These donors ranged in age from 23 to 67 years (median 47); 31% were male and 69% were female. All of the vaccinees had RBD IgG titers indicative of vaccination (Figure S1A).

PBMCs from unexposed donors were collected between May 2014 and March 2018 or in the March–May 2020 period, and were seronegative for RBD IgG. The unexposed donors ranged from 21 to 82 years old (median 34); 35% were male and 56% were female, while 9% were unknown (Table 1).

CD4+ and CD8+ T cell reactivity against ancestral S

We previously described the use of activation-induced marker (AIM) assays to measure CD4+ and CD8+ T cell responses to pools of overlapping peptides spanning SARS-CoV-2 antigens.[34,39,47,48] Here, we used the same AIM techniques using $OX40^+CD137^+$ and $CD69^+CD137^+$ markers for CD4+ and CD8+ T cell reactivity, respectively.

We compared the CD4+ and CD8+ T cell reactivity of the three cohorts against SARS-CoV-2 S by AIM and FluoroSPOT assays. The T cell responses to the ancestral S MP were significantly higher than the DMSO controls in unexposed ($p < 0.0001$ for CD4, $p = 0.0063$ for CD8, and $p = 0.0036$ for spot-forming cells [SFCs]/10^6 PBMCs by the Wilcoxon test), convalescent ($p < 0.0001$ for CD4, $p < 0.0001$ for CD8, and $p = 0.0008$ for SFC/10^6 PBMCs by the Wilcoxon test), and vaccinated donors ($p < 0.0001$ for CD4, $p < 0.0001$ for CD8, and $p < 0.0001$ for SFC/10^6 PBMCs by the Wilcoxon test) (Figures S1B–S1D).

When the AIM responses in the different donor cohorts were compared, we found significantly higher reactivity in COVID-19 convalescents compared to unexposed individuals (CD4: $p < 0.0001$; CD8: $p < 0.0001$ by the Mann-Whitney test) (Figures S1E and S1F), with 93% and 39% of donors positive (responses above the threshold of positivity as described in Method details) for CD4+ and CD8+ T cell responses, respectively. These rates of positivity are similar to what was observed in a previous study that analyzed responses to peptides spanning the whole SARS-CoV-2 proteome over a similar range of post-symptom onset (PSO) days,[39] in which 93% and 50% positivity rates were observed for CD4+ and CD8+ T cells, respectively. Likewise, the responses observed in the S MP in vaccinated donors were significantly higher than in unexposed donors (CD4: $p < 0.0001$; CD8: $p < 0.0001$ by Mann-Whitney test), with 100% and 48% of donors positive for CD4 and CD8 responses, respectively (Figures S1E and S1F).

With an interferon-γ (IFN-γ) FluoroSPOT assay, IFN-γ T cell responses were detected in 17%, 50%, and 76% of unexposed, convalescent, and vaccinated donors, respectively (unexposed versus convalescent: $p = 0.019$; unexposed versus vaccinees: $p < 0.0001$ by the Mann-Whitney test) (Figure S1G). These donor T cell responses provided the benchmarks for subsequent assessment of the impact of VOC mutations.

Please cite this article in press as: Tarke et al., Impact of SARS-CoV-2 variants on the total CD4+ and CD8+ T cell reactivity in infected or vaccinated individuals, Cell Reports Medicine (2021), https://doi.org/10.1016/j.xcrm.2021.100355

Table 1. Characteristics of donor cohorts

	COVID-19 (n = 28)	Vaccinees (n = 29)	Unexposed (n = 23)
Age, y	20–67 (median = 38, IQR = 32)	23–67 (median = 47, IQR = 36)	21–82 (median = 34, IQR = 24)
Gender, %			
Male	43 (12/28)	31 (9/29)	35 (8/23)
Female	57 (16/28)	69 (20/29)	56 (13/23)
Unknown	0 (0/28)	0 (0/29)	9 (2/23)
Sample collection date	May–October 2020	January–March 2021	May 2014–March 2018 (13/23) March–May 2020 (10/23)
SARS-CoV-2 PCR, %	positive 92 (12/13) not tested 54 (15/28)	N/A	N/A
S RBD IgG+ (%)	96 (27/28)	100 (29/29)	0 (10/10, collected in 2020)
Peak disease severity, %[a]			
Mild	86 (24/28)	N/A	N/A
Moderate	14 (4/28)		
Severe	0 (0/28)		
Critical	0 (0/28)		
Race-ethnicity, %			
White—not Hispanic or Latino	81 (23/28)	52 (15/29)	43 (10/23)
Hispanic or Latino	4 (1/28)	10 (3/29)	9 (2/23)
Asian	7 (2/28)	38 (11/29)	26 (6/23)
American Indian/Alaska Native	0 (0/28)	0 (0/29)	0 (0/23)
Black or African American	0 (0/28)	0 (0/29)	9 (2/23)
>1 race	4 (1/28)	0 (0/29)	4 (1/23)
Not reported	4 (1/28)	0 (0/29)	9 (2/23)
Days at collection	38–163 (28/28) (median = 78, IQR = 50)[b]	13–30 (14/29) Pfizer 12–16 (15/29) Moderna (median = 14, IQR = 14)[c]	N/A

[a]According to World Health Organization criteria.
[b]Post-symptom onset.
[c]2nd dose of vaccine.

CD4+ and CD8+ T cell reactivity against VOC S protein

We measured the CD4+ and CD8+ T cell responses to S MPs derived from the ancestral strain and corresponding MPs representing the B.1.1.7, B1.351, P.1, and CAL.20C variants. As shown in Figures 1A–1C, good CD4+ and CD8+ T cell reactivity was observed in convalescent donors with peptides spanning the S protein of the ancestral Wuhan sequence and the corresponding variant S peptides. Geomean S-specific responses ranged from 0.07 to 0.08 for CD4+ T cells and 0.08 to 0.10 for CD8+ T cells. No significant difference was observed between the ancestral and VOC S peptide pools (CD4: B.1.1.7 $p = 0.17$; B.1.351 $p = 0.30$; P.1 $p = 0.21$; CAL.20C $p = 0.06$ and CD8: B.1.1.7 $p = 0.15$; B.1.351 $p = 0.25$; P.1 $p = 0.47$; CAL.20C $p = 0.17$ by Wilcoxon test) (Figures 1A–1C). These values (here and in subsequent graphs) are not corrected for multiple comparisons, as a correction would decrease the statistical power for detecting significant differences; therefore, not performing multiple comparison corrections is the more conservative and stringent test.

To further test for potential differences in the recognition of VOCs by these T cells, we calculated fold change values per individual donor (Figures S2A and S2B). We then performed a non-inferiority test on one sample Wilcoxon signed-rank test compared to the lower bound fold change threshold calculated based on technical repeats (Table S3). No significant differences were observed, demonstrating that the decreases observed were within the range expected from technical repeat variation.

The VOC T cell analyses were extended using FluoroSPOT to measure the capacity of the various SARS-CoV-2 peptide pools to elicit functional responses in terms of secretion of IFN-γ and interleukin-5 (IL-5) cytokines (Figures 1D, 1E, and S2C). The results from the FluoroSPOT assay (Figures 1D and 1E) showed *ex vivo* IFN-γ reactivity using whole PBMCs, with a geomean of 45 SFCs/10^6 PBMCs (range 20–1,578) for the ancestral strain peptides. This overlaps with the 0–800 SFC range (median 110) detected in a previous study describing S reactivity in symptomatic donors[49] and with the 20–110 range for 2–5 months' symptomatic donors reported in a separate study.[50]

VOC reactivity was observed for the S MPs in convalescent donors, with geomean IFN-γ SFCs per million PBMCs ranging from 38 to 45 (Figure 1D). Compared to the ancestral strain, there were significant decreases of 12%, 6%, and 14% for the B.1.1.7, B.1.351, and CAL.20C variant pools (B.1.1.7 $p = 0.02$; B.1.351 $p = 0.03$; P.1 $p = 0.07$, and CAL.20C $p < 0.01$ by Wilcoxon test), while no difference was observed for P.1 (Figure 1D).

Please cite this article in press as: Tarke et al., Impact of SARS-CoV-2 variants on the total CD4+ and CD8+ T cell reactivity in infected or vaccinated individuals, Cell Reports Medicine (2021), https://doi.org/10.1016/j.xcrm.2021.100355

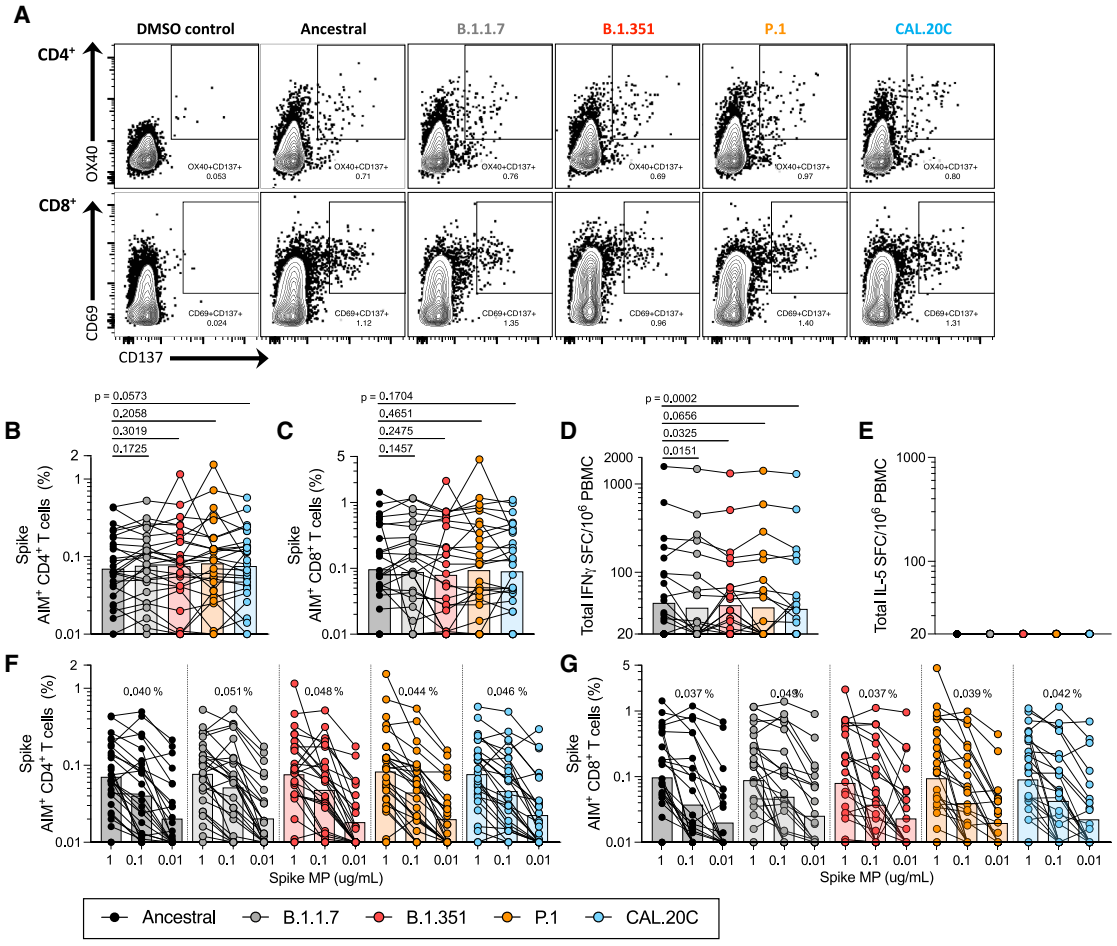


Figure 1. T cell responses of COVID-19 convalescent individuals against ancestral and variant SARS-CoV-2 spike (S)

PBMCs of COVID-19 convalescent individuals (n = 28) were stimulated with the S MPs corresponding to the ancestral reference strain (black) and the B.1.1.7 (gray), B.1.351 (red), P.1 (orange), and CAL.20C (light blue) SARS-CoV-2 variants.

(A) The gating strategy for the AIM assay is illustrated by representative graphs defining S-specific $CD4^+$ or $CD8^+$ T cells by expression of $OX40^+CD137^+$ and $CD69^+CD137^+$, respectively. These graphs depict 1 of the COVID-19 convalescent donors from this study tested with the S MPs corresponding to each of the VOCs tested.

(B) Percentages of AIM^+ ($OX40^+CD137^+$) $CD4^+$ T cells.

(C) Percentages of AIM^+ ($CD69^+CD137^+$) $CD8^+$ T cells.

(D) IFN-γ spot-forming cells (SFCs) per million PBMCs.

(E) IL-5 SFCs per million PBMCs.

(F and G) The data shown in (B) and (C) are plotted to show the titration of the S MPs (1 μg/mL, 0.1 μg/mL, and 0.01 μg/mL) for $CD4^+$ (F) and $CD8^+$ (G) T cells for each SARS-CoV-2 variant. The geometric mean of the 0.1 μg/mL condition is listed above each titration.

Paired comparisons of ancestral S MPs versus each of the variants were performed by 1-tailed Wilcoxon test and are indicated by the p values in (B)–(D). In all of the panels, the bars represent the geometric mean. See also Figures S1, S2, S4, and S5 and Tables S1–S3.

No significant differences were observed by fold change analysis, suggesting that the decreases observed were still within the technical fluctuation range (Figure S2C). No IL-5 reactivity was observed for any of the pools (Figure 1E).

To further assess the functionality of T cell recognition of these variants, we considered the variant peptide dose-response curves of the $CD4^+$ and $CD8^+$ T cells. As shown in Figures 1F and S2D, peptide concentration sensitivity of $CD4^+$ T cell responses to the ancestral and four variant pools was similar. The same pattern was also observed for $CD8^+$ T cell responses to S (Figures 1G and S2E).

$CD4^+$ and $CD8^+$ T cell total reactivity against VOCs

As shown in Table S1, mutations found in the variants studied herein were not limited to the S protein, but also occurred in several additional antigens encoded in the SARS-CoV-2 genome. To address the potential impact of non-S variant mutations on overall proteome-wide $CD4^+$ and $CD8^+$ T cell reactivity, we tested overlapping peptide MPs spanning the entire proteome of the ancestral Wuhan sequence in comparison with corresponding MPs representing the different variants.

Overall, reactivity to the peptide pools spanning the variant genomes was found to be similar to that against the ancestral

Please cite this article in press as: Tarke et al., Impact of SARS-CoV-2 variants on the total CD4+ and CD8+ T cell reactivity in infected or vaccinated individuals, Cell Reports Medicine (2021), https://doi.org/10.1016/j.xcrm.2021.100355

CellPress
OPEN ACCESS

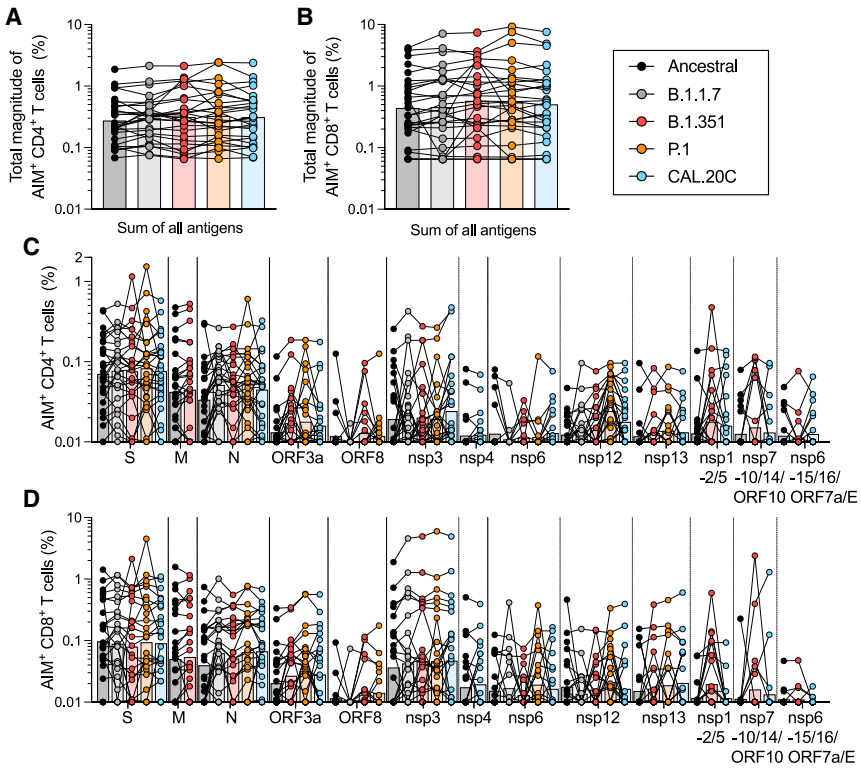


Figure 2. T cell responses of COVID-19 convalescent individuals against ancestral and variant SARS-CoV-2 proteomes

PBMCs of COVID-19 convalescent individuals (n = 28) were stimulated with MPs for the entire viral proteome corresponding to the ancestral reference strain (black) and the B.1.1.7 (gray), B.1.351 (red), P.1 (orange), and CAL.20C (light blue) SARS-CoV-2 variants.

(A) Percentages of AIM+ (OX40+CD137+) CD4+ T cells for the total reactivity.

(B) Percentages of AIM+ (CD69+CD137+) CD8+ T cells for the total reactivity.

(C) Percentages of AIM+ (OX40+CD137+) CD4+ T cells for each MP.

(D) Percentages of AIM+ (CD69+CD137+) CD8+ T cells for each MP.

All of the bars represent the geometric mean. See also Figure S1, S2, S4, and S5 and Tables S1–S3.

CD4+ and CD8+ T cell reactivity against VOCs by vaccinees

We also studied T cell responses by individuals who received US Food and Drug Administration (FDA)-authorized COVID-19 mRNA vaccines. We focused our analysis on T cell reactivity to the S antigen of the ancestral strain, which is the basis of the presently used vaccines.

Wuhan strain (Figures 2 and S2). When the sum total of reactivity throughout the genome was considered, no decreases in reactivity compared to the ancestral were noted for the variant pools (Figures 2A–2C, S2F, and S2G).

We previously showed in COVID-19 convalescent subjects that a set of 10 different antigens (nsp3, nsp4, nsp6, nsp12, nsp13, S, ORF3a, M, ORF8, and N) account for 83% and 81% of the total CD4+ and CD8+ T cell response, respectively.[51] Here, a similar overall pattern of dominant antigens was observed (Figures 2C and 2D). It is worth noting that this specific comparison is for illustration purposes only, as this study is not powered to rule out differences in individual antigens.

For comparison purposes, unexposed donors were also tested in the AIM assay, with MPs encompassing the ancestral and variant strains (Figures S2H–S2M). As expected, some T cell reactivity against the SARS-CoV-2 MPs was observed in unexposed donors, possibly due to previous exposure to common cold coronaviruses[47,48,52,53] or to other pathogens or autoantigens. The magnitude of the responses was lower than in COVID-19 convalescents (unexposed versus convalescent: CD4 $p < 0.0001$; CD8 $p < 0.0001$ by the Mann-Whitney test, comparison not shown in the graphs). Similar to the COVID-19 convalescents, no decrease in the geomean of AIM+ CD4+ or CD8+ T cell responses to the sum of all of the antigens of individual MPs tested was observed in unexposed donors (Figures S2H–S2K). These experiments suggest that CD4+ and CD8+ T cells from individuals infected with the ancestral SARS-CoV-2 strain recognize VOCs with similar efficiency.

For both CD4+ and CD8+ T cell reactivity, the magnitude of responses to pools encompassing the sequences from the ancestral Wuhan genome and the VOCs ranged from a geomean of 0.15–0.17 for CD4+ T cells and a geomean of 0.10–0.15 for CD8+ T cells (Figures 3A–3C). Comparison of the variant pools to the ancestral sequence showed no significant differences for CD4+ and CD8+ T cell reactivity in the AIM assay for B.1.1.7 and P.1 (CD4: B.1.1.7 $p = 0.41$, P.1 $p = 0.29$; CD8: B.1.1.7 $p = 0.10$, P.1 $p = 0.09$ by Wilcoxon test). Decreases of 14% and 22%, respectively, were observed with the B.1.351 pools for CD4+ and CD8+ T cells (B.1.351: $p < 0.01$ for both comparisons), and a 10% decrease with the CAL.20C pool for CD8+ T cells ($p = 0.04$) (Figures 3A–3C). The FluoroSPOT assay (Figures 3D–3F and S3C) showed IFN-γ reactivity, with geomeans ranging from 58 to the 74 SFCs per million PBMCs (Figures 3C and S3C). Given the number of comparisons made and that the reductions observed were on the low side (and others showed increases), we wanted to put these numbers into context by comparing them to what is expected based on the observed fold changes in repeat measurements of the same samples on different days for both AIM and FluoroSPOT assays, and asked whether the observed decreases were significantly higher (non-inferiority analysis; see Method details). No significant differences were observed, indicating that the magnitude of decreases for some strains was within the expected technical assay variability (Figures S3A–S3C). Minimal IL-5 responses were observed, with geomean reactivity ranging from 22 to 23 SFCs/10^6, which is slightly above the limit of detection (Figure 3E). On a per-donor basis, the IFN-γ response was found to account for >80% of the total response,

Please cite this article in press as: Tarke et al., Impact of SARS-CoV-2 variants on the total CD4+ and CD8+ T cell reactivity in infected or vaccinated individuals, Cell Reports Medicine (2021), https://doi.org/10.1016/j.xcrm.2021.100355

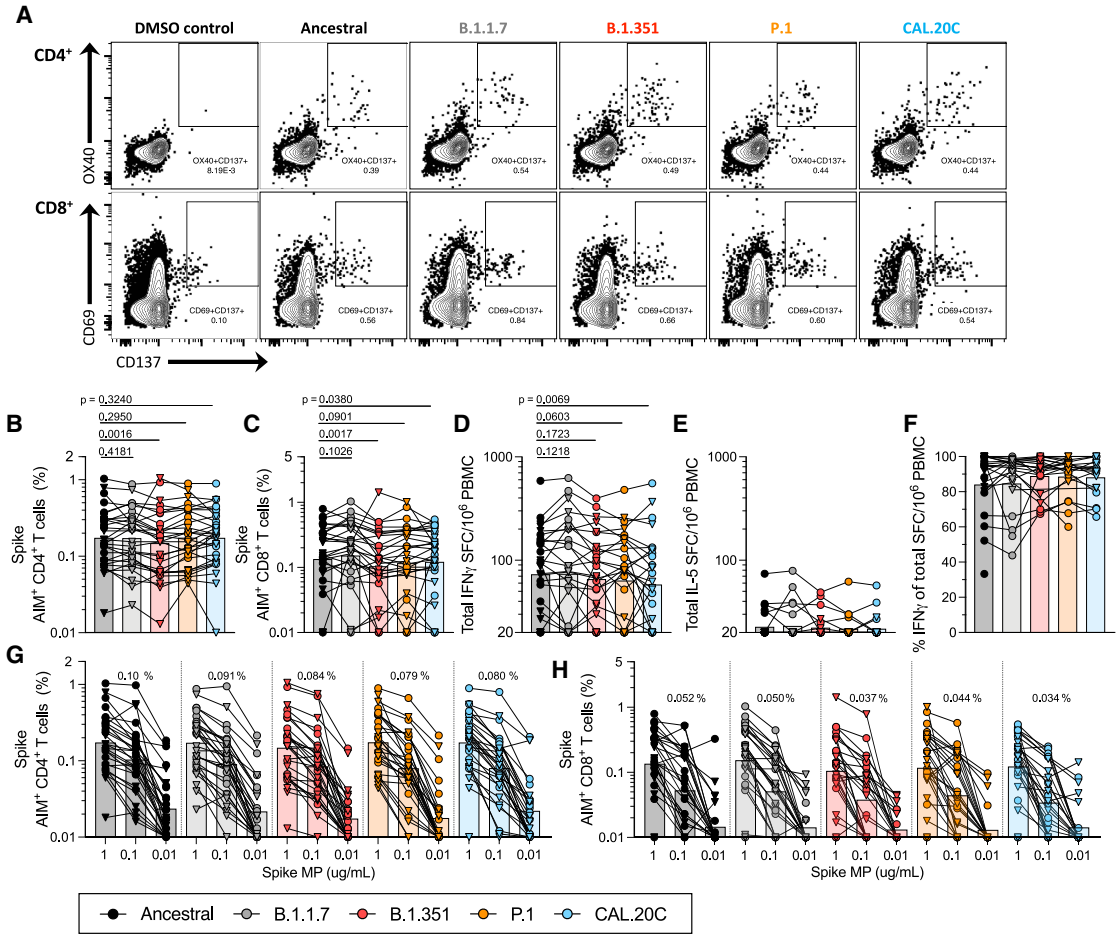


Figure 3. T cell responses of COVID-19 vaccinees against ancestral and variant SARS-CoV-2 S

PBMCs of Pfizer/BioNTech BNT162b2 (n = 14, triangles) and Moderna mRNA-1273 COVID-19 vaccinees (n = 15, circles) were stimulated with the S MPs corresponding to the ancestral reference strain (black) and the B.1.1.7 (gray), B.1.351 (red), P.1 (orange), and CAL.20C (light blue) SARS-CoV-2 variants.

(A) The gating strategy for the AIM assay is illustrated by representative graphs defining S-specific CD4+ or CD8+ T cells by the expression of OX40+CD137+ and CD69+CD137+, respectively. These graphs depict one of the COVID-19 vaccinated donors from this study tested with the S MPs corresponding to each of the VOCs tested.

(B) Percentages of AIM+ (OX40+CD137+) CD4+ T cells.

(C) Percentages of AIM+ (CD69+CD137+) CD8+ T cells.

(D) IFN-γ SFCs per million PBMCs.

(E) IL-5 SFCs per million PBMCs.

(F) Percentages of IFN-γ were calculated from the total IFN-γ and IL-5 SFCs per million PBMCs.

(G and H) The data shown in (B) and (C) are also plotted showing the S MPs titration (1, 0.1, and 0.01 μg/mL) for CD4+ (G) and CD8+ (H) T cells with each SARS-CoV-2 variant. The geometric mean of the 0.1 μg/mL condition is listed above each titration.

Paired comparisons of the ancestral reference strain-based S MP versus each of the variants were performed by one-tailed Wilcoxon test and are indicated by the p values in (B)–(D). In all of the panels, the bars represent the geometric mean. See also Figures S1 and S3–S5 and Tables S1–S3.

on average (range 84%–89%), irrespective of whether the ancestral strain or VOCs were considered (Figure 3F).

Similar to the experiments in convalescent donors, we also examined T cell dose responses in vaccinated donors. As shown in Figures 3G, 3H, S3D, and S3E, CD4+ and CD8+ T cell dose responses to the ancestral and VOC S pools were similar.

Phenotypes of CD4+ and CD8+ T cells

We further assessed the quality of AIM+ T cell responses by phenotyping S-specific CD4+ and CD8+ AIM+ T cell responses. Representative gating strategy is shown in Figure S5. Consistent with previous observations,[39,48] AIM+ T cell responses in COVID-19 convalescent donors above the threshold of positivity (CD4 >0.016% and CD8 >0.16%, as described in Method details; Figure S4), irrespective of the variant analyzed, had a memory phenotype, preferentially enriched for central (Tcm) and effector memory (Tem) phenotypes for CD4+ T cells (CD4+Tcm CD45RA+CCR7− and CD4+Tem CD45RA−CCR7−) (Figure S4A) and Tem phenotypes for CD8+ T cells (CD8+Temra CD45RA+CCR7−) (Figure S4B). Likewise, AIM+ T cell responses

Please cite this article in press as: Tarke et al., Impact of SARS-CoV-2 variants on the total CD4+ and CD8+ T cell reactivity in infected or vaccinated individuals, Cell Reports Medicine (2021), https://doi.org/10.1016/j.xcrm.2021.100355

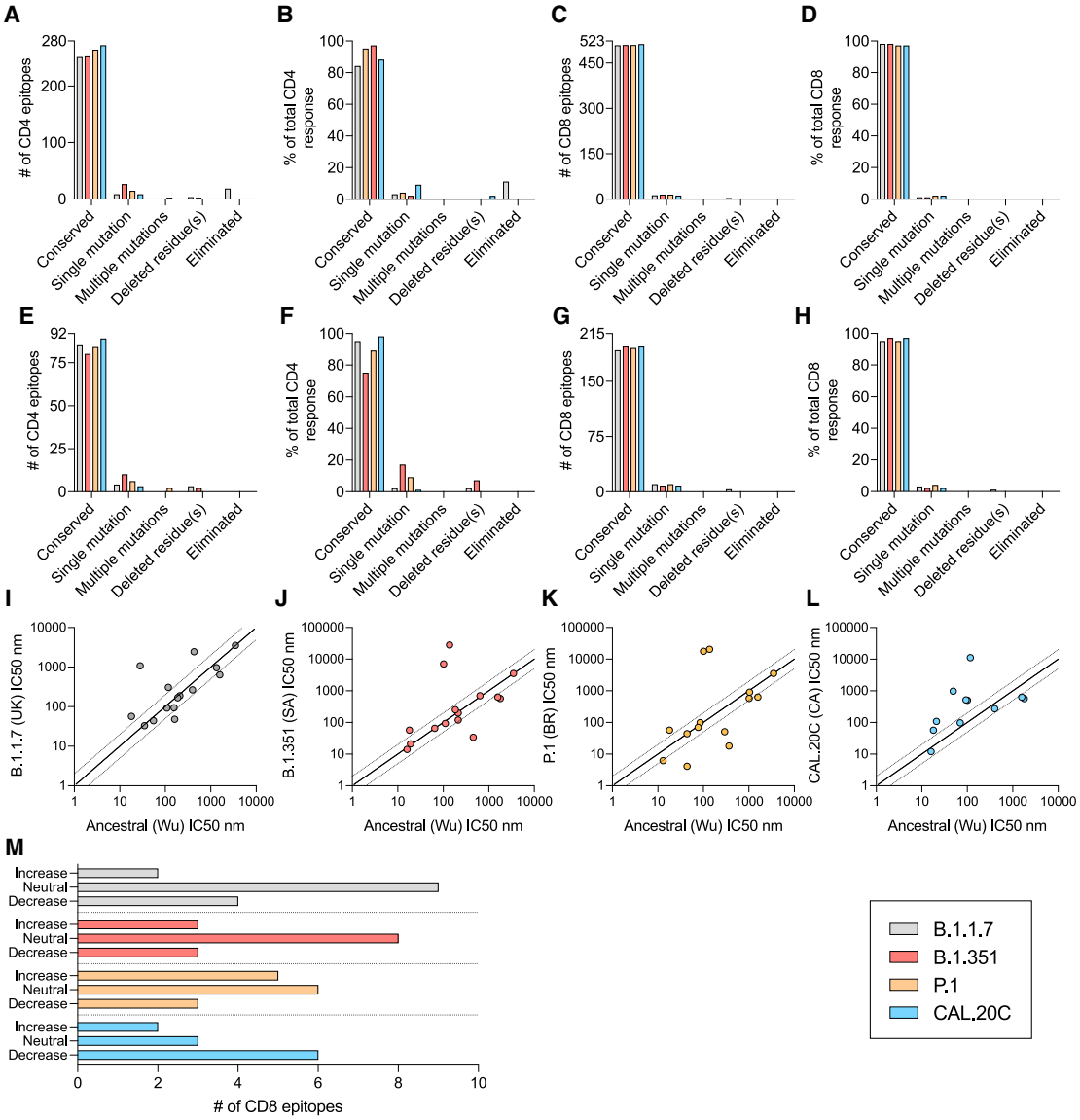


Figure 4. SARS-CoV-2 T cell epitope sequences affected by the variants

CD4+ and CD8+ T cell epitopes of the ancestral strain identified in a previous study[51] were analyzed as a function of the number and percentage of responses that are or are not conserved across the B.1.1.7 (gray), B.1.351 (red), P.1 (orange), and CAL.20C (light blue) SARS-CoV-2 variants.

(A–D) The SARS-CoV-2 epitopes for the most immunodominant SARS-CoV-2 proteins in terms of numbers and percentage of response are shown for CD4+ (A and B) and CD8+ (C and D) T cells.

(E–H) The SARS-CoV-2 epitopes for the S protein only in terms of numbers and percentage of response are shown for CD4+ (E and F) and CD8+ (G and H) T cells.

(I–M) The peptide-binding algorithm data of mutated versus wild-type epitopes are shown for CD8+ (I–L) and for the effect of the mutations on CD8+ T cell epitopes, in which each instance was categorized as a function of whether the binding capability of the mutated peptide was increased (>2-fold), neutral, or decreased (<2-fold) (M).

See also Table S4.

in COVID-19 vaccinees had a memory phenotype regardless of the origin of the SARS-CoV-2 sequences analyzed, with preferential enrichment for Tcm and Tem for CD4 (Figure S4C) and Tem and Temra for CD8 (Figure S4D). This provides additional evidence that donors primed by the ancestral strain S protein, by either infection or vaccination, mount a memory T cell response able to cross-recognize the SARS-CoV-2 VOCs.

Conservation analysis of sets of defined CD4+ and CD8+ T cell epitopes

An analysis of experimentally defined CD4 and CD8 epitopes is presented in Figure 4 and examines both the number of epitopes and the experimentally determined magnitude of responses associated with the epitopes. Specifically, we recently reported a comprehensive study of epitopes recognized in convalescent

subjects, leading to the identification of 280 different CD4+ T cell epitopes.[51] Here, we analyzed how many of those epitopes would be affected by mutations in the different variants. As shown in Figure 4A, we found that 89.6%, 90%, 94.3%, and 97.1% (average 93%) of the CD4+ T cell epitopes are conserved in the B.1.1.7, B1.351, P.1, and CAL.20C variants. A similar pattern is observed when the magnitude of T cell responses associated with the various epitopes is considered, rather than the simple number of epitopes. Fully conserved CD4+ T cell epitopes can be inferred to account for 84.4%, 88.1%, 95.7%, and 97.8% (average 91.5%) of the recognition of the B.1.1.7, B.1.351, P.1, and CAL.20C variants, respectively, based on the ancestral sequence (Figure 4B).

That same study also reported 523 CD8+ T cell epitopes.[51] Performing a similar analysis, 508 (97.1%) of these 523 CD8+ T cell epitopes are totally conserved within the B.1.1.7 variant, 509 (97.3%) within the B.1.351 and P.1 variants, and 512 (97.9%) within the CAL.20C variant (Figure 4C). Similarly, in terms of magnitude of the CD8+ T cell responses associated with the various epitopes, fully conserved CD8+ T cell epitopes can be inferred to account for 98.3%, 98.4%, 97.9%, and 97.8% of the recognition of the B.1.1.7, B.1.351, P.1, and CAL.20C variants, respectively, based on the ancestral sequence (Figure 4D, average of 98.1%).

Finally, we analyzed the degree of CD4+ and CD8+ T cell epitope conservation with an analysis restricted to epitopes in the S antigen. The number of S-derived epitopes conserved at 100% sequence identity was, on average, 84.5% for the CD4+ T cell epitopes (Figure 4E) and 95.3% for the CD8+ T cell epitopes (Figure 4G). Similarly, in terms of magnitude of CD4+ T cell responses to S epitopes, fully conserved CD4+ T cell epitopes can be inferred to account for 95.5%, 75.3%, 89.8%, and 98.3% of the recognition of B.1.1.7, B.1.351, P.1, and CAL.20C variants, respectively, with an average of 89.7% (Figure 4F). For CD8+ T cell responses, fully conserved epitopes can be inferred to account for 95.2%, 97.6%, 95.4%, and 97.3% of the recognition of B.1.1.7, B.1.351, P.1, and CAL.20C variants, respectively, with an average of 96.4% (Figure 4H).

While human leukocyte antigen (HLA) restriction of the class II epitopes[51] could not be unequivocally assigned, the restriction of the class I epitopes was implicitly inferred based on HLA allele-specific predictions and testing in HLA-matched donors. Accordingly, we analyzed the predicted binding affinity for each epitope and matching variant for the corresponding putative HLA class I restriction element. The predicted binding affinity for each ancestral epitope/matching variant for B.1.1.7, B.1.351, P.1, and CAL.20C variants is shown in Figures 4I–4L, tabulated in Table S4, and summarized in Figure 4M. The predicted binding capacity was determined using the NetMHCpan BA 4.1 tool provided by the Immune Epitope Database's (IEDB's) analysis resource.[54,55] In the case of the B.1.1.7, B.1.351, P.1, and CAL.20C variants, the number of mutations associated with binding capacity decreases (conservatively defined as a 2-fold reduction) was 4 out of 15, 3 out of 14, 3 out of 14, and 6 out of 11, respectively.

In conclusion, the analyses suggest that the vast majority of CD4+ and CD8+ T cell epitopes are unaffected by mutations found in all of the different variants. The corresponding mutations are also predicted to have minor effects on the total T cell response, thus providing a molecular basis for the overall impact on T cell reactivity by COVID-19 convalescent subjects and recipients of COVID-19 mRNA vaccines.

DISCUSSION

The present study addresses a key knowledge gap pertaining to the potential of emergent SARS-CoV-2 variants to evade overall recognition by human immune responses. We focused on T cell responses elicited by either natural infection or vaccination with the Pfizer/BioNTech and Moderna COVID-19 mRNA vaccines. We found similar total CD4+ or CD8+ T cell reactivity against the four variants investigated herein, B.1.1.7, B.1.351, P.1, and CAL.20C lineages found initially in the United Kingdom, South Africa, Brazil, and California, respectively. To assess the overall total T cell functionalities, the comparison between the original Wuhan isolate and the variants was performed using different T cell methodologies, such as the AIM assay (quantifying T cells with a range of functionalities), and the FluoroSPOT assay (quantifying cells with specific cytokine-secreting activity). We also tested whether any of the variant sequences may be associated with an altered cytokine polarization.

The data herein provide some positive news in light of justified concern over the impact of SARS-CoV-2 VOCs on the efforts to control and eliminate the present pandemic. Multiples of the VOCs are associated with increased transmissibility, and several have been associated with decreased susceptibility to neutralizing antibodies from infected or vaccinated individuals. In contrast, the data presented here suggest that the total T cell responses are not significantly disrupted by the VOCs. While it is not anticipated that circulating memory T cells would be effective in preventing SARS-CoV-2 infection, it is plausible that they can contribute in reducing COVID-19 severity.[33,56] Several lines of evidence support this notion, such as observations that early SARS-CoV-2 T cell responses are associated with milder COVID-19.[34,35] Thus, the T cell response may contribute to limiting COVID-19 severity induced by VOCs that partially or largely escape neutralizing antibodies. This is consistent with T cell-mediated immunity observed in humans against a different respiratory pathogen, influenza, for which heterologous immunity against diverse influenza strains is associated with memory T cells to conserved epitopes.[57–59]

Our data also provide insights on the predicted impact of the mutations associated with the variants analyzed on T cell responses in the context of the T cell epitopes recognized. Prior reports have identified a large number of T cell epitopes recognized throughout the SARS-CoV-2 proteome, including S.[40,53,60–63] We furthered this point by an analysis of the Tarke et al. dataset, showing that 93% of CD4+ and 97% of CD8+ T cell epitopes in SARS-CoV-2 are completely conserved in the variants. Furthermore, we found that even in the epitopes affected by single mutations, no negative effect on the predicted HLA-binding capacity in the majority of cases is expected. Overall, it is plausible to hypothesize that single amino acid substitutions or deletions across large peptidomes do not significantly affect a polyclonal memory T cell response. Nevertheless, the methodology applied in this study is not powered to reveal whether T cells specific to amino acid sequences that are changed in variant strains still retain functionality. The apparent higher conservation of CD8+ T cell epitopes is to be expected

Please cite this article in press as: Tarke et al., Impact of SARS-CoV-2 variants on the total CD4+ and CD8+ T cell reactivity in infected or vaccinated individuals, Cell Reports Medicine (2021), https://doi.org/10.1016/j.xcrm.2021.100355

based on the shorter length of HLA class I binding peptides (usually 9–10 amino acids) as compared to their class II counterparts (13–17 amino acids). This effect is counterbalanced by $CD8^+$ T cells being generally less tolerant of amino substitutions as compared to $CD4^+$ T cells.[47,64] Marginal IL-5 production was detected in all of the conditions tested. This is relevant, since it was reported that single amino acid replacements in an epitope sequence can lead to a change in the cytokines produced,[65,66] and a Th2-like response pattern was initially hypothesized to be linked to adverse outcomes in SARS-specific responses.[67] Overall, we observed that VOC mutations do not significantly disrupt the total $CD4^+$ and $CD8^+$ T cell responses.

VOC mutations could be reflective of adaptation in terms of optimizing replication or binding to ACE2, but also reflective of adaptation to escape immune recognition by antibodies.[13,14,22,68,69] Higher viral binding to a cellular receptor can be a mechanism of compensatory viral evolution in the presence of neutralizing antibodies.[70] In that respect, while mutations to escape antibody binding have been well documented for influenza[71–73] and SARS-CoV-2, immune escape at the level of T cell responses in human populations has not been reported for acute respiratory infections. Because of HLA polymorphism, the epitope repertoire recognized is likely to be substantially different from one individual to the next, greatly decreasing the likelihood of immune escape by an acute virus. An advantage conferred to the virus by a mutation in one person would not be linked to an immune response escape advantage in a non-HLA-matched individual. For SARS-CoV-2, this property of T cell recognition is further enhanced by the fact that the T cell responses against SARS-CoV-2 are highly multiantigenic and multi-specific, with tens of different epitopes recognized by $CD4^+$ and $CD8^+$ T cells in a given individual.[51,52,60,62] Nevertheless, our data do not rule out that certain individuals could be strongly affected by the mutations in specific VOCs.

These results here have potential implications for engineering coronavirus vaccines with broader protective immunity against VOCs. Clearly, the most straightforward path is to update the current vaccines to target a variant S, given how highly successful several COVID-19 vaccines have proven to be against the parental SARS-CoV-2 strain. Our results suggest that a parallel alternative approach could involve the inclusion of additional antigens and T cell epitopes, perhaps selected on the basis of low mutational propensity,[74] to ensure that neutralizing antibodies are complemented with T cell responses to minimize COVID-19 morbidity and mortality.

Limitations of the study

The present study did not assess decreases in antibody reactivity, as several other studies have already investigated this matter.[11,14–20,22,23] Our studies used overlapping peptide pools. As such, we cannot exclude that some of the mutations may involve alterations in terms of antigen processing for either class I or class II, which would be undetected by using pools of "preprocessed" peptides. While we have no reason to suspect that substantial differences may exist between the epitope specificity of responses elicited by different vaccines, our study did not address this point.

The statistical power of this study does not allow for the sensitivity needed to detect the loss of small populations of T cell clones that may be affected by variant sequences when sampled in the presence of the majority of conserved peptides. Alterations in major histocompatibility complex (MHC) binding do not necessarily confer changes in T cell receptor (TCR) affinity, as often several MHC-binding residues can be changed before changing TCR signaling, so we cannot exclude impaired T cell recognition of those specific mutated peptides. An additional limitation of the study in this respect is that one amino acid difference could influence TCR signal strength and memory T cell reactivation. However, each epitope can be recognized by multiple TCRs, and calculating the affinity would require cloning the multiple TCRs recognizing each of the epitopes seen in a polyclonal setting.

Our study was designed to test for differences in total T cell response against the different variants and was not powered or designed to investigate differences between the mRNA vaccines or at the single-peptide level. Furthermore, in our study, the sequences of the infecting virus were not determined, and therefore we cannot exclude that some of the donors may have been exposed to variants. To minimize this issue, the tested samples from convalescent donors were selected to be infected before October 2020; thus, it is unlikely that any of the donors would have been infected by any of the VOCs, as this date precedes their diffusion to an appreciable degree in the United States in general, and California in particular. While we could not clearly define the infecting strain in the COVID-19 convalescent, we showed similar results in vaccinated subjects in whom the response was induced by the ancestral S sequence. As also mentioned in the Results section, the cohorts investigated were predominantly White, reflective of the patient population available for recruitment.

Finally, we have only investigated whether the total T cell responses induced by the ancestral reference sequence are able to cross-recognize variant sequences, as this is relevant to the present situation. We have not examined whether responses induced by an infection with a variant sequence will be able to cross-recognize the ancestral reference sequence present in the approved vaccines.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Convalescent COVID-19 donors
  - COVID-19 vaccinees
  - Unexposed donors
- METHOD DETAILS
  - Isolation of peripheral blood mononuclear cells (PBMCs) and plasma
  - SARS-CoV-2 RBD ELISA
  - Mutation analysis of SARS-CoV-2 B.1.1.7, B.1.351, P.1 and CAL.20C variants
  - SARS-CoV-2 Wuhan and variant peptide synthesis and pooling

- Bioinformatic analysis of T cell epitopes
- Flow cytometry-based AIM assay
- FluoroSPOT assay

● QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.xcrm.2021.100355.

ACKNOWLEDGMENTS

This study has been funded by the NIH NIAID (SARS-CoV-2 Assessment of Viral Evolution program, award no. AI142742 to S.C. and A. Sette, contract no. 75N9301900065 to A. Sette and D.W., contract no. 75N93019C00001 to A. Sette and B.P., NIH grant U01 CA260541-01 to D.W., K08 award AI135078 to J.M.D., AI036214 to D.M.S., and HHSN75N93019C00076 to R.H.S.). Additional support has been provided by UCSD T32s (AI007036 and AI007384 to S.A.R. and S.I.R.) and the Jonathan and Mary Tu Foundation (to D.M.S.). A.T. was supported by a PhD student fellowship through the Clinical and Experimental Immunology Course at the University of Genoa, Italy. We thank Gina Levi and the LJI clinical core for assistance in sample coordination and blood processing. We gratefully thank the authors from the originating laboratories responsible for obtaining the specimens, as well as the submitting laboratories where the genome data were generated and shared via GISAID and on which this research is based. We would like to thank Vamseedhar Rayaprolu and Erica Ollmann Saphire for providing the recombinant SARS-CoV-2 RBD protein used in the ELISA assay.

AUTHOR CONTRIBUTIONS

Conceptualization, A.T., A.G., S.C., and A. Sette.; data curation and bioinformatic analysis, Y.Z., R.H.S., and B.P.; formal analysis, A.T., J.S., and A.G.; funding acquisition, S.C., A.S., D.W., S.I.R., S.A.R., and J.M.D.; investigation, A.T., N.M., A. Sutherland, B.G., J.S., E.D.Y., E.W., D.W., A. Sette, and A.G.; project administration, A.F.; resources, S.I.R., S.A.R., and J.M.D.; supervision, J.S., S.C., D.W., A. Sette, R.d.S.A., and A.G.; writing, A.T., D.W., S.C., A. Sette, and A.G.

DECLARATION OF INTERESTS

A. Sette is a consultant for Gritstone, Flow Pharma, CellCarta, Arcturus, Oxford-immunotech, and Avalia. S.C. is a consultant for Avalia. All of the other authors declare no competing interests. LJI has filed for patent protection for various aspects of vaccine design and identification of specific epitopes.

Received: February 27, 2021
Revised: June 7, 2021
Accepted: June 24, 2021
Published: July 2, 2021

REFERENCES

1. Kirby, T. (2021). New variant of SARS-CoV-2 in UK causes surge of COVID-19. Lancet Respir. Med. *9*, e20–e21.
2. Volz, E., Mishra, S., Chand, M., Barrett, J.C., Johnson, R., Geidelberg, L., Hinsley, W.R., Laydon, D.J., Dabrera, G., O'Toole, Á., et al.; COVID-19 Genomics UK (COG-UK) consortium (2021). Assessing transmissibility of SARS-CoV-2 lineage B.1.1.7 in England. Nature *593*, 266–269.
3. Davies, N.G., Abbott, S., Barnard, R.C., Jarvis, C.I., Kucharski, A.J., Munday, J.D., Pearson, C.A.B., Russell, T.W., Tully, D.C., Washburne, A.D., et al.; CMMID COVID-19 Working Group; COVID-19 Genomics UK (COG-UK) Consortium (2021). Estimated transmissibility and impact of SARS-CoV-2 lineage B.1.1.7 in England. Science *372*, eabg3055.
4. Tegally, H., Wilkinson, E., Lessells, R.J., Giandhari, J., Pillay, S., Msomi, N., Mlisana, K., Bhiman, J.N., von Gottberg, A., Walaza, S., et al. (2021). Sixteen novel lineages of SARS-CoV-2 in South Africa. Nat. Med. *27*, 440–446.
5. Voloch, C.M., da Silva Francisco, R., Jr., de Almeida, L.G.P., Cardoso, C.C., Brustolini, O.J., Gerber, A.L., Guimarães, A.P.C., Mariani, D., da Costa, R.M., Ferreira, O.C., Jr., et al.; Covid19-UFRJ Workgroup, LNCC Workgroup, Adriana Cony Cavalcanti (2021). Genomic characterization of a novel SARS-CoV-2 lineage from Rio de Janeiro, Brazil. J. Virol. *95*, e00119-21.
6. Faria, N.R., Mellan, T.A., Whittaker, C., Claro, I.M., Candido, D.D.S., Mishra, S., Crispim, M.A.E., Sales, F.C.S., Hawryluk, I., McCrone, J.T., et al. (2021). Genomics and epidemiology of the P.1 SARS-CoV-2 lineage in Manaus, Brazil. Science *372*, 815–821.
7. Zhang, W., Davis, B.D., Chen, S.S., Sincuir Martinez, J.M., Plummer, J.T., and Vail, E. (2021). Emergence of a Novel SARS-CoV-2 Variant in Southern California. JAMA *325*, 1324–1326.
8. Deng, X., Garcia-Knight, M.A., Khalid, M.M., Servellita, V., Wang, C., Morris, M.K., Sotomayor-González, A., Glasner, D.R., Reyes, K.R., Gliwa, A.S., et al. (2021). Transmission, infectivity, and antibody neutralization of an emerging SARS-CoV-2 variant in California carrying a L452R spike protein mutation. medRxiv, 2021.2003.2007.21252647.
9. Rambaut, A., Loman, N., Pybus, O., Barclay, W., Barrett, J., Carabelli, A., Connor, T., Peacock, T., Robertson, D.L., and Volz, E.; COVID-19 Genomics Consortium UK (CoG-UK (2020). Preliminary genomic characterisation of an emergent SARS-CoV-2 lineage in the UK defined by a novel set of spike mutations. https://virological.org/t/preliminary-genomic-characterisation-of-an-emergent-sars-cov-2-lineage-in-the-uk-defined-by-a-novel-set-of-spike-mutations/563.
10. Washington, N.L., Gangavarapu, K., Zeller, M., Bolze, A., Cirulli, E.T., Barrett, K.M.S., Larsen, B.B., Anderson, C., White, S., Cassens, T., et al. (2021). Genomic epidemiology identifies emergence and rapid transmission of SARS-CoV-2 B.1.1.7 in the United States. medRxiv. https://doi.org/10.1101/2021.2002.2006.21251159.
11. Greaney, A.J., Loes, A.N., Crawford, K.H.D., Starr, T.N., Malone, K.D., Chu, H.Y., and Bloom, J.D. (2021). Comprehensive mapping of mutations in the SARS-CoV-2 receptor-binding domain that affect recognition by polyclonal human plasma antibodies. Cell Host Microbe *29*, 463–476.e6.
12. Starr, T.N., Greaney, A.J., Addetia, A., Hannon, W.W., Choudhary, M.C., Dingens, A.S., Li, J.Z., and Bloom, J.D. (2021). Prospective mapping of viral mutations that escape antibodies used to treat COVID-19. Science *371*, 850–854.
13. Zahradník, J., Marciano, S., Shemesh, M., Zoler, E., Chiaravalli, J., Meyer, B., Dym, O., Elad, N., and Schreiber, G. (2021). SARS-CoV-2 RBD *in vitro* evolution follows contagious mutation spread, yet generates an able infection inhibitor. bioRxiv. https://doi.org/10.1101/2021.2001.2006.425392.
14. Wang, P., Nair, M.S., Liu, L., Iketani, S., Luo, Y., Guo, Y., Wang, M., Yu, J., Zhang, B., Kwong, P.D., et al. (2021). Antibody resistance of SARS-CoV-2 variants B.1.351 and B.1.1.7. Nature *593*, 130–135.
15. Skelly, D.T., Harding, A.C., Gilbert-Jaramillo, J., Knight, M.L., Longet, S., Brown, A., Adele, S., Adland, E., and Brown, H., et al.; Medawar Laboratory Team (2021). Vaccine-induced immunity provides more robust heterotypic immunity than natural infection to emerging SARS-CoV-2 variants of concern. Research Square, 10.21203/rs.3.rs-226857/v2.
16. Wu, K., Werner, A.P., Koch, M., Choi, A., Narayanan, E., Stewart-Jones, G.B.E., Colpitts, T., Bennett, H., Boyoglu-Barnum, S., Shi, W., et al. (2021). Serum Neutralizing Activity Elicited by mRNA-1273 Vaccine. N. Engl. J. Med. *384*, 1468–1470.
17. Muik, A., Wallisch, A.K., Sänger, B., Swanson, K.A., Mühl, J., Chen, W., Cai, H., Maurus, D., Sarkar, R., Türeci, Ö., et al. (2021). Neutralization of SARS-CoV-2 lineage B.1.1.7 pseudovirus by BNT162b2 vaccine-elicited human sera. Science *371*, 1152–1153.
18. Supasa, P., Zhou, D., Dejnirattisai, W., Liu, C., Mentzer, A.J., Ginn, H.M., Zhao, Y., Duyvesteyn, H.M.E., Nutalai, R., Tuekprakhon, A., et al. (2021).

Reduced neutralization of SARS-CoV-2 B.1.1.7 variant by convalescent and vaccine sera. Cell *184*, 2201–2211.e7.

19. Edara, V.V., Norwood, C., Floyd, K., Lai, L., Davis-Gardner, M.E., Hudson, W.H., Mantus, G., Nyhoff, L.E., Adelman, M.W., Fineman, R., et al. (2021). Infection- and vaccine-induced antibody binding and neutralization of the B.1.351 SARS-CoV-2 variant. Cell Host Microbe *29*, 516–521.e3.
20. Shen, X., Tang, H., McDanal, C., Wagh, K., Fischer, W., Theiler, J., Yoon, H., Li, D., Haynes, B.F., Sanders, K.O., et al. (2021). SARS-CoV-2 variant B.1.1.7 is susceptible to neutralizing antibodies elicited by ancestral spike vaccines. Cell Host Microbe *29*, 529–539.e3.
21. Stamatatos, L., Czartoski, J., Wan, Y.H., Homad, L.J., Rubin, V., Glantz, H., Neradilek, M., Seydoux, E., Jennewein, M.F., MacCamy, A.J., et al. (2021). mRNA vaccination boosts cross-variant neutralizing antibodies elicited by SARS-CoV-2 infection. Science *372*, 1413–1418.
22. Wang, Z., Schmidt, F., Weisblum, Y., Muecksch, F., Barnes, C.O., Finkin, S., Schaefer-Babajew, D., Cipolla, M., Gaebler, C., Lieberman, J.A., et al. (2021). mRNA vaccine-elicited antibodies to SARS-CoV-2 and circulating variants. Nature *592*, 616–622.
23. Wibmer, C.K., Ayres, F., Hermanus, T., Madzivhandila, M., Kgagudi, P., Oosthuysen, B., Lambson, B.E., de Oliveira, T., Vermeulen, M., van der Berg, K., et al. (2021). SARS-CoV-2 501Y.V2 escapes neutralization by South African COVID-19 donor plasma. Nat. Med. *27*, 622–625.
24. Emary, K.R.W., Golubchik, T., Aley, P.K., Ariani, C.V., Angus, B., Bibi, S., Blane, B., Bonsall, D., Cicconi, P., Charlton, S., et al.; COVID-19 Genomics UK Consortium; AMPHEUS Project; Oxford COVID-19 Vaccine Trial Group (2021). Efficacy of ChAdOx1 nCoV-19 (AZD1222) vaccine against SARS-CoV-2 variant of concern 202012/01 (B.1.1.7): an exploratory analysis of a randomised controlled trial. Lancet *397*, 1351–1362.
25. Cele, S., Gazy, I., Jackson, L., Hwa, S.H., Tegally, H., Lustig, G., Giandhari, J., Pillay, S., Wilkinson, E., Naidoo, Y., et al.; Network for Genomic Surveillance in South Africa; COMMIT-KZN Team (2021). Escape of SARS-CoV-2 501Y.V2 from neutralization by convalescent plasma. Nature *593*, 142–146.
26. Voysey, M., Clemens, S.A.C., Madhi, S.A., Weckx, L.Y., Folegatti, P.M., Aley, P.K., Angus, B., Baillie, V.L., Barnabas, S.L., Bhorat, Q.E., et al.; Oxford COVID Vaccine Trial Group (2021). Safety and efficacy of the ChAdOx1 nCoV-19 vaccine (AZD1222) against SARS-CoV-2: an interim analysis of four randomised controlled trials in Brazil, South Africa, and the UK. Lancet *397*, 99–111.
27. Hall, V.J., Foulkes, S., Saei, A., Andrews, N., Oguti, B., Charlett, A., Wellington, E., Stowe, J., Gillson, N., Atti, A., et al. (2021). Effectiveness of BNT162b2 mRNA Vaccine Against Infection and COVID-19 Vaccine Coverage in Healthcare Workers in England, Multicentre Prospective Cohort Study (the SIREN Study). Lancet *397*, 1725–1735.
28. Amit, S., Regev-Yochay, G., Afek, A., Kreiss, Y., and Leshem, E. (2021). Early rate reductions of SARS-CoV-2 infection and COVID-19 in BNT162b2 vaccine recipients. Lancet *397*, 875–877.
29. Dagan, N., Barda, N., Kepten, E., Miron, O., Perchik, S., Katz, M.A., Hernán, M.A., Lipsitch, M., Reis, B., and Balicer, R.D. (2021). BNT162b2 mRNA Covid-19 Vaccine in a Nationwide Mass Vaccination Setting. N. Engl. J. Med. *384*, 1412–1423.
30. Novavax Inc (2021). Novavax COVID-19 Vaccine Demonstrates 89.3% Efficacy in UK Phase 3 Trial. https://ir.novavax.com/2021-01-28-Novavax-COVID-19-Vaccine-Demonstrates-89-3-Efficacy-in-UK-Phase-3-Trial.
31. US Food and Drug Administration (2021). Vaccines and Related Biological Products Advisory Committee February 26, 2021. Sponsor Briefing Document Addendum. https://www.fda.gov/media/146218/download.
32. US Food and Drug Administration (2021). Vaccines and Related Biological Products Advisory Committee February 26, 2021 FDA Briefing Document—Addendum–Sponsor. https://www.fda.gov/media/146218/download.
33. Sette, A., and Crotty, S. (2021). Adaptive immunity to SARS-CoV-2 and COVID-19. Cell *184*, 861–880.
34. Rydyznski Moderbacher, C., Ramirez, S.I., Dan, J.M., Grifoni, A., Hastie, K.M., Weiskopf, D., Belanger, S., Abbott, R.K., Kim, C., Choi, J., et al. (2020). Antigen-Specific Adaptive Immunity to SARS-CoV-2 in Acute COVID-19 and Associations with Age and Disease Severity. Cell *183*, 996–1012.e19.
35. Tan, A.T., Linster, M., Tan, C.W., Le Bert, N., Chia, W.N., Kunasegaran, K., Zhuang, Y., Tham, C.Y.L., Chia, A., Smith, G.J.D., et al. (2021). Early induction of functional SARS-CoV-2-specific T cells associates with rapid viral clearance and mild disease in COVID-19 patients. Cell Rep. *34*, 108728.
36. Muñoz-Fontela, C., Dowling, W.E., Funnell, S.G.P., Gsell, P.S., Riveros-Balta, A.X., Albrecht, R.A., Andersen, H., Baric, R.S., Carroll, M.W., Cavaleri, M., et al. (2020). Animal models for COVID-19. Nature *586*, 509–515.
37. Soresina, A., Moratto, D., Chiarini, M., Paolillo, C., Baresi, G., Foca, E., Bezzi, M., Baronio, B., Giacomelli, M., and Badolato, R. (2020). Two X-linked agammaglobulinemia patients develop pneumonia as COVID-19 manifestation but recover. Pediatr. Allergy Immunol. *31*, 565–569.
38. Baker, D., Roberts, C.A.K., Pryce, G., Kang, A.S., Marta, M., Reyes, S., Schmierer, K., Giovannoni, G., and Amor, S. (2020). COVID-19 vaccine-readiness for anti-CD20-depleting therapy in autoimmune diseases. Clin. Exp. Immunol. *202*, 149–161.
39. Dan, J.M., Mateus, J., Kato, Y., Hastie, K.M., Yu, E.D., Faliti, C.E., Grifoni, A., Ramirez, S.I., Haupt, S., Frazier, A., et al. (2021). Immunological memory to SARS-CoV-2 assessed for up to 8 months after infection. Science *371*, eabf4063.
40. Peng, Y., Mentzer, A.J., Liu, G., Yao, X., Yin, Z., Dong, D., Dejnirattisai, W., Rostron, T., Supasa, P., Liu, C., et al.; Oxford Immunology Network Covid-19 Response T cell Consortium; ISARIC4C Investigators (2020). Broad and strong memory CD4+ and CD8+ T cells induced by SARS-CoV-2 in UK convalescent individuals following COVID-19. Nat. Immunol. *21*, 1336–1345.
41. Breton, G., Mendoza, P., Hägglöf, T., Oliveira, T.Y., Schaefer-Babajew, D., Gaebler, C., Turroja, M., Hurley, A., Caskey, M., and Nussenzweig, M.C. (2021). Persistent cellular immunity to SARS-CoV-2 infection. J. Exp. Med. *218*, e20202515.
42. Dowd, S.M., Zalta, A.K., Burgess, H.J., Adkins, E.C., Valdespino-Hayden, Z., and Pollack, M.H. (2020). Double-blind randomized controlled study of the efficacy, safety and tolerability of eszopiclone *vs* placebo for the treatment of patients with post-traumatic stress disorder and insomnia. World J. Psychiatry *10*, 21–28.
43. Sadoff, J., Le Gars, M., Shukarev, G., Heerwegh, D., Truyers, C., de Groot, A.M., Stoop, J., Tete, S., Van Damme, W., Leroux-Roels, I., et al. (2021). Interim Results of a Phase 1-2a Trial of Ad26.COV2.S Covid-19 Vaccine. N. Engl. J. Med. *384*, 1824–1835.
44. Baden, L.R., El Sahly, H.M., Essink, B., Kotloff, K., Frey, S., Novak, R., Diemert, D., Spector, S.A., Rouphael, N., Creech, C.B., et al.; COVE Study Group (2021). Efficacy and Safety of the mRNA-1273 SARS-CoV-2 Vaccine. N. Engl. J. Med. *384*, 403–416.
45. Keech, C., Albert, G., Cho, I., Robertson, A., Reed, P., Neal, S., Plested, J.S., Zhu, M., Cloney-Clark, S., Zhou, H., et al. (2020). Phase 1-2 Trial of a SARS-CoV-2 Recombinant Spike Protein Nanoparticle Vaccine. N. Engl. J. Med. *383*, 2320–2332.
46. Redd, A.D., Nardin, A., Kared, H., Bloch, E.M., Pekosz, A., Laeyendecker, O., Abel, B., Fehlings, M., Quinn, T.C., and Tobian, A.A.R. (2021). CD8+ T cell responses in COVID-19 convalescent individuals target conserved epitopes from multiple prominent SARS-CoV-2 circulating variants. medRxiv. https://doi.org/10.1101/2021.02.11.21251585.
47. Grifoni, A., Weiskopf, D., Ramirez, S.I., Mateus, J., Dan, J.M., Moderbacher, C.R., Rawlings, S.A., Sutherland, A., Premkumar, L., Jadi, R.S., et al. (2020). Targets of T Cell Responses to SARS-CoV-2 Coronavirus in Humans with COVID-19 Disease and Unexposed Individuals. Cell *181*, 1489–1501.e15.
48. Mateus, J., Grifoni, A., Tarke, A., Sidney, J., Ramirez, S.I., Dan, J.M., Burger, Z.C., Rawlings, S.A., Smith, D.M., Phillips, E., et al. (2020). Selective and cross-reactive SARS-CoV-2 T cell epitopes in unexposed humans. Science *370*, 89–94.

49. Zuo, J., Dowell, A.C., Pearce, H., Verma, K., Long, H.M., Begum, J., Aiano, F., Amin-Chowdhury, Z., Hoschler, K., Brooks, T., et al. (2021). Robust SARS-CoV-2-specific T cell immunity is maintained at 6 months following primary infection. Nat. Immunol. *22*, 620–626.

50. Le Bert, N., Clapham, H.E., Tan, A.T., Chia, W.N., Tham, C.Y.L., Lim, J.M., Kunasegaran, K., Tan, L.W.L., Dutertre, C.A., Shankar, N., et al. (2021). Highly functional virus-specific cellular immune response in asymptomatic SARS-CoV-2 infection. J. Exp. Med. *218*, e20202617. https://doi.org/10.1084/jem.20202617.

51. Tarke, A., Sidney, J., Kidd, C.K., Dan, J.M., Ramirez, S.I., Yu, E.D., Mateus, J., da Silva Antunes, R., Moore, E., Rubiro, P., et al. (2021). Comprehensive analysis of T cell immunodominance and immunoprevalence of SARS-CoV-2 epitopes in COVID-19 cases. Cell Rep. Med. *2*, 100204.

52. Braun, J., Loyal, L., Frentsch, M., Wendisch, D., Georg, P., Kurth, F., Hippenstiel, S., Dingeldey, M., Kruse, B., Fauchere, F., et al. (2020). SARS-CoV-2-reactive T cells in healthy donors and patients with COVID-19. Nature *587*, 270–274.

53. Le Bert, N., Tan, A.T., Kunasegaran, K., Tham, C.Y.L., Hafezi, M., Chia, A., Chng, M.H.Y., Lin, M., Tan, N., Linster, M., et al. (2020). SARS-CoV-2-specific T cell immunity in cases of COVID-19 and SARS, and uninfected controls. Nature *584*, 457–462.

54. Dhanda, S.K., Mahajan, S., Paul, S., Yan, Z., Kim, H., Jespersen, M.C., Jurtz, V., Andreatta, M., Greenbaum, J.A., Marcatili, P., et al. (2019). IEDB-AR: immune epitope database-analysis resource in 2019. Nucleic Acids Res. *47* (*W1*), W502–W506.

55. Reynisson, B., Alvarez, B., Paul, S., Peters, B., and Nielsen, M. (2020). NetMHCpan-4.1 and NetMHCIIpan-4.0: improved predictions of MHC antigen presentation by concurrent motif deconvolution and integration of MS MHC eluted ligand data. Nucleic Acids Res. *48* (*W1*), W449–W454.

56. Lipsitch, M., Grad, Y.H., Sette, A., and Crotty, S. (2020). Cross-reactive memory T cells and herd immunity to SARS-CoV-2. Nat. Rev. Immunol. *20*, 709–713.

57. Greenbaum, J.A., Kotturi, M.F., Kim, Y., Oseroff, C., Vaughan, K., Salimi, N., Vita, R., Ponomarenko, J., Scheuermann, R.H., Sette, A., and Peters, B. (2009). Pre-existing immunity against swine-origin H1N1 influenza viruses in the general human population. Proc. Natl. Acad. Sci. USA *106*, 20365–20370.

58. Sridhar, S., Begom, S., Bermingham, A., Hoschler, K., Adamson, W., Carman, W., Bean, T., Barclay, W., Deeks, J.J., and Lalvani, A. (2013). Cellular immune correlates of protection against symptomatic pandemic influenza. Nat. Med. *19*, 1305–1312.

59. Wilkinson, T.M., Li, C.K., Chui, C.S., Huang, A.K., Perkins, M., Liebner, J.C., Lambkin-Williams, R., Gilbert, A., Oxford, J., Nicholas, B., et al. (2012). Preexisting influenza-specific CD4+ T cells correlate with disease protection against influenza challenge in humans. Nat. Med. *18*, 274–280.

60. Ferretti, A.P., Kula, T., Wang, Y., Nguyen, D.M.V., Weinheimer, A., Dunlap, G.S., Xu, Q., Nabilsi, N., Perullo, C.R., Cristofaro, A.W., et al. (2020). Unbiased Screens Show CD8+ T Cells of COVID-19 Patients Recognize Shared Epitopes in SARS-CoV-2 that Largely Reside outside the Spike Protein. Immunity *53*, 1095–1107.e3.

61. Snyder, T.M., Gittelman, R.M., Klinger, M., May, D.H., Osborne, E.J., Taniguchi, R., Zahid, H.J., Kaplan, I.M., Dines, J.N., Noakes, M.N., et al. (2020). Magnitude and Dynamics of the T-Cell Response to SARS-CoV-2 Infection at Both Individual and Population Levels. medRxiv. https://doi.org/10.1101/2020.07.31.20165647.

62. Nelde, A., Bilich, T., Heitmann, J.S., Maringer, Y., Salih, H.R., Roerden, M., Lübke, M., Bauer, J., Rieth, J., Wacker, M., et al. (2021). SARS-CoV-2-derived peptides define heterologous and COVID-19-induced T cell recognition. Nat. Immunol. *22*, 74–85.

63. Keller, M.D., Harris, K.M., Jensen-Wachspress, M.A., Kankate, V.V., Lang, H., Lazarski, C.A., Durkee-Shock, J., Lee, P.H., Chaudhry, K., Webber, K., et al. (2020). SARS-CoV-2-specific T cells are rapidly expanded for therapeutic use and target conserved regions of the membrane protein. Blood *136*, 2905–2917.

64. Weiskopf, D., Angelo, M.A., Sidney, J., Peters, B., Shresta, S., and Sette, A. (2014). Immunodominance changes as a function of the infecting dengue virus serotype and primary versus secondary infection. J. Virol. *88*, 11383–11394.

65. Evavold, B.D., and Allen, P.M. (1991). Separation of IL-4 production from Th cell proliferation by an altered T cell receptor ligand. Science *252*, 1308–1310.

66. Sloan-Lancaster, J., and Allen, P.M. (1996). Altered peptide ligand-induced partial T cell activation: molecular mechanisms and role in T cell biology. Annu. Rev. Immunol. *14*, 1–27.

67. Peeples, L. (2020). News Feature: avoiding pitfalls in the pursuit of a COVID-19 vaccine. Proc. Natl. Acad. Sci. USA *117*, 8218–8221.

68. Starr, T.N., Greaney, A.J., Hilton, S.K., Ellis, D., Crawford, K.H.D., Dingens, A.S., Navarro, M.J., Bowen, J.E., Tortorici, M.A., Walls, A.C., et al. (2020). Deep Mutational Scanning of SARS-CoV-2 Receptor Binding Domain Reveals Constraints on Folding and ACE2 Binding. Cell *182*, 1295–1310.e20.

69. Andreano, E., Piccini, G., Licastro, D., Casalino, L., Johnson, N.V., Paciello, I., Monego, S.D., Pantano, E., Manganaro, N., Manenti, A., et al. (2020). SARS-CoV-2 escape < em > in vitro < /em > from a highly neutralizing COVID-19 convalescent plasma. bioRxiv. https://doi.org/10.1101/2020.12.28.424451.

70. Hensley, S.E., Das, S.R., Bailey, A.L., Schmidt, L.M., Hickman, H.D., Jayaraman, A., Viswanathan, K., Raman, R., Sasisekharan, R., Bennink, J.R., and Yewdell, J.W. (2009). Hemagglutinin receptor binding avidity drives influenza A virus antigenic drift. Science *326*, 734–736.

71. Doud, M.B., Lee, J.M., and Bloom, J.D. (2018). How single mutations affect viral escape from broad and narrow antibodies to H1 influenza hemagglutinin. Nat. Commun. *9*, 1386.

72. Andrews, S.F., Huang, Y., Kaur, K., Popova, L.I., Ho, I.Y., Pauli, N.T., Henry Dunand, C.J., Taylor, W.M., Lim, S., Huang, M., et al. (2015). Immune history profoundly affects broadly protective B cell responses to influenza. Sci. Transl. Med. *7*, 316ra192.

73. Krammer, F., García-Sastre, A., and Palese, P. (2018). Is It Possible to Develop a "Universal" Influenza Virus Vaccine? Potential Target Antigens and Critical Aspects for a Universal Influenza Vaccine. Cold Spring Harb. Perspect. Biol. *10*, a028845.

74. Gaiha, G.D., Rossin, E.J., Urbach, J., Landeros, C., Collins, D.R., Nwonu, C., Muzhingi, I., Anahtar, M.N., Waring, O.M., Piechocka-Trocha, A., et al. (2019). Structural topology defines protective CD8+ T cell epitopes in the HIV proteome. Science *364*, 480–484.

75. Stadlbauer, D., Amanat, F., Chromikova, V., Jiang, K., Strohmeier, S., Arunkumar, G.A., et al. (2020). SARS-CoV-2 Seroconversion in Humans: A Detailed Protocol for a Serological Assay, Antigen Production, and Test Setup. Curr Protoc Microbiol *57*, e100.

76. Vita, R., Mahajan, S., Overton, J.A., Dhanda, S.K., Martini, S., Cantrell, J.R., Wheeler, D.K., Sette, A., and Peters, B. (2019). The Immune Epitope Database (IEDB): 2018 update. Nucleic Acids Res. *47* (*D1*), D339–D343.

77. Jurtz, V., Paul, S., Andreatta, M., Marcatili, P., Peters, B., and Nielsen, M. (2017). NetMHCpan-4.0: Improved Peptide-MHC Class I Interaction Predictions Integrating Eluted Ligand and Peptide Binding Affinity Data. J Immunol *199* (*9*), 3360–3368.

78. Pickett, B.E., Sadat, E.L., Zhang, Y., Noronha, J.M., Squires, R.B., Hunt, V., Liu, M., Kumar, S., Zaremba, S., Gu, Z., et al. (2012). ViPR: an open bioinformatics database and analysis resource for virology research. Nucleic Acids Res. *40*, D593–D598.

79. Reiss, S., Baxter, A.E., Cirelli, K.M., Dan, J.M., Morou, A., Daigneault, A., Brassard, N., Silvestri, G., Routy, J.P., Havenar-Daughton, C., et al. (2017). Comparative analysis of activation induced marker (AIM) assays for sensitive identification of antigen-specific CD4 T cells. PLoS ONE *12*, e0186998.

80. da Silva Antunes, R., Pallikkuth, S., Williams, E., Esther, D.Y., Mateus, J., Quiambao, L., Wang, E., Rawlings, S.A., Stadlbauer, D., Jiang, K., et al. (2021). Differential T cell reactivity to endemic coronaviruses and SARS-CoV-2 in community and health care workers. J. Infect. Dis., jiab176.

Cell Reports Medicine
Article

CellPress
OPEN ACCESS

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
M5E2 (V500) [anti-CD14]	BD Biosciences	561391 (RRID:AB_10611856)
HIB19 (V500) [anti-CD19]	BD Biosciences	561121 (RRID:AB_10562391)
RPA-T4 (BV605) [anti-CD4]	BD Biosciences	562989 (RRID:AB_2737935)
RPA-T8 (BV650) [anti-CD8]	BioLegend	301042 (RRID:AB_2563505)
RPA-T8 (BUV496) [anti-CD8]	BD Biosciences	612942 (RRID:AB_2870223)
FN50 (PE) [anti-CD69]	BD Biosciences	555531 (RRID:AB_2737680)
Ber-ACT35 (PE-Cy7) [anti-OX40]	Biolegend	350012 (RRID:AB_10901161)
4B4-1 (APC) [anti-CD137]	BioLegend	309810 (RRID:AB_830672)
OKT3 (AF700) [anti-CD3]	Biolegend	317340 (RRID:AB_2563408)
UCHT1 (BUV805) [anti-CD3]	BD Biosciences	612895 (RRID:AB_2870183)
HI100 (BV421) [anti-CD45RA]	Biolegend	304130 (RRID:AB_10965547)
G043H7 (FITC) [anti-CCR7]	Biolegend	353216 (RRID:AB_10916386)
Biological samples		
Convalescent donor blood samples	UC San Diego Health	https://health.ucsd.edu/Pages/default.aspx
Convalescent donor blood samples	La Jolla Institute	https://www.lji.org
Convalescent donor blood samples	BioIVT	https://bioivt.com
COVID-19 vaccinee donor blood samples	La Jolla Institute	https://www.lji.org
Unexposed donor blood samples	San Diego Blood Bank	https://www.sandiegobloodbank.org
Unexposed donor blood samples	La Jolla Institute	https://www.lji.org
Unexposed donor blood samples	Carter Blood Care	https://www.carterbloodcare.org
Chemicals, peptides, and recombinant proteins		
Synthetic peptides	TC Peptide Lab	https://tcpeptidelab.com
SARS-CoV-2 Receptor Binding Domain (RBD) protein	Stadlbauer et al.[75]	N/A
Deposited data		
Ancestral	https://www.ncbi.nlm.nih.gov	NC_045512.2
B.1.1.7	https://www.GISAID.org	EPI_ISL_601443
B.1.351	https://www.GISAID.org	EPI_ISL_660629
B.1.351	https://www.GISAID.org	EPI_ISL_736930
B.1.351	https://www.GISAID.org	EPI_ISL_736932
B.1.351	https://www.GISAID.org	EPI_ISL_736944
B.1.351	https://www.GISAID.org	EPI_ISL_736971
B.1.351	https://www.GISAID.org	EPI_ISL_736966
B.1.351	https://www.GISAID.org	EPI_ISL_736973
B.1.351	https://www.GISAID.org	EPI_ISL_825104
B.1.351	https://www.GISAID.org	EPI_ISL_825120
B.1.351	https://www.GISAID.org	EPI_ISL_825131
P.1	https://www.GISAID.org	EPI_ISL_804823
CAL.20C	https://www.GISAID.org	EPI_ISL_847619
CAL.20C	https://www.GISAID.org	EPI_ISL_847621
CAL.20C	https://www.GISAID.org	EPI_ISL_847643

(*Continued on next page*)

Please cite this article in press as: Tarke et al., Impact of SARS-CoV-2 variants on the total CD4+ and CD8+ T cell reactivity in infected or vaccinated individuals, Cell Reports Medicine (2021), https://doi.org/10.1016/j.xcrm.2021.100355

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Software and algorithms		
IEDB	Vita et al.[76]	https://www.iedb.org
IEDB-AR (analysis resource)	Reynisson et al.[55]	http://tools.iedb.org/main
NetMHCpan BA 4.1	Jurtz et al.[77]	http://tools.iedb.org/mhci/
VIGOR4	Pickett et al.[78]	https://www.viprbrc.org//brc/home.spg?decorator=vipr
FlowJo 10	FlowJo, LLC	https://www.flowjo.com
GraphPad Prism 8.4	GraphPad	https://www.graphpad.com

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Dr. Alessandro Sette (alex@lji.org).

Materials availability

Aliquots of synthesized sets of peptides utilized in this study will be made available upon request. There are restrictions to the availability of the peptide reagents due to cost and limited quantity.

Data and code availability

The published article includes all data generated or analyzed during this study, and summarized in the accompanying tables, figures and supplemental materials.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Convalescent COVID-19 donors

Convalescent donors were enrolled at either a UC San Diego Health clinic under the approved IRB protocols of the University of California, San Diego (UCSD; 200236X), or at the La Jolla Institute (LJI; VD-214). All donors were California residents and samples were collected from May to October 2020, before any of the SARS-CoV-2 variants described herein had been detected in California. These donors were referred to the study by a health care provider or were self-referred. The CRO BioIVT provided additional cohorts of COVID-19 convalescent donors who had been confirmed positive for COVID-19 by PCR following the resolution of symptoms. The total cohort of convalescent donors represented both genders (43% male, 57% female) and ranged from 20 to 67 years of age (median 38 years). All samples were either PCR positive or seropositive against SARS-CoV-2 by ELISA, as described below. Details of this convalescent COVID-19 cohort are listed in Table 1. All convalescent COVID-19 donors provided informed consent to participate in the present and future studies at the time of enrollment.

COVID-19 vaccinees

The La Jolla Institute recruited 29 healthy adults who had received the first and second dose of the Pfizer/BioNTech BNT162b2 (n = 14) or Moderna mRNA-1273 COVID-19 vaccines (n = 15). Blood draws took place under IRB approved protocols (LJI; VD-214) two to four weeks after the second dose of the vaccine was administered. All donors had their SARS-CoV-2 antibody titers measured by ELISA, as described below. The cohort of vaccinees ranged from 23 to 67 years of age (median 47 years) and represented both genders (31% male, 69% female). At the time of enrollment in the study, all donors gave informed consent.

Unexposed donors

PBMCs from 23 healthy unexposed donors were used as controls. Healthy donors were recruited by the La Jolla Institute under IRB approved protocols (LJI; VD-214) or Carter Blood Care, where according to the blood center's criteria for eligibility the donors were also eligible for our study. All healthy donors were collected from May 2014 to March 2018 (n = 13). Additional healthy donors were recruited from March to May 2020 by the San Diego Blood Bank (SDBB (n = 10) and were confirmed seronegative for SARS-CoV-2 by ELISA. This cohort was 35% male, 56% female, and 9% unknown and ranged from 21 to 82 years of age (median 34). All donors gave informed consent and permission for their samples to be used for future studies.

Please cite this article in press as: Tarke et al., Impact of SARS-CoV-2 variants on the total CD4+ and CD8+ T cell reactivity in infected or vaccinated individuals, Cell Reports Medicine (2021), https://doi.org/10.1016/j.xcrm.2021.100355

METHOD DETAILS

Isolation of peripheral blood mononuclear cells (PBMCs) and plasma

Collection and processing of blood samples was performed as previously described.[39,51] Briefly, whole blood was collected in heparin coated blood bags or in ACD tubes and centrifuged for 15 minutes at 1850 rpm to separate the cellular fraction from the plasma. The plasma was then removed and stored at −20°C. The cellular fraction next underwent density-gradient sedimentation using Ficoll-Paque (Lymphoprep, Nycomed Pharma, Oslo, Norway) to separate the PBMCs as previously described.[64] Isolated PBMCs were cryopreserved in cell recovery media containing 10% DMSO (GIBCO), supplemented with 90% heat inactivated fetal bovine serum (FBS; Hyclone Laboratories, Logan UT) and stored in liquid nitrogen until used in the assays.

SARS-CoV-2 RBD ELISA

Serology to SARS-CoV-2 was determined for all donor cohorts as previously described.[34] Briefly, 96-well half-area plates (ThermoFisher 3690) were coated with 1 ug/mL SARS-CoV-2 Spike (S) Receptor Binding Domain (RBD) and incubated at 4°C overnight. The next day, plates were blocked at room temperature for 2 hours with 3% milk in phosphate buffered saline (PBS) containing 0.05% Tween-20. Heat-inactivated plasma was added to the plates for an additional 90-minute incubation at room temperature followed by incubation with the conjugated secondary antibody, detection, and subsequent data analysis by reading the plates on Spectramax Plate Reader at 450 nm using the SoftMax Pro. The limit of detection (LOD) was defined as 1:3. Limit of sensitivity (LOS) for SARS-CoV-2 infected individuals was established based on uninfected subjects, using plasma from normal healthy donors not exposed to SARS-CoV-2.

Mutation analysis of SARS-CoV-2 B.1.1.7, B.1.351, P.1 and CAL.20C variants

Genome sequences for the variant viruses were downloaded from GISAID. These sequences were screened to select those without ambiguous residues and generated from Illumina sequencing technologies using an in-house sequence QC script. The selected genomic sequences were then translated into protein amino acid sequences using the VIGOR4 tool available on the Virus Pathogen Resource (ViPR).[78] Sequence variations in the variant viruses were derived by comparison with Wuhan-1 (NC_045512.2). One or more representative sequences were considered for the B.1.1.7 (EPI_ISL_601443), B.1.351 (EPI_ISL_660629; EPI_ISL_736930; EPI_ISL_736932; EPI_ISL_736944; EPI_ISL_736966; EPI_ISL_736971; EPI_ISL_736973; EPI_ISL_825104; EPI_ISL_825120; EPI_ISL_825131), P.1 (EPI_ISL_804823), and CAL.20C (EPI_ISL_847619; EPI_ISL_847621; EPI_ISL_847643) variants. A summary of all the amino acids mutated in the different variants considered in this study with respect to the Wuhan sequence and is available in Table S1.

SARS-CoV-2 Wuhan and variant peptide synthesis and pooling

Peptides were synthesized that spanned entire SARS-CoV-2 proteins and corresponded to the ancestral Wuhan sequence or the B.1.1.7, B.1.351, P.1 and CAL.20C SARS-CoV-2 variants. Peptides were 15-mers overlapping by 10 amino acids and were synthesized as crude material (TC Peptide Lab, San Diego, CA). All peptides were individually resuspended in dimethyl sulfoxide (DMSO) at a concentration of 10–20 mg/mL. Megapools (MP) for each antigen were created by pooling aliquots of these individual peptides, undergoing another lyophilization, and resuspending in DMSO at 1 mg/mL.

Bioinformatic analysis of T cell epitopes

The binding capacity of SARS-CoV-2 T cell epitopes, and their corresponding variant-derived peptides, for their putative HLA class I restricting allele(s) was determined utilizing the NetMHCpan BA 4.1 algorithm,[55] as implemented by the IEDB's analysis resource.[54,76] Predicted binding is expressed in terms of IC_{50} nM. For each epitope-variant pair a ratio of affinities (variant IC_{50} nM/WT IC_{50} nM) was determined. Ratios > 2, indicating a 2-fold or greater decrease in affinity due to the mutation, were accordingly categorized as a decrease in binding capacity, and a ratio < 0.5 as an increase; ratios between 0.5 and 2 were designated as neutral.

Flow cytometry-based AIM assay

Activation induced cell marker (AIM) assay has previously been described in detail elsewhere.[39,79,80] In summary, PBMCs were cultured in the presence of SARS-CoV-2-specific (ancestral or VOC) MPs [1 μg/ml] in 96-well U-bottom plates at a concentration of $1x10^6$ PBMC per well. As a negative control, an equimolar amount of DMSO was used to stimulate the cells in triplicate wells and as positive controls phytohemagglutinin (PHA, Roche, 1 μg/ml) and a cytomegalovirus MP (CMV, combining CD4 and CD8 MPs, 1 μg/ml) were also included. After incubation for 20–24 hours at 37°C, 5% CO_2, the cells were stained with CD3 BUV805 or CD3 AF700 (4:100 or 4:100; BD Biosciences Cat# 612895 or Life Technologies Cat# 56-0038-42, respectively), CD4 BV605 (4:100; BD Biosciences Cat# 562658), CD8 BUV496 or BV650 (2:100 or 4:100; BD Biosciences Cat# 612942 or Biolegend Cat# 301042), CD14 V500 (2:100; BD Biosciences Cat# 561391), CD19 V500 (2:100; BD Biosciences Cat# 561121), and Live/Dead eFluor506 (5:1000; eBioscience Cat# 65-0866-14). Cells were also stained to measure activation with the following markers: CD137 APC (4:100; Biolegend Cat# 309810), OX40 PE-Cy7 (2:100; Biolegend Cat#350012), and CD69 PE (10:100; BD Biosciences Cat# 555531). The memory phenotyping was analyzed by staining with CD45RA (1:100; Biolegend Cat# 304130) and CCR7 (2:100; Biolegend Cat# 353216). All samples were acquired on a ZE5 5-laser or 4-laser cell analyzer (Bio-rad laboratories) and analyzed with

Please cite this article in press as: Tarke et al., Impact of SARS-CoV-2 variants on the total CD4+ and CD8+ T cell reactivity in infected or vaccinated individuals, Cell Reports Medicine (2021), https://doi.org/10.1016/j.xcrm.2021.100355

FlowJo software (Tree Star). In the resulting data generated from the AIM assays, the background was removed from the data by subtracting the average of the % of AIM+ cells plated in triplicate wells stimulated with DMSO. The Stimulation Index (SI) was calculated by dividing the % of AIM+ cells after SARS-CoV-2 stimulation with the average % of AIM+ cells in the negative DMSO control. An SI greater than 2 and a minimum of 0.016% or 0.16% AIM+ CD4+ or CD8+ cells, respectively, after background subtraction was considered to be a positive response based on the median twofold standard deviation of all negative controls measured in the study (> 80), as previously reported.[39] The gates for AIM+ cells were drawn relative to the negative and positive controls for each donor. A representative example of the gating strategy is depicted in Figure S5. Specifically, lymphocytes were gated, followed by single cells determination. T cells were considered to be positive for CD3 and negative for a Dump channel including in the same colors CD14, CD19 and Live/Dead staining. The CD3+CD4+ and CD3+CD8+ were further gated based on OX40+CD137+ and CD69+CD137+ AIM markers, respectively. Memory phenotype was gated based on the expression of CD45RA and CCR7 markers on the CD3+CD4+ and CD3+CD8+ T cells and the AIM+ T cells.

FluoroSPOT assay

96-well FluoroSpot plates were coated with anti-cytokine antibodies for IFNγ and IL-5 (mAbs 1-D1K and TRFK5, respectively; Mabtech, Stockholm, Sweden) at a concentration of 10 μg/mL. PBMCs were stimulated in triplicate at a density of $200x10^3$ cells/well with S MPs corresponding to each of the SARS-CoV-2 variants analyzed (1 μg/mL), PHA (1 μg/mL), and DMSO (0.1%), as positive and negative controls respectively. After 20 hours of incubation at 37°C, 5% CO_2, cells were discarded and plates were washed before the addition of cytokine antibodies (mAbs 7-B6-1-BAM and 5A10-WASP; Mabtech, Stockholm, Sweden). After a 2-hour incubation, plates were washed again with PBS/0.05% Tween20 and incubated for 1 hour with fluorophore-conjugated antibodies (Anti-BAM-490 and Anti-WASP-640). An AID iSPOT FluoroSpot reader (AIS-diagnostika, Germany) was used to count the fluorescent spots that resulted from cells secreting IFNγ and IL-5. Each peptide MP was considered positive compared to the DMSO negative control based on the following criteria: 20 or more spot forming cells (SFC) per 10^6 PBMC after background subtraction, a stimulation index (S.I.) greater than 2, and a p value < 0.05 by either a Poisson or t test calculated between the triplicates of the MP and the relative negative control.

QUANTIFICATION AND STATISTICAL ANALYSIS

Data and statistical analyses were performed in FlowJo 10 and GraphPad Prism 8.4, unless otherwise stated. Statistical details of the experiments are provided in the respective figure legends and in each methods section pertaining to the specific technique applied. Data plotted in logarithmic scales are expressed as geometric mean. Statistical analyses were performed using Wilcoxon matched-pairs signed rank test for paired comparisons. Multi-hypothesis testing corrections (MHTC) have not been applied in the study by design. The study is not designed or powered to address differences across different proteins. The primary hypothesis is that no significant differences are observed across the different variants, and this is more stringently addressed avoiding to correct for MHTC, since a difference that is not significant would remain so even after corrections. Therefore, reporting the data without applying MHTC is a more stringent criterion which is appropriately being applied in this case to avoid false negatives. Details pertaining to significance are also noted in the respective figure legends.

To determine the expected variability of the AIM assays, we also analyzed 23 donors for which CD4+ and CD8+ T cell responses to Spike, M and N megapools were evaluated in up to 4 repeat measurements. A total of 208 measurements were made for these donors for both CD4 and CD8, resulting in 69 independent assessments of repeat measures (comparing the initial measurement to a repeat in the same donor) (Table S3). We evaluated the variability in these repeat measures using a corrected log2 fold-change calculated as:

$$\text{Log2}\ \frac{(\text{measure } 1 + \text{LOD})}{(\text{measure } 2 + \text{LOD})}$$

where LOD is the level of detection threshold, which is 0.014% for CD4 and 0.06% for CD8 based on the twofold geometric 95%CI of all negative controls measured (> 80). The addition of the LOD ensures that variations between measurements close to- or below the level of detection do not result in inflated fold-change estimates. We evaluated the standard deviation of these corrected log-fold change values of repeat measurements (at an individual level), and used this value as a threshold for what would constitute a significant, non-random change at the population level. Similarly, FluoroSPOT assay variability was evaluated by analyzing 38 donors and a total of 301 measurements considering an LOD of 20 SFC/10^6 cells. For CD4-AIM assays, that corresponded to a corrected fold-change of 2.3, for CD8-AIM of 2.4 and for FluoroSPOT of 1.6 as shown in Table S3. Next, we performed a test of non-inferiority for the observed AIM assay or FluoroSPOT reactivities for any of the variants tested compared to the ancestral strain. A non-inferiority test determines if the observed values are significantly above what would be considered a biological meaningful reduction in responses, based on one sample Wilcoxon Signed Rank test compared to the lower bound fold change threshold. The test results in a p value expressing effectiveness of the comparison and a p value corresponding to the discrepancy. Confidence intervals for discrepancy are shown in all the graphs where the comparison was considered statistically significant ($p < 0.05$). The non-inferiority test found that all CD4 and CD8 responses or IFNγ responses to any of the variant peptide pools were significantly above the lower bound threshold. Therefore, by this approach at the population level, there is no biological meaningful reduction in the T cell response to the variant pools compared to the ancestral strain. Details pertaining to significance are also noted in the respective figure legends.

CellPress

Cell Host & Microbe

Review

SARS-CoV-2 human T cell epitopes: Adaptive immune response against COVID-19

Alba Grifoni,[1] John Sidney,[1] Randi Vita,[1] Bjoern Peters,[1,2] Shane Crotty,[1,2] Daniela Weiskopf,[1] and Alessandro Sette[1,2,*]
[1]Center for Infectious Disease and Vaccine Research, La Jolla Institute for Immunology (LJI), La Jolla, CA 92037, USA
[2]Department of Medicine, Division of Infectious Diseases and Global Public Health, University of California, San Diego (UCSD), La Jolla, CA 92037, USA
*Correspondence: alex@lji.org
https://doi.org/10.1016/j.chom.2021.05.010

SUMMARY

Over the past year, numerous studies in the peer reviewed and preprint literature have reported on the virological, epidemiological and clinical characteristics of the coronavirus, SARS-CoV-2. To date, 25 studies have investigated and identified SARS-CoV-2-derived T cell epitopes in humans. Here, we review these recent studies, how they were performed, and their findings. We review how epitopes identified throughout the SARS-CoV2 proteome reveal significant correlation between number of epitopes defined and size of the antigen provenance. We also report additional analysis of SARS-CoV-2 human CD4 and CD8 T cell epitope data compiled from these studies, identifying 1,400 different reported SARS-CoV-2 epitopes and revealing discrete immunodominant regions of the virus and epitopes that are more prevalently recognized. This remarkable breadth of epitope repertoire has implications for vaccine design, cross-reactivity, and immune escape by SARS-CoV-2 variants.

INTRODUCTION

Over the past year, a considerable amount of information has been produced by the scientific community on the SARS-CoV-2 virus and its associated disease, COVID-19, with studies in the peer-reviewed and pre-print literature investigating its different virological, epidemiological, and clinical characteristics. In particular, numerous studies have analyzed the immune response to the virus, its role in protection and disease, and its importance in the context of vaccine development and evaluation. Several excellent reviews—some in this special issue—cover these topics (DiPiazza et al., 2020; Jordan, 2021; Karlsson et al., 2020; Sette and Crotty, 2020b, 2021; Swadling and Maini, 2020).

Here, we focus on a specific topic: our current knowledge concerning the definition and recognition of SARS-CoV-2-derived T cell epitopes in humans. While the data related to this topic was initially sparse, 25 different studies have now been published as of March 15, 2021 (Chen et al., 2021; Ferretti et al., 2020; Gangaev et al., 2020; Habel et al., 2020; Joag et al., 2021; Kared et al., 2021; Keller et al., 2020; Le Bert et al., 2021, 2020; Lee et al., 2020; Mahajan et al., 2020; Mateus et al., 2020; Nelde et al., 2021; Nielsen et al., 2020; Peng et al., 2020; Poran et al., 2020b, 2020a; Prakash et al., 2020; Rha et al., 2021; Sahin et al., 2020; Saini et al., 2020, 2021, Schulien et al., 2021; Sekine et al., 2020; Shomuradova et al., 2020; Snyder et al., 2020; Tarke et al., 2021a), which collectively report data from 1,197 human subjects (870 COVID-19 and 327 unexposed controls), leading to the identification of over 1,400 different CD4 (n = 382) and CD8 (n = 1052) T cell epitopes. These studies are listed in Table 1, which also captures whether these studies defined class I/CD8 epitopes and/or class II/CD4 epitopes.

The list of papers we have reviewed is, to the best of our knowledge, exhaustive as of March 15th, 2021. Relevant papers were selected based on the objective curation process implemented over 20 years ago by the Immune Epitope Database (IEDB; www.iedb.org) based on the combined use of general broad PubMed queries combined with automated text classifiers and manual curation, as described in more detail elsewhere (Fleri et al., 2017; Salimi et al., 2012). In addition, the results of the IEDB curation were manually inspected by the coauthors to guard against papers missed by the IEDB curation workflow, but no additional papers were identified. This review focuses on SARS-CoV-2 epitopes recognized by human T cells, and thus does not discuss related topics, such as studies identifying T cell epitopes recognized in murine systems (Hassert et al., 2020; Takagi and Matsui, 2021), studies characterizing SARS-CoV-2 peptides eluted from HLA molecules (Knierman et al., 2020; Parker et al., 2020, Weingarten-Gabbay et al., 2020), or characterized by HLA binding in the absence of T cell recognition data (Prachar et al., 2020).

Here, we focus on cataloging and describing SARS-CoV-2 epitopes recognized by human T cells, from data collected from the 25 studies cited above. In this review, we have organized the data into a number of sections, initially describing epitope definitions, screening methodologies, and assay readouts. Subsequent sections describe the number of epitopes identified in the various studies, the antigens recognized, and the distribution of epitopes within them, which lead to the definition of immunodominant regions and immunodominant epitopes. Additional sections are devoted to discussion of epitope identification in different populations and cohorts and the related topics of HLA coverage and immunodominant HLA alleles. We also highlight how the breadth of the T cell repertoire informs

Table 1. Summary of results of epitope identification studies

Reference	Restriction	Screening strategy	Readout type	Assay Readout	# of epitopes	Antigens screened	# of donors		# of different restricting HLA molecules	
							COVID-19	unexposed	MHC class I	MHC class II
Chen, J Cell Mol Med, 2021 (Chen et al., 2021)	class I/CD8	predicted	*in vitro* expansion (*ex vivo* verification)	Proliferation, ICS	1	S		3	1	NI
Ferretti, Immunity, 2020 (Ferretti et al., 2020)	class I/CD8	predicted	*ex vivo*	ELISA, cytotoxicity, multimer staining	28	entire proteome	78		6	NI
Gangaev, Research Square, 2021 (Gangaev et al., 2020)	class I/CD8	predicted	*ex vivo*	multimer staining	9	entire proteome	18	4	4	NI
Habel, PNAS, 2020* (Habel et al., 2020)	class I/CD8	overlapping	*in vitro* expansion, *ex vivo*	ICS, multimer staining	14	S, N, M, ORF1ab	18	12	1	NI
Joag, J Immunol, 2021 (Joag et al., 2021)	class I/CD8	predicted	*in vitro* expansion	ICS	1	N	10		1	NI
Kared, J Clin Invest, 2021 (Kared et al., 2021)	class I/CD8	overlapping	*ex vivo*	multimer staining	45	entire proteome	30		6	NI
Keller, Blood, 2020 (Keller et al., 2020)	both	predicted	*in vitro* expansion	ELISpot	12	S, M, N, E	11		NI	12
Le Bert, bioRXiv, 2020 (Le Bert et al., 2021)	both	predicted	*in vitro* expansion	ICS	3	M, N, S	3		NI	NI
Le Bert, Nature, 2020 (Le Bert et al., 2020)	both	overlapping	*in vitro* expansion	ICS	9	N, nsp7, nsp13	36	37	NI	NI
Lee, J Virol, 2020 (Lee et al., 2020)	class I/CD8	predicted	*in vitro* expansion	degranulation, ICS	2	N	2		1	NI
Mahajan, bioRXiv, 2020 (Mahajan et al., 2020)	both	predicted	*in vitro* expansion	ICS, AIM	10	S		17	NI	NI
Mateus, Science, 2020 (Mateus et al., 2020)	both	overlapping and predicted	*in vitro* expansion	ELISpot	138	entire proteome		40	NI	30
Nelde, Nat. Immunol, 2021 (Nelde et al., 2021)	both	predicted	*in vitro* expansion	ELISpot, ICS	49	entire proteome	116	104	9	NI
Nielsen, bioRXiv, 2020 (Nielsen et al., 2020)	class I/CD8	predicted	*ex vivo*	multimer staining	9	M, N, S	106		1	NI

(*Continued on next page*)

Table 1. *Continued*

Reference	Restriction	Screening strategy	Readout type	Assay Readout	# of epitopes	Antigens screened	# of donors		# of different restricting HLA molecules	
							COVID-19	unexposed	MHC class I	MHC class II
Peng, Nat. Immunol, 2020 (Peng et al., 2020)	both	overlapping	*ex vivo*	ELISpot, multimer staining	16	S, N, M, E, ORF3a, ORF6, ORF7a, ORF8	42	16	6	NI
Poran, bioRXiv, 2020 (Poran et al., 2020b, Poran et al., 2020a)	class I/CD8	predicted	*in vitro* expansion	multimer staining	11	S, N, M, E, ORF1ab		3	1	NI
Prakash, bioRXiv, 2020 (Prakash et al., 2020)	both	predicted	*ex vivo*	ELISpot	27	entire proteome	63	10	1	NI
Rha, Immunity, 2021 (Rha et al., 2021)	class I/CD8	predicted	*ex vivo*	proliferation, ICS, multimer staining	2	S, M, N	116		1	NI
Sahin, medRXiv, 2020 (Sahin et al., 2020)	class I/CD8	predicted	*ex vivo*	multimer staining	8	S	3		3	NI
Saini, bioRXiv, 2020 (Saini et al., 2020, Saini et al., 2021)	class I/CD8	predicted	*ex vivo*	multimer staining	409	entire proteome	18	38	10	NI
Schulien, Nat. Med, 2021 (Schulien et al., 2021)	class I/CD8	predicted	*in vitro* expansion (*ex vivo* verification)	degranulation, ICS, multimer staining	40	entire proteome	26	8	9	NI
Sekine, Cell, 2020 (Sekine et al., 2020)	class I/CD8	predicted	*ex vivo*	multimer staining	2	Orf3a, ORF6, M, N, E, S	11	18	2	NI
Shomuradova, Immunity, 2020 (Shomuradova et al., 2020)	class I/CD8	predicted	*ex vivo*	multimer staining	12	S	17	17	1	NI
Snyder, medRXiv, 2020 (Snyder et al., 2020)	class I/CD8	predicted	*ex vivo*	AIM	235	entire proteome	47		NI	NI
Tarke, Cell Rep Med, 2021 (Tarke et al., 2021a)	both	overlapping (CD4), predicted (CD8)	*ex vivo*	AIM	734	entire proteome	99		26	35

The 25 different studies to date (2/28/2021) that have identified SARS-CoV-2 derived CD4 and CD8 epitopes are listed; in cases where the pre-print version analyzed has subsequently been published in the peer-reviewed literature, we have indicated both citations. Studies that, to date, are only available on pre-print servers are highlighted by italicized font. Respective columns summarize the scope and approach of each study, whether CD4 and/or CD8 epitopes were assayed, if predicted and/or overlapping peptide sets were used, and the types of T cell assay approaches utilized. Also tabulated are the number of unique epitopes identified and the specific antigens that were targeted for study. Additional columns show the number of COVID-19 positive and/or unexposed donors screened, and the number of unique HLA class I and class II restricting alleles identified. *In vitro* expansion refers to any assay that involved a period of *in vitro* culture before harvesting and assaying for T cell activity. An asterisk (*) highlights a study that also measured epitope specific responses in tissues. NI indicates not investigated.

discussions of pre-existing reactivity and cross-reactivity with common cold corona and other viruses, as well as cross-reactivity with MERS and SARS-CoV-1, and potential implications for immune escape by SARS-CoV-2 variants. This review is therefore relevant to the molecular definition of the targets of adaptive human T cell responses to SARS-CoV-2.

EPITOPE DEFINITIONS

A detailed review of the available epitope data requires a clear definition of the concepts and terminology that have been used, to permit different studies that have used different methodologies to be combined and integrated in a coherent fashion. According to classical textbook definitions, "A T-cell epitope is a short peptide derived from a protein antigen. It binds to an MHC molecule and is recognized by a particular T cell" (Murphey et al., 2012). And, similarly, "The parts of complex antigens that are specifically recognized by lymphocytes are called determinants or epitopes" (Abbas et al., 2007).

T cell epitopes are usually peptides composed of 20 naturally occurring amino acids, although the recognition of haptens, sugars, and post-translationally modified peptides has also been described (Petersen et al., 2009; Sun et al., 2016). (For more information on post-translationally modified epitopes, we refer readers to Petersen et al., 2009.) While many post-translationally modified epitopes have been described in cancer and in autoimmunity, few have been described in the case of viral antigens. However, one question of particular interest, also in the context of SARS-CoV-2, is whether glycosylated sites are differentially recognized and, in the context of N > D modifications, which are associated with the removal of the polysaccharide moiety in the course of cellular processing. However, thus far, in the case of SARS-CoV-2 no reports have appeared of post-translationally modified or glycosylated peptides being recognized by T cell responses.

T cells recognize a bimolecular complex of an epitope bound to a specific class I or class II MHC molecule (HLA in humans), which is called its restriction element. HLA class I restricted epitopes are generally 9–10 residues in size, with several also being 8 or 11 residues depending on HLA restriction, while class II restricted epitopes are typically 13–17 residues, although shorter and longer peptides have also been described (Peters et al., 2020; Gfeller et al., 2018; Trolle et al., 2016; Wang et al., 2008a; O'Brien et al., 2008) By the late 1980s, it was recognized that a given peptide can bind multiple HLA allelic variants, especially if those variants are structurally or genetically related (McMichael et al., 1988; O'Sullivan et al., 1991). The HLA variants or types associated with overlapping peptide-binding repertoires are classified into so-called HLA supertypes (Greenbaum et al., 2011; Sidney et al., 2008). Epitopes that bind multiple HLAs are referred to as promiscuous (Kilgus et al., 1991; Panina-Bordignon et al., 1989). In general, any given HLA-peptide complex can be recognized by a multitude of different T cell receptors, which often share a discernible pattern of sequence similarity (Dash et al., 2017; Glanville et al., 2017).

Viral genomes and proteomes are composed of multiple protein antigens. Each of these antigens is recognized in a human population to varying degrees (Sidney et al., 2020; Yewdell and Bennink, 1999). The concept of immunodominance usually refers to how strongly a given antigen is recognized, either in a given assay, individual, or population, while immunoprevalence refers to how often the antigen is recognized in a given population (Oseroff et al., 2008; Tan et al., 2014; Wang et al., 2008b), although in practice the two terms are frequently used interchangeably.

The immunodominance of a given antigen within a genome or proteome is influenced by variables such as levels of transcription and expression, stability, and patterns of expression in different cell types or anatomical sites. In the context of SARS-CoV-2, Poran et al. point out the potential of leveraging proteomic data to infer relative viral protein abundance (Poran et al., 2020b, 2020a). Several other studies have eluted SARS-CoV-2-derived peptides bound to HLA (Knierman et al., 2020; Parker et al., 2020; Weingarten-Gabbay et al., 2020) but have not shown that the epitopes are actually recognized by T cell responses. Future studies will examine the correspondence between eluted ligands and T cell recognition.

The fact that HLA binding is a necessary but not sufficient requisite for T cell recognition is well established (Assarsson et al., 2007; Kotturi et al., 2007; Yewdell and Bennink, 1999; Yewdell, 2006), as it does not guarantee that a peptide will be generated by antigen processing or ensure the availability of a repertoire of T cells capable of recognizing the corresponding epitope-HLA complex (Hataye et al., 2006; Kotturi et al., 2008). In the case of eluted ligands (Croft et al., 2019; Paul et al., 2020), factors to be considered are whether the assay used to detect eluted ligands has sensitivity comparable to T cell activation (a few epitope copies have been shown to be sufficient to activate T cells; Demotz et al., 1990; Sykulev et al., 1996) and the availability of TCR repertoire, which is also modulated by previous infection history, as discussed in more detail below.

Immunodominance and immunoprevalence within a given antigen indicates how frequently and vigorously a particular epitope is recognized given all possible peptide epitopes contained in the antigen (Sidney et al., 2020; Yewdell and Bennink, 1999). Immunodominance/prevalence hierarchies within an antigen are influenced by variables such as HLA binding capacity, antigen processing, and the repertoire of T Cell Receptor (TCR) recognizing a given HLA-epitope combination. Finally, the term breadth of responses is defined on the basis of how many antigens or epitopes are recognized, either at the level of a given individual or in a population as a whole (Sidney et al., 2020; Yewdell and Bennink, 1999).

A VARIETY OF SCREENING METHODOLOGIES

The process of epitope identification entails testing collections of candidate peptides in an assay of choice. The peptide collections utilized can span the entire genome or proteome or focus on selected antigens of interest. Peptide collections can also correspond to sets of overlapping peptides (a popular choice is 15-mers overlapping by 10 residues) that span a sequence or peptides predicted to bind to one or more different HLA types, as indicated in the third column of Table 1. In general, and in the case of SARS-CoV-2 in particular, overlapping peptides are more often used to define class II restricted epitopes (4 of 9 studies; 44%), partly due to the lower predictive efficacy of HLA class II predictions (Peters et al., 2020) relative to class I epitopes (6/25 studies; 24%), for which predicted binders are often used to

probe responses (21 of 25 studies; 85%). While the length of HLA class II restricted epitopes varies, the use of 15-mers overlapping by 10 residues ensures that any possible 10-mer is represented in the peptide set, with the addition of flanking residues at either or both ends. Given that the critical core of class II epitopes is 9 residues in size, this ensures that most, if not all, epitopes are identified without having to rely on bioinformatic predictions.

Another issue of relevance is whether responses are measured directly *ex vivo* or if an *in vitro* culture restimulation step is introduced. A restimulation step is often used to expand low frequency T cell specificity that would otherwise be difficult to detect. A number of different methodologies are used to detect or expand T cells, ranging from stimulation with whole antigens or antigen fragments, to the use of peptide pools or isolated individual peptides. However, *in vitro* restimulation is known to substantially alter the phenotypes and/or relative frequency of responding T cells. The expansion of naive T cells can also occur. In the case of SARS-CoV-2, studies have shown that when peripheral blood mononuclear cell (PBMCs) are expanded for 10–14 days before the assessment of SARS-CoV-2 responses, $CD4^+$ T cells expand to a much greater extent than do $CD8^+$ T cells (Habel et al., 2020; Mateus et al., 2020).

To overcome these caveats, it is preferable to assay T cells *ex vivo* whenever possible. In the case of SARS-CoV-2 T cell epitopes, 14 studies have used direct *ex vivo* assays (fourth column of Table 1) and 12 have utilized *in vitro* culture (one study utilized both *in vitro* and *ex vivo* approaches). Alternatively, once the epitopes are identified, they can be used to conduct secondary epitope validation experiments with direct *ex vivo* modalities, as shown by 2 studies (Chen et al., 2021; Schulien et al., 2021). Of particular note, Keller et al. showed that SARS-CoV-2 T cells can be expanded in controlled conditions and raised the possibility that epitope-expanded T cells could be used for adoptive therapy (Keller et al., 2020). The principle and conditions for adoptive therapy have been described and reviewed elsewhere (Riddell and Greenberg, 1995).

ASSAY READOUTS

Regardless of whether T cell responses are detected *ex vivo* or after *in vitro* expansion, a variety of different assay methodologies are available to investigate specific T cell responses. In selecting an approach, several considerations apply, including ease of implementation, throughput, and comprehensiveness and functionality. Certain assays, such as enzyme-linked immunospot (ELISpot), supernatant determination, and whole blood assays are easier to employ and more amenable to high-throughput testing. However, they are associated with less granular information. For example, the CD4 versus CD8 phenotype (and the expression of other cell markers) of the responding cells is not readily established by these approaches compared to other methods, such as intracellular cytokine staining (ICS) or activation-induced marker (AIM) assays. The methodologies utilized by the various studies reviewed here are listed in Table 1 and include AIM, degranulation, proliferation, ELISA, ELISpot, ICS, cytotoxicity, and multimer-based assays (for 3, 2, 2, 1, 5, 10, 1m and 13 studies, respectively).

Several studies (Kared et al., 2021; Nielsen et al., 2020; Poran et al., 2020b, 2020a; Prakash et al., 2020; Rha et al., 2021; Sekine et al., 2020; Shomuradova et al., 2020; Ferretti et al., 2020; Gangaev et al., 2020; Habel et al., 2020; Sahin et al., 2020; Schulien et al., 2021; Saini et al., 2020, 2021) performed high-resolution analysis of SARS-CoV-2-specific $CD8^+$ T cells using HLA multimers. However, none of the studies reported similar multimer analyses for $CD4^+$ T cells, despite the fact that, in general, HLA class II restricted SARS-CoV-2-specific T cell responses are more pronounced than HLA class I restricted T cell responses (Grifoni et al., 2020; Nelde et al., 2021). This reflects the relatively higher availability of HLA class I multimeric reagents as compared to their HLA class II counterparts. Some studies analyzed epitope-specific responses not only in blood but also in tissues such as tonsil and lung tissue from uninfected donors (Habel et al., 2020). The analysis of tissue-derived T cells can provide insight into disease—for example, by defining the characteristics of tissue resident memory T cells, which may differ from those circulating in the peripheral blood (Masopust and Soerens, 2019).

An issue encountered with ELISpot, ICS, and related assays is that while they, by definition, identify T cells capable of a functional response, they only (also by definition) detect T cells producing a cytokine of choice; therefore, they are "blind" to T cells that produce different cytokines or that do not produce cytokines in large amounts within the window of time of the assay (e.g., T follicular helper [Tfh] CD4 T cells generally produce very low amounts of cytokines). Both AIM (Dan et al., 2016; Locci et al., 2013; Reiss et al., 2017) and HLA tetramer/multimer assays are "agnostic" in this respect, as they detect all cells activated by the epitope (AIM), or all cells expressing a TCR capable of binding a given epitope-HLA complex (tetramer/multimer). Accordingly, it is often observed that AIM and tetramer assays have higher sensitivity because they detect larger numbers of T cells than ELISpot assays. Sahin et al. note that a comparison of data from MHC multimers with bulk $IFN\gamma^+$ $CD8^+$ T cell responses indicated that a functional T cell assay might underestimate the total cellular immune response (Sahin et al., 2020). Conversely, T cells captured by tetramers might not be functional or exhausted, and therefore might overestimate the cellular response that is relevant for immunity and infection control. However, for SARS-CoV-2, it has been observed that CD8 T cells identified by HLA multimers in COVID-19 subjects are functional and not exhausted (Rha et al., 2021). In conclusion, a variety of epitope screening and assay strategies have been utilized, each with its own features and potential advantages and disadvantages.

NUMBER OF EPITOPES IDENTIFIED IN THE DIFFERENT STUDIES

The sixth column of Table 1 lists the total number of characterized canonical CD4 and CD8 epitopes identified in each study, which ranged from 1 to 734 (median of 12). It is not possible to estimate the total number of unique identified epitopes by simply adding these numbers together, because the same epitope might have been identified independently in multiple studies (as addressed below in the immunodominance section). This is especially the case for CD4 epitope studies that have utilized overlapping peptides; essentially, the same epitope might have been identified by two largely overlapping peptides.

As such, to assess CD4 epitope redundancy, we refined the data further by taking advantage of the clustering tool provided

by the IEDB (Dhanda et al., 2018a), which automatically removes duplications and largely overlapping entries; we also performed additional manual curation. This clustering tool is an algorithm that generates clusters from a set of input epitopes based on representative or consensus sequences. It allows users to cluster peptide sequences on the basis of a specified level of identity by selecting among three different method options. For our purposes, we utilized the default "cluster-break" settings, which generate clusters in which all component epitopes share at minimum a specified level of homology (70% in our case) and no epitope is present in more than one cluster. Because of the closed ends of the class I MHC binding groove, and hence the incapacity of class I binding peptides to assume alternate frames, overlapping CD8 epitopes are considered unique epitopes by default.

For our analyses, we only considered epitopes of 8–14 residues for HLA class I and epitopes of 12–25 residues for class II. We used these parameters as they reflect the canonical sizes for class I and class II ligands and because of reports that overly short or long ligands can often represent "false positives" rather than being derived from peptides truly bound to MHC (Paul et al., 2018). We have not considered instances where the CD4/CD8 (class II/class I restriction) phenotype of responding T cells was not resolved or could not be reasonably inferred. These selection criteria did not lead to the exclusion of any studies, but rather to a few ambiguous epitopes being identified, accounting for a total of 81 unique sequences omitted from this analysis. Accordingly, we determined that the studies listed in Table 1 encompass 1,434 unique epitopes, which include 1,052 different class I and 382 different class II non-redundant epitopes (versus 416 when redundant epitopes were included).

Regarding limitations of the approach, in our review we have not considered data regarding HLA peptide binding (Prachar et al., 2020) or ligands eluted from HLA (Knierman et al., 2020; Parker et al., 2020; Weingarten-Gabbay et al., 2020) in absence of T cell recognition data. As more of this type of data is generated and reaches a critical mass, it will undoubtedly be of interest to correlate these data with T cell epitope recognition data. Our analyses have also not included epitopes defined in animal models. To date, few studies have described murine epitopes, and no data is available regarding the epitopes recognized by other species used in model systems such as Syrian hamsters or non-human primates (NHPs), even though some data has been reported suggesting that CD4 epitopes recognized in humans can be cross-recognized in NHPs (Shaan Lakshmanappa et al., 2021). Further experiments are required to enable the study of epitope-specific responses in SARS-CoV-2 animal studies. Finally, some of the information contained in this review is derived from preprint manuscripts that had not been formally peer reviewed at the time of analysis. The potential for variation in content between preprint and final versions of various studies is recognized by the curation process instituted by the IEDB team (of which B.P., A.S., and R.V. are part), in which each study originally curated at the preprint stage is re-curated when the study appears in the final published version.

ANTIGENIC TARGETS AND EPITOPE DISTRIBUTION

Ten of the 25 epitope identification studies (Ferretti et al., 2020; Gangaev et al., 2020; Kared et al., 2021;Mateus et al., 2020; Nelde et al., 2021; Saini et al., 2020, 2021; Schulien et al., 2021; Snyder et al., 2020; Tarke et al., 2021a;, Prakash et al., 2020) screened peptides derived from the entire SARS-CoV-2 proteome (seventh column of Table 1). The main antigenic targets of CD4 and CD8 SARS-CoV-2 T cell responses have been defined by several studies by utilizing overlapping peptides, rather than by resolving the actual epitopes (Grifoni et al., 2020; Tarke et al., 2021a), and are reviewed elsewhere (Altmann and Boyton, 2020; DiPiazza et al., 2020). These studies determined that structural proteins (S, M and N) are dominant targets of T cell responses, with ORF3, ORF8, and nsp3, 4, 6, 7, 12, and 13 (ORF1ab) also being frequently targeted. Other studies focused on specific subsets of SARS-CoV-2 antigens, as detailed in the seventh column of Table 1.

The various studies differ widely in the depth of screening, number of antigens tested, HLA alleles targeted, and number of peptides screened. For example, Peng et al. (2020) screened the whole proteome, with the exception of ORF1ab, using 423 peptides assayed in 42 infected and 16 non-exposed subjects, and they reported broad CD4 and CD8 responses. Conversely, Schulien et al. (2021) tested only 5 peptides predicted to bind to each of ten different HLAs. Tarke et al. (2021a) used PBMC from 99 donors and probed for CD4 responses using 1,925 peptides that spanned the entire SARS-CoV-2 proteome. To probe for CD8 responses, they tested an additional 5,600 peptides predicted to bind to one or more of 28 prominent HLA class I alleles. Snyder et al. (2020) screened 545 peptides distributed over the SARS-CoV-2 proteome for 26 class I alleles, testing about 20 peptides per allele. Nelde et al. (2021) screened a large number of donors (220 in total) with peptides spanning the breadth of antigens (i.e., the whole SARS-CoV-2 proteome) predicted to bind six HLA class I alleles or various HLA-DR class II. Le Bert et al. (2020) focused on peptides derived from N, nsp7, and nsp13, while Ferretti et al. (2020) screened predicted peptides from the entire proteome for 6 HLA alleles in 5 to 9 donors per HLA.

The epitope distribution along the SARS-CoV-2 proteome is analyzed in more detail in Figures 1A and 1B, in which the number of epitopes identified in each antigen is shown for CD4 and CD8 epitopes, respectively. Figures 1C and 1D show the correlation between the number of epitopes and the total number of residues (size) of each antigen. A significant correlation exists between antigen size and the number of epitopes identified for both CD4 ($p = 0.0015$ and $r^2 = 0.36$) and CD8 epitopes ($p < 0.0001$ and $r^2 = 0.76$). Certain antigens (N, M, S, and E) were studied in more detail (more studies focused on those antigen targets instead of considering the entire SARS-CoV-2 proteome) (Figures 1E and 1F). This is a significant factor, in addition to antigen length, in influencing the number of epitopes identified. Additionally, we recognized early on that the immunodominance pattern of the CD4 and CD8 T cell response to SARS-CoV-2 largely tracks with the expression level of each of the 25 viral proteins (Grifoni et al., 2020). S, M, and N sgRNAs are highly expressed by SARS-CoV-2 infected cells, and those three proteins are the most immunodominant targets of human CD4 and CD8 T cell responses to SARS-CoV-2 (Grifoni et al., 2020).

In conclusion, T cell responses are multi-antigenic, with the structural antigens being broadly recognized but with other proteins, such as nsp3, nsp4, nsp12 and ORF3a, also being

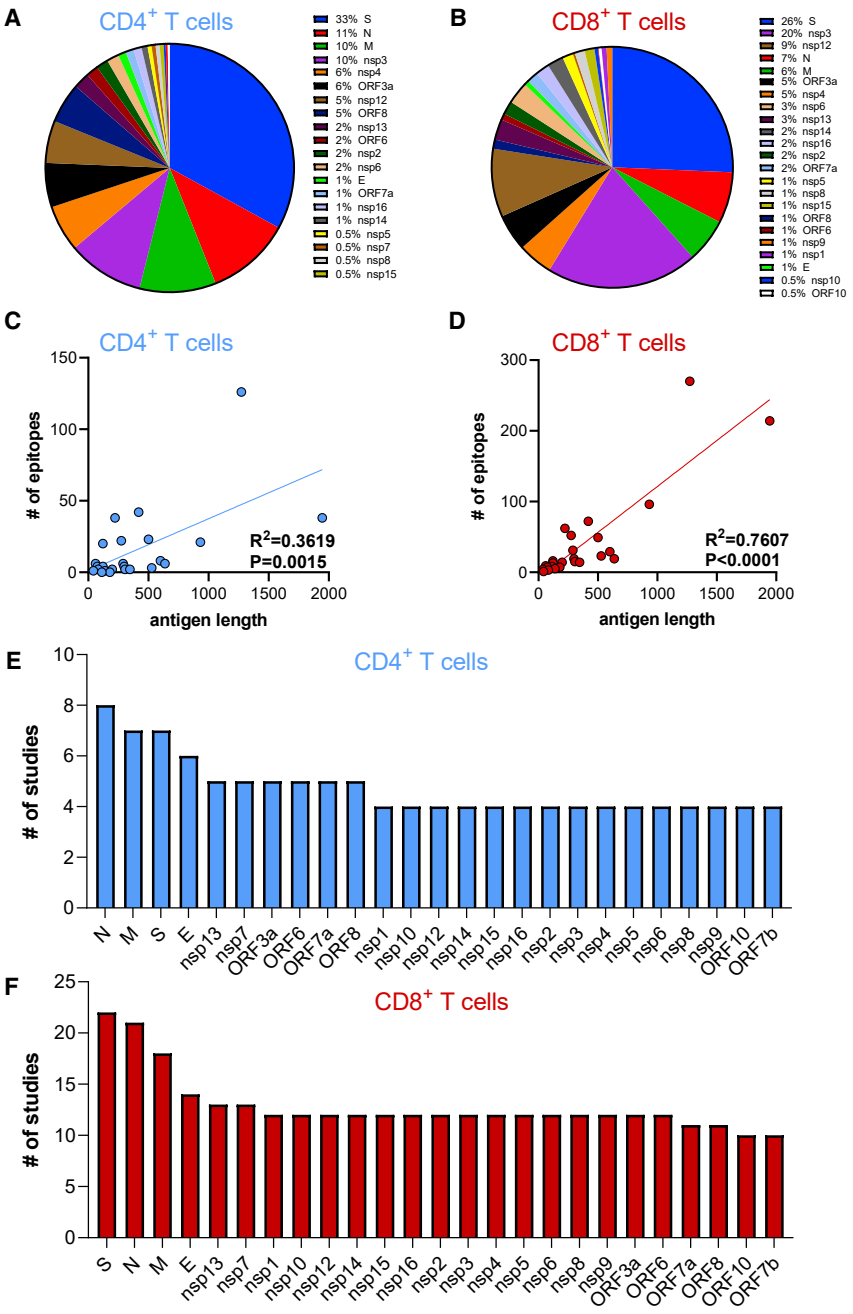


Figure 1. Distribution of CD4 and CD8 epitopes by SARS-CoV-2 antigen

The fraction of known CD4 and CD8 epitopes derived from recognized SARS-CoV-2 antigens is shown in (A) and (B), respectively. The number of epitopes derived from each antigen as a function of antigen size is plotted in (C) and (D) for CD4$^+$ (light blue) and CD8$^+$(red) T cells, respectively; p values were calculated using a simple linear regression. (E) and (F) show the number of studies that probed responses to each antigen. All the source data used in these analyses were derived from the papers cited within Tables 1 and S1.

vigorously recognized. This difference is not unexpected, since structural proteins are present in high concentrations in the virus and are accessible to the exogenous processing pathway and to HLA class II molecules. Non-structural proteins, which are produced in infected cells, also have access to the endogenous processing pathway and to HLA class I molecules.

IMMUNOME BROWSER ANALYSIS IDENTIFIES PATTERNS OF IMMUNODOMINANCE

We also assessed whether discrete immunodominant regions would become apparent when we took a global view of the reported epitope data. We utilized the Immunome Browser tool (Vita et al., 2019; Dhanda et al., 2018b), developed and hosted by the IEDB (www.iedb.org). This tool allows patterns of immunodominance to be visualized across the entire SARS-CoV-2 proteome by plotting the 95% confidence interval (CI) of the Response Frequency (RF) for each residue, which is defined as the number of individuals and assays reporting positive responses to a peptide encompassing the particular residue. The lower bound RF values, using an average across a sliding 10-residue window, are plotted for human CD4 and CD8 epitopes in Figure 2 for the antigens S, M, N, nsp3, and nsp12. These antigens were chosen as their epitopes were described in sufficient number to allow us to delineate discrete immunodominant regions.

In the case of the spike protein, several immunodominant regions were observed for CD4 (residues 154–254, 296–370 and 682–925; Figure 2A), compared to a more homogeneous distribution for CD8 (Figure 2B). For the other structural proteins, namely the membrane and nucleocapsid, similar immunodominant regions for CD4 (Figure 2C) and CD8 (Figure 2D) were noted, with the 7–101 and 131–213 residue ranges being more prominent for the membrane protein and the 31–173 and 201–371 ranges for the nucleocapsid. More marked differences in CD4 and CD8 immunodominant regions, and in overall response frequency, were observed in the cases of nsp3 (Figure 2E) and nsp12 (Figure 2F). For both these proteins, defined immunodominant regions for CD4 (789–843, 1118–1158, and 1873–1903 for nsp3 and 863–903 for nsp12) were evident, versus more homogeneous patterns of CD8 recognition similar to that noted for the spike protein (Figure 2B). In conclusion, CD4 $^+$ T cells, in general, recognize more defined immunodominant regions than do their corresponding CD8$^+$ counterparts.

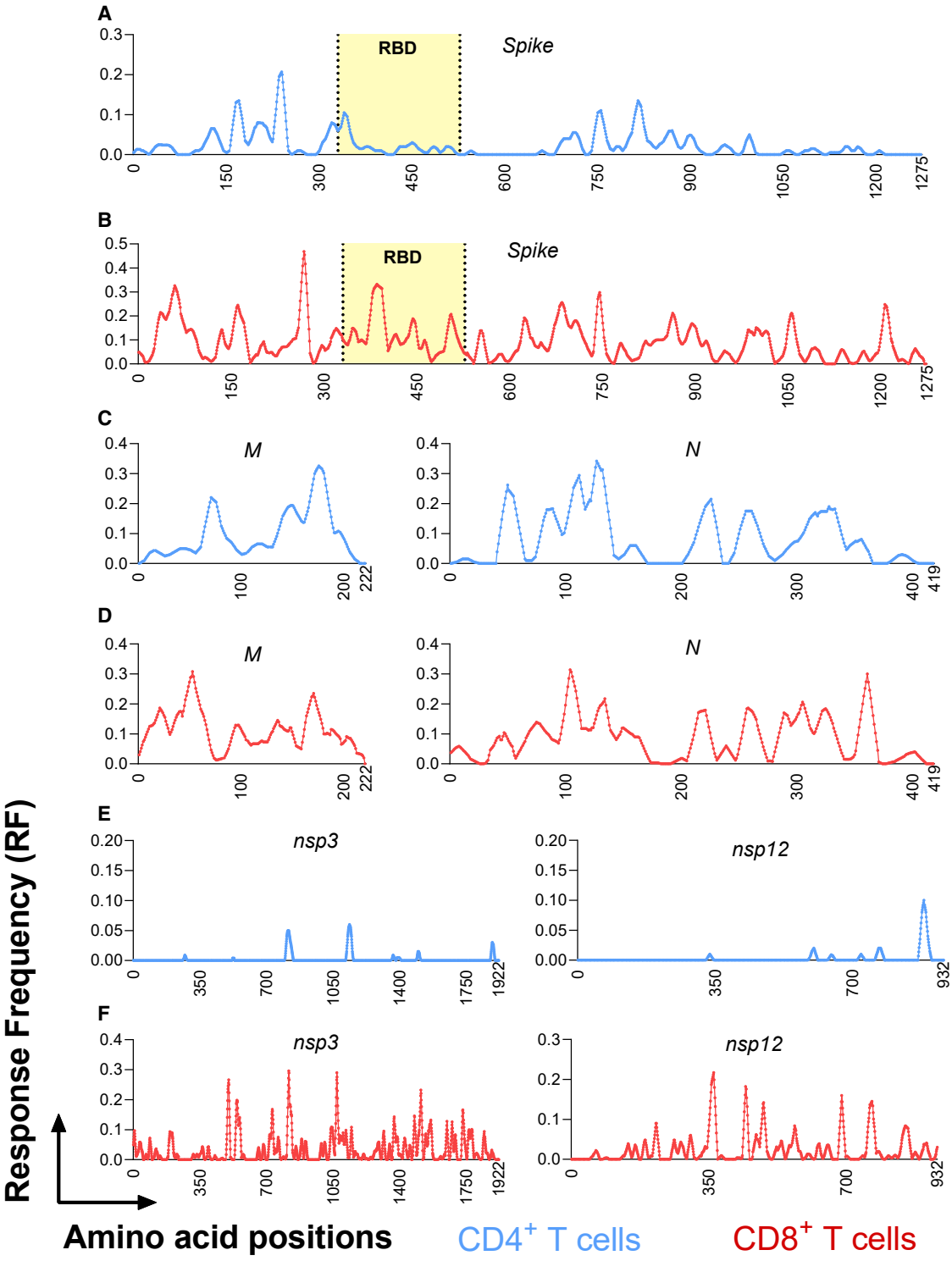


Figure 2. Identification of immunodominant antigenic regions

The IEDB's Immunome Browser tool was utilized to identify potential antigenic regions across the entire SARS-CoV-2 proteome. After searching for SARS-CoV-2-derived $CD4^+$ (light blue) and $CD8^+$(red) T cell epitopes, individual antigens were selected for further evaluation. From the antigen-specific Immunome Browser link, data was downloaded as an Excel file to obtain position-specific lower bound response frequency scores (RF), defined as the number of individuals and assays reporting positive responses to a peptide including that particular residue. For visualization, RF scores for each residue were recalculated to represent a sliding 10-residue window. Position-specific RF values for CD4 (light blue) and CD8 (red) epitopes are shown for the most dominant viral antigens: spike (A and B); M and N (C and D); nsp3 and nsp12 (E and F). The receptor binding domain region of the spike protein, is indicated in yellow in A and B because it is critically recognized by neutralizing antibodies and implicated in viral cell entry.

EPITOPE IDENTIFICATION IN DIFFERENT POPULATIONS AND COHORTS

As a whole, the different studies considered here have reported epitope identification results from a total of 1,197 donors (median = 34, range 2 to 220; see the eighth and ninth columns of Table 1). Of those, 870 donors were SARS-CoV-2 infected and 327 were unexposed.

It should be noted that these figures reflect the maximum number of donors utilized in each epitope identification and characterization study, as some assays and some epitopes have been tested in a different number of donors. For example, in some cases 20 donors were tested in ELISpot, but only 10 were evaluated using MHC multimers. Similarly, in several instances, because of the need to match peptide candidates to specific predicted HLA alleles (e.g., HLA-A*02:01 candidate epitopes may only have been tested in HLA-A*02:01-positive donors), the actual number of donors in which each peptide was tested might be significantly lower in comparison to other peptides.

Several studies have analyzed differences between the infected and unexposed cohorts and also in the context of potential cross-reactivity of SARS-CoV-2 epitopes with homologous sequences from common cold coronaviruses or other viruses, as discussed in more detail below. Also, as noted elsewhere (Sette and Crotty, 2021), considerable heterogeneity exists in SARS-CoV-2 infection and immune responses as a function of different variables such as age, gender, disease severity, ethnicity, co-morbidities, and time since symptom onset. As yet, the epitope identification studies do not answer the question as to whether differences in the types of epitopes recognized exist as a function of these variables. However, the epitopes defined in these studies, together with data generated from peptide pools, will undoubtedly be key to probing these variables and their role in the differences observed in terms of SARS-CoV-2-specific immune responses by evaluating the overall pattern of reactivity instead of focusing on few antigens or epitope candidates.

One issue to consider in future studies, and touched on further below, is to ensure that different ethnicities are adequately represented in SARS-CoV-2 studies. Thus far, most studies have been performed in donor cohorts that mostly consist of Caucasians and in which other ethnic groups are relatively under-represented.

HLA COVERAGE AND EPITOPE IDENTIFICATION RESULTS

It is well appreciated that HLA molecules are associated with an outstanding degree of diversity. Class I molecules are encoded by 3 main HLA loci (A, B, and C), and class II molecules are encoded by four main loci (DRB1, DRB3/4/5, DP, and DQ). Each locus is highly polymorphic, and because of heterozygosity, each individual might express close to 14 different HLA molecules and a minimum of 7 (if homozygous at all loci). Not only are the various HLA loci highly polymorphic, but the frequencies of their respective alleles vary, sometimes dramatically, across different ethnicities (Gonzalez-Galarza et al., 2020; Robinson et al., 2020). Establishing the extent to which epitope identification studies adequately cover the worldwide population is both a key and non-trivial issue (Greenbaum et al., 2011; McKinney et al., 2013; Sette and Sidney, 1999).

To meaningfully discuss population coverage of HLA allelic variants in the context of epitope identification efforts, we need to define what is meant by population coverage. The total phenotypic coverage provided by a set of HLA alleles represents the fraction of individuals that express at least one of a given set of alleles, while genotypic coverage corresponds to the fraction of genes at a specific locus the set of allelic variants covers. By way of example, an analysis targeting the HLA-A*01:01, B*07:02 and DRB1*01:01 molecules will give a phenotypic coverage (probability that an individual in the average worldwide population will express at least one of these alleles) of approximately 35%. However, these three allelic variants represent only about 5%–10% of the gene variants at each of these three different HLA loci. This is important because in an individual that is ''covered,'' in the sense of expressing one HLA, the bulk of the T cell response will likely be directed to the other (up to thirteen) class I and class II alleles, leading to a gross misrepresentation of the total response magnitude and target specificity.

In previous studies, we have devoted significant efforts to analyzing the number of different HLA alleles associated with good genotypic and phenotypic coverage, and found that ~25 different HLA class II and ~25 different HLA class I alleles are required to cover 90% or more individuals in an idealized population (43, 61, 62). In the case of SARS-CoV-2 epitope identification studies, HLA restricted epitopes have been identified for 30 HLA class I and for 45 HLA class II alleles (Figures 3A and 3B), including, in both cases, the vast majority of the most common specificities in the general worldwide population (Gonzalez-Galarza et al., 2020; Weiskopf et al., 2013; Greenbaum et al., 2011).

The median number of epitopes per allele is 35 (range 1 to 219) for class I, and 12 for class II (range 1 to 82). In the case of class I, as might be expected, most restrictions have been identified in the contexts of A*02:01, A*24:02, A*01:01, and B*07:02, as these are the most common class I alleles worldwide. Similarly, most class II restrictions are for DRB1*07:01 and DRB1*15:01, the most common DRB1 specificities worldwide. In both cases, the number of restrictions generally corresponds to overall allele frequency in the respective cohorts. This data exemplifies how the number of epitopes associated with a particular allelic specificity may not necessarily reflect immunodominance, but rather bias due to the availability of corresponding donor samples. Thus, the limited number of epitopes identified for several alleles is because they are rarer and therefore reflective of investigational bias. Additional studies are required to provide fully unbiased investigations of SARS-CoV-2 on a global scale. The number of allelic restrictions identified by the different studies is summarized in the tenth and eleventh columns of Table 1.

Overall, the 25 different studies mapped or inferred 1,191 class I restrictions, including 1,019 unique epitope-allele combinations (Table S1), with individual studies defining between 1 and 523 (median 8). For class II, 783 restrictions were mapped or inferred, with 760 representing unique epitope-allele combinations (Table S1). Only 9 studies investigated CD4 responses, with just 3 identifying class II restrictions (see Table 1). Thus, the number of experimentally defined HLA restrictions are fewer for class II relative to class I, which is consistent with the fact that class I restrictions are more easily inferred or determined and

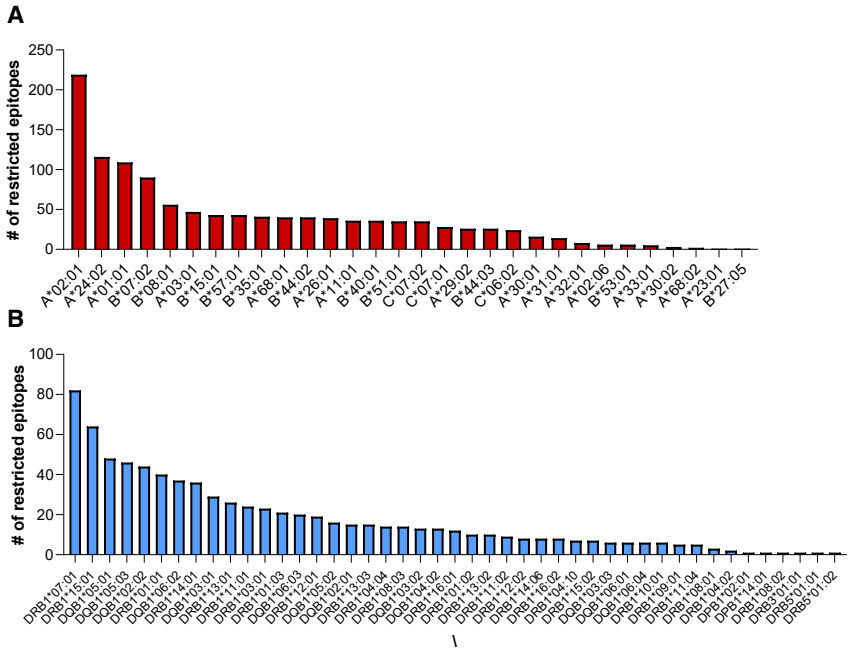


Figure 3. Defined HLA class I and class II restrictions

HLA-restricted epitopes have been identified for 30 class I (red, A) and 45 class II (light blue, B) molecules. The charts shows the number of restricted epitopes associated with each allele (alleles shown on the horizontal axis).

that multimers/tetramers (which implicitly assign restriction) are more broadly available for HLA class I than for HLA class II.

IMMUNODOMINANCE AT THE LEVEL OF SPECIFIC EPITOPES AND ALLELES

Different studies report numerous peptides as being immunodominant, although each study also used different subjective definitions of immunodominance. While some peptides are repeatedly and independently identified, differences among these studies all contribute to the differences in their outcomes. These include differences in screening procedures, in HLA alleles considered, in the antigens targeted, the sampling of small numbers of individuals, and in how ''immunodominance'' was defined. For example, Peng et al., (2020) report several immunodominant peptides that they defined as being recognized by 6 or more of the up to 16 subjects screened. Tarke et al., (2021a) also highlight some epitopes as being more dominant, with 49 class II epitopes being recognized in 3 or more donors from an average of 10 donors tested and 41 class I epitopes recognized in 50% or more of the HLA matched donors tested. The same study also found that the response is broad and multi-specific, with ~8–9 different antigens required to cover about 80% of the total CD4 and CD8 response (Tarke et al., 2021a). Nielsen et al. also concluded that the response is broad, since the top three immunogenic epitopes derived from separate SARS CoV-2 proteins (Nielsen et al., 2020). Keller et al. reported immunodominant epitopes defined as epitopes being recognized in multiple donors from M, N, and S (Keller et al., 2020).

Some specific epitopes are highlighted as being immunodominant in multiple studies. For example, in the context of the HLA-A*02:01 class I molecule, which is the most studied for CD8 SARS-CoV-2 responses, the S 269–277 epitope (sequence YLQPRTFLL) is detected in 81% of HLA-A2+ individuals in the Nielsen study (Nielsen et al., 2020). The same A2 dominant epitope is also reported by Shomuradova et al., who tested 13 A2 peptides in total, and also identified a less strongly recognized epitope (Shomuradova et al., 2020). In the Habel et al. study, of the 14 peptides screened, S 269–277 generated the strongest IFN-γ^+ response, with S 976–984 and ORF1ab 3183–3191 less prominently recognized (Habel et al., 2020). Ferretti et al. identified 3 epitopes recognized in 3 or more subjects (67% of the subjects tested), including S 269–277 (Ferretti et al., 2020). The study by Sahin et al. reports S 269–277 as the most dominant epitope and also identifies epitopes strongly recognized in the context of HLA-A*24:02 and HLA-B*35:01 (Sahin et al., 2020). Rha et al. detected S 269–277 responses in 37 of 112 (33%) patients, while S 1220–1228 was detected in only 2 of 40 (5%) patients (Rha et al., 2021), although other studies have observed higher response rates for this latter epitope. Overall, the S 269–277 epitope was found to be positive in 11 independent studies.

Another example of an immunodominant epitope is provided by the HLA-A*01:01-restricted nsp3 819–828 epitope (sequence TTDPSFLGRY). This epitope was reported by Nelde et al. as being positive in 83% of the donors tested (Nelde et al., 2021). This study also identified a large number of additional dominant CD4 and CD8 restricted epitopes. The same A1-restricted epitope was also reported by Saini et al., who tested over 3,000 peptides for 10 alleles (Saini et al., 2020; Saini et al., 2021) and found 214 peptides that were recognized in 16 out of the 18 samples analyzed. Two additional HLA-A*01:01 epitopes that overlap with TTDPSFLGRY (nsp3 818–828, sequence HTTDPSFLGRY, and nsp3 819–829, sequence TTDPSFLGRYM) were also identified as particularly dominant. The study by Gangaev et al. screened 50 epitopes for 10 alleles using tetramers (500 total) in 18 donors and identified nine epitopes in total, including the immunodominant nsp3 epitope restricted by HLA-A*01:01 (Gangaev et al., 2020).

GLOBAL ANALYSIS OF IMMUNODOMINANT EPITOPES

We further assessed published epitope data to determine whether particular HLA alleles and epitopes are dominantly recognized. In the case of HLA class II, because of the technical issues discussed above, dominant alleles are less readily assigned as restriction elements. In the case of HLA class I, certain alleles, such as HLA-A*01:01, B*07:02, B*08:01 and B*44:01 were associated with dominant responses (Tarke et al., 2021a).

Other alleles, such as HLA A*02:01, were associated with numerous epitopes but with responses of lower magnitude on average, and alleles such as A*30:01 and A*32:01 were associated with weak and infrequent responses. This HLA-allele-specific variation in response frequency/magnitude has been observed previously in the contexts of HIV and Dengue virus, where responses mediated by particular HLA allelic variants were associated with protection or susceptibility to disease (Goulder and Walker, 2012; Weiskopf et al., 2013). Whether HLA types play a role in influencing disease severity in the context of SARS-CoV-2 will have to be established as larger datasets become available.

For present purposes, we have defined the most dominant CD4 and CD8 epitopes as those recognized in 3 or more donors/studies, consistent with the definitions utilized by Mateus et al. and Tarke et al. (Mateus et al., 2020; Tarke et al., 2021a). We utilized this threshold based on previous experience in this matter. By selecting epitopes that have been recognized in multiple different experiments in separate donors allow, we can narrow the number of epitopes and focus on more dominant or prevalent responses while still preserving the goal of representing epitopes presented by a wide variety of HLA alleles. That is because less common HLAs are found, by definition, in fewer individuals, and the studies considered involved a median of 34 donors. Therefore, raising the "bar" further would restrict "immunodominant epitopes" to just those restricted by alleles that are very common in Caucasians.

The immunodominant epitopes identified in this way are highlighted in Table S1 and total 399 epitopes (110 CD4 epitopes and 289 CD8 epitopes). It is important to note that no epitope was not recognized in 100% of the cases/donors it was tested in, as has been observed in other viral systems (e.g., HBV, HIV, Poxviruses, Flu). This is relevant because it argues against the use of single-epitope tetramers to measure responses because of the likelihood of false negative results. Instead, the results argue for the use of peptide pools or multiplexing strategies (Kared et al., 2021; Nelde et al., 2021; Sekine et al., 2020; Shomuradova et al., 2020) to ensure the broad coverage of responses.

Another important consideration, as noted above, is the influence of investigational bias. It is apparent that epitopes from the spike protein, and those restricted by the most common HLA alleles, are overrepresented, which is likely a reflection that the spike antigen and those particular HLA alleles are more frequently studied (Figures 1E and 1F).

BREADTH OF THE T CELL REPERTOIRE

As summarized above in Figure 1, a total of 1,434 unique, non-redundant CD4 and CD8 epitopes have been defined, with the top 10 antigens accounting for 86% of the total. In these 10 most dominant antigens, a median of 87 epitopes (range of 33 to 396) is recognized. The data presented above demonstrates that T cell responses are multi-antigenic, with structural antigens being broadly recognized, but with other proteins such as nsp3, nsp12, ORF3a, and ORF8 also being vigorously recognized. Furthermore, data from Tarke et al. show that each individual is conservatively estimated to recognize, on average, 19 different CD4 and 17 different CD8 epitopes (Tarke et al., 2021a). Although individuals in our experience target multiple epitopes, the efficacy of the responses and number of epitopes targeted may vary substantially, dependent on HLA, the severity of disease, and other factors.

This breadth of response is apparently at variance with other reports describing only a limited number of epitopes for SARS-COV-2 (Chen et al., 2021; Le Bert et al., 2020; Lee et al., 2020; Nielsen et al., 2020; Rha et al., 2021; Sekine et al., 2020; Kared et al., 2021; Sahin et al., 2020). In some cases, *in vitro* expansion with artificial antigens has been utilized, and/or a limited number of subjects, cells, and/or epitope candidates were screened. Furthermore, several of the reported narrow repertoire epitopes differ among the different studies, consistent with a stochastic selection effect. Overall, the data curated in the IEDB as of March 15, 2021, reveals that over 1,400 different SARS-CoV-2-derived peptide sequences are reported as being recognized by human T cell responses, and which consist of 382 CD4 and 1,052 CD8 epitopes based on the meta-analysis performed in the current review.

PRE-EXISTING REACTIVITY AND CROSS-REACTIVITY WITH COMMON COLD CORONA AND OTHER VIRUSES

Several studies have detected responses to SARS-CoV-2 sequences in unexposed controls (Sette and Crotty, 2020b, 2021). In some cases, these responses might correspond to infections associated with a lack of antibodies or to a transient antibody response (Sekine et al., 2020; Nelde et al., 2021). However, in other cases, these responses appear to be linked to pre-existing memory responses, which, in some instances, have been mapped to the cross-reactive recognition of the SARS-CoV-2 sequences by T cells induced by endemic "common cold" coronaviruses (17) and potentially other viral species (Bacher et al., 2020; Le Bert et al., 2020). This phenomenon has received considerable attention because of its potential to influence disease severity and vaccination outcomes and because of its potential implications for herd immunity (Bacher et al., 2020; Sette and Crotty, 2020b, 2021; Lipsitch et al., 2020; Sagar et al., 2021).

Epitopes recognized in non-exposed individuals have been defined in 12 studies. In some cases, these SARS-CoV-2 epitopes had significant homology to common cold coronavirus sequences, with cross-reactivity demonstrated at the molecular level in several instances (Mateus et al., 2020). Other studies, as discussed in more detail below, have examined whether SARS-CoV-2 specific T cells might cross-react with other more closely related viruses, such as SARS-CoV-1 and the Middle East Respiratory Syndrome virus (MERS) (see also below). This issue is of relevance in the context of developing vaccines that can elicit T cell responses that broadly recognize coronaviruses of pandemic potential.

The topic of pre-existing immune responses and cross-reactivity with common cold coronaviruses was addressed by several studies that reported a range of findings. Schulien et al. detected cross-reactive T cells in longitudinal samples pre- and-post SARS-CoV-2 infection and reported that these cells were expanded after *in vitro* restimulation (Schulien et al., 2021). Sekine et al. also detected widespread reactivity in non-exposed individuals using peptide pools (Sekine et al., 2020). Shomuradova et al. detected pre-existing T cell reactivity in

unexposed donors using HLA-A2 tetramers but at much lower levels compared to those seen in exposed individuals (Shomuradova et al., 2020). Nelde et al. tested the reactivity of non-exposed donors to epitopes identified in exposed individuals and detected reactivity, albeit at lower levels, for several epitopes (Nelde et al., 2021). Keller et al. detected T cells with minimal cross reactivity with two homologous nucleocapsid peptides from NL63 and OC43 (Keller et al., 2020). Ferretti detected reactivity to OC43 and HKU1 sequences for 2 of 29 dominant epitopes and no reactivity for NL63 and 229E (Ferretti et al., 2020). Rha et al. reported that the SARS-CoV-2 S 269-277 and S 1220-1228 epitopes had low homology to OC43, HKU1, 229E, and NL63 and that MHC class I multimer+ cells were not detected in unexposed subjects (Rha et al., 2021). Prakash identified 24 epitopes, and of those, 11 recalled memory CD8+ T cells from unexposed healthy individuals (Prakash et al., 2020).

A potential explanation for the differences observed in the degree of cross-reactivity of epitope repertoires detected in infected and unexposed subjects is provided by the studies of Mateus et al. (Mateus et al., 2020) and Tarke et al. (Tarke et al., 2021a). These studies demonstrated that, overall, 50% of the epitopes defined in unexposed donors were also recognized in SARS-CoV-2-infected subjects (Mateus et al., 2020; Tarke et al., 2021a), but also that the viral infection created a new repertoire of epitopes recognized only in infected subjects. Conversely, over 80% of the epitopes defined in SARS-CoV-2-infected subjects were not recognized in unexposed donors. This suggests that a pre-existing repertoire of cross-reactive T cells is present in unexposed donors, but that the SARS-CoV-2 infection generates a largely novel repertoire of T cells in addition to the pre-existing one. Consistent with this view, the antigens dominantly recognized in exposed donors tend to only partially overlap with those dominant in non-exposed donors (Le Bert et al., 2020).

The issue of how preexisting memory reactivity might influence immunity has been debated, and a firm conclusion has not been reached as yet (Lipsitch et al., 2020; Sette and Crotty, 2020a, 2020b). While it is not expected that preexisting T cell reactivity might protect against infection, it is possible that pre-existing SARS-CoV-2 cross-reactive T cells might modulate disease severity, as reported by a recent study (Sagar et al., 2021), or might even modulate vaccine responsiveness, allowing for a faster or more vigorous response.

The study of protective versus detrimental T cell responses is important for determining the optimal T cell engagement strategies for vaccines. In addition to understanding the relationship between pre-existing immunity to human coronaviruses and host defense against SARS-CoV-2, it is relevant to also consider the contribution of COVID-19-vaccine-boosted cross-reactive immune responses to vaccine-induced protective immunity.

CROSS-REACTIVITY WITH MERS AND SARS-CoV-1

As mentioned above, several studies have addressed whether SARS-CoV-2 T cells might cross-react with more closely related viruses such as SARS-CoV-1 and MERS, an issue that is important for the development of vaccines that can elicit T cell responses to coronaviruses of pandemic potential.

As might be expected on the basis of the higher degree of sequence homology, cross-reactivity between SARS-CoV-2 responses and SARS-CoV-1 and MERS was more frequently detected than cross-reactivity between SARS-CoV-2 responses and common cold coronaviruses. More specifically, Le Bert et al. analyzed a cohort of 23 patients who recovered from SARS-1 and found long lasting memory T cells 17 years after the SARS-1 outbreak of 2003 (Le Bert et al., 2020). Habel et al. reported that T cells recognizing selected A2/SARS-CoV-2 CD8+ T cell epitopes can cross-react with SARS-CoV-1 and MERS, while they did not share homology with the common cold coronaviruses (Habel et al., 2020). Rha et al. reported that the S 269–277 epitope was specific to SARS-CoV-2, whereas the S 1220–1228 epitope was conserved in SARS-CoV-1 (Rha et al., 2021). In the study of Gangaev, of the 9 CD8 T cell epitopes they identified, 5 were unique for SARS-CoV-2 and 4 were shared between SARS-CoV-2 and SARS-CoV-1 (Gangaev et al., 2020). Prakash et al. also studied conserved pan-species epitope sequences for all coronaviruses, including those responsible for zoonotic infections (Prakash et al., 2020).

POTENTIAL FOR IMMUNE ESCAPE BY SARS-CoV-2 VARIANTS

Another topic of relevance is the effect of naturally occurring mutations on epitope recognition. SARS-CoV-2 does mutate, and a key question, particularly for vaccine programs, is whether it will mutate to escape T cell responses. The large breadth of T cell epitopes recognized, and the fact that each individual tends to recognize their own unique sets of epitopes, depending on their HLA polymorphisms, has profound implications in terms of immune escape. A recent study showed that SARS-CoV-2 mutations predicted to have a negative impact on epitope binding to HLA were indeed associated with reduced T cell activity (Agerer et al., 2021). Other analyses of mutations associated with several variants of concern (VOCs) suggest that the vast majority of defined epitopes are conserved in SARS-CoV-2 variants (Tarke et al., 2021b; Redd et al., 2021).

The topic of potential immune escape by variants has been elevated by the observation that several recent SARS-CoV-2 VOCs have accumulated unusually large numbers of mutations and exhibit significant evidence of escape from neutralizing antibodies (Tegally et al., 2021; Wang et al., 2021; Thomson et al., 2021). This evolution appears to be due to the virus's extended replication in immunocompromised individuals, at least in some cases (Avanzato et al., 2020). Given that immunity against COVID-19 consists of both antibody and T cell responses, there has been concern as to whether these variants escape T cell immunity.

The study of sequence variation and epitope recognition is of particular importance in the context of several well-described VOCs. Two independent studies (Tarke et al., 2021b; Redd et al., 2021) have shown that most of the epitopes defined by Tarke et al. (Tarke et al., 2021a) or Kared et al., (2021) are conserved within VOCs. Consistent with these observations, it has been shown that the antigens containing the sequence variations pertaining to the B.1.1.7, B.1.351, P.1, and CAL.20C variants are cross-recognized by individuals previously infected with the SARS-CoV-2 ancestral strain or that received

COVID-19 vaccination. While the frequency of response across the different variants is kept, a decrease in magnitude of 30% or less is observed in terms of T cell reactivity for specific VOCs/assay combinations, suggesting an overall negligible impact of the VOCs in in the context of the T cell responses in the groups of vaccinated and convalescent donors tested thus far (Tarke et al., 2021b; Redd et al., 2021). Because of the high number of different epitopes reported, as noted above, and because of the large breadth of epitopes recognized in any given individual (estimated to be an average of 19 class II and 17 class I epitopes per person, genome-wide, and 9 if only the spike protein is considered), as suggested by one study (Tarke et al., 2021a), it appears unlikely that the new variants will escape T cell recognition at either the population or individual level.

In light of the data that indicate that T cell escape is not occurring (Tarke et al., 2021b), it is also relevant to consider the immunological and virological features that make T cell escape by SARS-CoV-2 unlikely. First, as noted, the broader the T cell response, in terms of epitopes, the less likely viral escape becomes, because any individual epitope that can escape through viral mutation would represent a small fraction of the overall immunity and thus represent a small selective pressure. Given that SARS-CoV-2 is a large RNA virus, the breadth of the CD4 and CD8 T cell responses is not surprising, per se.

Second, there are few examples in the literature of T cell epitope escape in humans for a virus that causes acute infections. In contrast, viruses that cause chronic viral infections, such as HIV and HCV, are well known to escape T cell epitope recognition. This is due to a fundamental difference in selective pressure. Within a single person, there is strong selective pressure for a chronic viral infection to escape T cell responses over time. In contrast, in a population of people, the diversity of HLA alleles presents a fundamental challenge for viral escape. This phenomenon is a basic premise in the evolutionary value of human HLA diversity. The escape of one or more T cell epitopes in one individual is unlikely to give the virus a selection advantage in the next host; indeed, escape mutations are more likely to be disadvantageous, because the original viral protein sequence was selected for functionality. However, in the influenza virus context (Rimmelzwaan et al., 2005), multiple compensatory co-mutations in the nucleoprotein have been observed to restore viral fitness. It remains possible that SARS-CoV-2 cytotoxic T-lymphocyte escape mutants might survive by a similar mechanism. The potential selection of viral T cell escape variants will depend on how well the spread of SARS-CoV-2 is controlled and, although selection for T cell escape variants may be highly restricted (owing to the factors discussed above), it cannot be ruled out at this time.

Third, a cornerstone feature of SARS-CoV-2 is the rapidity of replication and transmission within the human upper respiratory tract. Approximately half of SARS-CoV-2 transmissions occur in the pre-symptomatic phase of infection, before a T cell response has been mounted (in a previously unexposed or unvaccinated individual). The kinetics of SARS-CoV-2 replication and transmission are inconsistent with T cell pressure being a major component of intra-host selection in most individuals nor an evolutionarily relevant pressure, even though viral escape mutations may arise quickly, in acute infection, during the viremic phase. Combined, these virological, immunological, and epidemiological factors make it unlikely that SARS-CoV-2 will escape human T cell responses at the population level. Nevertheless, it is still possible that escape from T cell epitope recognition could occur in immunocompromised patients, some of whom have high levels of viral replication for > 120 days; therefore, it could be speculated that SARS-CoV-2 could/can undergo extensive mutation in such individuals during this time.

As mentioned above, it is important to evaluate SARS-CoV-2 epitope recognition in convalescents over time. Indeed, Bilich et al. (Bilich et al., 2021) published a recent study (which just missed the analysis time-point of March 15, 2021) in which they evaluated the T cell recognition of specific SARS-CoV-2 epitopes in a six-month follow-up of 51 convalescent individuals after mild or moderate SARS-CoV-2 infection. They detected epitopes capable of mediating long-term T cell responses, while responses to other T cell epitopes got lost over time.

STUDIES ADDRESSING TCR REPERTOIRES

Several studies have also investigated TCR repertoires and attempted to establish a link between epitope recognition and particular TCR sequences. More specifically, a seminal study by Gittelman et al. (Gittelman et al., 2021) obtained TCR sequence information from the entire municipality of Vò (Italy) during the initial surge of SARS-CoV-2 infections and detected notable correlations with disease severity and other characteristics. Snyder et al. (Snyder et al., 2020) expanded these findings by inferring several epitopes that may be able to be recognized by specific TCRs. They also built a classifier to diagnose infection based solely on TCR sequencing from blood samples. Along the same lines, Shomuradova et al. (Shomuradova et al., 2020) observed specific TCR motifs in the subjects they analyzed, in some cases shared across multiple donors, and Ferretti et al. (Ferretti et al., 2020) sorted epitope-specific T cells and used single-cell sequencing to define paired TCR α and TCR β chains expressed by these T cells. Gangaev et al. have also reported TCR sequences that recognize a defined SARS-CoV-2 epitope (Gangaev et al., 2020).

In conclusion, given the large number of different epitopes recognized in the context of a myriad of different HLA types, it will be necessary to compile an extensive catalog of TCR sequences to completely capture the TCR repertoire associated with SARS-CoV-2 responses in humans. Early reports indicate that the study of TCR repertoires might lead to interesting diagnostic applications and could yield additional insights into the pathogenesis of SARS-CoV-2, particularly given the recent Emergency Use Authorization of a TCR-based diagnostic developed by Adaptive Biotech (see: https://www.fda.gov/media/146478/download).

CONCLUSIONS

Here we reviewed 25 different studies describing the identification of over 1,400 different unique epitopes (382 for CD4 and 1052 for CD8) SARS-CoV-2 epitopes recognized by human T cells, herein annotated in terms of available metadata. This review highlights several key findings and also raises outstanding questions for future SARS-COV-2 research to address.

First, the epitope data described here derives in aggregate, from studies with 1,197 human subjects (870 COVID-19 and 327 unexposed controls). These cohorts represent considerable heterogeneity as a function of age, gender, disease severity (with severe disease less represented), and time since symptoms onset. However, different ethnicities were not broadly represented; this will be an important knowledge gap to be addressed in future investigations.

Second, and related to the above issue, HLA-restricted epitopes were identified for 30 class I and 45 class II molecules, which provides good coverage of a number of different loci and alleles. However, while the median number of epitopes per allele is 15, it ranged from 1 to 219, with a large bias toward the HLA alleles that are more frequently encountered in the general population.

Third, we note that while twenty studies defined class I/CD8 epitopes, only 9 defined class II/CD4 epitopes. Also, given the prominent role of CD4 responses in immune responses to SARS CoV2 in the context of natural infection and vaccination, this observation suggests that a more balanced study of both CD4 and CD8 epitopes remains an outstanding issue for future research.

Fourth, in terms of the antigens targeted by epitope identification studies, ten studies screened peptides derived from the entire proteome but fifteen studies concentrated on specific subsets of antigens, mostly based on the fact that the main SARS-CoV-2 T cell antigenic targets have been independently defined utilizing pools of overlapping peptides. Structural proteins (S, M, and N) are dominant targets of T cell responses, but ORF3, ORF8, nsp3, nsp4, and nsp12 are also frequently targeted. Within the main antigens, immunodominant regions are typically pronounced in the case of CD4 recognition but less so in the case of CD8 responses, which are more evenly distributed across the dominant antigens. The precise identification of immunodominant antigens and regions is of interest also for its potential in the context of the identification of immunogenic regions of the SARS CoV2 proteome, conserved in different coronavirus species of pandemic potential.

Finally, the fact that already more than 1,400 epitopes have been identified—also considering that many HLA alleles and regions of the SARS CoV2 proteome are relatively less studied—highlights that a large breadth of epitopes are recognized in human populations, making it unlikely that SARS CoV2 variants might escape T cell recognition at the population level.

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.chom.2021.05.010.

ACKNOWLEDGMENTS

This work has been supported by NIH contracts 75N93019C00001(A.S. and B.P.) and 75N9301900065 (A.S. and D.W.).

AUTHOR CONTRIBUTIONS

Conceptualization, D.W. and A.S.; Investigation, D.W., A.G., J.S., and A.S.; Resources, R.V., B.P., and A.S.; Data curation, D.W., A.G., J.S., R.V., and A.S.; Writing–original draft, D.W., J.S., and A.S.; Writing–review & editing, D.W., A.G., J.S., R.V., B.P., S.C., and A.S.; Data visualization, D.W., A.G., and J.S.; Supervision, D.W. and A.S.; Funding acquisition, D.W., B.P., and A.S.

DECLARATION OF INTERESTS

A.S. is a consultant for Gritstone, Flow Pharma, Merck, Epitogenesis, Gilead, and Avalia. S.C. is a consultant for Avalia. L.J.I. has filed for patent protection for various aspects of T cell epitope and vaccine design work.

REFERENCES

Abbas, A.K., Lichtman, A.H., and Pillai, S. (2007). Cellular and Molecular Immunology (Saunders Elsevier).

Agerer, B., Koblischke, M., Gudipati, V., Montaño-Gutierrez, L.F., Smyth, M., Popa, A., Genger, J.W., Endler, L., Florian, D.M., Mühlgrabner, V., et al. (2021). SARS-CoV-2 mutations in MHC-I-restricted epitopes evade CD8$^+$ T cell responses. Sci. Immunol. *6*, 6.

Altmann, D.M., and Boyton, R.J. (2020). SARS-CoV-2 T cell immunity: Specificity, function, durability, and role in protection. Sci. Immunol. *5*, 5.

Assarsson, E., Sidney, J., Oseroff, C., Pasquetto, V., Bui, H.H., Frahm, N., Brander, C., Peters, B., Grey, H., and Sette, A. (2007). A quantitative analysis of the variables affecting the repertoire of T cell specificities recognized after vaccinia virus infection. J. Immunol. *178*, 7890–7901.

Avanzato, V.A., Matson, M.J., Seifert, S.N., Pryce, R., Williamson, B.N., Anzick, S.L., Barbian, K., Judson, S.D., Fischer, E.R., Martens, C., et al. (2020). Case Study: Prolonged Infectious SARS-CoV-2 Shedding from an Asymptomatic Immunocompromised Individual with Cancer. Cell *183*, 1901–1912.e9.

Bacher, P., Rosati, E., Esser, D., Martini, G.R., Saggau, C., Schiminsky, E., Dargvainiene, J., Schröder, I., Wieters, I., Khodamoradi, Y., et al. (2020). Low-Avidity CD4$^+$ T Cell Responses to SARS-CoV-2 in Unexposed Individuals and Humans with Severe COVID-19. Immunity *53*, 1258–1271C.e5.

Bilich, T., Nelde, A., Heitmann, J.S., Maringer, Y., Roerden, M., Bauer, J., Rieth, J., Wacker, M., Peter, A., Hörber, S., et al. (2021). T cell and antibody kinetics delineate SARS-CoV-2 peptides mediating long-term immune responses in COVID-19 convalescent individuals. Sci. Transl. Med. *13*, eabf7517.

Chen, Z., Ruan, P., Wang, L., Nie, X., Ma, X., and Tan, Y. (2021). T and B cell Epitope analysis of SARS-CoV-2 S protein based on immunoinformatics and experimental research. J. Cell. Mol. Med. *25*, 1274–1289.

Croft, N.P., Smith, S.A., Pickering, J., Sidney, J., Peters, B., Faridi, P., Witney, M.J., Sebastian, P., Flesch, I.E.A., Heading, S.L., et al. (2019). Most viral peptides displayed by class I MHC on infected cells are immunogenic. Proc. Natl. Acad. Sci. USA *116*, 3112–3117.

Dan, J.M., Lindestam Arlehamn, C.S., Weiskopf, D., da Silva Antunes, R., Havenar-Daughton, C., Reiss, S.M., Brigger, M., Bothwell, M., Sette, A., and Crotty, S. (2016). A Cytokine-Independent Approach To Identify Antigen-Specific Human Germinal Center T Follicular Helper Cells and Rare Antigen-Specific CD4+ T Cells in Blood. J. Immunol. *197*, 983–993.

Dash, P., Fiore-Gartland, A.J., Hertz, T., Wang, G.C., Sharma, S., Souquette, A., Crawford, J.C., Clemens, E.B., Nguyen, T.H.O., Kedzierska, K., et al. (2017). Quantifiable predictive features define epitope-specific T cell receptor repertoires. Nature *547*, 89–93.

Demotz, S., Grey, H.M., and Sette, A. (1990). The minimal number of class II MHC-antigen complexes needed for T cell activation. Science *249*, 1028–1030.

Dhanda, S.K., Vaughan, K., Schulten, V., Grifoni, A., Weiskopf, D., Sidney, J., Peters, B., and Sette, A. (2018a). Development of a novel clustering tool for linear peptide sequences. Immunology *155*, 331–345.

Dhanda, S.K., Vita, R., Ha, B., Grifoni, A., Peters, B., and Sette, A. (2018b). ImmunomeBrowser: a tool to aggregate and visualize complex and heterogeneous epitopes in reference proteins. Bioinformatics *34*, 3931–3933.

DiPiazza, A.T., Graham, B.S., and Ruckwardt, T.J. (2020). T cell immunity to SARS-CoV-2 following natural infection and vaccination. Biochem. Biophys. Res. Commun. *538*, 211–217.

Ferretti, A.P., Kula, T., Wang, Y., Nguyen, D.M.V., Weinheimer, A., Dunlap, G.S., Xu, Q., Nabilsi, N., Perullo, C.R., Cristofaro, A.W., et al. (2020). Unbiased Screens Show CD8+ T Cells of COVID-19 Patients Recognize Shared Epitopes in SARS-CoV-2 that Largely Reside outside the Spike Protein. Immunity *53*, 1095–1107.e3.

Fleri, W., Vaughan, K., Salimi, N., Vita, R., Peters, B., and Sette, A. (2017). The Immune Epitope Database: How Data Are Entered and Retrieved. J. Immunol. Res. *2017*, 5974574.

Gangaev, A., Ketelaars, S.L.C., Isaeva, O.I., Patiwael, S., Dopler, A., Hoefakker, K.D.E., Biasi, S., Gibellini, L., Mussini, C., Guaraldi, G., et al. (2020). Identification and characterization of an immunodominant SARS-CoV-2-specific CD8 T cell response. J. Immunother. Cancer *8*, A513–A514.

Gfeller, D., Guillaume, P., Michaux, J., Pak, H.S., Daniel, R.T., Racle, J., Coukos, G., and Bassani-Sternberg, M. (2018). The Length Distribution and Multiple Specificity of Naturally Presented HLA-I Ligands. J. Immunol. *201*, 3705–3716.

Gittelman, R.M., Lavezzo, E., Snyder, T.M., Zahid, H.J., Elyanow, R., Dalai, S., Kirsch, I., Baldo, L., Manuto, L., Franchin, E., et al. (2021). Diagnosis and Tracking of SARS-CoV-2 Infection By T-Cell Receptor Sequencing. medRxiv. https://doi.org/10.1101/2020.11.09.20228023.

Glanville, J., Huang, H., Nau, A., Hatton, O., Wagar, L.E., Rubelt, F., Ji, X., Han, A., Krams, S.M., Pettus, C., et al. (2017). Identifying specificity groups in the T cell receptor repertoire. Nature *547*, 94–98.

Gonzalez-Galarza, F.F., McCabe, A., Santos, E.J.M.D., Jones, J., Takeshita, L., Ortega-Rivera, N.D., Cid-Pavon, G.M.D., Ramsbottom, K., Ghattaoraya, G., Alfirevic, A., et al. (2020). Allele frequency net database (AFND) 2020 update: gold-standard data classification, open access genotype data and new query tools. Nucleic Acids Res. *48*, D783–D788.

Goulder, P.J., and Walker, B.D. (2012). HIV and HLA class I: an evolving relationship. Immunity *37*, 426–440.

Greenbaum, J., Sidney, J., Chung, J., Brander, C., Peters, B., and Sette, A. (2011). Functional classification of class II human leukocyte antigen (HLA) molecules reveals seven different supertypes and a surprising degree of repertoire sharing across supertypes. Immunogenetics *63*, 325–335.

Grifoni, A., Weiskopf, D., Ramirez, S.I., Mateus, J., Dan, J.M., Moderbacher, C.R., Rawlings, S.A., Sutherland, A., Premkumar, L., Jadi, R.S., et al. (2020). Targets of T Cell Responses to SARS-CoV-2 Coronavirus in Humans with COVID-19 Disease and Unexposed Individuals. Cell *181*, 1489–1501.e15.

Habel, J.R., Nguyen, T.H.O., van de Sandt, C.E., Juno, J.A., Chaurasia, P., Wragg, K., Koutsakos, M., Hensen, L., Jia, X., Chua, B., et al. (2020). Suboptimal SARS-CoV-2-specific CD8+ T cell response associated with the prominent HLA-A*02:01 phenotype. Proc. Natl. Acad. Sci. USA *117*, 24384–24391.

Hassert, M., Geerling, E., Stone, E.T., Steffen, T.L., Feldman, M.S., Dickson, A.L., Class, J., Richner, J.M., Brien, J.D., and Pinto, A.K. (2020). mRNA induced expression of human angiotensin-converting enzyme 2 in mice for the study of the adaptive immune response to severe acute respiratory syndrome coronavirus 2. PLoS Pathog. *16*, e1009163.

Hataye, J., Moon, J.J., Khoruts, A., Reilly, C., and Jenkins, M.K. (2006). Naive and memory CD4+ T cell survival controlled by clonal abundance. Science *312*, 114–116.

Joag, V., Wijeyesinghe, S., Stolley, J.M., Quarnstrom, C.F., Dileepan, T., Soerens, A.G., Sangala, J.A., O'Flanagan, S.D., Gavil, N.V., Hong, S.W., et al. (2021). Cutting Edge: Mouse SARS-CoV-2 Epitope Reveals Infection and Vaccine-Elicited CD8 T Cell Responses. J Immunol. *206*, 931–935.

Jordan, S.C. (2021). Innate and adaptive immune responses to SARS-CoV-2 in humans: relevance to acquired immunity and vaccine responses. Clin. Exp. Immunol. *204*, 310–320.

Kared, H., Redd, A.D., Bloch, E.M., Bonny, T.S., Sumatoh, H., Kairi, F., Carbajo, D., Abel, B., Newell, E.W., Bettinotti, M.P., et al. (2021). SARS-CoV-2-specific CD8+ T cell responses in convalescent COVID-19 individuals. J. Clin. Invest. *131*, 145476.

Karlsson, A.C., Humbert, M., and Buggert, M. (2020). The known unknowns of T cell immunity to COVID-19. Sci. Immunol. *5*, 5.

Keller, M.D., Harris, K.M., Jensen-Wachspress, M.A., Kankate, V., Lang, H., Lazarski, C.A., Durkee-Shock, J.R., Lee, P.H., Chaudhry, K., Webber, K., et al. (2020). SARS-CoV-2 specific T-cells Are Rapidly Expanded for Therapeutic Use and Target Conserved Regions of Membrane Protein. Blood. https://doi.org/10.1182/blood.2020008488.

Kilgus, J., Jardetzky, T., Gorga, J.C., Trzeciak, A., Gillessen, D., and Sinigaglia, F. (1991). Analysis of the permissive association of a malaria T cell epitope with DR molecules. J. Immunol. *146*, 307–315.

Knierman, M.D., Lannan, M.B., Spindler, L.J., McMillian, C.L., Konrad, R.J., and Siegel, R.W. (2020). The Human Leukocyte Antigen Class II Immunopeptidome of the SARS-CoV-2 Spike Glycoprotein. Cell Rep. *33*, 108454.

Kotturi, M.F., Peters, B., Buendia-Laysa, F., Jr., Sidney, J., Oseroff, C., Botten, J., Grey, H., Buchmeier, M.J., and Sette, A. (2007). The CD8+ T-cell response to lymphocytic choriomeningitis virus involves the L antigen: uncovering new tricks for an old virus. J. Virol. *81*, 4928–4940.

Kotturi, M.F., Scott, I., Wolfe, T., Peters, B., Sidney, J., Cheroutre, H., von Herrath, M.G., Buchmeier, M.J., Grey, H., and Sette, A. (2008). Naive precursor frequencies and MHC binding rather than the degree of epitope diversity shape CD8+ T cell immunodominance. J. Immunol. *181*, 2124–2133.

Le Bert, N., Tan, A.T., Kunasegaran, K., Tham, C.Y.L., Hafezi, M., Chia, A., Chng, M.H.Y., Lin, M., Tan, N., Linster, M., et al. (2020). SARS-CoV-2-specific T cell immunity in cases of COVID-19 and SARS, and uninfected controls. Nature *584*, 457–462.

Le Bert, N., Clapham, H.E., Tan, A.T., Chia, W.N., Tham, C.Y.L., Lim, J.M., Kunasegaran, K., Tan, L.W.L., Dutertre, C.A., Shankar, N., et al. (2021). Highly functional virus-specific cellular immune response in asymptomatic SARS-CoV-2 infection. J. Exp. Med. *218*, e20202617.

Lee, E., Sandgren, K., Duette, G., Stylianou, V.V., Khanna, R., Eden, J.S., Blyth, E., Gottlieb, D., Cunningham, A.L., and Palmer, S. (2020). Identification of SARS-CoV-2 Nucleocapsid and Spike T-cell Epitopes for Assessing T-cell Immunity. J Virol.

Lipsitch, M., Grad, Y.H., Sette, A., and Crotty, S. (2020). Cross-reactive memory T cells and herd immunity to SARS-CoV-2. Nat. Rev. Immunol. *20*, 709–713.

Locci, M., Havenar-Daughton, C., Landais, E., Wu, J., Kroenke, M.A., Arlehamn, C.L., Su, L.F., Cubas, R., Davis, M.M., International AIDS Vaccine Initiative Protocol C Principal Investigators, et al. (2013). Human circulating PD-1+CXCR3-CXCR5+ memory Tfh cells are highly functional and correlate with broadly neutralizing HIV antibody responses. Immunity *39*, 758–769.

Mahajan, S., Kode, V., Bhojak, K., Magdalene, C.M., Lee, K., Manoharan, M., Ramesh, A., Sudheendra, H., Srivastava, A., Sathian, R., et al. (2020). Immunodominant T-cell epitopes from the SARS-CoV-2 spike antigen reveal robust pre-existing T-cell immunity in unexposed individuals. bioRxiv, 2020.11.03.367375.

Masopust, D., and Soerens, A.G. (2019). Tissue-Resident T Cells and Other Resident Leukocytes. Annu. Rev. Immunol. *37*, 521–546.

Mateus, J., Grifoni, A., Tarke, A., Sidney, J., Ramirez, S.I., Dan, J.M., Burger, Z.C., Rawlings, S.A., Smith, D.M., Phillips, E., et al. (2020). Selective and cross-reactive SARS-CoV-2 T cell epitopes in unexposed humans. Science *370*, 89–94.

McKinney, D.M., Southwood, S., Hinz, D., Oseroff, C., Arlehamn, C.S., Schulten, V., Taplitz, R., Broide, D., Hanekom, W.A., Scriba, T.J., et al. (2013). A strategy to determine HLA class II restriction broadly covering the DR, DP, and DQ allelic variants most commonly expressed in the general population. Immunogenetics *65*, 357–370.

McMichael, A.J., Gotch, F.M., Santos-Aguado, J., and Strominger, J.L. (1988). Effect of mutations and variations of HLA-A2 on recognition of a virus peptide epitope by cytotoxic T lymphocytes. Proc. Natl. Acad. Sci. USA *85*, 9194–9198.

Murphey, K., Travers, P., Walport, M., and Janeway, C. (2012). Janeway's Immunobiology (Garland Science).

Nelde, A., Bilich, T., Heitmann, J.S., Maringer, Y., Salih, H.R., Roerden, M., Lübke, M., Bauer, J., Rieth, J., Wacker, M., et al. (2021). SARS-CoV-2-derived peptides define heterologous and COVID-19-induced T cell recognition. Nat. Immunol. *22*, 74–85.

Nielsen, S.S., Vibholm, L.K., Monrad, I., Olesen, R., Frattari, G.S., Pahus, M.H., Højen, J.F., Gunst, J.D., Erikstrup, C., Holleufer, A., et al. (2020). SARS-CoV-2 Elicits Robust Adaptive Immune Responses Regardless of Disease Severity. bioRxiv, 2020.10.08.331645.

O'Brien, C., Flower, D.R., and Feighery, C. (2008). Peptide length significantly influences in vitro affinity for MHC class II molecules. Immunome Res. *4*, 6.

O'Sullivan, D., Arrhenius, T., Sidney, J., Del Guercio, M.F., Albertson, M., Wall, M., Oseroff, C., Southwood, S., Colón, S.M., Gaeta, F.C., et al. (1991). On the interaction of promiscuous antigenic peptides with different DR alleles. Identification of common structural motifs. J. Immunol. *147*, 2663–2669.

Oseroff, C., Peters, B., Pasquetto, V., Moutaftsi, M., Sidney, J., Panchanathan, V., Tscharke, D.C., Maillere, B., Grey, H., and Sette, A. (2008). Dissociation between epitope hierarchy and immunoprevalence in CD8 responses to vaccinia virus western reserve. J. Immunol. *180*, 7193–7202.

Panina-Bordignon, P., Tan, A., Termijtelen, A., Demotz, S., Corradin, G., and Lanzavecchia, A. (1989). Universally immunogenic T cell epitopes: promiscuous binding to human MHC class II and promiscuous recognition by T cells. Eur. J. Immunol. *19*, 2237–2242.

Parker, R., Partridge, T., Wormald, C., Kawahara, R., Stalls, V., Aggelakopoulou, M., Parker, J., Doherty, R.P., Morejon, Y.A., Lee, E., et al. (2020). Mapping the SARS-CoV-2 spike glycoprotein-derived peptidome presented by HLA class II on dendritic cells. bioRxiv, 2020.08.19.255901.

Paul, S., Karosiene, E., Dhanda, S.K., Jurtz, V., Edwards, L., Nielsen, M., Sette, A., and Peters, B. (2018). Determination of a Predictive Cleavage Motif for Eluted Major Histocompatibility Complex Class II Ligands. Front. Immunol. *9*, 1795.

Paul, S., Croft, N.P., Purcell, A.W., Tscharke, D.C., Sette, A., Nielsen, M., and Peters, B. (2020). Benchmarking predictions of MHC class I restricted T cell epitopes in a comprehensively studied model system. PLoS Comput. Biol. *16*, e1007757.

Peng, Y., Mentzer, A.J., Liu, G., Yao, X., Yin, Z., Dong, D., Dejnirattisai, W., Rostron, T., Supasa, P., Liu, C., et al.; Oxford Immunology Network Covid-19 Response T cell Consortium; ISARIC4C Investigators (2020). Broad and strong memory CD4+ and CD8+ T cells induced by SARS-CoV-2 in UK convalescent individuals following COVID-19. Nat. Immunol. *21*, 1336–1345.

Peters, B., Nielsen, M., and Sette, A. (2020). T Cell Epitope Predictions. Annu. Rev. Immunol. *38*, 123–145.

Petersen, J., Purcell, A.W., and Rossjohn, J. (2009). Post-translationally modified T cell epitopes: immune recognition and immunotherapy. J. Mol. Med. (Berl.) *87*, 1045–1051.

Poran, A., Harjanto, D., Malloy, M., Arieta, C.M., Rothenberg, D.A., Lenkala, D., van Buuren, M.M., Addona, T.A., Rooney, M.S., Srinivasan, L., and Gaynor, R.B. (2020a). Sequence-based prediction of SARS-CoV-2 vaccine targets using a mass spectrometry-based bioinformatics predictor identifies immunogenic T cell epitopes. Genome Med. *12*, 70.

Poran, A., Harjanto, D., Malloy, M., Rooney, M.S., Srinivasan, L., and Gaynor, R.B. (2020b). Sequence-based prediction of vaccine targets for inducing T cell responses to SARS-CoV-2 utilizing the bioinformatics predictor RECON. bioRxiv, 2020.04.06.027805.

Prachar, M., Justesen, S., Steen-Jensen, D.B., Thorgrimsen, S., Jurgons, E., Winther, O., and Bagger, F.O. (2020). Identification and validation of 174 COVID-19 vaccine candidate epitopes reveals low performance of common epitope prediction tools. Sci. Rep. *10*, 20465.

Prakash, S., Srivastava, R., Coulon, P.-G., Dhanushkodi, N.R., Chentoufi, A.A., Tifrea, D.F., Edwards, R.A., Figueroa, C.J., Schubl, S.D., Hsieh, L., et al. (2020). Genome-Wide Asymptomatic B-Cell, CD4+ and CD8+ T-Cell Epitopes, that are Highly Conserved Between Human and Animal Coronaviruses, Identified from SARS-CoV-2 as Immune Targets for Pre-Emptive Pan-Coronavirus Vaccines. bioRxiv,, 2020.09.27.316018.

Redd, A.D., Nardin, A., Kared, H., Bloch, E.M., Pekosz, A., Laeyendecker, O., Abel, B., Fehlings, M., Quinn, T.C., and Tobian, A.A. (2021). CD8+ T cell responses in COVID-19 convalescent individuals target conserved epitopes from multiple prominent SARS-CoV-2 circulating variants. medRxiv.

Reiss, S., Baxter, A.E., Cirelli, K.M., Dan, J.M., Morou, A., Daigneault, A., Brassard, N., Silvestri, G., Routy, J.P., Havenar-Daughton, C., et al. (2017). Comparative analysis of activation induced marker (AIM) assays for sensitive identification of antigen-specific CD4 T cells. PLoS One *12*, e0186998.

Rha, M.S., Jeong, H.W., Ko, J.H., Choi, S.J., Seo, I.H., Lee, J.S., Sa, M., Kim, A.R., Joo, E.J., Ahn, J.Y., et al. (2021). PD-1-Expressing SARS-CoV-2-Specific CD8$^+$ T Cells Are Not Exhausted, but Functional in Patients with COVID-19. Immunity *54*, 44–52.e3.

Riddell, S.R., and Greenberg, P.D. (1995). Principles for adoptive T cell therapy of human viral diseases. Annu. Rev. Immunol. *13*, 545–586.

Rimmelzwaan, G.F., Berkhoff, E.G.M., Nieuwkoop, N.J., Smith, D.J., Fouchier, R.A.M., and Osterhaus, A.D.M.E. (2005). Full restoration of viral fitness by multiple compensatory co-mutations in the nucleoprotein of influenza A virus cytotoxic T-lymphocyte escape mutants. J. Gen. Virol. *86*, 1801–1805.

Robinson, J., Barker, D.J., Georgiou, X., Cooper, M.A., Flicek, P., and Marsh, S.G.E. (2020). IPD-IMGT/HLA Database. Nucleic Acids Res. *48* (D1), D948–D955.

Sagar, M., Reifler, K., Rossi, M., Miller, N.S., Sinha, P., White, L.F., and Mizgerd, J.P. (2021). Recent endemic coronavirus infection is associated with less-severe COVID-19. J. Clin. Invest. *131*, 131.

Sahin, U., Muik, A., Vogler, I., Derhovanessian, E., Kranz, L.M., Vormehr, M., Quandt, J., Bidmon, N., Ulges, A., Baum, A., et al. (2020). BNT162b2 induces SARS-CoV-2-neutralising antibodies and T cells in humans. medRxiv.

Saini, S.K., Hersby, D.S., Tamhane, T., Povlsen, H.R., Hernandez, S.P.A., Nielsen, M., Gang, A.O., and Hadrup, S.R. (2020). SARS-CoV-2 genome-wide mapping of CD8 T cell recognition reveals strong immunodominance and substantial CD8 T cell activation in COVID-19 patients. bioRxiv, 2020.10.19.344911.

Saini, S.K., Hersby, D.S., Tamhane, T., Povlsen, H.R., Amaya Hernandez, S.P., Nielsen, M., Gang, A.O., and Hadrup, S.R. (2021). SARS-CoV-2 genome-wide T cell epitope mapping reveals immunodominance and substantial CD8$^+$ T cell activation in COVID-19 patients. Sci. Immunol. *6*, 6.

Salimi, N., Fleri, W., Peters, B., and Sette, A. (2012). The immune epitope database: a historical retrospective of the first decade. Immunology *137*, 117–123.

Schulien, I., Kemming, J., Oberhardt, V., Wild, K., Seidel, L.M., Killmer, S., Sagar, F.D., Daul, F., Salvat Lago, M., Decker, A., et al. (2021). Characterization of pre-existing and induced SARS-CoV-2-specific CD8$^+$ T cells. Nat. Med. *27*, 78–85.

Sekine, T., Perez-Potti, A., Rivera-Ballesteros, O., Strålin, K., Gorin, J.B., Olsson, A., Llewellyn-Lacey, S., Kamal, H., Bogdanovic, G., Muschiol, S., et al.; Karolinska COVID-19 Study Group (2020). Robust T Cell Immunity in Convalescent Individuals with Asymptomatic or Mild COVID-19. Cell *183*, 158–168.e14.

Sette, A., and Crotty, S. (2020a). Author Correction: Pre-existing immunity to SARS-CoV-2: the knowns and unknowns. Nat. Rev. Immunol. *20*, 644.

Sette, A., and Crotty, S. (2020b). Pre-existing immunity to SARS-CoV-2: the knowns and unknowns. Nat. Rev. Immunol. *20*, 457–458.

Sette, A., and Crotty, S. (2021). Adaptive immunity to SARS-CoV-2 and COVID-19. Cell.

Sette, A., and Sidney, J. (1999). Nine major HLA class I supertypes account for the vast preponderance of HLA-A and -B polymorphism. Immunogenetics *50*, 201–212.

Shaan Lakshmanappa, Y., Elizaldi, S.R., Roh, J.W., Schmidt, B.A., Carroll, T.D., Weaver, K.D., Smith, J.C., Verma, A., Deere, J.D., Dutra, J., et al. (2021). SARS-CoV-2 induces robust germinal center CD4 T follicular helper cell responses in rhesus macaques. Nat. Commun. *12*, 541.

Shomuradova, A.S., Vagida, M.S., Sheetikov, S.A., Zornikova, K.V., Kiryukhin, D., Titov, A., Peshkova, I.O., Khmelevskaya, A., Dianov, D.V., Malasheva, M., et al. (2020). SARS-CoV-2 Epitopes Are Recognized by a Public and Diverse Repertoire of Human T Cell Receptors. Immunity *53*, 1245–1257.e5.

Sidney, J., Peters, B., Frahm, N., Brander, C., and Sette, A. (2008). HLA class I supertypes: a revised and updated classification. BMC Immunol. *9*, 1.

Sidney, J., Peters, B., and Sette, A. (2020). Epitope prediction and identification- adaptive T cell responses in humans. Semin. Immunol. *50*, 101418.

Snyder, T.M., Gittelman, R.M., Klinger, M., May, D.H., Osborne, E.J., Taniguchi, R., Zahid, H.J., Kaplan, I.M., Dines, J.N., Noakes, M.N., et al. (2020). Magnitude and Dynamics of the T-Cell Response to SARS-CoV-2 Infection at Both Individual and Population Levels. medRxiv, 2020.07.31.20165647.

Sun, L., Middleton, D.R., Wantuch, P.L., Ozdilek, A., and Avci, F.Y. (2016). Carbohydrates as T-cell antigens with implications in health and disease. Glycobiology *26*, 1029–1040.

Swadling, L., and Maini, M.K. (2020). T cells in COVID-19 - united in diversity. Nat. Immunol. *21*, 1307–1308.

Sykulev, Y., Joo, M., Vturina, I., Tsomides, T.J., and Eisen, H.N. (1996). Evidence that a single peptide-MHC complex on a target cell can elicit a cytolytic T cell response. Immunity *4*, 565–571.

Takagi, A., and Matsui, M. (2021). Identification of HLA-A*02:01-Restricted Candidate Epitopes Derived from the Nonstructural Polyprotein 1a of SARS-CoV-2 That May Be Natural Targets of CD8$^+$ T Cell Recognition In Vivo. J. Virol. *95*, e01837-e20.

Tan, A.T., Sodsai, P., Chia, A., Moreau, E., Chng, M.H., Tham, C.Y., Ho, Z.Z., Banu, N., Hirankarn, N., and Bertoletti, A. (2014). Immunoprevalence and immunodominance of HLA-Cw*0801-restricted T cell response targeting the hepatitis B virus envelope transmembrane region. J. Virol. *88*, 1332–1341.

Tarke, A., Sidney, J., Kidd, C.K., Dan, J.M., Ramirez, S.I., Yu, E.D., Mateus, J., Da Silva Antunes, R., Moore, E., Rubiro, P., Methot, N., et al. (2021a). Comprehensive analysis of T cell immunodominance and immunoprevalence of SARS-CoV-2 epitopes in COVID-19 cases. Cell Rep Med *2*, 100204.

Tarke, A., Sidney, J., Methot, N., Zhang, Y., Dan, J.M., Goodwin, B., Rubiro, P., Sutherland, A., da Silva Antunes, R., Frazier, A., et al. (2021b). Negligible impact of SARS-CoV-2 variants on CD4 + and CD8 + T cell reactivity in COVID-19 exposed donors and vaccinees. bioRxiv. https://doi.org/10.1101/2021.02.27.433180.

Tegally, H., Wilkinson, E., Giovanetti, M., Iranzadeh, A., Fonseca, V., Giandhari, J., Doolabh, D., Pillay, S., San, E.J., Msomi, N., et al. (2021). Emergence of a SARS-CoV-2 variant of concern with mutations in spike glycoprotein. Nature *592*. https://doi.org/10.1038/s41586-021-03402-9.

Thomson, E.C., Rosen, L.E., Shepherd, J.G., Spreafico, R., da Silva Filipe, A., Wojcechowskyj, J.A., Davis, C., Piccoli, L., Pascall, D.J., Dillen, J., et al.; ISARIC4C Investigators; COVID-19 Genomics UK (COG-UK) Consortium (2021). Circulating SARS-CoV-2 spike N439K variants maintain fitness while evading antibody-mediated immunity. Cell *184*, 1171–1187.e20.

Trolle, T., McMurtrey, C.P., Sidney, J., Bardet, W., Osborn, S.C., Kaever, T., Sette, A., Hildebrand, W.H., Nielsen, M., and Peters, B. (2016). The Length Distribution of Class I-Restricted T Cell Epitopes Is Determined by Both Peptide Supply and MHC Allele-Specific Binding Preference. J. Immunol. *196*, 1480–1487.

Vita, R., Mahajan, S., Overton, J.A., Dhanda, S.K., Martini, S., Cantrell, J.R., Wheeler, D.K., Sette, A., and Peters, B. (2019). The Immune Epitope Database (IEDB): 2018 update. Nucleic Acids Res. *47*, D339–D343.

Wang, P., Sidney, J., Dow, C., Mothé, B., Sette, A., and Peters, B. (2008a). A systematic assessment of MHC class II peptide binding predictions and evaluation of a consensus approach. PLoS Comput. Biol. *4*, e1000048.

Wang, X.F., Kerzerho, J., Adotevi, O., Nuyttens, H., Badoual, C., Munier, G., Oudard, S., Tu, S., Tartour, E., and Maillère, B. (2008b). Comprehensive analysis of HLA-DR- and HLA-DP4-restricted CD4+ T cell response specific for the tumor-shared antigen survivin in healthy donors and cancer patients. J. Immunol. *181*, 431–439.

Wang, P., Nair, M.S., Liu, L., Iketani, S., Luo, Y., Guo, Y., Wang, M., Yu, J., Zhang, B., Kwong, P.D., et al. (2021). Antibody Resistance of SARS-CoV-2 Variants B.1.351 and B.1.1.7. Nature *593*, 130–135.

Weingarten-Gabbay, S., Klaeger, S., Sarkizova, S., Pearlman, L.R., Chen, D.Y., Bauer, M.R., Taylor, H.B., Conway, H.L., Tomkins-Tinch, C.H., Finkel, Y., et al. (2020). SARS-CoV-2 infected cells present HLA-I peptides from canonical and out-of-frame ORFs. bioRxiv, 2020.10.02.324145.

Weiskopf, D., Angelo, M.A., de Azeredo, E.L., Sidney, J., Greenbaum, J.A., Fernando, A.N., Broadwater, A., Kolla, R.V., De Silva, A.D., de Silva, A.M., et al. (2013). Comprehensive analysis of dengue virus-specific responses supports an HLA-linked protective role for CD8+ T cells. Proc. Natl. Acad. Sci. USA *110*, E2046–E2053.

Yewdell, J.W. (2006). Confronting complexity: real-world immunodominance in antiviral CD8+ T cell responses. Immunity *25*, 533–543.

Yewdell, J.W., and Bennink, J.R. (1999). Immunodominance in major histocompatibility complex class I-restricted T lymphocyte responses. Annu. Rev. Immunol. *17*, 51–88.

RESEARCH ARTICLE SUMMARY

CORONAVIRUS

Immunological memory to SARS-CoV-2 assessed for up to 8 months after infection

Jennifer M. Dan*, Jose Mateus*, Yu Kato*, Kathryn M. Hastie, Esther Dawen Yu, Caterina E. Faliti, Alba Grifoni, Sydney I. Ramirez, Sonya Haupt, April Frazier, Catherine Nakao, Vamseedhar Rayaprolu, Stephen A. Rawlings, Bjoern Peters, Florian Krammer, Viviana Simon, Erica Ollmann Saphire, Davey M. Smith, Daniela Weiskopf†, Alessandro Sette†, Shane Crotty†

INTRODUCTION: Immunological memory is the basis for durable protective immunity after infections or vaccinations. Duration of immunological memory after severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection and COVID-19 is unclear. Immunological memory can consist of memory B cells, antibodies, memory $CD4^+$ T cells, and/or memory $CD8^+$ T cells. Knowledge of the kinetics and interrelationships among those four types of memory in humans is limited. Understanding immune memory to SARS-CoV-2 has implications for understanding protective immunity against COVID-19 and assessing the likely future course of the COVID-19 pandemic.

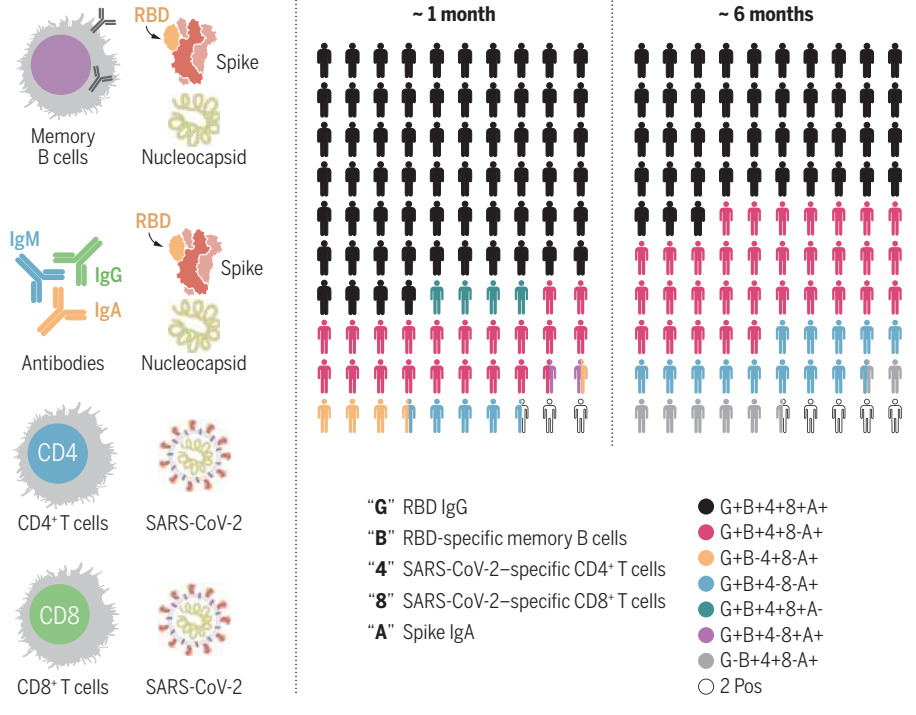
RATIONALE: Assessing virus-specific immune memory over at least a 6-month period is likely necessary to ascertain the durability of immune memory to SARS-CoV-2. Given the evidence that antibodies, $CD4^+$ T cells, and $CD8^+$ T cells can all participate in protective immunity to SARS-CoV-2, we measured antigen-specific antibodies, memory B cells, $CD4^+$ T cells, and $CD8^+$ T cells in the blood from subjects who recovered from COVID-19, up to 8 months after infection.

RESULTS: The study involved 254 samples from 188 COVID-19 cases, including 43 samples at 6 to 8 months after infection. Fifty-one subjects in the study provided longitudinal blood samples, allowing for both cross-sectional and longitudinal analyses of SARS-CoV-2–specific immune memory. Antibodies against SARS-CoV-2 spike and receptor binding domain (RBD) declined moderately over 8 months, comparable to several other reports. Memory B cells against SARS-CoV-2 spike actually increased between 1 month and 8 months after infection. Memory $CD8^+$ T cells and memory $CD4^+$ T cells declined with an initial half-life of 3 to 5 months. This is the largest antigen-specific study to date of the four major types of immune memory for any viral infection.

Among the antibody responses, spike immunoglobulin G (IgG), RBD IgG, and neutralizing antibody titers exhibited similar kinetics. Spike IgA was still present in the large majority of subjects at 6 to 8 months after infection. Among the memory B cell responses, IgG was the dominant isotype, with a minor population of IgA memory B cells. IgM memory B cells appeared to be short-lived. $CD8^+$ T cell and $CD4^+$ T cell memory was measured for all SARS-CoV-2 proteins. Although ~70% of individuals possessed detectable $CD8^+$ T cell memory at 1 month after infection, that proportion declined to ~50% by 6 to 8 months after infection. For $CD4^+$ T cell memory, 93% of subjects had detectable SARS-CoV-2 memory at 1 month after infection, and the proportion of subjects positive for $CD4^+$ T cells (92%) remained high at 6 to 8 months after infection. SARS-CoV-2 spike-specific memory $CD4^+$ T cells with the specialized capacity to help B cells [T follicular helper (T_{FH}) cells] were also maintained.

The different types of immune memory each had distinct kinetics, resulting in complex interrelationships between the abundance of T cell, B cell, and antibody immune memory over time. Additionally, substantially heterogeneity in memory to SARS-CoV-2 was observed.

CONCLUSION: Substantial immune memory is generated after COVID-19, involving all four major types of immune memory. About 95% of subjects retained immune memory at ~6 months after infection. Circulating antibody titers were not predictive of T cell memory. Thus, simple serological tests for SARS-CoV-2 antibodies do not reflect the richness and durability of immune memory to SARS-CoV-2. This work expands our understanding of immune memory in humans. These results have implications for protective immunity against SARS-CoV-2 and recurrent COVID-19. ■



Immunological memory consists of antibodies, memory B cells, memory $CD8^+$ T cells, and memory $CD4^+$ T cells. This study examined all of the types of virus-specific immune memory against SARS-CoV-2 in COVID-19 subjects. Robust immune memory was observed in most individuals.

The list of author affiliations is available in the full article online.
*These authors contributed equally to this work.
†Corresponding author. Email: shane@lji.org (S.C.); alex@lji.org (A.S.); daniela@lji.org (D.W.)

Cite this article as J. M. Dan *et al.*, *Science* **371**, eabf4063 (2021). DOI: 10.1126/science.abf4063

READ THE FULL ARTICLE AT
https://doi.org/10.1126/science.abf4063

RESEARCH ARTICLE

CORONAVIRUS

Immunological memory to SARS-CoV-2 assessed for up to 8 months after infection

Jennifer M. Dan[1,2]*, Jose Mateus[1]*, Yu Kato[1]*, Kathryn M. Hastie[1], Esther Dawen Yu[1], Caterina E. Faliti[1], Alba Grifoni[1], Sydney I. Ramirez[1,2], Sonya Haupt[1], April Frazier[1], Catherine Nakao[1], Vamseedhar Rayaprolu[1], Stephen A. Rawlings[2], Bjoern Peters[1,3], Florian Krammer[4], Viviana Simon[4,5,6], Erica Ollmann Saphire[1,2], Davey M. Smith[2], Daniela Weiskopf[1]†, Alessandro Sette[1,2]†, Shane Crotty[1,2]†

Understanding immune memory to severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is critical for improving diagnostics and vaccines and for assessing the likely future course of the COVID-19 pandemic. We analyzed multiple compartments of circulating immune memory to SARS-CoV-2 in 254 samples from 188 COVID-19 cases, including 43 samples at ≥6 months after infection. Immunoglobulin G (IgG) to the spike protein was relatively stable over 6+ months. Spike-specific memory B cells were more abundant at 6 months than at 1 month after symptom onset. SARS-CoV-2–specific $CD4^+$ T cells and $CD8^+$ T cells declined with a half-life of 3 to 5 months. By studying antibody, memory B cell, $CD4^+$ T cell, and $CD8^+$ T cell memory to SARS-CoV-2 in an integrated manner, we observed that each component of SARS-CoV-2 immune memory exhibited distinct kinetics.

Coronavirus disease 2019 (COVID-19), caused by the novel severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), is a serious disease that has resulted in widespread global morbidity and mortality. Humans make SARS-CoV-2–specific antibodies, $CD4^+$ T cells, and $CD8^+$ T cells in response to SARS-CoV-2 infection (*1*–*4*). Studies of acute and convalescent COVID-19 patients have observed that T cell responses are associated with reduced disease (*5*–*7*), suggesting that SARS-CoV-2–specific $CD4^+$ T cell and $CD8^+$ T cell responses may be important for control and resolution of primary SARS-CoV-2 infection. Ineffective innate immunity has been strongly associated with a lack of control of primary SARS-CoV-2 infection and a high risk of fatal COVID-19 (*8*–*12*), accompanied by innate cell immunopathology (*13*–*18*). Neutralizing antibodies have generally not correlated with lessened COVID-19 disease severity (*5*, *19*, *20*), which was also observed for Middle Eastern respiratory syndrome (MERS), caused by MERS-CoV (*21*). Instead, neutralizing antibodies are associated with protective immunity against secondary infection with SARS-CoV-2 or SARS-CoV in nonhuman primates (*3*, *22*–*25*). Passive transfer of neutralizing antibodies in advance of infection (mimicking preexisting conditions upon secondary exposure) effectively limits upper respiratory tract (URT) infection, lower respiratory tract (lung) infection, and symptomatic disease in animal models (*26*–*28*). Passive transfer of neutralizing antibodies provided after initiation of infection in humans has had more limited effects on COVID-19 (*29*, *30*), consistent with a substantial role for T cells in control and clearance of an ongoing SARS-CoV-2 infection. Thus, studying antibody, memory B cell, $CD4^+$ T cell, and $CD8^+$ T cell memory to SARS-CoV-2 in an integrated manner is likely important for understanding the durability of protective immunity against COVID-19 generated by primary SARS-CoV-2 infection (*1*, *19*, *31*).

Whereas sterilizing immunity against viruses can only be accomplished by high-titer neutralizing antibodies, successful protection against clinical disease or death can be accomplished by several other immune memory scenarios. Possible mechanisms of immunological protection can vary according to the relative kinetics of the immune memory responses and infection. For example, clinical hepatitis after hepatitis B virus (HBV) infection is prevented by vaccine-elicited immune memory even in the absence of circulating antibodies, because of the relatively slow course of HBV disease (*32*, *33*). The relatively slow course of severe COVID-19 in humans [median 19 days post–symptom onset (PSO) for fatal cases (*34*)] suggests that protective immunity against symptomatic or severe secondary COVID-19 may involve memory compartments such as circulating memory T cells and memory B cells (which can take several days to reactivate and generate recall T cell responses and/or anamnestic antibody responses) (*19*, *21*, *31*).

Immune memory, from either primary infection or immunization, is the source of protective immunity from a subsequent infection (*35*–*37*). Thus, COVID-19 vaccine development relies on immunological memory (*1*, *3*). Despite intensive study, the kinetics, duration, and evolution of immune memory in humans to infection or immunization are not in general predictable on the basis of the initial effector phase, and immune responses at short time points after resolution of infection are not very predictive of long-term memory (*38*–*40*). Thus, assessing responses over an interval of 6 months or more is usually required to ascertain the durability of immune memory.

A thorough understanding of immune memory to SARS-CoV-2 requires evaluation of its various components, including B cells, $CD8^+$ T cells, and $CD4^+$ T cells, as these different cell types may have immune memory kinetics that are relatively independent of each other. Understanding the complexities of immune memory to SARS-CoV-2 is key to gaining insights into the likelihood of durability of protective immunity against reinfection with SARS-CoV-2 and secondary COVID-19 disease. In this study, we assessed immune memory of all three branches of adaptive immunity ($CD4^+$ T cell, $CD8^+$ T cell, and humoral immunity) in a predominantly cross-sectional study of 188 recovered COVID-19 cases, extending up to 8 months after infection. The findings have implications for immunity against secondary COVID-19, and thus the potential future course of the pandemic (*41*, *42*).

COVID-19 cohort

We recruited 188 individuals with COVID-19 for this study. Subjects (80 male, 108 female) represented a range of asymptomatic, mild, moderate, and severe COVID-19 cases (Table 1) and were recruited from multiple sites throughout the United States. The majority of subjects were from California or New York. Most subjects had a "mild" case of COVID-19, not requiring hospitalization. Ninety-three percent of subjects were never hospitalized for COVID-19; 7% of subjects were hospitalized, some of whom required intensive care unit (ICU) care (Table 1). This case severity distribution was consistent with the general distribution of symptomatic disease severity among COVID-19 cases in the United States. The study primarily consisted of symptomatic disease cases (97%, Table 1), owing to the nature of the study recruitment design. Subject ages ranged from 19 to 81 years old (Table 1). Most subjects provided a blood

[1]Center for Infectious Disease and Vaccine Research, La Jolla Institute for Immunology (LJI), La Jolla, CA 92037, USA. [2]Department of Medicine, Division of Infectious Diseases and Global Public Health, University of California, San Diego (UCSD), La Jolla, CA 92037, USA. [3]Department of Medicine, University of California, San Diego (UCSD), La Jolla, CA 92037, USA. [4]Department of Microbiology, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA. [5]Division of Infectious Diseases, Department of Medicine, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA. [6]The Global Health and Emerging Pathogens Institute, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA.
*These authors contributed equally to this work.
†Corresponding author. Email: shane@lji.org (S.C.); alex@lji.org (A.S.); daniela@lji.org (D.W.)

Table 1. Participant characteristics.

	COVID-19 (*n* = 188)
Age (years)	19 to 81 (median = 40, IQR* = 18.75)
Gender	
Male (%)	43% (80/188)
Female (%)	57% (108/188)
Race	
African American or Black (%)	3% (5/188)
Alaskan Native or American Indian (%)	1% (1/188)
Asian (%)	7% (14/188)
Native Hawaiian or Pacific Islander (%)	0% (0/188)
Multiracial (%)	1% (2/188)
Other (%)	1% (1/188)
Unknown (%)	10% (19/188)
White (%)	78% (146/188)
Ethnicity	
Hispanic or Latino (%)	15% (28/188)
Non-Hispanic (%)	80% (150/188)
Unknown (%)	5% (10/188)
Hospitalization status	
Never hospitalized (%)	93% (174/188)
Hospitalized (%)	7% (13/188)
Unknown if hospitalized (%)	1% (1/188)
Sample collection dates	March-October 2020
SARS-CoV-2 PCR positivity	
Positive	77% (145/188)
Negative	1% (2/188)
Not performed	20% (37/188)
Unknown	2% (4/188)
Peak disease severity (%) [Female (F), Male (M)]	
Asymptomatic (score 1)	2% (4/188) (2F, 2M)
Mild (nonhospitalized; score 2–3)	90% (170/188) (100F, 70M)
Moderate (hospitalized; score 4–5)	3% (6/188) (3F, 3M)
Severe (hospitalized; Score 6+)	4% (7/188) (3F, 4M)
Unknown	1% (1/188) (0F, 1M)
Days post–symptom onset at collection; *n* = 254	6–240 (median 88, IQR 97.75)
Blood collection frequency	
Multiple time point Donors (two to four times)	27% (51/188)
Single–time point donors	73% (137/188)

*IQR, interquartile range.

sample at a single time point, between 6 and 240 days PSO (Table 1), with 43 samples at ≥6 months PSO (178 days or longer). Additionally, 51 subjects in the study provided longitudinal blood samples over a duration of several months (two to four time points; Table 1), allowing for longitudinal assessment of immune memory in a subset of the cohort.

SARS-CoV-2 circulating antibodies over time

The vast majority of SARS-CoV-2–infected individuals seroconvert, at least for a duration of months (*1*, *2*, *4*, *43*–*45*). Seroconversion rates range from 91 to 99% in large studies (*44*, *45*). Durability assessments of circulating antibody titers in Fig. 1 were based on data ≥20 days PSO, with the plot of the best-fitting curve fit model shown in blue (see materials and methods). SARS-CoV-2 spike immunoglobulin G (IgG) endpoint enzyme-linked immunosorbent assay (ELISA) titers in plasma were measured for all subjects of this cohort (Fig. 1, A and B). Spike receptor binding domain (RBD) IgG was also measured (Fig. 1, C and D), as RBD is the target of most neutralizing antibodies against SARS-CoV-2 (*4*, *27*, *46*, *47*). SARS-CoV-2 pseudovirus (PSV)–neutralizing antibody titers were measured in all subjects (Fig. 1, E and F). Nucleocapsid (N) IgG endpoint ELISA titers were also measured for all subjects (Fig. 1, G and H), as nucleocapsid is a common antigen in commercial SARS-CoV-2 serological test kits.

SARS-CoV-2 spike IgG titers were relatively stable from 20 to 240 days PSO, when assessing all COVID-19 subjects by cross-sectional analysis (half-life $t_{1/2}$ = 140 days, Fig. 1A). Spike IgG titers were heterogeneous among subjects (range 5 to 73,071; median 575), as has been widely observed (*45*, *47*). This gave a wide confidence interval (CI) for the spike IgG $t_{1/2}$ (95% CI: 89 to 325 days). Although the antibody responses may have more complex underlying decay kinetics, the best fit curve was a continuous decay, likely related to heterogeneity between individuals. SARS-CoV-2 nucleocapsid IgG kinetics were similar to those of spike IgG over 8 months ($t_{1/2}$ 68 days; 95% CI: 50 to 106 days, Fig. 1G). As a complementary approach, using paired samples from the subset of subjects who donated at two or more time points, the calculated spike IgG titer average $t_{1/2}$ was 103 days, (95% CI: 66 to 235 days; Fig. 1B) and the nucleocapsid IgG titer average $t_{1/2}$ was 68 days, (95% CI: 55 to 90 days; Fig. 1H). The percentage of subjects seropositive for spike IgG at 1 month PSO (20 to 50 days) was 98% (54 out of 55). The percentage of subjects seropositive for spike IgG at 6 to 8 months PSO (≥178 days) was 90% (36 out of 40).

Cross-sectional analysis of SARS-CoV-2 RBD IgG titers from 20 to 240 days PSO gave an estimated $t_{1/2}$ of 83 days (95% CI: 62 to 126 days; Fig. 1C). As a complementary approach, we again used paired samples, which gave an average $t_{1/2}$ of 69 days (95% CI: 58 to 87 days; Fig. 1D). The percentage of subjects seropositive for RBD IgG at 6 to 8 months PSO was 88% (35 out of 40). Thus, RBD IgG titer maintenance largely matched that of spike IgG. SARS-CoV-2 PSV neutralization titers in the full cohort largely matched the results of SARS-CoV-2 RBD IgG ELISA binding titers (Fig. 1, E and F). A one-phase decay model was the best fit (P = 0.015, F test; initial decay $t_{1/2}$ 27 days, followed by an extended plateau phase, Fig. 1E), whereas a continuous decay fit gave an estimated $t_{1/2}$ of 114 days (Fig. 1E, black line). Paired time points analysis of the PSV neutralization titers gave an estimated $t_{1/2}$ of 90 days, (95% CI: 70 to 125 days; Fig. 1F). The percentage of subjects seropositive for SARS-CoV-2–neutralizing antibodies (titer ≥20) at 6 to 8 months PSO was 90% (36 out of 40). Notably, even low circulating neutralizing antibody titers (≥1:20) were associated with a substantial degree of protection against COVID-19 in nonhuman primates (*24*, *48*). Thus, modest amounts of circulating SARS-CoV-2 neutralizing antibodies are of biological interest in humans.

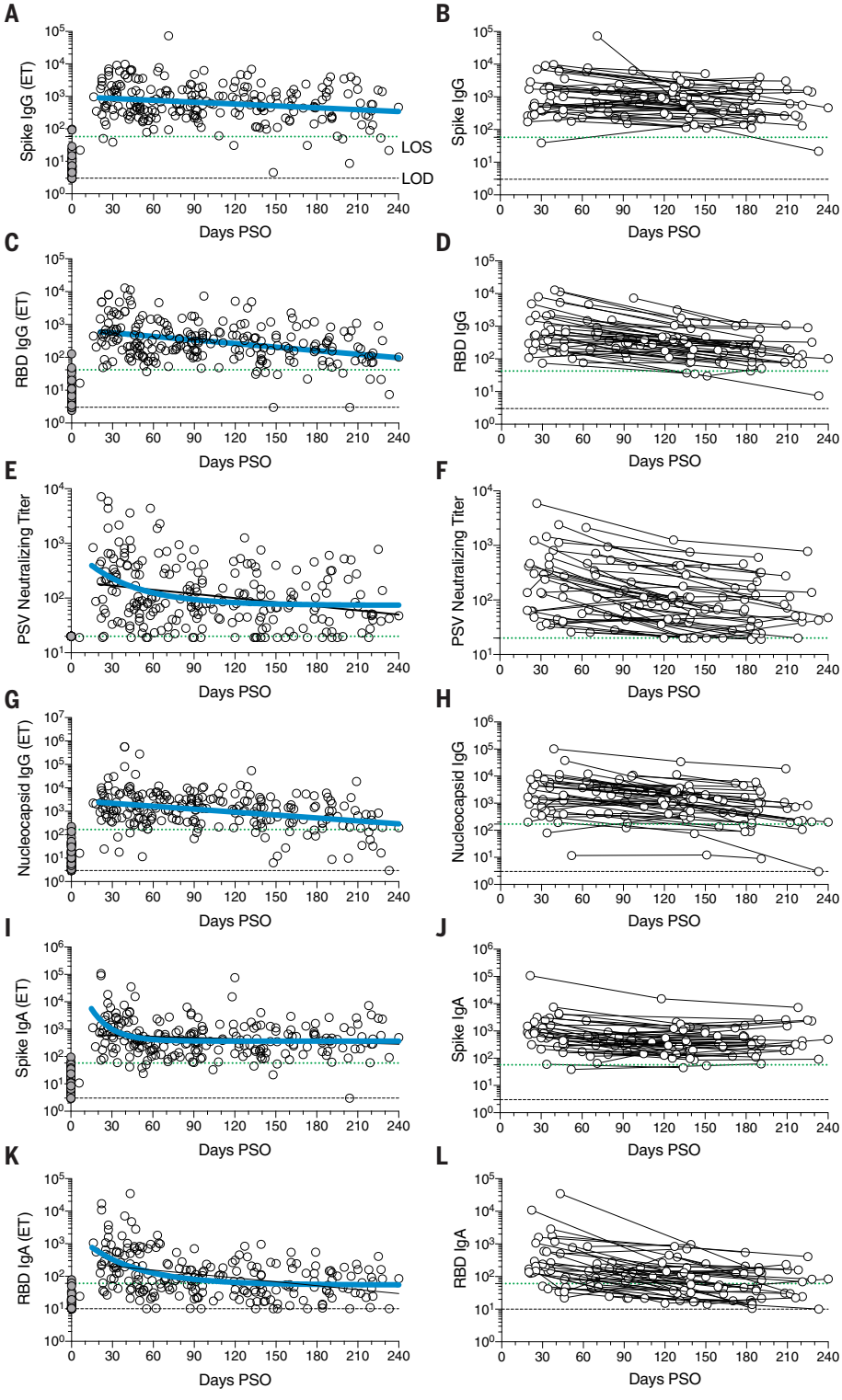


Fig. 1. Circulating antibodies to SARS-CoV-2 over time. (**A**) Cross-sectional spike IgG from COVID-19 subject plasma samples (n = 228). Continuous decay preferred model for best fit curve, $t_{1/2}$ = 140 days; 95% CI: 89 to 325 days. R = −0.23, p = 0.0006. (**B**) Longitudinal spike IgG (n = 51), average $t_{1/2}$ = 103 days; 95% CI: 65 to 235 days. (**C**) Cross-sectional RBD IgG. Continuous decay preferred model for best fit curve, $t_{1/2}$ = 83 days; 95% CI: 62 to 126 days. R = −0.36, p < 0.0001. (**D**) Longitudinal RBD IgG, average $t_{1/2}$ = 69 days; 95% CI: 58 to 87 days. (**E**) Cross-sectional SARS-CoV-2 PSV-neutralizing titers. One-phase decay (blue line) preferred model for best fit curve, initial $t_{1/2}$ = 27 days; 95% CI: 11 to 157 days. R = −0.32. Continuous decay fit line shown as black line. (**F**) Longitudinal PSV-neutralizing titers of SARS-CoV-2–infected subjects, average $t_{1/2}$ = 90 days; 95% CI: 70 to 125 days. (**G**) Cross-sectional nucleocapsid IgG. Continuous decay preferred model for best fit curve, $t_{1/2}$ = 68 days; 95% CI: 50 to 106 days. R = −0.34, p < 0.0001. (**H**) Longitudinal nucleocapsid IgG, average $t_{1/2}$ = 68 days; 95% CI: 55 to 90 days. (**I**) Cross-sectional spike IgA titers. One-phase decay (blue line) preferred model for best fit curve, initial $t_{1/2}$ = 11 days; 95% CI: 5 to 25 days. R = −0.30. Continuous decay fit shown as black line. (**J**) Longitudinal spike IgA, $t_{1/2}$ = 210 days, 95% CI 126 to 627 days. (**K**) Cross-sectional RBD IgA. One-phase decay (blue line) preferred model for best fit curve, initial $t_{1/2}$ = 27 days; 95% CI: 15 to 59 days. R = −0.45. Continuous decay line fit shown in black. (**L**) Longitudinal RBD IgA, average $t_{1/2}$ = 74 days; 95% CI: 56 to 107 days. For cross-sectional analyses, SARS-CoV-2–infected subjects (white circles, n = 238) and unexposed subjects (gray circles, n = 51). For longitudinal samples, SARS-CoV-2 subjects (n = 51). The dotted black line indicates limit of detection (LOD). The dotted green line indicates limit of sensitivity (LOS) above uninfected controls. Unexposed subjects are depicted in gray, COVID subjects in white. Log data analyzed in all cases. Thick blue line represents best fit curve. When two fit curves are shown, the thin black line represents the alternative fit curve.

SARS-CoV-2 spike IgA (Fig. 1, I and J) and RBD IgA (Fig. 1, K and L) titers were also assessed. Paired time points analysis of spike IgA titers yielded an estimated $t_{1/2}$ of 210 days (95% CI 126 to 703 days, Fig. 1J). Cross-sectional analysis of spike IgA fit a short one-phase decay model with an extended plateau phase (initial $t_{1/2}$ of 14 days, Fig. 1I). Circulating RBD IgA had an estimated initial $t_{1/2}$ of 27 days, decaying by ~90 days in most COVID-19 cases to levels indistinguishable from those of uninfected controls (Fig. 1K), consistent with observations 3 months PSO (*44*, *49*). By paired sample analysis, long-lasting RBD IgA was made in some subjects, but often near the limit of sensitivity (LOS) (Fig. 1L).

SARS-CoV-2 memory B cells

To identify SARS-CoV-2–specific memory B cells, we used fluorescently labeled multimerized probes to detect B cells specific to spike, RBD, and nucleocapsid (Fig. 2A and fig. S1). Antigen-binding memory B cells (defined as IgD^- and/or $CD27^+$) were further distinguished according to surface

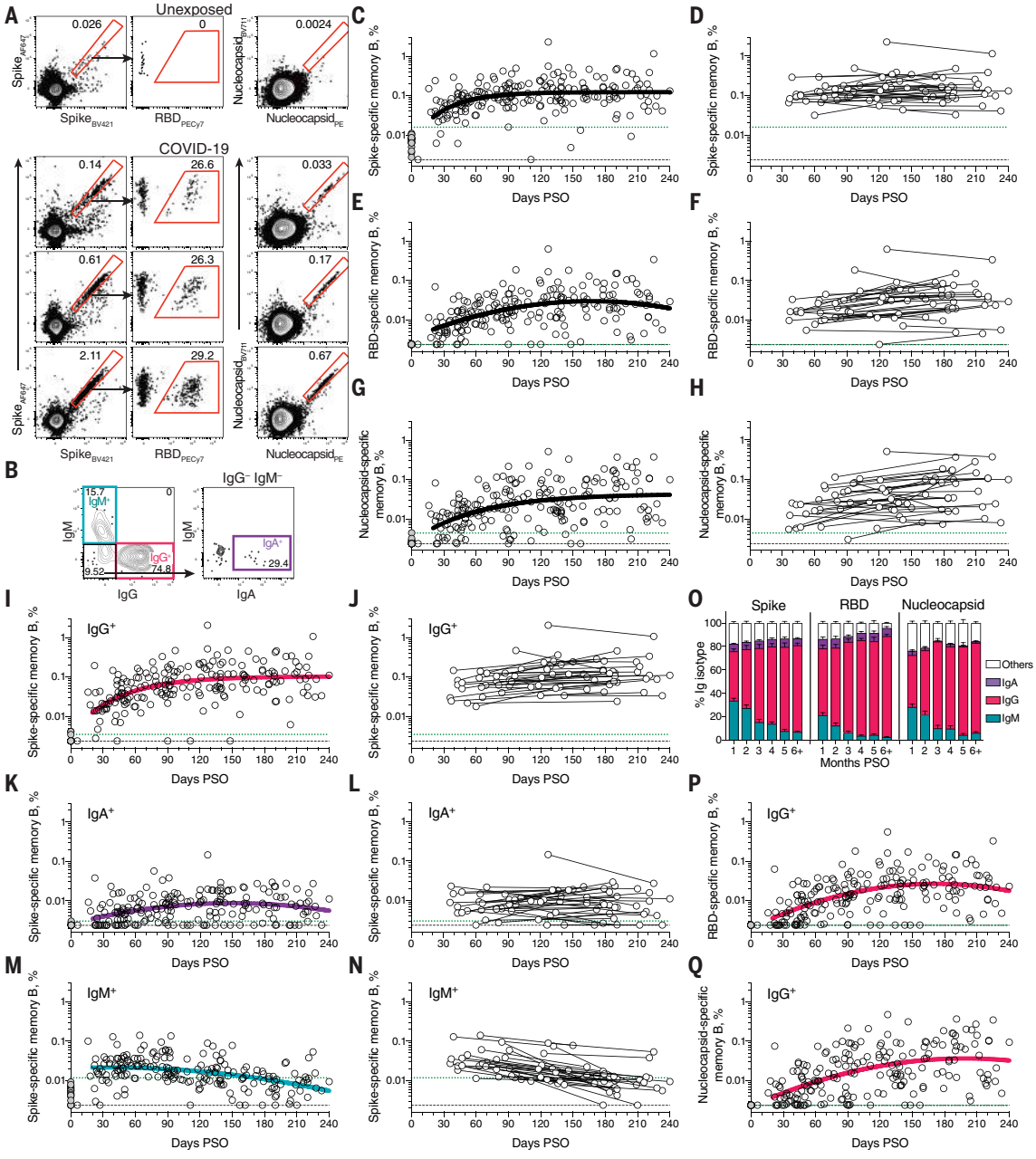


Fig. 2. Kinetics of SARS-CoV-2 memory B cell responses. (**A**) Example flow cytometry plots showing staining patterns of SARS-CoV-2 antigen probes on memory B cells (see fig. S1 for gating). One unexposed donor and three convalescent COVID-19 subjects are shown. Numbers indicate percentages. (**B**) Gating strategies to define IgM$^+$, IgG$^+$, or IgA$^+$ SARS-CoV-2 spike-specific memory B cells. The same gating strategies were used for RBD- or nucleocapsid-specific B cells. (**C**) Cross-sectional analysis of frequency (percentage of CD19$^+$ CD20$^+$ B cells) of SARS-CoV-2 spike-specific total (IgG$^+$, IgM$^+$, or IgA$^+$) memory B cells. Pseudo–first-order kinetic model for best fit curve (R = 0.38). (**D**) Longitudinal analysis of SARS-CoV-2 spike-specific memory B cells. (**E**) Cross-sectional analysis of SARS-CoV-2 RBD-specific total (IgG$^+$, IgM$^+$, or IgA$^+$) memory B cells. Second-order polynomial model for best fit curve (R = 0.46). (**F**) Longitudinal analysis of SARS-CoV-2 RBD-specific memory B cells. (**G**) Cross-sectional analysis of SARS-CoV-2 nucleocapsid-specific total (IgG$^+$, IgM$^+$, or IgA$^+$) memory B cells. Pseudo–first-order kinetic model for best fit curve (R = 0.44). (**H**) Longitudinal analysis of IgG$^+$ SARS-CoV-2 nucleocapsid-specific memory B cells. (**I**) Cross-sectional analysis of SARS-CoV-2 spike-specific IgG$^+$ memory B cells. Pseudo–first-order kinetic model for best fit curve (R = 0.49). (**J**) Longitudinal analysis of SARS-CoV-2 spike-specific IgG$^+$ memory B cells. (**K**) Cross-sectional analysis of SARS-CoV-2 spike-specific IgA$^+$ memory B cells. Second-order polynomial model for best fit curve ($|R|$ = 0.32). (**L**) Longitudinal analysis of SARS-CoV-2 spike-specific IgA$^+$ memory B cells. (**M**) Cross-sectional analysis of SARS-CoV-2 spike-specific IgM$^+$ memory B cells. Second-order polynomial model for best fit curve ($|R|$ = 0.41). (**N**) Longitudinal analysis of SARS-CoV-2 spike-specific IgM$^+$ memory B cells. (**O**) Fraction of SARS-CoV-2 antigen-specific memory B cells that belong to indicated Ig isotypes at 1 to 8 months PSO. Mean ± SEM. (**P**) Cross-sectional analysis of SARS-CoV-2 RBD-specific IgG$^+$ memory B cells. Second-order polynomial model for best fit curve ($|R|$ = 0.51). (**Q**) Cross-sectional analysis of SARS-CoV-2 nucleocapsid-specific IgG$^+$ memory B cells. Second-order polynomial model for best fit curve ($|R|$ = 0.51). n = 20 unexposed subjects (gray circles) and n = 160 COVID-19 subjects (n = 197 data points, white circles) for cross-sectional analysis. n = 36 COVID-19 subjects (n = 73 data points, white circles) for longitudinal analysis. The dotted black line indicates LOD. The dotted green line indicates LOS.

Ig isotypes: IgM, IgG, or IgA (Fig. 2B and fig. S1).

Cross-sectional analysis of COVID-19 subjects revealed that frequencies of SARS-CoV-2 spike-specific memory B cells increased over the first ~120 days PSO and then plateaued (pseudo–first- order model for best fit curve, R = 0.38; better fit than second-order polynomial model by Akaike's information criterion; Fig. 2C and fig. S2A). Spike-specific memory B cell frequencies increased from the first time point (36 to 163 days) to the second time point (111 to 240 days) in paired samples from 24 of 36 longitudinally tracked donors (Fig. 2D). Spike-specific memory B cells in SARS-CoV-2–unexposed

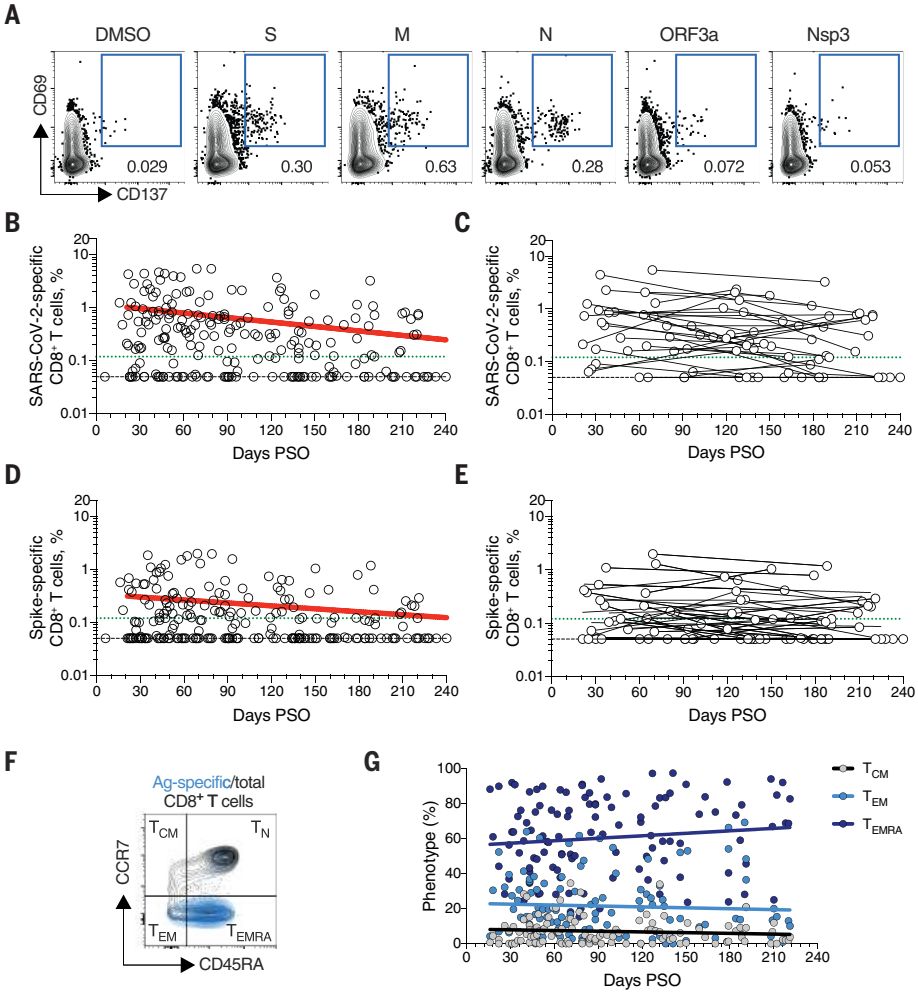


Fig. 3. SARS-CoV-2 circulating memory CD8$^+$ T cells. (**A**) Representative flow cytometry plots of SARS-CoV-2–specific CD8$^+$ T cells (CD69$^+$ CD137$^+$; see fig. S3 for gating) after overnight stimulation with S, N, M, ORF3a, or nsp3 peptide pools, compared to negative control (DMSO). (**B**) Cross-sectional analysis of frequency (percentage of CD8$^+$ T cells) of total SARS-CoV-2–specific CD8$^+$ T cells. Continuous decay preferred fit model, $t_{1/2}$ = 125 days. R = −0.24, p = 0.0003. (**C**) Longitudinal analysis of total SARS-CoV-2–specific CD8$^+$ T cells in paired samples. (**D**) Cross-sectional analysis of spike-specific CD8$^+$ T cells. Linear decay preferred model, $t_{1/2}$ = 225 days. R = −0.18, p = 0.007. (**E**) Longitudinal analysis of spike-specific CD8$^+$ T cells in paired samples. (**F** and **G**) Distribution of central memory (T_{CM}), effector memory (T_{EM}), and terminally differentiated effector memory cells (T_{EMRA}) among total SARS-CoV-2–specific CD8$^+$ T cells. n = 169 COVID-19 subjects (n = 215 data points, white circles) for cross-sectional analysis. n = 37 COVID-19 subjects (n = 83 data points, white circles) for longitudinal analysis. The dotted black line indicates LOD. The dotted green line indicates LOS.

subjects were rare (median 0.0078%; Fig. 2, A and C).

RBD-specific memory B cells displayed similar kinetics to spike-specific memory B cells. RBD-specific memory B cells were undetectable in SARS-CoV-2–unexposed subjects (Fig. 2E and fig. S2C), as expected. RBD-specific memory B cells appeared as early as 16 days PSO, and the frequency steadily increased in the following 4 to 5 months (Fig. 2E and fig. S2, B and C). Twenty-nine of 36 longitudinally tracked individuals had higher frequencies of RBD-specific memory B cells at the later time point (Fig. 2F), again showing an increase in SARS-CoV-2–specific memory B cells several months after infection. About 10 to 30% of spike-specific memory B cells from SARS-CoV-2 convalescent donors were specific for the RBD domain (Fig. 2A and fig. S2B).

SARS-CoV-2 nucleocapsid-specific memory B cells were also detected after SARS-CoV-2 infection (Fig. 2A). Similar to spike- and RBD-specific memory B cells, nucleocapsid-specific memory B cell frequency steadily increased during the first ~4 to 5 months PSO (Fig. 2, G and H, and fig. S2D). Antibody affinity maturation could potentially explain the increased frequencies of SARS-CoV-2–specific memory B cells detected by the antigen probes. However, geometric mean fluorescent intensity (MFI) of probe binding was stable over time (fig. S2, I and J), not supporting an affinity maturation explanation for the increased memory B cell frequencies.

Representation of Ig isotypes among the SARS-CoV-2 spike-specific memory B cell population shifted with time (Fig. 2, I to O). During the earliest phase of memory (20 to 60 days PSO), IgM$^+$ and IgG$^+$ isotypes were similarly represented (Fig. 2O), but IgM$^+$ memory B cells then declined (Fig. 2, M to O), and IgG$^+$ spike-specific memory B cells then dominated by 6 months PSO (Fig. 2O). IgA$^+$ spike-specific memory B cells were detected as a small fraction of the total spike-specific memory B cells (~5%, Fig. 2O). IgG$^+$ spike-specific memory B cell frequency increased, whereas IgA$^+$ frequency was low and stable over the 8-month period (Fig. 2, I to L). Similar patterns of increasing IgG$^+$ memory, short-lived IgM$^+$ memory, and stable IgA$^+$ memory were observed for RBD- and nucleocapsid-specific memory B cells over the 8-month period (Fig. 2, O to Q, and fig. S2, E to H).

There is limited knowledge of memory B cell kinetics following primary acute viral infection in humans. A recently published SARS-CoV-2 study found RBD-specific memory B cells up to ~90 days PSO, with increasing frequencies (and a low frequency of IgA$^+$ cells) (*50*), consistent with observations reported here. For other acute infectious diseases, we are not currently aware of other cross-sectional or longitudinal analyses of antigen-specific memory B cells by flow cytometry covering a 6+-month window after infection, except for four individuals with Ebola (*51*) and two individuals studied after yellow fever virus immunization (*52*) (we exclude influenza vaccines for comparison here, because people have numerous exposures and complex immune history to influenza). In the yellow fever study, short-lived IgM$^+$ memory and longer-lasting isotype-switched memory B cells were observed in the two individuals. Overall, on the basis of the observations here, development of B cell memory to SARS-CoV-2 was robust and is likely long-lasting.

SARS-CoV-2 memory CD8$^+$ T cells

SARS-CoV-2 memory CD8$^+$ T cells were measured in 169 COVID-19 subjects using a series of 23 peptide pools covering the entirety of the SARS-CoV-2 ORFeome (*2*, *5*). The most commonly recognized open reading frames (ORFs) were spike, membrane (M), nucleocapsid, and ORF3a (CD69$^+$ CD137$^+$; Fig. 3A and fig. S3, A and B), consistent with our previous study (*2*). The percentage of subjects with detectable circulating SARS-CoV-2 memory CD8$^+$ T cells at 1 month PSO (20 to 50 days) was 70% (40 out of 57, Fig. 3B).

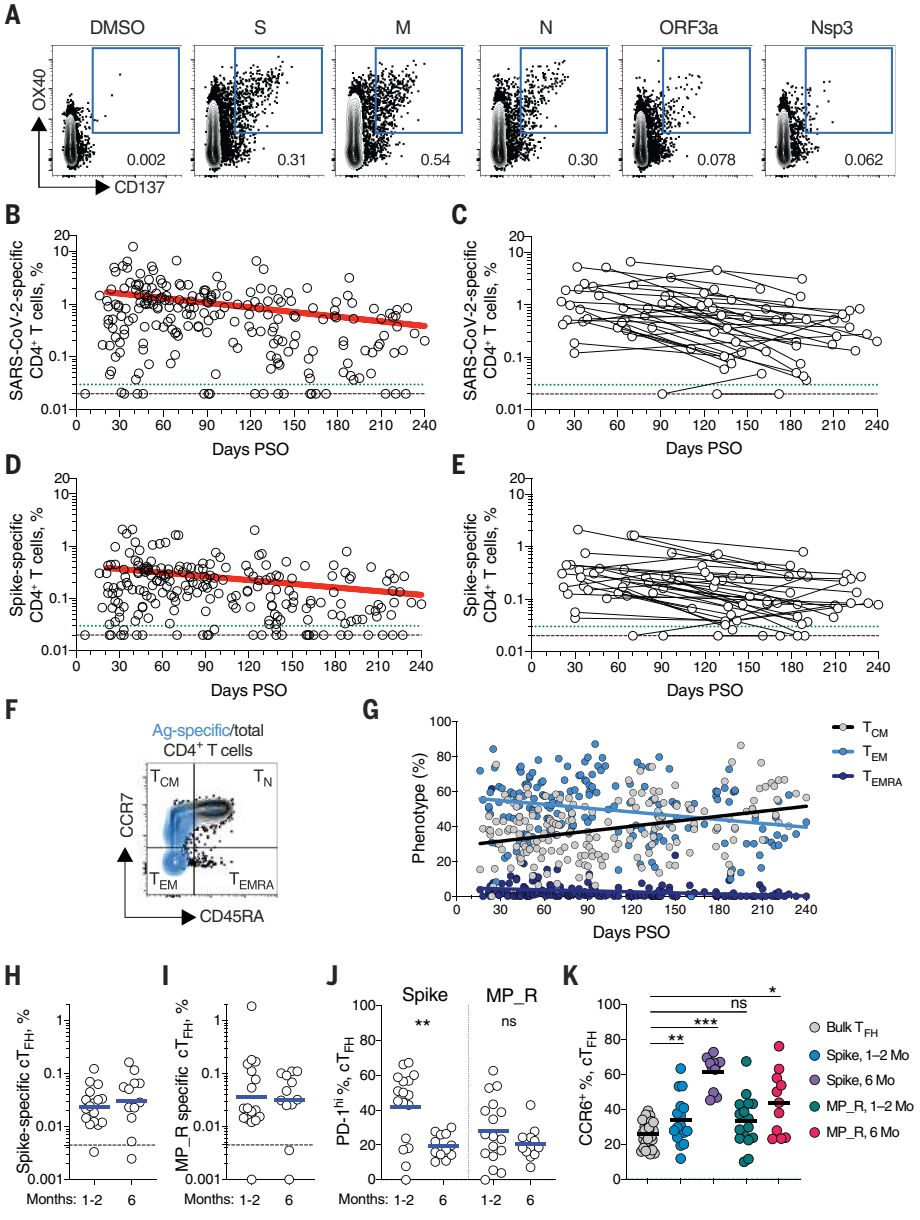


Fig. 4. SARS-CoV-2 circulating memory CD4⁺ T cells. (**A**) Representative flow cytometry plots of SARS-CoV-2–specific CD4⁺ T cells (CD137⁺ OX40⁺; see fig. S4 for gating) after overnight stimulation with S, N, M, ORF3a, or nsp3 peptide pools, compared to negative control (DMSO). (**B**) Cross-sectional analysis of frequency (percentage of CD4⁺ T cells) of total SARS-CoV-2–specific CD4⁺ T cells. Continuous decay preferred fit model, $t_{1/2}$ = 94 days. R = −0.29, p < 0.0001. (**C**) Longitudinal analysis of total SARS-CoV-2–specific CD4⁺ T cells in paired samples from the same subjects. (**D**) Cross-sectional analysis of spike-specific CD4⁺ T cells. Linear decay preferred model, $t_{1/2}$ = 139 days. R = −0.26, p < 0.0001. (**E**) Longitudinal analysis of spike-specific CD4⁺ T cells in paired samples from the same subjects. (**F** and **G**) Distribution of central memory (T_{CM}), effector memory (T_{EM}), and terminally differentiated effector memory cells (T_{EMRA}) among total SARS-CoV-2–specific CD4⁺ T cells. (**H** and **I**) Quantitation of SARS-CoV-2–specific circulating T follicular helper (cT_{FH}) cells (surface CD40L⁺ OX40⁺, as percentage of CD4⁺ T cells; see fig. S5 for gating) after overnight stimulation with (H) spike (S) or (I) MP_R peptide pools. (**J**) PD-1^{hi} SARS-CoV-2–specific T_{FH} at 1 to 2 months (mo) and 6 mo PSO. (**K**) CCR6⁺ SARS-CoV-2–specific cT_{FH} in comparison to bulk cT_{FH} cells in blood. For (A) to (E), n = 169 COVID-19 subjects (n = 215 data points, white circles) for cross-sectional analysis, n = 37 COVID-19 subjects (n = 83 data points, white circles) for longitudinal analysis. The dotted black line indicates limit of detection. The dotted green line indicates LOS. For (H) to (J), n = 29 COVID-19 subject samples (white circles), n = 17 COVID-19 subjects at 1 to 2 mo, n = 12 COVID-19 subjects at 6 mo. The dotted black line indicates LOD. Statistics by (J) Mann-Whitney U test and (K) Wilcoxon signed-rank test. *p < 0.05, **p < 0.01, ***p < 0.001. ns, not statistically significant.

The proportion of subjects positive for SARS-CoV-2 memory CD8⁺ T cells at ≥6 months PSO was 50% (18 out of 36). This could potentially underestimate CD8⁺ T cell memory, as 15-mer peptides can be suboptimal for detection of some antigen-specific CD8⁺ T cells (*53*); however, pools of predicted SARS-CoV-2 class I epitope of optimal size also detected virus-specific CD8⁺ T cells in ~70% of individuals 1 to 2 months PSO, indicating consistency between the two experimental approaches (*2*).

SARS-CoV-2 memory CD8⁺ T cells declined with an apparent $t_{1/2}$ of 125 days in the full cohort (Fig. 3B) and $t_{1/2}$ 190 days among 29 paired samples (Fig. 3C). Spike-specific memory CD8⁺ T cells exhibited similar kinetics to the overall SARS-CoV-2–specific memory CD8⁺ T cells ($t_{1/2}$ 225 days for the full cohort and 185 days among paired samples; Fig. 3, D and E, respectively). Phenotypic markers indicated that the majority of SARS-CoV-2–specific memory CD8⁺ T cells were terminally differentiated effector memory cells (T_{EMRA}) (*54*), with small populations of central memory (T_{CM}) and effector memory (T_{EM}) (Fig. 3, F and G). In the context of influenza, CD8⁺ T_{EMRA} cells were associated with protection against severe disease in humans (*55*). The memory CD8⁺ T cell half-lives observed here were comparable to the 123 days $t_{1/2}$ observed for memory CD8⁺ T cells after yellow fever immunization (*56*). Thus, the kinetics of circulating SARS-CoV-2–specific CD8⁺ T cell were consistent with what has been reported for another virus that causes acute infections in humans.

SARS-CoV-2 memory CD4⁺ T cells

SARS-CoV-2 memory CD4⁺ T cells were identified in 169 subjects using the same series of 23 peptide pools covering the SARS-CoV-2 ORFeome (*2*, *5*). The most commonly recognized ORFs were spike, M, nucleocapsid, ORF3a, and nsp3 (CD137⁺ OX40⁺; Fig. 4A and fig. S4, A and B), consistent with our previous study (*2*). Circulating SARS-CoV-2 memory CD4⁺ T cell responses were quite robust (Fig. 4B); 42% (24 out of 57) of COVID-19 cases at 1 month PSO had >1.0% SARS-CoV-2–specific CD4⁺ T cells. SARS-CoV-2 memory CD4⁺ T cells declined with an apparent $t_{1/2}$ of 94 days in the full cohort (Fig. 4B) and 64 days among 36 paired samples (Fig. 4C). The percentage of subjects with detectable circulating SARS-CoV-2 memory CD4⁺ T cells at 1 month PSO (20 to 50 days) was 93% (53 out 57, Fig. 4B). The proportion of subjects positive for SARS-CoV-2 memory CD4⁺ T cells at ≥6 months PSO was 92% (33 out of 36).

Spike-specific and M-specific memory CD4⁺ T cells exhibited similar kinetics to the overall SARS-CoV-2–specific memory CD4⁺ T cells (whole cohort $t_{1/2}$ of 139 days and 153 days, respectively; Fig. 4, D and E, and fig. S4D). A plurality of the SARS-CoV-2 memory CD4⁺

T cells present at ≥6 months PSO had a T_{CM} phenotype (Fig. 4F).

T follicular helper (T_{FH}) cells are the specialized subset of $CD4^+$ T cells required for B cell help (*57*) and are therefore critical for the generation of neutralizing antibodies and long-lived humoral immunity in most contexts. Thus, we examined circulating T_{FH} (cT_{FH}) memory $CD4^+$ T cells, with particular interest in spike-specific memory cT_{FH} cells, owing to the importance of antibody responses against spike. Memory cT_{FH} cells specific for predicted epitopes across the remainder of the SARS-CoV-2 genome were also measured, using the MP_R megapool. Memory cT_{FH} cells specific for SARS-CoV-2 spike and MP_R were detected in the majority of COVID-19 cases at early time points (16 out of 17; Fig. 4, H and I, and fig. S5, A to D). cT_{FH} memory appeared to be stable, with almost all subjects positive for spike and MP_R memory cT_{FH} cells at 6 months PSO (11 out of 12 and 10 out of 12, respectively; Fig. 4, H and I). Recently activated cT_{FH} cells are PD-1^{hi} (*57*). Consistent with conversion to resting memory cT_{FH} cells, the percentage of PD-1^{hi} SARS-CoV-2–specific memory cT_{FH} dropped over time (Fig. 4J). $CCR6^+$ SARS-CoV-2–specific cT_{FH} cells have been associated with reduced COVID-19 disease severity (*5*) and have been reported to be a major fraction of spike-specific cT_{FH} cells in some studies (*5*, *50*, *58*). Here we confirmed that a significant fraction of both spike-specific and MP_R memory cT_{FH} cells were $CCR6^+$. We also observed increases in $CCR6^+$ cT_{FH} memory over time (p = 0.001 and p = 0.014 at ≥6 months PSO compared to bulk cT_{FH}, Fig. 4K). Overall, substantial cT_{FH} memory was observed after SARS-CoV-2 infection, with a durability ≥6 months PSO.

Immune memory relationships

Immune memory to SARS-CoV-2 was considered, including relationships between the compartments of immune memory. Males had higher spike IgG [analysis of covariance (ANCOVA) p = 0.00018, Fig. 5A] and RBD and nucleocapsid IgG (ANCOVA p = 0.00077 and p = 0.018; fig. S6, A and B), consistent with other studies (*46*, *47*). Higher spike IgG was also observed in males when only nonhospitalized cases were considered (ANCOVA p = 0.00025, fig. S6C). By contrast, no differences were observed in IgA or PSV neutralization titers (fig. S6, D to F), and no differences were detected in SARS-CoV-2 memory B cell, memory $CD8^+$ T cell, or memory $CD4^+$ T cell frequencies between males and females (fig. S6, G to K).

Immune memory was examined for associations between magnitude of memory and COVID-19 disease severity. The number of previously hospitalized COVID-19 cases (n = 13) limited analysis options. However, the cases were well distributed between males and females (Table 1), data from large numbers of nonhospitalized cases were available for comparison, and the analyses in Figs. 1 to 4 demonstrated that immune memory was relatively stable over the time window analyzed. Therefore, we could simplify the disease severity analysis by grouping all samples from 120+ days PSO [also limiting data to a single sample per subject (figs. S7 to S9); most of the previously hospitalized subjects were sampled at two time points, fig. S7A] and then comparing nonhospitalized and hospitalized subjects. Spike and RBD IgG titers in hospitalized cases were higher than in nonhospitalized cases (Fig. 5B), consistent with other studies (*46*, *47*). Spike and RBD-specific memory B cell frequencies were also higher in hospitalized cases (~1.7-fold and ~2.5-fold, respectively; Fig. 5C and fig. S8). By contrast, memory $CD8^+$ T cell frequencies were not higher in hospitalized cases compared to nonhospitalized cases (Fig. 5D and fig. S9), and memory $CD4^+$ T cell frequencies trended lower in hospitalized cases compared to nonhospitalized cases (Fig. 5E and fig. S9). Therefore, although our conclusions are limited by the number of hospitalized subjects, increased spike IgG titers were consistent across three independent studies, and increased memory B cells among hospitalized cases were observed here (not measured in other studies), indicating that both compartments of long-term humoral immunity to SARS-CoV-2 are higher in individuals who experienced a more severe COVID-19 disease course. T cell memory did not follow the same pattern, consistent with indications that hospitalized cases of COVID-19 can be associated with poorer T cell responses in the acute phase (*5*, *59*). Additionally, these data show that, although gender and COVID-19 disease severity contribute to differences in immune memory to SARS-CoV-2, neither factor could account for the majority of the heterogeneity in immune memory to this virus.

Very few published data sets compare antigen-specific antibody, B cell, $CD8^+$ T cell, and $CD4^+$ T cell memory to an acute viral infection in the same individuals. We therefore made use of this combined data set to examine interrelationships between compartments of immune memory. We focused on RBD IgG, RBD memory B cells, spike IgA, total SARS-CoV-2–specific $CD8^+$ T cells, and total SARS-CoV-2–specific $CD4^+$ T cells, owing to their putative potential roles in protective immunity. The majority (64%) of COVID-19 cases were positive for all five of these immune memory compartments at 1 to 2 months PSO (Fig. 5, F and G), with the incomplete responses largely reflecting individuals with no detectable $CD8^+$ T cell memory and/or poor IgA responses (Fig. 5G). At 5 to 8 months after COVID-19 infection, the proportion of individuals positive for all five of these immune memory compartments had dropped to 43%; nevertheless, 95% of individuals were still positive for at least three out of five SARS-CoV-2 immune memory responses (Fig. 5G). Immune memory at 5 to 8 months PSO represented contributions from different immune memory compartments in different individuals (Fig. 5G). Similar results were obtained if RBD IgG was replaced by neutralizing antibodies (fig. S10A). Overall, these findings again highlight heterogeneity of immune memory, with different patterns of immune memory in different individuals.

Interrelationships between the components of memory were next examined by assessing ratios between immune memory compartments over time. The ratio of SARS-CoV-2 $CD4^+$ T cell memory to SARS-CoV-2 $CD8^+$ T cell memory was largely stable over time (Fig. 5H and fig. S10B). Given that serological measurements are the simplest measurements of immune memory at a population scale, we examined how well such serological measurements may serve as surrogate markers of other components of SARS-CoV-2 immune memory over time. The relationship between circulating RBD IgG and RBD-specific memory B cells changed ~20-fold over the time range studied (R = 0.60, Fig. 5H and fig. S10C). The changing relationship between circulating spike IgA and RBD-specific memory B cells was even larger (R = 0.55, Fig. 5H and fig. S10D). The relationship between RBD IgG and SARS-CoV-2 $CD4^+$ T cell memory was relatively flat over the time range studied (Fig. 5H); however, variation spanned a ~1000-fold range (fig. S10E). Thus, predictive power of circulating RBD IgG for assessing T cell memory was poor because of the heterogeneity between individuals (R = 0.046). In sum, although heterogeneity of immune responses is a defining feature of COVID-19, immune memory to SARS-CoV-2 develops in almost all subjects, with complex relationships between the individual immune memory compartments.

Concluding remarks

In this study, we aimed to fill gaps in our basic understanding of immune memory after COVID-19. This required simultaneous measurement of circulating antibodies, memory B cells, $CD8^+$ T cells, and $CD4^+$ T cells specific for SARS-CoV-2, in a group of subjects with a full range of disease, and distributed from short time points after infection to 8 months later. By studying these multiple compartments of adaptive immunity in an integrated manner, we observed that each component of SARS-CoV-2 immune memory exhibited distinct kinetics.

The spike IgG titers were durable, with modest declines in titers at 6 to 8 months PSO at the population level. RBD IgG and SARS-CoV-2 PSV–neutralizing antibody titers

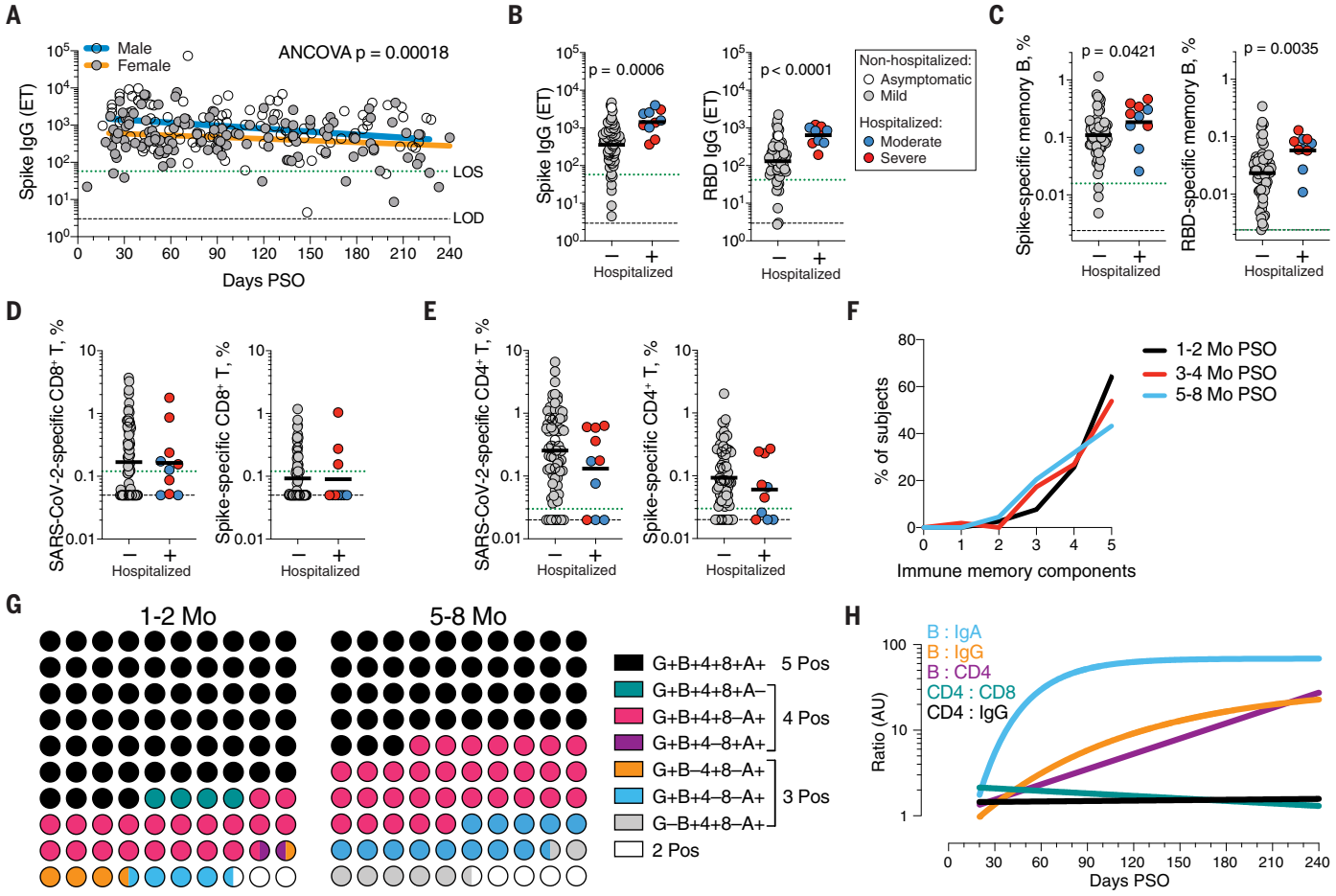


Fig. 5. Immune memory relationships. (**A**) Relationship between gender and spike IgG titers over time. Males: Linear decay preferred model, $t_{1/2}$ = 110 days; 95% CI: 65 to 349 days, R = −0.27, p = 0.0046. Females: linear decay preferred model, $t_{1/2}$ = 159 days; 95% CI: 88 to 846 days, R = −0.22, p = 0.016. ANCOVA p = 0.00018. Test for homogeneity of regressions F = 1.51, p = 0.22. (**B** to **E**) Immune memory at 120+ days PSO in COVID-19 nonhospitalized and hospitalized subjects. Symbol colors represent peak disease severity (white: asymptomatic, gray: mild, blue: moderate, red: severe.) For subjects with multiple sample time points, only the final time point was used for these analyses. (B) Spike-specific IgG (left) and RBD-specific IgG (right) binding titers. n = 64 (nonhospitalized), n = 10 (hospitalized). Mann-Whitney U tests. (C) Frequency of memory B cells specific to spike (left) and RBD (right) at 120+ days PSO. n = 66 (nonhospitalized), n = 10 (hospitalized). Mann-Whitney U tests. (D) Frequency of total SARS-CoV-2–specific $CD8^+$ T cells (left) and spike-specific $CD8^+$ T cells (right). p = 0.72 for total SARS-2-CoV–specific, p = 0.60 for spike-specific by Mann-Whitney U tests. n = 72 (nonhospitalized), n = 10 (hospitalized). (E) Frequency of total SARS-CoV-2–specific $CD4^+$ T cells (left) and spike-specific $CD4^+$ T cells (right). p = 0.23 for total SARS-CoV-2–specific, p = 0.24 for spike-specific by Mann-Whitney U tests. (**F**) Immune memory to SARS-CoV-2 during the early phase (1 to 2 mo, black line), medium phase (3 to 4 mo, red line), or late phase (5 to 8 mo, blue line). For each individual, a score of 1 was assigned for each response above LOS for RBD IgG, spike IgA, RBD-specific memory B cells, SARS-CoV-2–specific $CD4^+$ T cells, and SARS-CoV-2–specific $CD8^+$ T cells, giving a maximum total of five components of SARS-CoV-2 immune memory. Only COVID-19 convalescent subjects with all five immunological parameters tested were included in the analysis. n = 78 (1 to 2 mo), n = 52 (3 to 4 mo), n = 44 (5 to 8 mo). (**G**) Percentage dot plots showing frequencies (normalized to 100%) of subjects with indicated immune memory components as described in (B) during the early (1 to 2 mo) or late (5 to 8 mo) phase. G, RBD-specific IgG; B, RBD-specific memory B cells; 4, SARS-CoV-2–specific $CD4^+$ T cells; 8, SARS-CoV-2–specific $CD8^+$ T cells; A, spike-specific IgA. n = 78 (1 to 2 mo), n = 44 (5 to 8 mo). (**H**) Relationships between immune memory compartments in COVID-19 subjects over time, as ratios (full curves and data shown in fig. S10, B to F). AU, arbitrary units, scaled from fig. S10, B to F; B:IgA, RBD-specific memory B cell ratio to spike IgA antibodies; B:IgG, RBD-specific memory B cell ratio to RBD IgG antibodies; B:CD4, RBD-specific memory B cell ratio to SARS-CoV-2–specific $CD4^+$ T cells; CD4:CD8, SARS-CoV-2–specific $CD4^+$ T cells ratio to SARS-CoV-2–specific $CD8^+$ T cells; CD4:IgG, SARS-CoV-2–specific $CD4^+$ T cells ratio to RBD IgG antibodies.

were potentially similarly stable, consistent with the RBD domain of spike being the dominant neutralizing antibody target. We collected data at two time points for most longitudinal individuals herein. It is well recognized that the magnitude of the antibody response against SARS-CoV-2 is highly heterogeneous between individuals. We observed that heterogeneous initial antibody responses did not collapse into a homogeneous circulating antibody memory; rather, heterogeneity is also a central feature of immune memory to this virus. For antibodies, the responses spanned a ~200-fold range. Additionally, this heterogeneity means that long-term longitudinal studies will be required to precisely define antibody kinetics to SARS-CoV-2. We are reporting the simplest statistical models that explain the data. These curve fits do not disprove more complex kinetics such as overlapping kinetics, but those models would require much denser longitudinal sampling in future studies. Biologically, IgG antibodies having a half-life of ~21 days, and the

magnitude of the antibody response over time, reflect antibodies produced first by short-lived plasma cells and then long-lived plasma cells, with affinity maturation also affecting the apparent magnitude in conventional binding assays and neutralization assays. Overall, at 5 to 8 months PSO, almost all individuals were positive for SARS-CoV-2 spike and RBD IgG.

Notably, memory B cells specific for the spike protein or RBD were detected in almost all COVID-19 cases, with no apparent half-life at 5 to 8 months after infection. Other studies of RBD memory B cells report similar findings (*50*, *60*). B cell memory to some other infections has been observed to be long-lived, including 60+ years after smallpox vaccination (*61*), or 90+ years after infection with influenza (*62*). The memory T cell half-lives observed over 6+ months PSO in this cohort (~125 to 225 days for $CD8^+$ and ~94 to 153 days for $CD4^+$ T cells) were comparable to the 123 days $t_{1/2}$ observed for memory $CD8^+$ T cells after yellow fever immunization (*56*). SARS-CoV-2 T cell memory at 6 months has also now been reported in another study (*63*). Notably, the durability of a fraction of the yellow fever virus–specific memory $CD8^+$ T cells possessed an estimated $t_{1/2}$ of 485 days by deuterium labeling (*56*). Using different approaches, studies determined the long-term durability of memory $CD4^+$ T cells to smallpox, over a period of many years, to be an estimated $t_{1/2}$ of ~10 years (*61*, *64*), which is also consistent with recent detection of SARS-CoV-specific T cells 17 years after the initial infection (*65*). These data suggest that T cell memory might reach a more stable plateau, or slower decay phase, beyond the first 8 months after infection.

Although immune memory is the source of long-term protective immunity, direct conclusions about protective immunity cannot be made on the basis of quantifying SARS-CoV-2 circulating antibodies, memory B cells, $CD8^+$ T cells, and $CD4^+$ T cells, because mechanisms of protective immunity against SARS-CoV-2 or COVID-19 are not defined in humans. Nevertheless, some reasonable interpretations can be made. Antibodies are the only component of immune memory that can provide truly sterilizing immunity. Immunization studies in nonhuman primates have indicated that circulating neutralization titers of ~200 may provide sterilizing immunity against a relatively high-dose URT challenge (*66*), and neutralizing titers of ~3400 may provide sterilizing immunity against a very high dose URT challenge (*67*), although direct comparisons are not possible because the neutralizing antibody assays have not been standardized (*3*). Conclusions are also constrained by the limited overall amount of data on protective immunity to SARS-CoV-2.

Beyond sterilizing immunity, immune responses that confine SARS-CoV-2 to the URT and oral cavity would minimize COVID-19 disease severity to that of a "common cold" or asymptomatic disease. This outcome is the primary goal of current COVID-19 vaccine clinical trials (*3*, *68*). Such an outcome could potentially be mediated by a mixture of memory $CD4^+$ T cells, memory $CD8^+$ T cells, and memory B cells specific for RBD-producing anamnestic neutralizing antibodies, based on mechanisms of action in mouse models of other viral infections (*69*–*71*). In human COVID-19 infections, SARS-CoV-2–specific $CD4^+$ T cells and $CD8^+$ T cells are associated with less COVID-19 disease severity during an ongoing SARS-CoV-2 infection (*5*). Rapid seroconversion was associated with substantially reduced viral loads in acute disease over 14 days (*29*). Both of those associations are consistent with the hypothesis that SARS-CoV-2 memory T cells and B cells would be capable of substantially limiting SARS-CoV-2 dissemination and/or cumulative viral load, resulting in reduced COVID-19 disease severity. The likelihood of such outcomes is also closely tied to the kinetics of the infection, as memory B and T cell responses can take 3 to 5 days to successfully respond to an infection. As noted above, given the relatively slow course of severe COVID-19 in humans, resting immune memory compartments can potentially contribute in meaningful ways to protective immunity against pneumonia or severe secondary COVID-19. The presence of substerilizing neutralizing antibody titers at the time of SARS-CoV-2 exposure would blunt the size of the initial infection, and may provide an added contribution to limiting COVID-19 severity, on the basis of observations of protective immunity for other human respiratory viral infections (*37*, *72*–*74*) and observations of SARS-CoV-2 vaccines in nonhuman primates (*48*, *67*, *75*).

The current study has some limitations. Longitudinal data for each subject, with at least three time points per subject, would be required for more precise understanding of the kinetics of durability of SARS-CoV-2–specific antibodies. Nevertheless, the current cross-sectional data describe well the dynamics of SARS-CoV-2 memory B cells, $CD8^+$ T cell, and $CD4^+$ T cell over 8 months PSO. This study was not sufficiently powered to control for many variables simultaneously. Additionally, circulating memory was assessed here; it is possible that local URT immune memory is a minimal, moderate, or large component of immune memory after a primary infection with SARS-CoV-2. This remains to be determined.

Individual case reports show that reinfections with SARS-CoV-2 are occurring (*76*, *77*). However, a 2800-person study found no symptomatic re-infections over a ~118-day window (*78*), and a 1246-person study observed no symptomatic reinfections over 6 months (*79*). We observed heterogeneity in the magnitude of adaptive immune responses to SARS-CoV-2 persisting into the immune memory phase. It is therefore possible that a fraction of the SARS-CoV-2–infected population with low immune memory would become susceptible to reinfection relatively soon. Although gender and disease severity both contribute to the heterogeneity of immune memory reported here, the source of much of the heterogeneity in immune memory to SARS-CoV-2 is unknown and worth further examination. Perhaps heterogeneity derives from low cumulative viral load or a small initial inoculum in some individuals. Nevertheless, our data show immune memory in at least three immunological compartments was measurable in ~95% of subjects 5 to 8 months PSO, indicating that durable immunity against secondary COVID-19 disease is a possibility in most individuals.

Materials and methods

Human subjects

The Institutional Review Boards of the University of California, San Diego (UCSD; 200236X) and the La Jolla Institute for Immunology (LJI; VD-214) approved the protocols used for blood collection for subjects with COVID-19 who donated at all sites other than Mount Sinai. The Icahn School of Medicine at Mount Sinai IRB approved the samples collected at this institution in New York City (IRB-16-00791). All human subjects were assessed for medical decision-making capacity using a standardized, approved assessment and voluntarily gave informed consent before being enrolled in the study. Study inclusion criteria included a diagnosis of COVID-19 or suspected COVID-19, age of 18 years or greater, and willingness and ability to provide informed consent. Although not a strict inclusion criterion, evidence of positive polymerase chain reaction (PCR)–based testing for SARS-CoV-2 was requested from subjects before participation. A total of 145 cases were confirmed SARS-CoV-2 positive by PCR-based testing (Table 1). Two subjects tested negative by SARS-CoV-2 PCR (Table 1). The remainder were not tested or did not have test results available for review (Table 1). Subjects who had a medical history and/or symptoms consistent with COVID-19, but lacked positive PCR-based testing for SARS-CoV-2 and subsequently had negative laboratory-based serologic testing for SARS-CoV-2, were then excluded; i.e., all COVID-19 cases in this study were confirmed cases by SARS-CoV-2 PCR or SARS-CoV-2 serodiagnostics, or both. Adults of all races, ethnicities, ages, and genders were eligible to participate. Study exclusion criteria included lack of willingness to participate, lack of ability to provide informed consent, or a medical contraindication to blood donation (e.g., severe anemia). Subject samples at LJI were obtained from

individuals in California and at least seven other states.

Blood collection and processing methods at LJI were performed as previously described (*5*). Briefly, whole blood was collected via phlebotomy in acid citrate dextrose (ACD) serum separator tubes (SST) or ethylenediaminetetraacetic acid (EDTA) tubes and processed for peripheral blood mononuclear cells (PBMCs), serum, and plasma isolation. Most donors were screened for symptoms prior to scheduling blood draws and had to be symptom-free and approximately 3 to 4 weeks out from symptom onset at the time of the initial blood draw at UCSD or LJI, respectively. Samples were coded, and then deidentified before analysis. Other efforts to maintain the confidentiality of participants included the labeling samples with coded identification numbers. An overview of the characteristics of subjects with COVID-19 is provided in Table 1.

COVID-19 disease severity was scored from 0 to 10 using a numerical scoring system based on the NIH ordinal scale (*5*, *80*). A categorical descriptor was applied based on this scoring system: "asymptomatic" for a score of 1, "mild" for a score of 2 to 3, "moderate" for a score of 4 to 5, and "severe" for a score of 6 or more. Subjects with a numerical score of 4 or higher required hospitalization (including admission for observation) for management of COVID-19. Only one of 13 hospitalized subjects is shared from the previous study of acute COVID-19 (*5*). The days PSO was determined based on the difference between the date of the blood collection and the date of first reported symptoms consistent with COVID-19. For asymptomatic subjects, the day from first positive SARS-CoV-2 PCR-based testing was used in place of the date of first reported COVID-19 symptoms.

Recombinant proteins

Stabilized spike protein [2P (*81*)] and the RBD were expressed in HEK293F cells. Briefly, DNA expressing stabilized spike protein and RBD were subcloned into separate phCMV vectors and transfected into HEK293F cells at a ratio of 1 mg of DNA to 1 liter of cells. The cells were cultured at 37°C in a shaker incubator set to 125 rpm, 80% humidity, and 8% CO_2. When cell viability dropped below 80% (typically 4 to 5 days), media was harvested and centrifuged to remove cells. Biolock reagent was added to the supernatant media to remove any excess biotin. The media was then filtered through a 0.22-μm filter to remove Biolocked-aggregates. Proteins were purified using Streptrap HP 5 ml columns (Cytiva) using 100 mM Tris, 100 mM NaCl as the wash buffer and 100 mM Tris, 100 mM NaCl, 2.5 mM d-Desthiobiotin as the elution buffer. The eluted fractions for spike proteins were concentrated on 100-kDa Amicon filters and the RBDs were concentrated on 10-kDa filters. The samples were further purified using S6increase columns for the spike variants and S200increase column for the RBD.

SARS-CoV-2 ELISAs

SARS-CoV-2 ELISAs were performed as previously described (*2*, *5*, *82*). Briefly, Corning 96-well half-area plates (ThermoFisher 3690) were coated with 1 μg/ml of antigen overnight at 4°C. Antigens included recombinant SARS-CoV-2 RBD protein, recombinant spike protein, and recombinant nucleocapsid protein (GenScript Z03488) [recombinant nucleocapsid antigens were also tested from Sino Biological (40588-V07E) and Invivogen (his-sars2-n) and yielded comparable results to GenScript nucleocapsid]. The following day, plates were blocked with 3% milk in phosphate-buffered saline (PBS) containing 0.05% Tween-20 for 1.5 hours at room temperature. Plasma was heat inactivated at 56°C for 30 to 60 min. Plasma was diluted in 1% milk containing 0.05% Tween-20 in PBS starting at a 1:3 dilution followed by serial dilutions by three and incubated for 1.5 hours at room temperature. Plates were washed five times with 0.05% PBS-Tween-20. Secondary antibodies were diluted in 1% milk containing 0.05% Tween-20 in PBS. For IgG, anti-human IgG peroxidase antibody produced in goat (Sigma A6029) was used at a 1:5,000 dilution. For IgA, anti-human IgA horseradish peroxidase antibody (Hybridoma Reagent Laboratory HP6123-HRP) was used at a 1:1,000 dilution. The HP6123 monoclonal anti-IgA was used because of its CDC- and WHO-validated specificity for human IgA1 and IgA2 and lack of cross-reactivity with non-IgA isotypes (*83*).

End-point titers were plotted for each sample, using background-subtracted data. Negative and positive controls were used to standardize each assay and normalize across experiments. A positive control standard was created by pooling plasma from six convalescent COVID-19 donors to normalize between experiments. The limit of detection (LOD) was defined as 1:3 for IgG, 1:10 for IgA. Limit of sensitivity (LOS) for SARS-CoV-2–infected individuals was established on the basis of uninfected subjects, using plasma from normal healthy donors never exposed to SARS-CoV-2. For cross-sectional analyses, modeling for the best fit curve (e.g., one phase decay versus simple linear regression) was performed using GraphPad Prism 8.0. Best curve fit was defined by an extra sum-of-squares F Test, selecting the simpler model unless $P < 0.05$ (*84*). Continuous decay (linear regression), one-phased decay, or two-phased decay of log data were assessed in all cases, with the best fitting statistical model chosen on the basis of the F test; in several cases, a quadratic equation fit was also considered. To calculate the $t_{1/2}$, $\log_2$-transformed data were utilized. Using the best fit curve, either a one-phase decay nonlinear fit or a simple linear regression (continuous decay) was utilized. For simple linear regressions, Pearson's R was calculated for correlation using $\log_2$-transformed data. For one-phase decay nonlinear fit, R was reported. For longitudinal samples, a simple linear regression was performed, with $t_{1/2}$ calculated from $\log_2$-transformed data for each pair. For gender analyses, modeling and $t_{1/2}$ were performed similar to cross-sectional analyses; ANCOVA (VassarStats or GraphPad Prism 8.4) was then performed between male and female data sets. ANCOVA p-values of the adjusted means were reported and considered significant if the test for homogeneity of regressions was not significant.

Neutralizing antibody assays

The pseudovirus-neutralizing antibody assay was performed as previously described (*5*). Briefly, Vero cells were seeded in 96-well plates to produce a monolayer at the time of infection. Pretitrated amounts of rVSV-SARS-Cov-2 [phCMV3-SARS-CoV-2 spike SARS-CoV-2-pseudotyped VSV-ΔG-GFP (green fluorescent protein) were generated by transfecting HEK293T cells, ATCC CRL-3216] were incubated with serially diluted human plasma at 37°C for 1 hour before addition to confluent Vero cell monolayers (ATCC CCL-81) in 96-well plates. Cells were incubated for 12 to 16 hours at 37°C in 5% CO_2. Cells were then fixed in 4% paraformaldehyde, stained with 1 μg/ml Hoechst, and imaged using a CellInsight CX5 imager to quantify the total number of cells expressing GFP. Infection was normalized to the average number of cells infected with rVSV-SARS-CoV-2 incubated with normal human plasma. The LOD was established as <1:20 on the basis of plasma samples from a series of unexposed control subjects. Negative signals were set to 1:19. Neutralization IC_{50} (median inhibitory concentration) titers were calculated using One-Site Fit LogIC50 regression in GraphPad Prism 8.0.

Detection of antigen-specific memory B cells

To detect SARS-CoV-2 specific B cells, biotinylated protein antigens were individually multimerized with fluorescently labeled streptavidin at 4°C for 1 hour. Full-length SARS-CoV-2 spike (2P-stabilized, double Strep-tagged) and RBD were generated in-house. Biotinylation was performed using biotin-protein ligase standard reaction kit (Avidity, catalog no. Bir500A) following the manufacturer's standard protocol and dialyzed overnight against PBS. Biotinylated spike was mixed with streptavidin BV421 (BioLegend, catalog no. 405225) and streptavidin Alexa Fluor 647 (Thermo Fisher Scientific, catalog no. S21374) at 20:1 ratio (~6:1 molar ratio). Biotinylated RBD was

mixed with streptavidin phycoerythrin (PE)/Cyanine7 (BioLegend, catalog no. 405206) at 2.2:1 ratio (~4:1 molar ratio). Biotinylated SARS-CoV-2 full-length nucleocapsid (Avi- and His-tagged; Sino Biological, catalog no. 40588-V27B-B) was multimerized using streptavidin PE (BioLegend, catalog no. 405204) and streptavidin BV711 (BioLegend, catalog no. 405241) at 5.5:1 ratio (~6:1 molar ratio). Streptavidin PE/Cyanine5.5 (Thermo Fisher Scientific, catalog no. SA1018) was used as a decoy probe to gate out SARS-CoV-2 non-specific streptavidin-binding B cells. The antigen probes prepared individually as above were then mixed in Brilliant Buffer (BD Bioscience, catalog no. 566349) containing 5 μM free d-biotin (Avidity, catalog no. Bir500A). Free d-biotin ensured minimal cross-reactivity of antigen probes. About 10^7 previously frozen PBMC samples were prepared in U-bottom 96-well plates and stained with 50 μl of antigen probe cocktail containing 100 ng of spike per probe (total 200ng), 27.5 ng of RBD, 40 ng of nucleocapsid per probe (total 80 ng), and 20 ng of streptavidin PE/Cyanine5.5 at 4°C for 1 hour to ensure maximal staining quality before surface staining with antibodies as listed in table S1 was performed in Brilliant Buffer at 4°C for 30 min. Dead cells were stained using LIVE/DEAD Fixable Blue Stain Kit (Thermo Fisher Scientific, catalog no. L34962) in DPBS at 4°C for 30 min. About 80% of antigen-specific memory (IgD^- and/or $CD27^+$) B cells detected using this method were IgM^+, IgG^+, or IgM^- IgG^- IgA^+, which were comparable to nonspecific memory B cells. On the basisof these observations, we concluded that the antigen probes did not substantially affect the quality of surface immunoglobulin staining. Stained PBMC samples were acquired on Cytek Aurora and analyzed using FlowJo10.7.1 (BD Bioscience).

The frequency of antigen-specific memory B cells was expressed as a percentage of total B cells ($CD19^+$ $CD20^+$ $CD38^{int/-}$, $CD3^-$, $CD14^-$, $CD16^-$, $CD56^-$, LIVE/$DEAD^-$, lymphocytes), or as number per 10^6 PBMCs (LIVE/$DEAD^-$ cells). LOD was set on the basis of median + 2 × standard deviation (SD) of [1/(number of total B cells recorded)] or median + 2 × SD of [10^6/(number of PBMCs recorded)]. LOS was set as the median + 2 × SD of the results in unexposed donors. Phenotype analysis of antigen-specific B cells was performed only in subjects with at least 10 cells detected in the respective antigen-specific memory B cell gate. In each experiment, PBMCs from a known positive control (COVID-19 convalescent subject) and unexposed subjects were included to ensure consistent sensitivity and specificity of the assay. For each data set, second-order polynomial, simple linear regression, and pseudo–first- order kinetic models were considered. The model with a lower Akaike's information criterion value was determined to be a better fit and visualized.

Activation-induced markers (AIM) T cell assay

Antigen-specific $CD4^+$ T cells were measured as a percentage of AIM^+ ($OX40^+CD137^+$) $CD4^+$ T and ($CD69^+CD137^+$) $CD8^+$ T cells after stimulation of PBMCs with overlapping peptide megapools (MPs) spanning the entire SARS-CoV-2 ORFeome, as previously described (*2*). Cells were cultured for 24 hours in the presence of SARS-CoV-2–specific MPs (1 μg/ml) or 5 μg/ml phytohemagglutinin (PHA, Roche) in 96-wells U-bottom plates at 1×10^6 PBMCs per well. Stimulation with an equimolar amount of dimethyl sulfoxide (DMSO) was performed as a negative control. PHA and stimulation with a combined $CD4^+$ and $CD8^+$ cytomegalovirus epitope MP (CMV, 1 μg/ml) were included as positive controls. Any sample with a low PHA signal was excluded as a quality control.

Antigen-specific $CD4^+$ and $CD8^+$ T cells were measured as background (DMSO)–subtracted data, with a minimal DMSO level set to 0.005%. All positive ORFs (>0.02% for $CD4^+$, >0.05% for $CD8^+$) were then aggregated into a combined sum of SARS-CoV-2–specific $CD4^+$ or $CD8^+$ T cells. The threshold for positivity for antigen-specific $CD4^+$ T cell responses (0.03%) and antigen-specific $CD8^+$ T cell responses (0.12%) was calculated using the median twofold standard deviation of all negative controls measured (>150). The antibody panel utilized in the ($OX40^+CD137^+$) $CD4^+$ T and ($CD69^+$ $CD137^+$) $CD8^+$ T cells AIM staining is shown in table S2. A consistency analysis was performed for multiple measurements of AIM T cell assays by two different operators. Before merging, we compared the protein immunodominance, total SARS-CoV-2–specific $CD4^+$ and $CD8^+$ T cell responses, and half-life calculations between the two groups of experimental data. In longitudinal analyses, half-life calculations excluded any samples that were negative at both time points (because a half-life could not be calculated), though all data were included in the graphs.

For surface $CD40L^+OX40^+$ $CD4^+$ T cell AIM assays, experiments were performed as previously described (*5*), with the following modifications. Cells were cultured in complete RPMI containing 5% human AB serum (Gemini Bioproducts), β-mercaptoethanol, penicillin/streptomycin, sodium pyruvate (NaPy), and nonessential amino acids. Prior to addition of peptide MPs, cells were blocked at 37°C for 15 min with 0.5 μg/ml of anti-CD40 mAb (Miltenyi Biotec). A stimulation with an equimolar amount of DMSO was performed to determine background subtraction, and activation from staphylococcal enterotoxin B (SEB) at 1 μg/ml was used as a (positive) quality control. LOD for antigen-specific cT_{FH} among $CD4^+$ T cells was based on the LOD for antigen-specific $CD4^+$ T cells (described above) multiplied by the average % cT_{FH} in the bulk CD4 T cells among control samples. An inclusion threshold of ten events after the cT_{FH} $CXCR5^+$ gate was used for PD-1^{hi} and $CCR6^+$ calculations, and Mann-Whitney nonparametric and Wilcoxon signed-rank statistical tests were applied for the respective comparisons.

REFERENCES AND NOTES

1. D. S. Stephens, M. J. McElrath, COVID-19 and the Path to Immunity. *JAMA* **324**, 1279–1281 (2020). doi: 10.1001/jama.2020.16656; pmid: 32915201
2. A. Grifoni *et al*., Targets of T Cell Responses to SARS-CoV-2 Coronavirus in Humans with COVID-19 Disease and Unexposed Individuals. *Cell* **181**, 1489–1501.e15 (2020). doi: 10.1016/j.cell.2020.05.015; pmid: 32473127
3. F. Krammer, SARS-CoV-2 vaccines in development. *Nature* **586**, 516–527 (2020). doi: 10.1038/s41586-020-2798-3; pmid: 32967006
4. M. S. Suthar *et al*., Rapid generation of neutralizing antibody responses in COVID-19 patients. *Cell Rep. Med.* **1**, 100040 (2020). doi: 10.1016/j.xcrm.2020.100040; pmid: 32835303
5. C. Rydyznski Moderbacher *et al*., Antigen-specific adaptive immunity to SARS-CoV-2 in acute COVID-19 and associations with age and disease severity. *Cell* **183**, 996–1012.e19 (2020). doi: 10.1016/j.cell.2020.09.038; pmid: 33010815
6. R. Zhou *et al*., Acute SARS-CoV-2 infection impairs dendritic cell and T cell responses. *Immunity* **53**, 864–877.e5 (2020). doi: 10.1016/j.immuni.2020.07.026; pmid: 32791036
7. M. Liao *et al*., Single-cell landscape of bronchoalveolar immune cells in patients with COVID-19. *Nat. Med.* **26**, 842–844 (2020). doi: 10.1038/s41591-020-0901-9; pmid: 32398875
8. A. G. Laing *et al*., A dynamic COVID-19 immune signature includes associations with poor prognosis. *Nat. Med.* **26**, 1623–1635 (2020). doi: 10.1038/s41591-020-1038-6; pmid: 32807934
9. D. Blanco-Melo *et al*., Imbalanced Host Response to SARS-CoV-2 Drives Development of COVID-19. *Cell* **181**, 1036–1045.e9 (2020). doi: 10.1016/j.cell.2020.04.026; pmid: 32416070
10. P. S. Arunachalam *et al*., Systems biological assessment of immunity to mild versus severe COVID-19 infection in humans. *Science* **369**, 1210–1220 (2020). doi: 10.1126/science.abc6261; pmid: 32788292
11. P. Bastard *et al*., Autoantibodies against type I IFNs in patients with life-threatening COVID-19. *Science* **370**, eabd4585 (2020). doi: 10.1126/science.abd4585; pmid: 32972996
12. Q. Zhang *et al*., Inborn errors of type I IFN immunity in patients with life-threatening COVID-19. *Science* **370**, eabd4570 (2020). doi: 10.1126/science.abd4570; pmid: 32972995
13. D. M. Del Valle *et al*., An inflammatory cytokine signature predicts COVID-19 severity and survival. *Nat. Med.* **26**, 1636–1643 (2020). doi: 10.1038/s41591-020-1051-9; pmid: 32839624
14. L. Kuri-Cervantes *et al*., Comprehensive mapping of immune perturbations associated with severe COVID-19. *Sci. Immunol.* **5**, eabd7114 (2020). doi: 10.1126/sciimmunol.abd7114; pmid: 32669287
15. S. Li *et al*., Clinical and pathological investigation of patients with severe COVID-19. *JCI Insight* **5**, (2020). doi: 10.1172/jci.insight.138070; pmid: 32427582
16. C. Radermecker *et al*., Neutrophil extracellular traps infiltrate the lung airway, interstitial, and vascular compartments in severe COVID-19. *J. Exp. Med.* **217**, e20201012 (2020). doi: 10.1084/jem.20201012; pmid: 32926097
17. B. Schurink *et al*., Viral presence and immunopathology in patients with lethal COVID-19: A prospective autopsy cohort study. *Lancet Microbe* **1**, e290–e299 (2020). doi: 10.1016/S2666-5247(20)30144-0; pmid: 33015653
18. M. Aid *et al*., Vascular Disease and Thrombosis in SARS-CoV-2-Infected Rhesus Macaques. *Cell* **183**, 1354–1366.e13 (2020). doi: 10.1016/j.cell.2020.10.005; pmid: 33065030
19. N. Baumgarth, J. Nikolich-Žugich, F. E.-H. Lee, D. Bhattacharya, Antibody Responses to SARS-CoV-2: Let's Stick to Known Knowns. *J. Immunol.* **205**, 2342–2350 (2020). doi: 10.4049/jimmunol.2000839; pmid: 32887754
20. A. Wajnberg *et al*., Robust neutralizing antibodies to SARS-CoV-2 infection persist for months. *Science* **370**, 1227–1230 (2020). doi: 10.1126/science.abd7728; pmid: 33115920

21. A. Sariol, S. Perlman, Lessons for COVID-19 immunity from other coronavirus infections. *Immunity* **53**, 248–263 (2020). doi: 10.1016/j.immuni.2020.07.005; pmid: 32717182
22. K. Subbarao, SARS-CoV-2: A New Song Recalls an Old Melody. *Cell Host Microbe* **27**, 692–694 (2020). doi: 10.1016/j.chom.2020.04.019; pmid: 32407706
23. W. Deng *et al*., Primary exposure to SARS-CoV-2 protects against reinfection in rhesus macaques. *Science* **369**, 818–823 (2020). doi: 10.1126/science.abc5343; pmid: 32616673
24. Q. Gao *et al*., Development of an inactivated vaccine candidate for SARS-CoV-2. *Science* **369**, 77–81 (2020). doi: 10.1126/science.abc1932; pmid: 32376603
25. A. Chandrashekar *et al*., SARS-CoV-2 infection protects against rechallenge in rhesus macaques. *Science* **369**, 812–817 (2020). doi: 10.1126/science.abc4776; pmid: 32434946
26. S. J. Zost *et al*., Potently neutralizing and protective human antibodies against SARS-CoV-2. *Nature* **584**, 443–449 (2020). doi: 10.1038/s41586-020-2548-6; pmid: 32668443
27. T. F. Rogers *et al*., Isolation of potent SARS-CoV-2 neutralizing antibodies and protection from disease in a small animal model. *Science* **369**, 956–963 (2020). doi: 10.1126/science.abc7520; pmid: 32540903
28. A. Baum *et al*., REGN-COV2 antibodies prevent and treat SARS-CoV-2 infection in rhesus macaques and hamsters. *Science* **370**, 1110–1115 (2020). doi: 10.1126/science.abe2402; pmid: 33037066
29. D. M. Weinreich *et al*., REGN-COV2, a Neutralizing Antibody Cocktail, in Outpatients with Covid-19. *N. Engl. J. Med.* NEJMoa2035002 (2020). doi: 10.1056/NEJMoa2035002; pmid: 33332778
30. Lilly announces proof of concept data for neutralizing antibody LY-CoV555 in the COVID-19 outpatient setting | Eli Lilly and Company; https://investor.lilly.com/news-releases/news-release-details/lilly-announces-proof-concept-data-neutralizing-antibody-ly.
31. D. M. Altmann, R. J. Boyton, SARS-CoV-2 T cell immunity: Specificity, function, durability, and role in protection. *Sci. Immunol.* **5**, eabd6160 (2020). doi: 10.1126/sciimmunol.abd6160; pmid: 32680954
32. P. Van Damme, K. Van Herck, A review of the long-term protection after hepatitis A and B vaccination. *Travel Med. Infect. Dis.* **5**, 79–84 (2007). doi: 10.1016/j.tmaid.2006.04.004; pmid: 17298912
33. M. M. Rosado *et al*., Switched memory B cells maintain specific memory independently of serum antibodies: The hepatitis B example. *Eur. J. Immunol.* **41**, 1800–1808 (2011). doi: 10.1002/eji.201041187; pmid: 21469123
34. F. Zhou *et al*., Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: A retrospective cohort study. *Lancet* **395**, 1054–1062 (2020). doi: 10.1016/S0140-6736(20)30566-3; pmid: 32171076
35. W. A. Orenstein, R. Ahmed, Simply put: Vaccination saves lives. *Proc. Natl. Acad. Sci. U.S.A.* **114**, 4031–4033 (2017). doi: 10.1073/pnas.1704507114; pmid: 28396427
36. P. Piot *et al*., Immunization: Vital progress, unfinished agenda. *Nature* **575**, 119–129 (2019). doi: 10.1038/s41586-019-1656-7; pmid: 31695203
37. S. Plotkin, W. Orenstein, P. Offit, *Plotkin's Vaccines* (Elsevier, ed. 7, 2018).
38. F. Sallusto, A. Lanzavecchia, K. Araki, R. Ahmed, From vaccines to memory and back. *Immunity* **33**, 451–463 (2010). doi: 10.1016/j.immuni.2010.10.008; pmid: 21029957
39. S. Crotty, R. Ahmed, Immunological memory in humans. *Semin. Immunol.* **16**, 197–203 (2004). doi: 10.1016/j.smim.2004.02.008; pmid: 15130504
40. F. Weisel, M. Shlomchik, Memory B Cells of Mice and Humans. *Annu. Rev. Immunol.* **35**, 255–284 (2017). doi: 10.1146/annurev-immunol-041015-055531; pmid: 28142324
41. S. M. Kissler, C. Tedijanto, E. Goldstein, Y. H. Grad, M. Lipsitch, Projecting the transmission dynamics of SARS-CoV-2 through the postpandemic period. *Science* **368**, 860–868 (2020). doi: 10.1126/science.abb5793; pmid: 32291278
42. C. M. Saad-Roy *et al*., Immune life history, vaccination, and the dynamics of SARS-CoV-2 over the next 5 years. *Science* **370**, 811–818 (2020). doi: 10.1126/science.abd7343; pmid: 32958581
43. Q.-X. Long *et al*., Clinical and immunological assessment of asymptomatic SARS-CoV-2 infections. *Nat. Med.* **26**, 1200–1204 (2020). doi: 10.1038/s41591-020-0965-6; pmid: 32555424
44. D. F. Gudbjartsson *et al*., Humoral Immune Response to SARS-CoV-2 in Iceland. *N. Engl. J. Med.* **383**, 1724–1734 (2020). doi: 10.1056/NEJMoa2026116; pmid: 32871063
45. A. Wajnberg *et al*., Robust neutralizing antibodies to SARS-CoV-2 infection persist for months. *Science* **370**, 1227–1230 (2020). doi: 10.1126/science.abd7728; pmid: 33115920
46. L. Piccoli *et al*., Mapping neutralizing and immunodominant sites on the SARS-CoV-2 spike receptor-binding domain by structure-guided high-resolution serology. *Cell* **183**, 1024–1042.e21 (2020). doi: 10.1016/j.cell.2020.09.037; pmid: 32991844
47. D. F. Robbiani *et al*., Convergent antibody responses to SARS-CoV-2 in convalescent individuals. *Nature* **584**, 437–442 (2020). doi: 10.1038/s41586-020-2456-9; pmid: 32555388
48. J. Yu *et al*., DNA vaccine protection against SARS-CoV-2 in rhesus macaques. *Science* **369**, 806–811 (2020). doi: 10.1126/science.abc6284; pmid: 32434945
49. B. Isho *et al*., Persistence of serum and saliva antibody responses to SARS-CoV-2 spike antigens in COVID-19 patients. *Sci. Immunol.* **5**, eabe5511 (2020). doi: 10.1126/sciimmunol.abe5511; pmid: 33033173
50. L. B. Rodda *et al*., Functional SARS-CoV-2-specific immune memory persists after mild COVID-19. *Cell* S0092-8674(20)31565-8 (2020). doi: 10.1016/j.cell.2020.11.029; pmid: 33296701
51. C. W. Davis *et al*., Longitudinal Analysis of the Human B Cell Response to Ebola Virus Infection. *Cell* **177**, 1566–1582.e17 (2019). doi: 10.1016/j.cell.2019.04.036; pmid: 31104840
52. A. Z. Wec *et al*., Longitudinal dynamics of the human B cell response to the yellow fever 17D vaccine. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 6675–6685 (2020). doi: 10.1073/pnas.1921388117; pmid: 32152119
53. M. F. Kotturi *et al*., The CD8+ T-cell response to lymphocytic choriomeningitis virus involves the L antigen: Uncovering new tricks for an old virus. *J. Virol.* **81**, 4928–4940 (2007). doi: 10.1128/JVI.02632-06; pmid: 17329346
54. J. Neidleman *et al*., SARS-CoV-2-specific T cells exhibit phenotypic features of helper function, lack of terminal differentiation, and high proliferation potential. *Cell Rep. Med.* **1**, 100081 (2020). doi: 10.1016/j.xcrm.2020.100081; pmid: 32839763
55. S. Sridhar *et al*., Cellular immune correlates of protection against symptomatic pandemic influenza. *Nat. Med.* **19**, 1305–1312 (2013). doi: 10.1038/nm.3350; pmid: 24056771
56. R. S. Akondy *et al*., Origin and differentiation of human memory CD8 T cells after vaccination. *Nature* **552**, 362–367 (2017). doi: 10.1038/nature24633; pmid: 29236685
57. S. Crotty, T Follicular Helper Cell Biology: A Decade of Discovery and Diseases. *Immunity* **50**, 1132–1148 (2019). doi: 10.1016/j.immuni.2019.04.011; pmid: 31117010
58. J. A. Juno *et al*., Humoral and circulating follicular helper T cell responses in recovered patients with COVID-19. *Nat. Med.* **26**, 1428–1434 (2020). doi: 10.1038/s41591-020-0995-0; pmid: 32661393
59. A. T. Tan *et al*., Early induction of SARS-CoV-2 specific T cells associates with rapid viral clearance and mild disease in COVID-19 patients. *bioRxiv* 2020.10.15.341958 [Preprint]. 16 October 2020. https://doi.org/10.1101/2020.10.15.341958
60. C. Gaebler *et al*., Evolution of Antibody Immunity to SARS-CoV-2. *bioRxiv* 2020.11.03.367391 (2020). doi: 10.1101/2020.11.03.367391; pmid: 33173867
61. S. Crotty *et al*., Cutting edge: Long-term B cell memory in humans after smallpox vaccination. *J. Immunol.* **171**, 4969–4973 (2003). doi: 10.4049/jimmunol.171.10.4969; pmid: 14607890
62. X. Yu *et al*., Neutralizing antibodies derived from the B cells of 1918 influenza pandemic survivors. *Nature* **455**, 532–536 (2008). doi: 10.1038/nature07231; pmid: 18716625
63. J. Zuo *et al*., Robust SARS-CoV-2-specific T-cell immunity is maintained at 6 months following primary infection. *bioRxiv* 2020.11.01.362319 [Preprint]. doi: 10.1101/2020.11.01.362319
64. E. Hammarlund *et al*., Duration of antiviral immunity after smallpox vaccination. *Nat. Med.* **9**, 1131–1137 (2003). doi: 10.1038/nm917; pmid: 12925846
65. N. Le Bert *et al*., SARS-CoV-2-specific T cell immunity in cases of COVID-19 and SARS, and uninfected controls. *Nature* **584**, 457–462 (2020). doi: 10.1038/s41586-020-2550-z; pmid: 32668444
66. N. B. Mercado *et al*., Single-shot Ad26 vaccine protects against SARS-CoV-2 in rhesus macaques. *Nature* **586**, 583–588 (2020). doi: 10.1038/s41586-020-2607-z; pmid: 32731257
67. K. S. Corbett *et al*., Evaluation of the mRNA-1273 Vaccine against SARS-CoV-2 in Nonhuman Primates. *N. Engl. J. Med.* **383**, 1544–1555 (2020). doi: 10.1056/NEJMoa2024671; pmid: 32722908
68. L. Corey, J. R. Mascola, A. S. Fauci, F. S. Collins, A strategic approach to COVID-19 vaccine R&D. *Science* **368**, 948–950 (2020). doi: 10.1126/science.abc5312; pmid: 32393526
69. J. Zhao *et al*., Airway Memory CD4$^+$ T Cells Mediate Protective Immunity against Emerging Respiratory Coronaviruses. *Immunity* **44**, 1379–1391 (2016). doi: 10.1016/j.immuni.2016.05.006; pmid: 27287409
70. W. E. Purtha, T. F. Tedder, S. Johnson, D. Bhattacharya, M. S. Diamond, Memory B cells, but not long-lived plasma cells, possess antigen specificities for viral escape mutants. *J. Exp. Med.* **208**, 2599–2606 (2011). doi: 10.1084/jem.20110740; pmid: 22162833
71. D. Masopust, A. G. Soerens, Tissue-Resident T Cells and Other Resident Leukocytes. *Annu. Rev. Immunol.* **37**, 521–546 (2019). doi: 10.1146/annurev-immunol-042617-053214; pmid: 30726153
72. H. C. Whittle *et al*., Effect of subclinical infection on maintaining immunity against measles in vaccinated children in West Africa. *Lancet* **353**, 98–102 (1999). doi: 10.1016/S0140-6736(98)02364-2; pmid: 10023894
73. S. A. Plotkin, Vaccines: Correlates of vaccine-induced immunity. *Clin. Infect. Dis.* **47**, 401–409 (2008). doi: 10.1086/589862; pmid: 18558875
74. N. Burdin, L. K. Handy, S. A. Plotkin, What Is Wrong with Pertussis Vaccine Immunity? The Problem of Waning Effectiveness of Pertussis Vaccines. *Cold Spring Harb. Perspect. Biol.* **9**, a029454 (2017). doi: 10.1101/cshperspect.a029454; pmid: 28289064
75. N. van Doremalen *et al*., ChAdOx1 nCoV-19 vaccine prevents SARS-CoV-2 pneumonia in rhesus macaques. *Nature* **586**, 578–582 (2020). doi: 10.1038/s41586-020-2608-y; pmid: 32731258
76. R. L. Tillett *et al*., Genomic evidence for reinfection with SARS-CoV-2: A case study. *Lancet Infect. Dis.* **21**, 52–58 (2021). doi: 10.1016/S1473-3099(20)30764-7; pmid: 33058797
77. K. K.-W. To *et al*., COVID-19 re-infection by a phylogenetically distinct SARS-coronavirus-2 strain confirmed by whole genome sequencing. *Clin. Infect. Dis.* ciaa1275 (2020). doi: 10.1093/cid/ciaa1275; pmid: 32840608
78. D. Wyllie *et al*., SARS-CoV-2 responsive T cell numbers are associated with protection from COVID-19: A prospective cohort study in keyworkers. *medRxiv* 2020.11.02.20222778 [Preprint]. 4 November 2020. doi: 10.1101/2020.11.02.20222778
79. S. F. Lumley *et al*., Antibodies to SARS-CoV-2 are associated with protection against reinfection. *medRxiv* 2020.11.18.20234369 [Preprint]. 19 November 2020. doi: 10.1101/2020.11.18.20234369
80. J. H. Beigel *et al*., Remdesivir for the Treatment of Covid-19 - Final Report. *N. Engl. J. Med.* **383**, 1813–1826 (2020). doi: 10.1056/NEJMoa2007764; pmid: 32445440
81. D. Wrapp *et al*., Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science* **367**, 1260–1263 (2020). doi: 10.1126/science.abb2507; pmid: 32075877
82. F. Amanat *et al*., A serological assay to detect SARS-CoV-2 seroconversion in humans. *Nat. Med.* **26**, 1033–1036 (2020). doi: 10.1038/s41591-020-0913-5; pmid: 32398876
83. J. Mestecky *et al*., Evaluation of monoclonal antibodies with specificity for human IgA, IgA subclasses and allotypes and secretory component. Results of an IUIS/WHO collaborative study. *J. Immunol. Methods* **193**, 103–148 (1996). doi: 10.1016/0022-1759(95)00289-8; pmid: 8699027
84. GraphPad, GraphPad Prism 8 Curve Fitting Guide (2020); www.graphpad.com/guides/prism/8/curve-fitting/index.htm.

ACKNOWLEDGMENTS

We thank the LJI Clinical Core, specifically G. Levi and B. Schwan for healthy donor enrollment and blood sample procurement. We thank C. Moderbacher for input on data analysis. We are also grateful to the Mt. Sinai Personalized Virology Initiative for sharing banked samples from study participants with COVID-19. We are grateful to A. Wajnberg for study participant referrals and to the Personalized Virology Initiative (G. Kleiner, L. C. F. Mulder, M. Saksena, K. Srivastava, C. Gleason, C. M. Bermúdez-González, K. Beach, K. Russo, L. Sominsky, E. Ferreri, R. Chernet, L. Eaker, A. Salimbangon, D. Jurczyszak, H. Alshammary, W. Mendez, A. Amoako, S. Fabre, S. Suthakaran, M. Awawda, E. Hirsch, A. Shin) for sharing banked samples from study participants with COVID-19. **Funding:** This work was funded by the NIH NIAID under awards AI142742 (Cooperative Centers for Human Immunology) (A.S., S.C.), NIH contract no. 75N9301900065 (D.W., A.S.), U01 AI141995-03 (A.S., P.B.), and U01 CA260541-01 (D.W). This work

was additionally supported in part by LJI Institutional Funds, the John and Mary Tu Foundation (D.S.), the NIAID under K08 award AI135078 (J.M.D.), UCSD T32s AI007036 and AI007384 Infectious Diseases Division (S.I.R., S.A.R.), and the Bill and Melinda Gates Foundation INV-006133 from the Therapeutics Accelerator, Mastercard, Wellcome, private philanthropic contributions (K.M.H., E.O.S., S.C.), and a FastGrant from Emergent Ventures in aid of COVID-19 research. This work was partially supported by the NIAID Centers of Excellence for Influenza Research and Surveillance (CEIRS) contract HHSN272201400008C (F.K., for reagent generation), the Collaborative Influenza Vaccine Innovation Centers (CIVIC) contract 75N93019C00051, the JPB foundation (F.K., V.S.), the Cohen Foundation (V.S., F.K.), and the Open Philanthropy Project (no. 2020-215611; F.K., V.S.), as well as by other philanthropic donations. We also thank all of the COVID-19 and healthy human subjects who made this research possible through their generous blood donations. **Author contributions:** Conceptualization, S.C., A.S., and D.W.; Investigation, J.M.D., J.M., Y.K., K.M.H., E.D.Y., C.E.F., A.G., S.H., and C.N.; Formal Analysis, J.M.D., J.M., Y.K., K.M.H., C.E.F., S.H., B.P., D.W., A.S., and S.C.; Patient Recruitment and Samples, S.I.R., A.F., S.A.R., F.K., V.S., D.M.S., and D.W.; Material Resources, F.K., V.S., V.R., E.O.S., D.W., A.S., and S.C.; Data Curation, Y.K., J.M.D., J.M., and S.H.; Writing, Y.K., J.M.D., J.M., S.I.R., D.W., A.S., and S.C.; Supervision, D.W., A.S., and S.C., Project Administration, A.F. **Competing interests:** A.S. is a consultant for Gritstone, Flow Pharma, Merck, Epitogenesis, Gilead, and Avalia. S.C. is a consultant for Avalia. L.J.I. has filed for patent protection for various aspects of T cell epitope and vaccine design work. Mount Sinai has licensed serological assays to commercial entities and has filed for patent protection for serological assays. D.S., F.A., V.S., and F.K. are listed as inventors on the pending patent application (F.K., V.S.), and Newcastle disease virus (NDV)–based SARS-CoV-2 vaccines that name F.K. as inventor. All other authors declare no conflict of interest. **Data and materials availability:** All data are provided in the supplementary materials. Epitope pools used in this paper will be made available to the scientific community upon request and execution of a material transfer agreement. This work is licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0) license, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. To view a copy of this license, visit https://creativecommons.org/licenses/by/4.0/. This license does not apply to figures/photos/artwork or other content included in the article that is credited to a third party; obtain authorization from the rights holder before using such material.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/371/6529/eabf4063/suppl/DC1
Figs. S1 to S10
Tables S1 and S2
MDAR Reproducibility Checklist
Data File S1

View/request a protocol for this paper from *Bio-protocol*.

23 October 2020; accepted 23 December 2020
Published online 6 January 2021
10.1126/science.abf4063

Science

Immunological memory to SARS-CoV-2 assessed for up to 8 months after infection

Jennifer M. DanJose MateusYu KatoKathryn M. HastieEsther Dawen YuCaterina E. FalitiAlba GrifoniSydney I. RamirezSonya HauptApril FrazierCatherine NakaoVamseedhar RayaproluStephen A. RawlingsBjoern PetersFlorian KrammerViviana SimonErica Ollmann SaphireDavey M. SmithDaniela WeiskopfAlessandro SetteShane Crotty

Science, 371 (6529), eabf4063. • DOI: 10.1126/science.abf4063

Variable memory

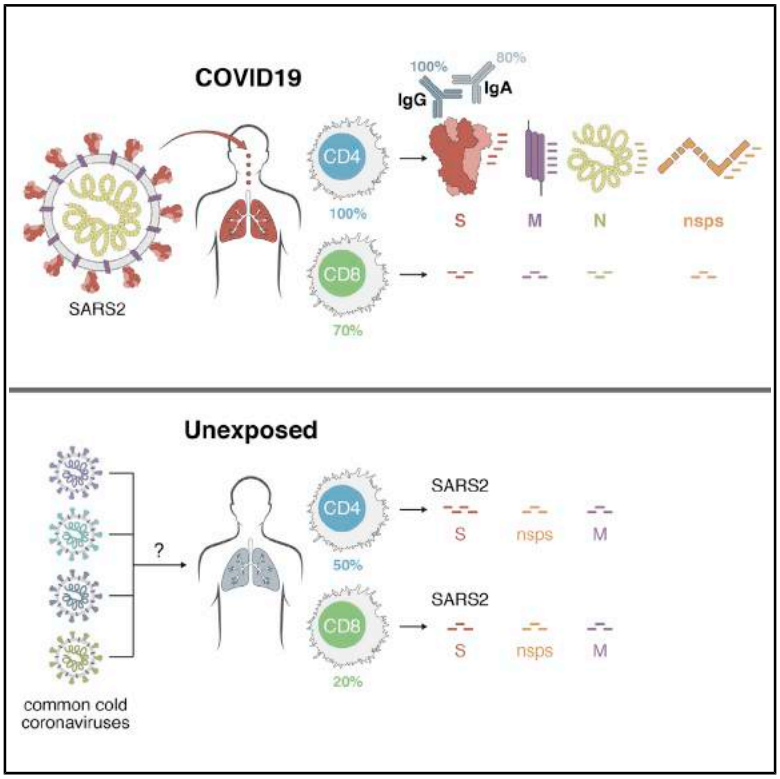
Immune memory against severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) helps to determine protection against reinfection, disease risk, and vaccine efficacy. Using 188 human cases across the range of severity of COVID-19, Dan *et al.* analyzed cross-sectional data describing the dynamics of SARS-CoV-2 memory B cells, CD8 T cells, and CD4 T cells for more than 6 months after infection. The authors found a high degree of heterogeneity in the magnitude of adaptive immune responses that persisted into the immune memory phase to the virus. However, immune memory in three immunological compartments remained measurable in greater than 90% of subjects for more than 5 months after infection. Despite the heterogeneity of immune responses, these results show that durable immunity against secondary COVID-19 disease is a possibility for most individuals.

Science, this issue p. eabf4063

Cell

Targets of T Cell Responses to SARS-CoV-2 Coronavirus in Humans with COVID-19 Disease and Unexposed Individuals

Graphical Abstract



Highlights

- Measuring immunity to SARS-CoV-2 is key for understanding COVID-19 and vaccine development
- Epitope pools detect $CD4^+$ and $CD8^+$ T cells in 100% and 70% of convalescent COVID patients
- T cell responses are focused not only on spike but also on M, N, and other ORFs
- T cell reactivity to SARS-CoV-2 epitopes is also detected in non-exposed individuals

Authors

Alba Grifoni, Daniela Weiskopf, Sydney I. Ramirez, ..., Davey M. Smith, Shane Crotty, Alessandro Sette

Correspondence

shane@lji.org (S.C.),
alex@lji.org (A.S.)

In Brief

An analysis of immune cell responses to SARS-CoV-2 from recovered patients identifies the regions of the virus that is targeted and also reveals cross-reactivity with other common circulating coronaviruses

Grifoni et al., 2020, Cell *181*, 1489–1501
June 25, 2020
https://doi.org/10.1016/j.cell.2020.05.015

Cell

CellPress

Article

Targets of T Cell Responses to SARS-CoV-2 Coronavirus in Humans with COVID-19 Disease and Unexposed Individuals

Alba Grifoni,[1] Daniela Weiskopf,[1] Sydney I. Ramirez,[1,2] Jose Mateus,[1] Jennifer M. Dan,[1,2] Carolyn Rydyznski Moderbacher,[1] Stephen A. Rawlings,[2] Aaron Sutherland,[1] Lakshmanane Premkumar,[3] Ramesh S. Jadi,[3] Daniel Marrama,[1] Aravinda M. de Silva,[3] April Frazier,[1] Aaron F. Carlin,[2] Jason A. Greenbaum,[1] Bjoern Peters,[1,2] Florian Krammer,[4] Davey M. Smith,[2] Shane Crotty,[1,2,5,*] and Alessandro Sette[1,2,5,6,*]

[1]Center for Infectious Disease and Vaccine Research, La Jolla Institute for Immunology, La Jolla, CA 92037, USA

[2]Department of Medicine, Division of Infectious Diseases and Global Public Health, University of California, San Diego, La Jolla, CA 92037, USA

[3]Department of Microbiology and Immunology, University of North Carolina School of Medicine, Chapel Hill, NC 27599-7290, USA

[4]Department of Microbiology, Icahn School of Medicine at Mount Sinai, New York, NY, USA

[5]These authors contributed equally

[6]Lead Contact

*Correspondence: shane@lji.org (S.C.), alex@lji.org (A.S.)

https://doi.org/10.1016/j.cell.2020.05.015

SUMMARY

Understanding adaptive immunity to SARS-CoV-2 is important for vaccine development, interpreting coronavirus disease 2019 (COVID-19) pathogenesis, and calibration of pandemic control measures. Using HLA class I and II predicted peptide "megapools," circulating SARS-CoV-2-specific $CD8^+$ and $CD4^+$ T cells were identified in ~70% and 100% of COVID-19 convalescent patients, respectively. $CD4^+$ T cell responses to spike, the main target of most vaccine efforts, were robust and correlated with the magnitude of the anti-SARS-CoV-2 IgG and IgA titers. The M, spike, and N proteins each accounted for 11%–27% of the total $CD4^+$ response, with additional responses commonly targeting nsp3, nsp4, ORF3a, and ORF8, among others. For $CD8^+$ T cells, spike and M were recognized, with at least eight SARS-CoV-2 ORFs targeted. Importantly, we detected SARS-CoV-2-reactive $CD4^+$ T cells in ~40%–60% of unexposed individuals, suggesting cross-reactive T cell recognition between circulating "common cold" coronaviruses and SARS-CoV-2.

INTRODUCTION

COVID-19 is a worldwide emergency. The first cases occurred in December 2019, and now more than 240,000 deaths and 3,000,000 cases of SARS-CoV-2 infection have been reported worldwide as of May 1st (Dong et al., 2020; Wu and McGoogan, 2020). Vaccines against SARS-CoV-2 are just beginning development (Amanat and Krammer, 2020; Thanh Le et al., 2020). An understanding of human T cell responses to SARS-CoV2 is lacking, due to the rapid emergence of the pandemic. There is an urgent need for foundational information about T cell responses to this virus.

The first steps for such an understanding are the ability to quantify the virus-specific $CD4^+$ and $CD8^+$ T cells. Such knowledge is of immediate relevance, as it will provide insights into immunity and pathogenesis of SARS-CoV-2 infection, and the same knowledge will assist vaccine design and evaluation of candidate vaccines. Estimations of immunity are also central to epidemiological model calibration of future social distancing pandemic control measures (Kissler et al., 2020). Such projections are dramatically different depending on whether SARS-CoV-2 infection creates substantial immunity, and whether any cross-reactive immunity exists between SARS-CoV-2 and circulating seasonal "common cold" human coronaviruses. Definition and assessment of human antigen-specific SARS-CoV2 T cell responses are best made with direct *ex vivo* T cell assays using broad-based epitope pools and assays capable of detecting T cells of any cytokine polarization. Herein, we have completed such an assessment with blood samples from COVID-19 patients.

There is also great uncertainty about whether adaptive immune responses to SARS-CoV-2 are protective or pathogenic, or whether both scenarios can occur depending on timing, composition, or magnitude of the adaptive immune response. Hypotheses range the full gamut (Peeples, 2020), based on available clinical data from severe acute respiratory disease syndrome (SARS) or Middle East respiratory syndrome (MERS) (Alshukairi et al., 2018; Wong et al., 2004; Zhao et al., 2017) or animal model data with SARS in mice (Zhao et al., 2009, 2010, 2016), SARS in non-human primates (NHPs) (Liu et al., 2019; Takano et al., 2008) or feline infectious peritonitis virus (FIPV) in cats

(Vennema et al., 1990). Protective immunity, immunopathogenesis, and vaccine development for COVID-19 are each briefly discussed below, related to introducing the importance of defining T cell responses to SARS-CoV-2.

Based on data from SARS patients in 2003–2004 (caused by SARS-CoV, the most closely related human betacoronavirus to SARS-CoV-2), and based on the fact that most acute viral infections result in development of protective immunity (Sallusto et al., 2010), a likely possibility has been that substantial $CD4^+$ T cell, $CD8^+$ T cell, and neutralizing antibody responses develop to SARS-CoV-2, and all contribute to clearance of the acute infection, and, as a corollary, some of the T and B cells are retained long term (i.e., multiple years) as immunological memory and protective immunity against SARS-CoV-2 infection (Guo et al., 2020b; Li et al., 2008). However, a contrarian viewpoint is also legitimate. While most acute infections result in the development of protective immunity, available data for human coronaviruses suggest the possibility that substantive adaptive immune responses can fail to occur (Choe et al., 2017; Okba et al., 2019; Zhao et al., 2017) and robust protective immunity can fail to develop (Callow et al., 1990). A failure to develop protective immunity could occur due to a T cell and/or antibody response of insufficient magnitude or durability, with the neutralizing antibody response being dependent on the $CD4^+$ T cell response (Crotty, 2019; Zhao et al., 2016). Thus, there is urgent need to understand the magnitude and composition of the human $CD4^+$ and $CD8^+$ T cell responses to SARS-CoV-2. If natural infection with SARS-CoV-2 elicits potent $CD4^+$ and $CD8^+$ T cell responses commonly associated with protective antiviral immunity, COVID-19 is a strong candidate for rapid vaccine development.

Immunopathogenesis in COVID-19 is a serious concern (Cao, 2020; Peeples, 2020). It is most likely that an early $CD4^+$ and $CD8^+$ T cell response against SARS-CoV-2 is protective, but an early response is difficult to generate because of efficient innate immune evasion mechanisms of SARS-CoV-2 in humans (Blanco-Melo et al., 2020). Immune evasion by SARS-CoV-2 is likely exacerbated by reduced myeloid cell antigen-presenting cell (APC) function or availability in the elderly (Zhao et al., 2011). In such cases, it is conceivable that late T cell responses may instead amplify pathogenic inflammatory outcomes in the presence of sustained high viral loads in the lungs, by multiple hypothetical possible mechanisms (Guo et al., 2020a; Li et al., 2008; Liu et al., 2019). Critical (ICU) and fatal COVID-19 (and SARS) outcomes are associated with elevated levels of inflammatory cytokines and chemokines, including interleukin-6 (IL-6) (Giamarellos-Bourboulis et al., 2020; Wong et al., 2004; Zhou et al., 2020)

Vaccine development against acute viral infections classically focuses on vaccine-elicited recapitulation of the type of protective immune response elicited by natural infection. Such foundational knowledge is currently missing for COVID-19, including how the balance and the phenotypes of responding cells vary as a function of disease course and severity. Such knowledge can guide selection of vaccine strategies most likely to elicit protective immunity against SARS-CoV-2. Furthermore, knowledge of the T cell responses to COVID-19 can guide selection of appropriate immunological endpoints for COVID-19 candidate vaccine clinical trials, which are already starting.

Limited information is also available about which SARS-CoV-2 proteins are recognized by human T cell immune responses. In some infections, T cell responses are strongly biased toward certain viral proteins, and the targets can vary substantially between $CD4^+$ and $CD8^+$ T cells (Moutaftsi et al., 2010; Tian et al., 2019). Knowledge of SARS-CoV-2 proteins and epitopes recognized by human T cell responses is of immediate relevance, as it will allow for monitoring of COVID-19 immune responses in laboratories worldwide. Epitope knowledge will also assist candidate vaccine design and facilitate evaluation of vaccine candidate immunogenicity. Almost all of the current COVID-19 vaccine candidates are focused on the spike protein.

A final key issue to consider in the study of SARS-CoV-2 immunity is whether some degree of cross-reactive coronavirus immunity exists in a fraction of the human population, and whether this might influence susceptibility to COVID-19 disease. This issue is also relevant for vaccine development, as cross-reactive immunity could influence responsiveness to candidate vaccines (Andrews et al., 2015).

In sum, the ability to measure and understand the human $CD4^+$ and $CD8^+$ T cell responses to SARS-CoV-2 infection is a major knowledge gap currently impeding COVID-19 vaccine development, interpretation of COVID-19 disease pathogenesis, and calibration of future social distancing pandemic control measures.

RESULTS

SARS-CoV-2 Peptides and Predicted Class I and Class II Epitopes

We recently predicted SARS-CoV-2 T cell epitopes utilizing the Immune Epitope Database and Analysis Resource (IEDB) (Dhanda et al., 2019; Vita et al., 2019). Utilizing bioinformatic approaches, we identified specific peptides in SARS-CoV-2 with increased probability of being T cell targets (Grifoni et al., 2020). We previously developed the megapool (MP) approach to allow simultaneous testing of large numbers of epitopes. By this technique, numerous epitopes are solubilized, pooled, and re-lyophilized to avoid cell toxicity problems (Carrasco Pro et al., 2015). These MPs have been used in human T cell studies of a number of indications, including allergies (Hinz et al., 2016), tuberculosis (Lindestam Arlehamn et al., 2016), tetanus, pertussis (Bancroft et al., 2016; da Silva Antunes et al., 2017), and dengue virus, for both $CD4^+$ and $CD8^+$ T cell epitopes (Grifoni et al., 2017; Weiskopf et al., 2015). Here, we generated MPs based on predicted SARS-CoV-2 epitopes. Specifically, one MP corresponds to 221 predicted HLA class II $CD4^+$ T cell epitopes (Grifoni et al., 2020) covering all proteins in the viral genome, apart from the spike (S) antigen (CD4_R MP). The prediction strategy utilized is geared to capture ~50% of the total response (Dhanda et al., 2018; Paul et al., 2015) and was designed and validated to predict dominant epitopes independently of ethnicity and HLA polymorphism. This approach takes advantage of the extensive cross-reactivity and repertoire overlap between different HLA class II loci and allelic variants to predict promiscuous epitopes, capable of binding many of the most common HLA class II prototypic specificities (Greenbaum

Table 1. Participant Characteristics

	Unexposed (n = 20)	COVID-19 (n = 20)
Age (years)	20–66 (median = 31, IQR = 21)	20–64 (median = 44, IQR = 9)
Gender		
Male (%)	35% (7/20)	45% (9/20)
Female (%)	65% (13/20)	55% (11/20)
Residency		
California (%)	95% (19/20)	100% (20/20)
USA, Non-California (%)	5% (1/20)	0% (0/20)
Sample Collection Date	March 2015–March 2018	March–April 2020
SARS-CoV-2 PCR Positivity	N/A	100% (16/16 tested)
Antibody Test Positivity[a]	N/A	90% (18/20)
Disease Severity[b]		
Mild	N/A	70% (14/20)
Moderate	N/A	20% (4/20)
Severe	N/A	10% (2/20)
Critical	N/A	0% (0/20)
Symptoms		
Cough	N/A	79% (15/19)
Fatigue	N/A	42% (8/19)
Fever	N/A	37% (7/19)
Anosmia	N/A	21% (4/19)
Dyspnea	N/A	16% (3/19)
Diarrhea	N/A	5% (1/19)
Days Post Symptom Onset at Collection	N/A	20–36 (18/20) (median = 26, IQR = 7)
Past Medical History		
No known	N/A	65% (13/20)
Hyperlipidemia	N/A	15% (3/20)
Hypertension	N/A	10% (2/20)
Asthma	N/A	10% (2/20)
Known or suspected sick contact/exposure	N/A	75% (15/20)

[a]Commercial skin prick lateral flow assay.
[b]WHO criteria.

et al., 2011; O'Sullivan et al., 1991; Sidney et al., 2010a, 2010b; Southwood et al., 1998).

For the spike protein, to ensure that all T cell reactivity against this important antigen can be detected, we generated a separate MP covering the entire antigen with 253 15-mer peptides overlapping by 10-residues (MP_S, Table S1). As stated above, the MP used to probe the non-spike regions is expected to capture ~50% of the total response. The use of overlapping peptides spanning entire open reading frames (ORFs) instead allows for a more complete characterization but also requires more cells. This factor should be kept in mind in terms of comparison of the magnitude of the $CD4^+$ T cell responses to those pools.

In the case of CD8 epitopes, since the overlap between different HLA class I allelic variants and loci is more limited to specific groups of alleles, or supertypes (Sidney et al., 2008), we targeted a set of the 12 most prominent HLA class I A and B alleles, which together allow broad coverage (>85%) of the general population. Two class I MPs were synthesized based on epitope predictions for those 12 most common HLA A and B alleles (Grifoni et al., 2020), which collectively encompass 628 predicted HLA class I $CD8^+$ T cell epitopes from the entire SARS-CoV-2 proteome (CD8 MP-A and MP-B).

Immunological Phenotypes of Recovered COVID-19 Patients

To test for the generation of SARS-CoV-2 $CD4^+$ and $CD8^+$ T cell responses following infection, we initially recruited 20 adult patients who had recovered from COVID-19 disease (Table 1). We also utilized peripheral blood mononuclear cell (PBMC) and plasma samples from local healthy control donors collected in 2015–2018 (see STAR Methods). Blood samples were collected at 20–35 days post-symptoms onset from non-hospitalized COVID-19 patients who were no longer symptomatic. SARS-CoV-2 infection was determined by swab test viral PCR during the acute phase of the infection. Verification of SARS-CoV-2 exposure was attempted both by lateral flow serology and SARS-CoV-2 spike protein receptor binding domain (RBD) ELISA (Stadlbauer et al., 2020), using plasma from the convalescence stage blood draw. Most patients were confirmed positive by lateral flow immunoglobulin (Ig) tests (Table 1). All patients were confirmed COVID-19 cases by SARS-CoV-2 RBD ELISA (Figures 1 and S1). All cases were IgG positive; anti-RBD IgM and IgA was also detected in the large majority of cases (Figures 1 and S1).

We defined a 21-color flow cytometry panel of mononuclear leukocyte lineage and phenotypic markers (Table S2) to broadly assess the immunological cellular profile of recovered COVID-19 patients (Figures 1 and S2). The frequency of $CD3^+$ cells was slightly increased in recovered COVID-19 patients relative to non-exposed controls, while no significant differences overall were observed in the frequencies of $CD4^+$ or $CD8^+$ T cells between the two groups. Frequencies of $CD19^+$ cells were somewhat decreased, while no differences were observed in the frequencies of $CD3^-CD19^-$ cells or $CD14^+CD16^-$ monocytes (Figures 1 and S2). No evidence of general lymphopenia was observed in the convalescing patients, consistent with the literature. Next, we utilized the SARS-CoV-2 MPs to probe $CD4^+$ and $CD8^+$ T cell responses.

Identification and Quantitation of SARS-CoV-2-Specific $CD4^+$ T Cell Responses

We utilized T cell receptor (TCR) dependent activation induced marker (AIM) assays to identify and quantify SARS-CoV-2-specific $CD4^+$ T cells in recovered COVID-19 patients. Initial definition and assessment of human antigen-specific SARS-CoV-2 T cell responses are best made with direct *ex vivo* T cell assays using broad-based epitope pools, such as MPs, and assays capable of detecting T cells of unknown cytokine

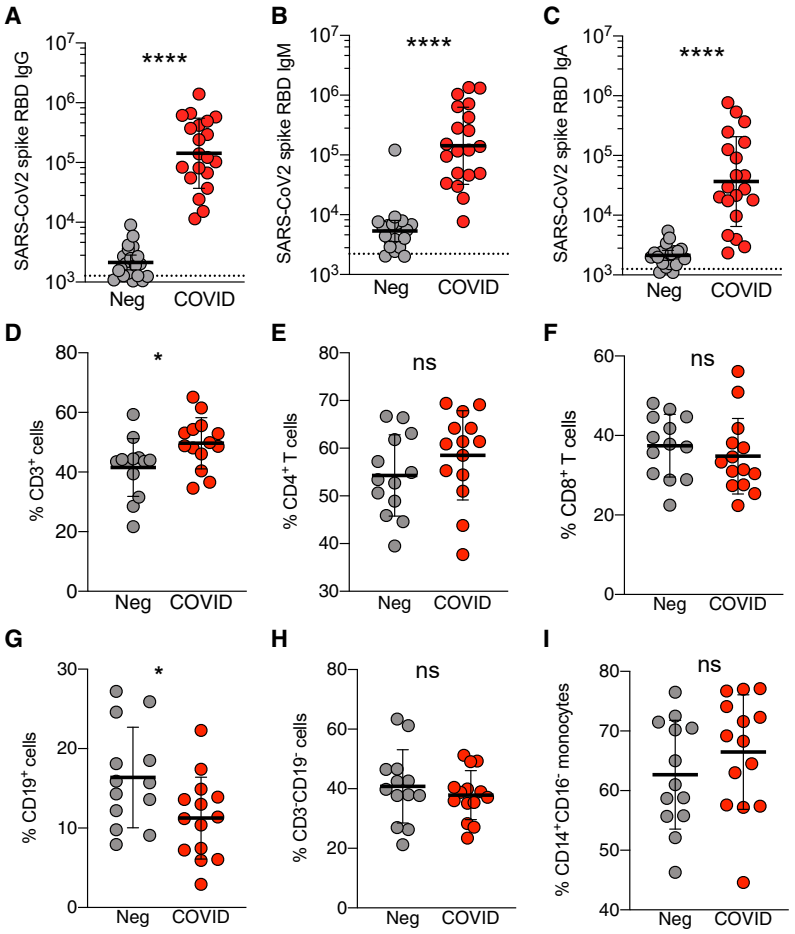


Figure 1. SARS-CoV-2 IgM, IgA, and IgG Responses of Recovered COVID-19 Patients

(A–C) Plasma ELISA titers to SARS-CoV-2 spike RBD. (A) IgG. (B) IgM. (C) IgA. Neg, unexposed donors from 2015–2018 (n = 20); COVID, convalescing COVID-19 patients (n = 20). All data are shown as ELISA titers based on a standard. The dotted line indicates limit of detection. Geometric mean titers with geometric SDs are indicated.

(D–I) Immunophenotyping of mononuclear leukocytes. Frequency of (D) $CD3^+$ total T cells, (E) $CD4^+$ T cells ($CD4^+CD3^+$), (F) $CD8^+$ T cells ($CD8^+CD3^+$), (G) $CD19^+$ B cells ($CD19^+CD3^-$), (H) $CD3^-CD19^-$ cells, and (I) $CD14^+CD16^-$ monocytes ($CD3^-CD19^-CD56^-$) from the PBMCs of unexposed donors (Neg, n = 13) or convalescing COVID-19 patients (COVID, n = 14). Data were analyzed using the Mann-Whitney test with mean and standard deviation shown.

*p < 0.05, ****p < 0.0001. See also Figures S1 and S2.

polarization and functional attributes. AIM assays are cytokine-independent assays to identify antigen-specific $CD4^+$ T cells (Havenar-Daughton et al., 2016; Reiss et al., 2017). AIM assays have been successfully used to identify virus-specific, vaccine-specific, or tuberculosis-specific $CD4^+$ T cells in a range of studies (Dan et al., 2016, 2019; Herati et al., 2017; Morou et al., 2019).

We stimulated PBMCs from 10 COVID-19 cases and 11 healthy controls (SARS-CoV-2 unexposed, collected in 2015–2018) with a spike MP (MP_S) and the class II MP covering the remainder of the SARS-CoV-2 orfeome ("non-spike," MP CD4_R). A CMV MP was used as a positive control, while DMSO was used as the negative control (Figures 2 and S3). SARS-CoV-2 spike-specific $CD4^+$ T cell responses ($OX40^+CD137^+$) were detected in 100% of COVID-19 cases ($p < 0.0001$ versus unexposed donors spike MP, Figures 2A and 2B. $p = 0.002$ versus DMSO control, Figure 2C). $CD4^+$ T cell responses to the remainder of the SARS-CoV-2 orfeome were also detected in 100% of COVID-19 cases ($p < 0.0079$ versus unexposed donors non-spike MP, Figures 2A and 2B. $p = 0.002$, non-spike versus DMSO control, Figure 2C). The magnitude of the SARS-CoV-2-specific $CD4^+$ T cell responses measured was similar to that of the CMV MP (Figure S3C). The concordance between SARS-CoV-2-specific $CD4^+$ T cell measurements in independent experiments was high ($p < 0.0002$, Figure S3D). To assess functionality and polarization of the SARS-CoV-2-specific $CD4^+$ T cell response, we measured cytokines secreted in response to MP stimulation. The SARS-CoV-2-specific $CD4^+$ T cells were functional, as the cells produced IL-2 in response to non-spike and spike MPs (Figure 2D). Polarization of the cells appeared to be a classic T_H1 type, as substantial interferon (IFN)-γ was produced (Figure 2E), while little to no IL-4, IL-5, IL-13, or IL-17α was expressed (Figures S3G–S3J).

Thus, recovered COVID-19 patients consistently generated a substantial $CD4^+$ T cell response against SARS-CoV-2. Similar conclusions were reached using stimulation index as the metric (Figures S3E and S3F). In terms of total $CD4^+$ T cell response per donor (Figure 2A), on average ~50% of the detected response was directed against the spike protein, and ~50% was directed against the MP representing the remainder of the SARS-CoV-2 orfeome (Figure 2A). This is of significance, since the SARS-CoV-2 spike protein is a key component of the vast majority of candidate COVID-19 vaccines under development. Of note, given the nature of the MP_R peptide predictions, the actual $CD4^+$ T cell response to be ascribed to non-spike ORFs was likely to be higher, addressed in further experiments below.

Identification and Quantitation of SARS-CoV-2-Specific $CD8^+$ T Cell Responses

To measure SARS-CoV-2-specific $CD8^+$ T cells in the recovered COVID-19 patients, we utilized two complementary methodologies, AIM assays and intracellular cytokine staining (ICS). The two SARS-CoV-2 class I MPs were used, CD8-A and CD8-B, with CMV MP and DMSO serving as positive and negative controls, respectively (Figures 3 and S4). $CD8^+$ T cell responses were detected by AIM ($CD69^+CD137^+$) in 70% of COVID-19 cases ($p < 0.0011$ versus unexposed donors "CD8 total,"

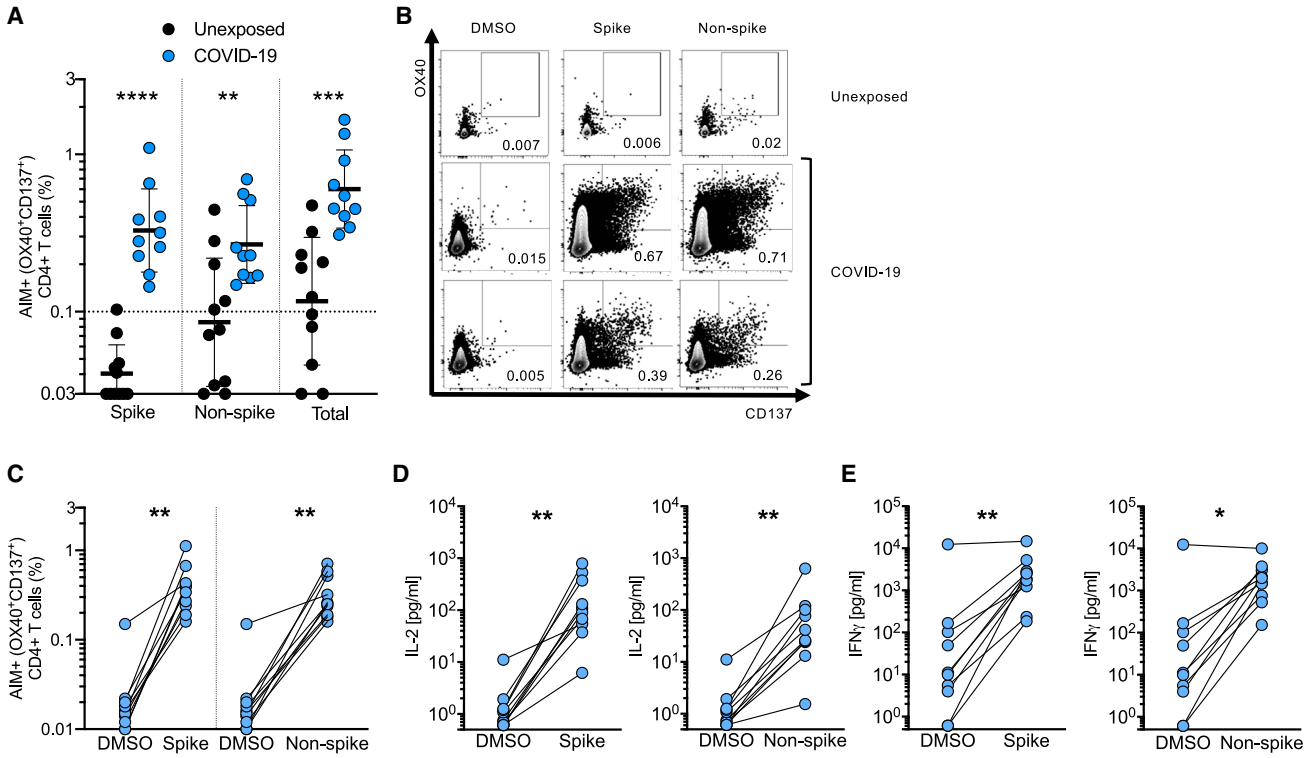


Figure 2. SARS-CoV-2-Specific CD4$^+$ T Cell Responses of Recovered COVID-19 Patients

(A) SARS-CoV-2-specific CD4$^+$ T cells measured as percentage of AIM$^+$ (OX40$^+$CD137$^+$) CD4$^+$ T cells after stimulation of PBMCs with peptide pools encompassing spike only (Spike) MP or the CD4_R MP representing all the proteome without spike (Non-spike). Data were background subtracted against DMSO negative control and are shown with geometric mean and geometric standard deviation. Samples were from unexposed donors (Unexposed, n = 11) and recovered COVID-19 patients (COVID-19, n = 10).

(B) Fluorescence-activated cell sorting (FACS) plot examples, gated on total CD4$^+$ T cells.

(C) AIM$^+$ CD4$^+$ T cell reactivity in COVID-19 cases between the negative control (DMSO) and antigen-specific stimulations.

(D and E) Cytokine levels in the supernatant of PBMCs from COVID-19 donors after stimulation with peptide pools (Spike and Non-spike) or the negative control (DMSO). (D) IL-2. (E) IFN-γ.

Statistical comparisons across cohorts were performed with the Mann-Whitney test. Pairwise comparisons (C–E) were performed with the Wilcoxon test. **p < 0.01; ***p < 0.001. See also Figure S3 and Table S6.

Figures 3A and 3B; p = 0.002, CD8-A or CD8-B versus DMSO control, Figure S4B). MP CD8-A contains spike epitopes, among epitopes to other proteins. The magnitude of the SARS-CoV-2 reactive CD8$^+$ T cell responses measured by AIM was somewhat lower than the CMV MP (Figure S4C). Similar conclusions were reached using stimulation index (Figures S3D and S3E).

Independently, ICS assays detected IFN-γ$^+$ SARS-CoV-2-specific CD8$^+$ T cells in the majority of COVID-19 cases (Figures 3C and 3D). The majority of IFN-γ$^+$ cells co-expressed granzyme B (Figures 3D and 3E). A substantial fraction of the IFN-γ$^+$ cells expressed tumor necrosis factor (TNF) but not IL-10 (Figure 3D). Thus, the majority of recovered COVID-19 patients generated a CD8$^+$ T cell response against SARS-CoV-2.

Relationship between SARS-CoV-2-Specific CD4$^+$ T Cell Responses and IgG and IgA Titers

Most protective antibody responses are dependent on CD4$^+$ T cell help. Therefore, we assessed whether stronger SARS-CoV-2-specific CD4$^+$ T cell responses were associated with higher antibody titers in COVID-19 cases. Given that spike is the primary target of SARS neutralizing antibodies, we examined spike-specific CD4$^+$ T cells. Spike-specific CD4$^+$ T cell responses correlated well with the magnitude of the anti-spike RBD IgG titers (R = 0.81; p < 0.0001; Figure 4A). Similar results were obtained using stimulation index (Figure S5A). The non-spike SARS-CoV-2-specific CD4$^+$ T cell response did not correlate as well with anti-spike RBD IgG titers (Figures 4B and S5B), consistent with a common requirement for intramolecular CD4$^+$ T cell help (Sette et al., 2008). Anti-spike IgA titers also correlated with spike-specific CD4$^+$ T cells (p < 0.0002, Figure S5). Thus, COVID-19 patients make anti-spike RBD antibody responses commensurate with the magnitude of their spike-specific CD4$^+$ T cell response. We then assessed the relationship between the CD4$^+$ and CD8$^+$ T cell responses to SARS-CoV-2. SARS-CoV-2-specific CD4$^+$ and CD8$^+$ T cell responses were well correlated (R = 0.62. p = 0.0025, Figures 4C and S5). Thus, antibody, CD4$^+$, and CD8$^+$ T cell responses to SARS-CoV-2 were generally well correlated.

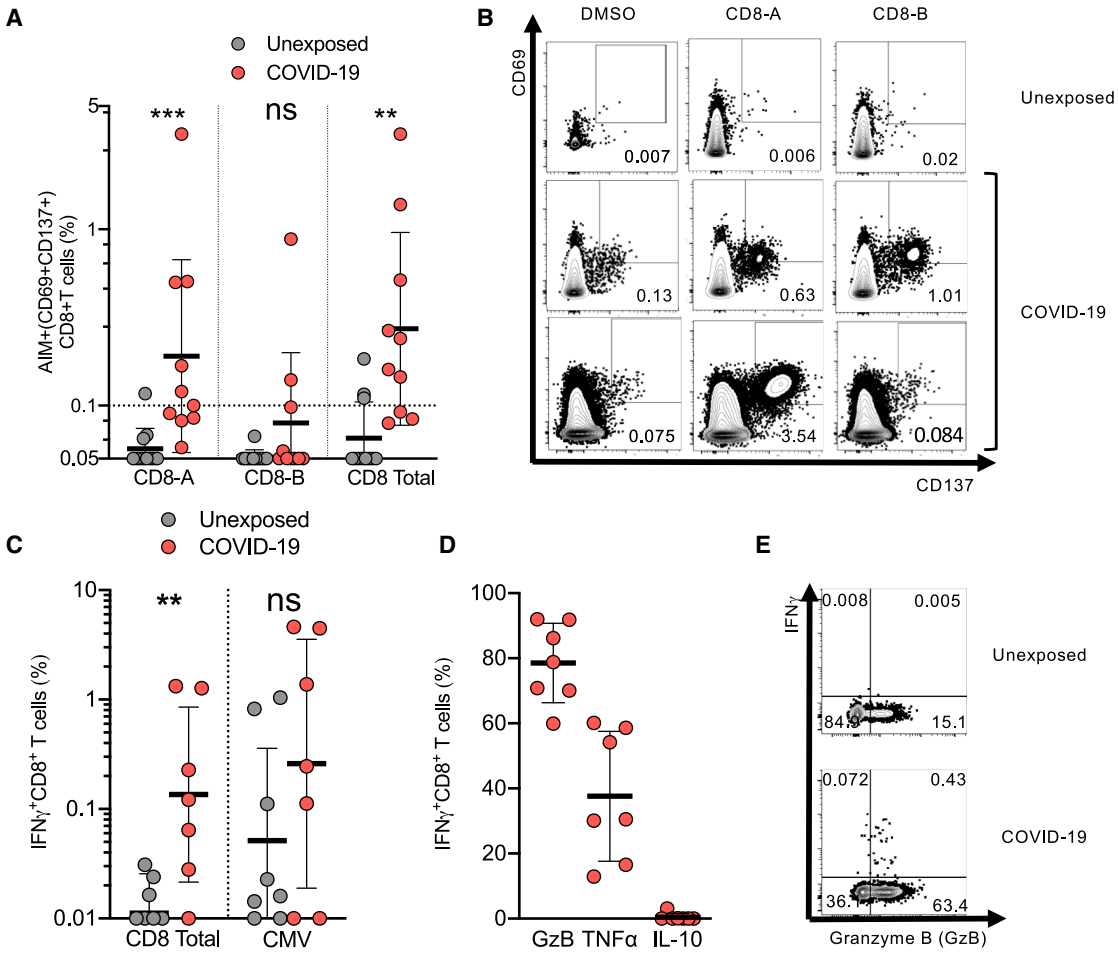


Figure 3. SARS-CoV-2-Specific CD8⁺ T Cell Responses by Recovered COVID-19 Patients

(A) SARS-CoV-2-specific CD8⁺ T cells measured as percentage of AIM⁺ (CD69⁺CD137⁺) CD8⁺ T cells after stimulation of PBMCs with class I MPs (CD8-A, CD8-B, and the combined data [Total]). Data were background subtracted against DMSO negative control and are shown with geometric mean and geometric standard deviation. Samples were from unexposed donors (Unexposed, n = 11) and recovered COVID-19 patients (COVID-19, n = 10).

(B) FACS plot examples.

(C) Percentage of CD8⁺ T cells producing IFN-γ in response to SARS-CoV-2 MPs, or CMV MP, in PBMCs from COVID-19 and unexposed donors after background subtraction. Data are shown with geometric mean and geometric standard deviation.

(D) Functional profile of IFN-γ⁺CD8⁺ T cells producing granzyme B (GzB), TNF-α (TNF), or IL-10 in response to SARS-CoV-2 MPs. Mean and SD are shown.

(E) FACS plot examples of IFN-γ and granzyme B co-expression.

Statistical comparisons across cohorts were performed with the Mann-Whitney test. *p < 0.05; **p < 0.01.; ns not significant. See also Figure S4 and Table S6.

Pre-existing Cross-Reactive Coronavirus-Specific T Cells

While spike- and non-spike-specific CD4⁺ T cell responses were detectable in all COVID-19 cases, cells were also detected in unexposed individuals (Figures 3A and 3B). These responses were statistically significant for non-spike-specific CD4⁺ T cell reactivity (non-spike, p = 0.039; spike, p = 0.067; Figures 5A and 5B). Non-spike-specific CD4⁺ T cell responses were above the limit of detection in 50% of donors based on stimulation index (SI) (Figure S3E). All of the donors were recruited between 2015 and 2018, excluding any possibility of exposure to SARS-CoV-2. Four human coronaviruses are known causes of seasonal "common cold" upper-respiratory tract infections: HCoV-OC43, HCoV-HKU1, HCoV-NL63, and HCoV-229E. We tested the SARS-CoV-2 unexposed donors for seroreactivity to HCoV-OC43 and HCoV-NL63 as a representative betacoronavirus and alphacoronavirus, respectively. All donors were IgG seropositive to HCoV-OC43 and HCoV-NL63 RBD, to varying degrees (Figure 5C), consistent with the endemic nature of these viruses (Gorse et al., 2010; Huang et al., 2020; Severance et al., 2008). We therefore examined whether these represented true pan-coronavirus T cells capable of recognizing SARS-CoV-2 epitopes.

SARS-CoV-2 ORF Targets of CD4⁺ and CD8⁺ T Cells

A most pressing, yet unresolved, set of issues in understanding SARS-CoV-2 immune responses is what antigens are targeted by CD4⁺ and CD8⁺ T cells, whether the corresponding antigens

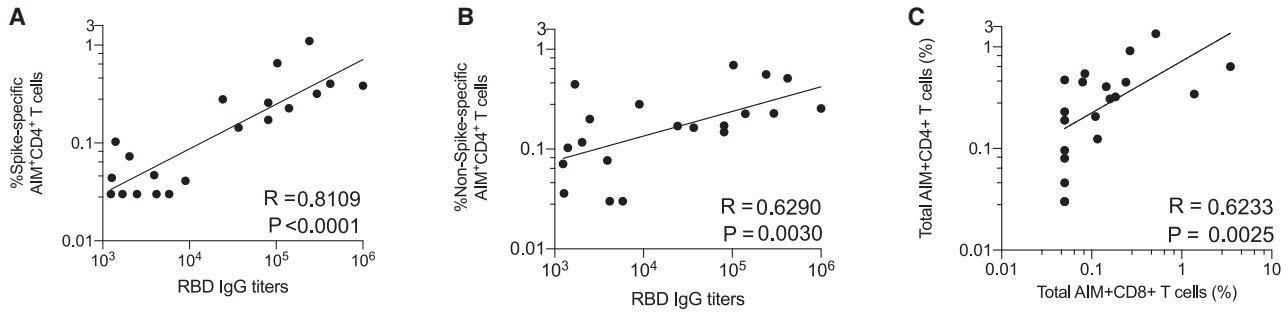


Figure 4. Correlations between SARS-CoV-2-Specific CD4$^+$ T Cells, Antibodies, and CD8$^+$ T Cells

(A) Correlation between SARS-CoV-2 spike-specific CD4$^+$ T cells (%) and anti-spike RBD IgG.

(B) Correlation between SARS-CoV-2 non-spike-specific CD4$^+$ T cells (%) and anti-spike RBD IgG.

(C) Correlation between SARS-CoV-2-specific CD4$^+$ T cells and SARS-CoV-2-specific CD8$^+$ T cells. Total MP responses per donor were used in each case ("Non-spike" + "spike" (CD4_R + MP_S) for CD4$^+$ T cells, CD8_A + CD8_B for CD8$^+$ T cells).

Statistical comparisons were performed using Spearman correlation. See also Figure S5.

are the same or different, and how do they reflect the antigens currently considered for COVID-19 vaccine development. We synthesized sets of overlapping peptides spanning the entire sequence of SARS-CoV-2 and pooled them separately so that each pool would represent one antigen (with the exception of nsp3, for which two pools were made; Table S1).

In the case of CD4$^+$ T cell responses, no obvious pattern of antigen specificity was observed based on SARS-CoV-2 genome organization; however, coronaviruses increase protein synthesis of certain ORFs in infected cells via subgenomic RNAs. Accounting for the relative abundance of subgenomic RNAs (Figure 6A) (Irigoyen et al., 2016; Snijder et al., 2003; Xie et al., 2020), the ORFs were re-ordered based on predicted protein abundance (Figure 6B). A clear hierarchy of SARS-CoV-2-specific CD4$^+$ T cell targets was then apparent, with the majority of the CD4$^+$ T cell response in COVID-19 cases directed against highly expressed SARS-CoV-2 ORFs spike, M, and N. On average, these antigens accounted for 27%, 21%, and 11% of the total CD4$^+$ T cell response, respectively. Most COVID-19 cases also had CD4$^+$ T cells specific for SARS-CoV-2 nsp3, nsp4, and ORF8 (Figure 6B), on average each accounting for ~5% of the total CD4$^+$ T cell response (Figure 6C). E, ORF6, hypothetical ORF10, and nsp1 are all small antigens (or potentially not expressed, in the case of ORF10) and were most likely predominantly unrecognized as a result. These results are somewhat unexpected, because data for other coronaviruses, from 27 different studies curated in the IEDB, reported that spike accounted for nearly two-thirds of reported CD4$^+$ T cell reactivity (Table S3). N accounted for most of the remaining epitopes in the published literature, although human N-specific CD4$^+$ T cell responses were not observed in one of the most comprehensive studies of human SARS-CoV-1 T cell responses (Li et al., 2008). Coronavirus M has not previously been described as a prominent target of CD4$^+$ T cell responses (Table S3). In sum, these results, fully scanning the SARS2 orfeome, demonstrate a pattern of robust and diverse SARS-CoV-2-specific CD4$^+$ T cell reactivity in convalescing COVID-19 cases that correlated largely with predicted viral protein abundance in infected cells.

When examining the non-exposed donors, the pattern of CD4$^+$ T cell targets changed. While S was still a relatively prominent target (23% of total, on average), there was no, or marginal, reactivity against SARS-CoV-2 N and M. Among donors with detectable CD4$^+$ T cells, a shift in reactivity was observed toward SARS-CoV-2 nsp14 (25%), nsp4 (15%) and nsp6 (14%) (Figures 6B and 6C). SARS-CoV-2-reactive CD4$^+$ T cells were detected in at least six different unexposed donors, demonstrating that the cross-reactivity is relatively widely distributed (Figure S6A).

Having scanned the full SARS-CoV-2 orfeome for CD4$^+$ T cell reactivity in multiple donors, it was possible to assess whether the epitope prediction MP approach successfully enriched for SARS-CoV-2 epitopes targeted by human CD4$^+$ T cells. When the total reactivity observed with the CD4_R MP was plotted versus the sum total of all antigen pools (excluding spike, given that spike predictions were not included in the CD4_R MP), a significant correlation was observed ($p < 0.0002$, Figure S6C). The single MP-R captured ~50% (44% +/− range 28%–80%) of the non-spike response per COVID-19 donor, demonstrating the success of the prediction approach, which, as mentioned above, was devised to attempt to capture approximately 50% of the total response (Dhanda et al., 2018; Paul et al., 2015).

In the case of CD8$^+$ T cell responses, the data in the literature from other coronaviruses (57 different studies curated in the IEDB; Table S3) reported spike accounting for 50% and N accounting for 36% of the defined epitopes. In a large study of human SARS-CoV-1 responses, spike was reported as essentially the only target of CD8$^+$ T cell responses (Li et al., 2008), while in a study of MERS CD8$^+$ T cells, responses were noted for spike, N and a pool of M/E peptides (Zhao et al., 2017). Few epitopes have been reported from other coronavirus antigens (Table S3). Here, we scanned the full SARS-CoV-2 orfeome for CD8$^+$ T cell recognition. Our data indicate a somewhat different pattern of immunodominance for SARS-CoV-2 CD8$^+$ T cell reactivity (Figures 6D and 6E), with spike protein accounting for ~26% of the reactivity, and N accounting for ~12%. Significant reactivity in COVID-19 recovered subjects was derived from other antigens, such as M (22%), nsp6 (15%), ORF8 (10%), and ORF3a (7%) (Figures 6D and 6E). In unexposed donors, SARS-

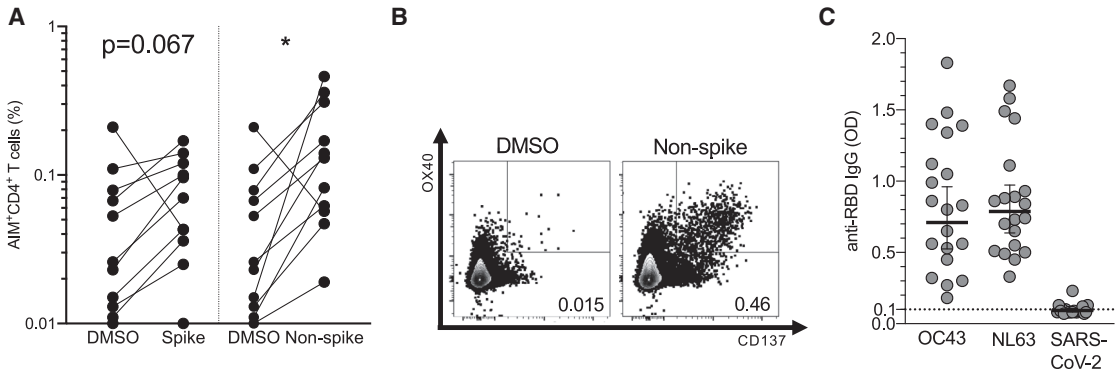


Figure 5. SARS-CoV-2 Epitope Reactivity in Unexposed Individuals

(A) SARS-CoV-2-reactive CD4$^+$ T cells measured as percentage of AIM$^+$ (OX40$^+$CD137$^+$) CD4$^+$ T cells in unexposed (n = 11) donors.

(B) FACS plot examples, gated on total CD4$^+$ T cells.

(C) Plasma IgG ELISAs for seroreactivity to RBD of HCoV-OC43 or HCoV-NL63. Data are expressed as geometric mean and geometric SD.

Pairwise statistical comparisons (A) were performed with the Wilcoxon test. *p < 0.05; ns, not significant.

CoV-2-reactive CD8$^+$ T cells were detected in at least four different donors (Figure S7), with less clear targeting of specific SARS-CoV-2 proteins than was observed for CD4$^+$ T cells, suggesting that coronavirus CD8$^+$ T cell cross-reactivity exists but is less widespread than CD4$^+$ T cell cross-reactivity.

DISCUSSION

There is a critical need for foundational knowledge about T cell responses to SARS-CoV-2. Here, we report functional validation of predicted epitopes when arranged in epitope MPs, utilizing PBMCs derived from convalescing COVID-19 cases. The experiments also used protein-specific peptide pools to determine which SARS-CoV-2 proteins are the predominant targets of human SARS-CoV-2-specific CD4$^+$ and CD8$^+$ T cells generated during COVID-19 disease. Importantly, we utilized the exact same series of experimental techniques with blood samples from healthy control donors (PBMCs collected in the 2015–2018 time frame), and substantial cross-reactive coronavirus T cell memory was observed.

Our results demonstrate that the epitope MPs are reagents well suited to analyze and detect SARS-CoV-2-specific T cell responses with limited sample material. We also developed and tested peptide pools corresponding to each of the 25 proteins encoded in the SARS-CoV-2 genome. Data from both the epitope MPs and protein peptide pool experiments can be interpreted in the context of previously reported T cell response immunodominance patterns observed for other coronaviruses, particularly the SARS and MERS viruses, which have been studied in humans, HLA-transgenic mice, wild-type mice, and other species. In the case of CD4$^+$ T cell responses, data for other coronaviruses found that spike accounted for nearly two-thirds of reported CD4$^+$ T cell reactivity, with N and M accounting for limited reactivity, and no reactivity in one large study of human SARS-CoV-1 responses (Li et al., 2008). Our SARS-CoV-2 data reveal that the pattern of immunodominance in COVID-19 is different. In particular, M, spike, and N proteins were clearly co-dominant, each recognized by 100% of COVID-19 cases studied here. Significant CD4$^+$ T cell responses were also directed against nsp3, nsp4, ORF3s, ORF7a, nsp12, and ORF8. These data suggest that a candidate COVID-19 vaccine consisting only of SARS-CoV-2 spike would be capable of eliciting SARS-CoV-2-specific CD4$^+$ T cell responses of similar representation to that of natural COVID-19 disease, but the data also indicate that there are many potential CD4$^+$ T cell targets in SARS-CoV-2, and inclusion of additional SARS-CoV-2 structural antigens such as M and N would better mimic the natural SARS-CoV-2-specific CD4$^+$ T cell response observed in mild to moderate COVID-19 disease.

Regarding SARS-CoV-2 CD8$^+$ T cell responses, the pattern of immunodominance found here differed from the literature for other coronaviruses. However, stringent comparisons are not possible, as some earlier studies were not similarly comprehensive and did not utilize the same experimental strategy. The spike protein was a target of human SARS-CoV-2 CD8$^+$ T cell responses, but it is not dominant. SARS-CoV-2 M was just as strongly recognized, and significant reactivity was noted for other antigens, mostly nsp6, ORF3a, and N, which comprised nearly 50% of the total CD8$^+$ T cell response, on average. Thus, these data indicate that candidate COVID-19 vaccines endeavoring to elicit CD8$^+$ T cell responses against the spike protein will be eliciting a relatively narrow CD8$^+$ T cell response compared to the natural CD8$^+$ T cell response observed in mild to moderate COVID-19 disease. An optimal vaccine CD8$^+$ T cell response to SARS-CoV-2 might benefit from additional class I epitopes, such as the ones derived from the M, nsp6, ORF3a, and/or N.

There have been concerns regarding vaccine enhancement of disease by certain candidate COVID-19 vaccine approaches, via antibody-dependent enhancement (ADE) or development of a T_H2 responses (Peeples, 2020). Herein, we saw predominant T_H1 responses in convalescing COVID-19 cases, with little to no T_H2 cytokines. Clearly more studies are required, but the data here appear to predominantly represent a classic T_H1 response to SARS-CoV-2.

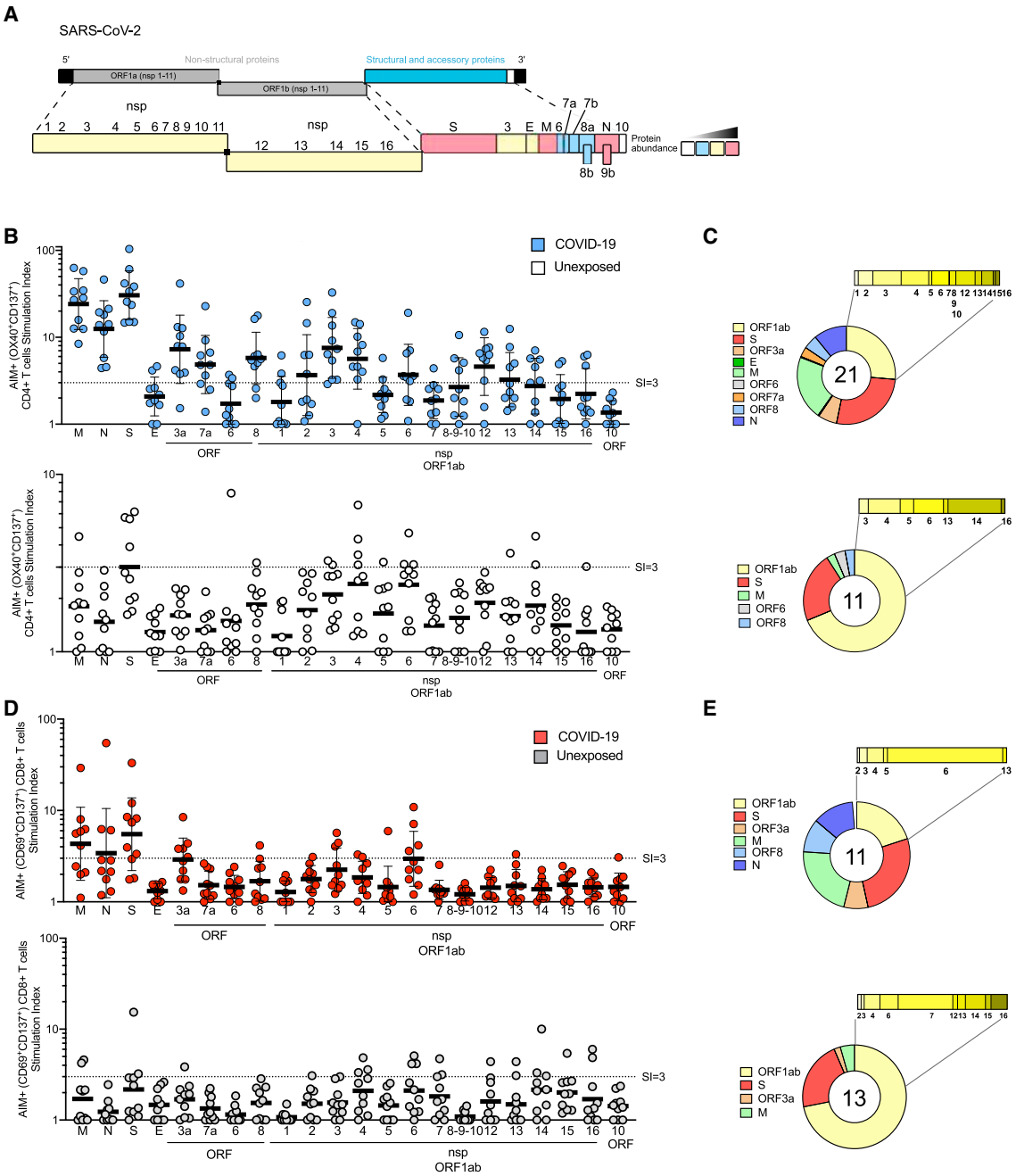


Figure 6. Protein Immunodominance of SARS-CoV-2-Specific CD4$^+$ and CD8$^+$ T Cells in COVID-19 Cases and Unexposed Donors

(A) SARS-CoV-2 genome organization and predicted viral protein abundance in infected cells.

(B) SARS-CoV-2 antigen-specific CD4$^+$ T cells (AIM$^+$, OX40$^+$CD137$^+$) quantified by stimulation index, using a peptide pool for each viral protein (with two exceptions, see Table S1). COVID-19 cases (top, in blue. n = 10) and unexposed donors (bottom, in white. n = 10). Data are expressed as geometric mean and geometric SD.

(C) Fraction of SARS-CoV-2 proteins recognized by CD4$^+$ T cells in COVID-19 cases (top) and unexposed donors (bottom).

(D) SARS-CoV-2 antigen-specific CD4$^+$ T cells (AIM$^+$, OX40$^+$CD137$^+$) quantified by stimulation index, using a peptide pool for each viral protein (with two exceptions, see Table S1). COVID-19 cases (top, in red. n = 10) and unexposed donors (bottom, in gray. n = 10). Data are expressed as geometric mean and geometric SD.

(E) Fraction of SARS-CoV-2 proteins recognized by CD8$^+$ T cells in COVID-19 cases (top) and unexposed donors (bottom).

See also Figures S6 and S7 and Table S6.

While it was important to identify antigen-specific T cell responses in COVID-19 cases, it is also of great interest to understand whether cross-reactive immunity exists between coronaviruses to any degree. A key step in developing that understanding is to examine antigen-specific $CD4^+$ and $CD8^+$ T cells in COVID-19 cases and in unexposed healthy controls, utilizing the exact same antigens and series of experimental techniques. $CD4^+$ T cell responses were detected in 40%–60% of unexposed individuals. This may be reflective of some degree of cross-reactive, preexisting immunity to SARS-CoV-2 in some, but not all, individuals. Whether this immunity is relevant in influencing clinical outcomes is unknown—and cannot be known without T cell measurements before and after SARS-CoV-2 infection of individuals—but it is tempting to speculate that the cross-reactive $CD4^+$ T cells may be of value in protective immunity, based on SARS mouse models (Zhao et al., 2016). Clear identification of the cross-reactive peptides, and their sequence homology relation to other coronaviruses, requires deconvolution of the positive peptide pools, which is not feasible with the cell numbers presently available, and time frame of the present study.

Regarding the value of cross-reactive T cells, influenza (flu) immunology in relationship to pandemics may be instructive. In the context of the 2009 H1N1 influenza pandemic, preexisting T cell immunity existed in the adult population, which focused on the more conserved internal influenza viral proteins (Greenbaum et al., 2009). The presence of cross-reactive T cells was found to correlate with less severe disease (Sridhar et al., 2013; Wilkinson et al., 2012). The frequent availability of cross-reactive memory T cell responses might have been one factor contributing to the lesser severity of the H1N1 flu pandemic (Hancock et al., 2009). Cross-reactive immunity to influenza strains has been modeled to be a critical influencer of susceptibility to newly emerging, potentially pandemic, influenza strains (Gostic et al., 2016). Given the severity of the ongoing COVID-19 pandemic, it has been modeled that any degree of cross-protective coronavirus immunity in the population could have a very substantial impact on the overall course of the pandemic, and the dynamics of the epidemiology for years to come (Kissler et al., 2020).

Limitations and Future Directions

Caveats of this study include the sample size and the focus on non-hospitalized COVID-19 cases. Sample size was limited by expediency. The focus on non-hospitalized cases of COVID-19 is a strength, in that these donors had uncomplicated disease of moderate duration, and thus it was encouraging that substantial $CD4^+$ T cell and antibody responses were detected in all cases, and $CD8^+$ T cell responses in the majority of cases. Complementing these data with MP T cell data from acute patients and patients with complicated disease course will also be of clear value, as will studies on the longevity of SARS-CoV-2 immunological memory. Additionally, lack of detailed information on common cold history or matched blood samples pre-exposure to SARS-CoV-2 prevents conclusions regarding the abundance of cross-reactive coronavirus T cells before exposure to SARS-CoV-2 and any potential protective efficacy of such cells. Finally, full epitope mapping in the future will add important detailed resolution of the human coronavirus-specific T cell responses.

In sum, we measured SARS-CoV-2-specific $CD4^+$ and $CD8^+$ T cells responses in COVID-19 cases. Using multiple experimental approaches, SARS-CoV-2-specific $CD4^+$ T cell and antibody responses were observed in all COVID-19 cases, and $CD8^+$ T cell responses were observed in most. Importantly, pre-existing SARS-CoV-2-cross-reactive T cell responses were observed in healthy donors, indicating some potential for pre-existing immunity in the human population. ORF mapping of T cell specificities revealed valuable targets for incorporation in candidate vaccine development and revealed distinct specificity patterns between COVID-19 cases and unexposed healthy controls.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead Contact
  - Materials Availability
  - Data and Code Availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Human Subjects
- METHOD DETAILS
  - Peptide Pools
  - PBMC isolation
  - SARS-CoV-2 RBD ELISA
  - OC43 and NL63 coronavirus RBD ELISA
  - Flow Cytometry
  - Cytokine bead assays
  - Identification of coronavirus epitopes and associated literature references
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.cell.2020.05.015.

ACKNOWLEDGMENTS

We would like to thank Cheryl Kim, director of the LJI flow cytometry core facility for outstanding expertise. We thank Prof. Peter Kim, Abigail Powell, PhD, and colleagues (Stanford) for RBD protein synthesized from Prof. Florian Krammer (Mt. Sinai) constructs. J.M. was supported by PhD student fellowships from the Departamento Administrativo de Ciencia, Tecnología e Innovación (COLCIENCIAS), and Pontificia Universidad Javeriana. This work was funded by the NIH NIAID under awards AI142742 (Cooperative Centers for Human Immunology) (S.C. and A.S.), National Institutes of Health contract Nr. 75N9301900065 (A.S. and D.W.), and U19 AI118626 (A.S. and B.P.). The BD FACSymphony purchase was partially funded by the Bill and Melinda Gates Foundation and LJI Institutional Funds (S.C. and A.S.). This work was additionally supported in part by the Johnathan and Mary Tu Foundation (D.M.S.), the NIAID under K08 award AI135078 (J.D.), and UCSD T32s AI007036 and AI007384 Infectious Diseases Division (S.I.R. and S.A.R.).

AUTHOR CONTRIBUTIONS

Conceptualization, A.G., D.W., S.C., and A.S.; Investigation, A.G., D.W., J.M., C.R.M., J.M.D., D.M., L.P., R.S.J., A.S., and D.W.; Formal Analysis, A.G., D.W., C.R.M., J.M.D., J.M., and S.C.; Resources, S.I.R., S.A.R., D.M.S., A.F.C., F.K., S.C., and A.S.; Data Curation, J.A.G. and B.P.; Writing, S.C., A.S., A.G., and D.W.; Supervision, B.P., A.M.d.S., S.C., and A.S.; Project Administration, A.F.; Funding Acquisition, S.C., A.S., D.W., D.S., and J.D.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: April 20, 2020
Revised: May 4, 2020
Accepted: May 7, 2020
Published: May 14, 2020

REFERENCES

Alshukairi, A.N., Zheng, J., Zhao, J., Nehdi, A., Baharoon, S.A., Layqah, L., Bokhari, A., Al Johani, S.M., Samman, N., Boudjelal, M., et al. (2018). High Prevalence of MERS-CoV Infection in Camel Workers in Saudi Arabia. MBio *9*, e01985–e01918.

Amanat, F., and Krammer, F. (2020). SARS-CoV-2 Vaccines: Status Report. Immunity *52*, 583–589.

Andrews, S.F., Huang, Y., Kaur, K., Popova, L.I., Ho, I.Y., Pauli, N.T., Henry Dunand, C.J., Taylor, W.M., Lim, S., Huang, M., et al. (2015). Immune history profoundly affects broadly protective B cell responses to influenza. Sci. Transl. Med. *7*, 316ra192.

Bancroft, T., Dillon, M.B., da Silva Antunes, R., Paul, S., Peters, B., Crotty, S., Lindestam Arlehamn, C.S., and Sette, A. (2016). Th1 versus Th2 T cell polarization by whole-cell and acellular childhood pertussis vaccines persists upon re-immunization in adolescence and adulthood. Cell. Immunol. *304-305*, 35–43.

Blanco-Melo, D., Nilsson-Payant, B.E., Liu, W.-C., Møller, R., Panis, M., Sachs, D., Albrecht, R.A., and tenOever, B.R. (2020). SARS-CoV-2 launches a unique transcriptional signature from in vitro, ex vivo, and in vivo systems. bioRxiv. https://doi.org/10.1101/2020.03.24.004655.

Callow, K.A., Parry, H.F., Sergeant, M., and Tyrrell, D.A. (1990). The time course of the immune response to experimental coronavirus infection of man. Epidemiol. Infect. *105*, 435–446.

Cao, X. (2020). COVID-19: immunopathology and its implications for therapy. Nat. Rev. Immunol. *20*, 269–270.

Carrasco Pro, S., Sidney, J., Paul, S., Lindestam Arlehamn, C., Weiskopf, D., Peters, B., and Sette, A. (2015). Automatic Generation of Validated Specific Epitope Sets. J. Immunol. Res. *2015*, 763461.

Choe, P.G., Perera, R.A.P.M., Park, W.B., Song, K.H., Bang, J.H., Kim, E.S., Kim, H.B., Ko, L.W.R., Park, S.W., Kim, N.J., et al. (2017). MERS-CoV Antibody Responses 1 Year after Symptom Onset, South Korea, 2015. Emerg. Infect. Dis. *23*, 1079–1084.

Crotty, S. (2019). T Follicular Helper Cell Biology: A Decade of Discovery and Diseases. Immunity *50*, 1132–1148.

da Silva Antunes, R., Paul, S., Sidney, J., Weiskopf, D., Dan, J.M., Phillips, E., Mallal, S., Crotty, S., Sette, A., and Lindestam Arlehamn, C.S. (2017). Definition of Human Epitopes Recognized in Tetanus Toxoid and Development of an Assay Strategy to Detect Ex Vivo Tetanus CD4+ T Cell Responses. PLoS ONE *12*, e0169086.

Dan, J.M., Lindestam Arlehamn, C.S., Weiskopf, D., da Silva Antunes, R., Havenar-Daughton, C., Reiss, S.M., Brigger, M., Bothwell, M., Sette, A., and Crotty, S. (2016). A Cytokine-Independent Approach To Identify Antigen-Specific Human Germinal Center T Follicular Helper Cells and Rare Antigen-Specific CD4+ T Cells in Blood. J. Immunol. *197*, 983–993.

Dan, J.M., Havenar-Daughton, C., Kendric, K., Al-Kolla, R., Kaushik, K., Rosales, S.L., Anderson, E.L., LaRock, C.N., Vijayanand, P., Seumois, G., et al. (2019). Recurrent group A Streptococcus tonsillitis is an immunosusceptibility disease involving antibody deficiency and aberrant TFH cells. Sci. Transl. Med. *11*, eaau3776.

Dhanda, S.K., Karosiene, E., Edwards, L., Grifoni, A., Paul, S., Andreatta, M., Weiskopf, D., Sidney, J., Nielsen, M., Peters, B., and Sette, A. (2018). Predicting HLA CD4 Immunogenicity in Human Populations. Front. Immunol. *9*, 1369.

Dhanda, S.K., Mahajan, S., Paul, S., Yan, Z., Kim, H., Jespersen, M.C., Jurtz, V., Andreatta, M., Greenbaum, J.A., Marcatili, P., et al. (2019). IEDB-AR: immune epitope database-analysis resource in 2019. Nucleic Acids Res. *47* (W1), W502–W506.

Dong, E., Du, H., and Gardner, L. (2020). An interactive web-based dashboard to track COVID-19 in real time. Lancet Infect. Dis. *20*, 533–534.

Giamarellos-Bourboulis, E.J., Netea, M.G., Rovina, N., Akinosoglou, K., Antoniadou, A., Antonakos, N., Damoraki, G., Gkavogianni, T., Adami, M.E., Katsaounou, P., et al. (2020). Complex Immune Dysregulation in COVID-19 Patients with Severe Respiratory Failure. Cell Host Microbe. Published online April 17, 2020. https://doi.org/10.1016/j.chom.2020.04.009.

Gorse, G.J., Patel, G.B., Vitale, J.N., and O'Connor, T.Z. (2010). Prevalence of antibodies to four human coronaviruses is lower in nasal secretions than in serum. Clin. Vaccine Immunol. *17*, 1875–1880.

Gostic, K.M., Ambrose, M., Worobey, M., and Lloyd-Smith, J.O. (2016). Potent protection against H5N1 and H7N9 influenza via childhood hemagglutinin imprinting. Science *354*, 722–726.

Greenbaum, J.A., Kotturi, M.F., Kim, Y., Oseroff, C., Vaughan, K., Salimi, N., Vita, R., Ponomarenko, J., Scheuermann, R.H., Sette, A., and Peters, B. (2009). Pre-existing immunity against swine-origin H1N1 influenza viruses in the general human population. Proc. Natl. Acad. Sci. USA *106*, 20365–20370.

Greenbaum, J., Sidney, J., Chung, J., Brander, C., Peters, B., and Sette, A. (2011). Functional classification of class II human leukocyte antigen (HLA) molecules reveals seven different supertypes and a surprising degree of repertoire sharing across supertypes. Immunogenetics *63*, 325–335.

Grifoni, A., Angelo, M.A., Lopez, B., O'Rourke, P.H., Sidney, J., Cerpas, C., Balmaseda, A., Silveira, C.G.T., Maestri, A., Costa, P.R., et al. (2017). Global Assessment of Dengue Virus-Specific CD4$^+$ T Cell Responses in Dengue-Endemic Areas. Front. Immunol. *8*, 1309.

Grifoni, A., Sidney, J., Zhang, Y., Scheuermann, R.H., Peters, B., and Sette, A. (2020). A Sequence Homology and Bioinformatic Approach Can Predict Candidate Targets for Immune Responses to SARS-CoV-2. Cell Host Microbe *27*, 671–680.

Guo, T., Fan, Y., Chen, M., Wu, X., Zhang, L., He, T., Wang, H., Wan, J., Wang, X., and Lu, Z. (2020a). Cardiovascular Implications of Fatal Outcomes of Patients With Coronavirus Disease 2019 (COVID-19). JAMA Cardiol. Published online March 27, 2020. https://doi.org/10.1001/jamacardio.2020.1017.

Guo, X., Guo, Z., Duan, C., Chen, Z., Wang, G., Lu, Y., Li, M., and Lu, J. (2020b). Long-Term Persistence of IgG Antibodies in SARS-CoV Infected Healthcare Workers. medRxiv. https://doi.org/10.1101/2020.02.12.20021386.

Hancock, K., Veguilla, V., Lu, X., Zhong, W., Butler, E.N., Sun, H., Liu, F., Dong, L., DeVos, J.R., Gargiullo, P.M., et al. (2009). Cross-reactive antibody responses to the 2009 pandemic H1N1 influenza virus. N. Engl. J. Med. *361*, 1945–1952.

Havenar-Daughton, C., Reiss, S.M., Carnathan, D.G., Wu, J.E., Kendric, K., Torrents de la Peña, A., Kasturi, S.P., Dan, J.M., Bothwell, M., Sanders, R.W., et al. (2016). Cytokine-Independent Detection of Antigen-Specific Germinal Center T Follicular Helper Cells in Immunized Nonhuman Primates Using a Live Cell Activation-Induced Marker Technique. J. Immunol. *197*, 994–1002.

Herati, R.S., Muselman, A., Vella, L., Bengsch, B., Parkhouse, K., Del Alcazar, D., Kotzin, J., Doyle, S.A., Tebas, P., Hensley, S.E., et al. (2017). Successive annual influenza vaccination induces a recurrent oligoclonotypic memory response in circulating T follicular helper cells. Sci Immunol *2*, eaag2152.

Hinz, D., Seumois, G., Gholami, A.M., Greenbaum, J.A., Lane, J., White, B., Broide, D.H., Schulten, V., Sidney, J., Bakhru, P., et al. (2016). Lack of allergy to timothy grass pollen is not a passive phenomenon but associated with the allergen-specific modulation of immune reactivity. Clin. Exp. Allergy *46*, 705–719.

Huang, A.T., Garcia-Carreras, B., Hitchings, M.D.T., Yang, B., Katzelnick, L., Rattigan, S.M., Borgert, B., Moreno, C., Solomon, B.D., Rodriguez-Barraquer, I., et al. (2020). A systematic review of antibody mediated immunity to coronaviruses: antibody kinetics, correlates of protection, and association of antibody responses with severity of disease. medRxiv, 2020.2004.2014.20065771.

Irigoyen, N., Firth, A.E., Jones, J.D., Chung, B.Y., Siddell, S.G., and Brierley, I. (2016). High-Resolution Analysis of Coronavirus Gene Expression by RNA Sequencing and Ribosome Profiling. PLoS Pathog. *12*, e1005473.

Jurtz, V., Paul, S., Andreatta, M., Marcatili, P., Peters, B., and Nielsen, M. (2017). NetMHCpan-4.0: Improved Peptide-MHC Class I Interaction Predictions Integrating Eluted Ligand and Peptide Binding Affinity Data. J. Immunol. *199*, 3360–3368.

Kissler, S.M., Tedijanto, C., Goldstein, E., Grad, Y.H., and Lipsitch, M. (2020). Projecting the transmission dynamics of SARS-CoV-2 through the postpandemic period. Science, eabb5793.

Li, C.K., Wu, H., Yan, H., Ma, S., Wang, L., Zhang, M., Tang, X., Temperton, N.J., Weiss, R.A., Brenchley, J.M., et al. (2008). T cell responses to whole SARS coronavirus in humans. J. Immunol. *181*, 5490–5500.

Lindestam Arlehamn, C.S., McKinney, D.M., Carpenter, C., Paul, S., Rozot, V., Makgotlho, E., Gregg, Y., van Rooyen, M., Ernst, J.D., Hatherill, M., et al. (2016). A Quantitative Analysis of Complexity of Human Pathogen-Specific CD4 T Cell Responses in Healthy M. tuberculosis Infected South Africans. PLoS Pathog. *12*, e1005760.

Liu, L., Wei, Q., Lin, Q., Fang, J., Wang, H., Kwok, H., Tang, H., Nishiura, K., Peng, J., Tan, Z., et al. (2019). Anti-spike IgG causes severe acute lung injury by skewing macrophage responses during acute SARS-CoV infection. JCI Insight *4*, 123158.

Morou, A., Brunet-Ratnasingham, E., Dubé, M., Charlebois, R., Mercier, E., Darko, S., Brassard, N., Nganou-Makamdop, K., Arumugam, S., Gendron-Lepage, G., et al. (2019). Altered differentiation is central to HIV-specific $CD4^+$ T cell dysfunction in progressive disease. Nat. Immunol. *20*, 1059–1070.

Moutaftsi, M., Tscharke, D.C., Vaughan, K., Koelle, D.M., Stern, L., Calvo-Calle, M., Ennis, F., Terajima, M., Sutter, G., Crotty, S., et al. (2010). Uncovering the interplay between CD8, CD4 and antibody responses to complex pathogens. Future Microbiol. *5*, 221–239.

O'Sullivan, D., Arrhenius, T., Sidney, J., Del Guercio, M.F., Albertson, M., Wall, M., Oseroff, C., Southwood, S., Colón, S.M., Gaeta, F.C., et al. (1991). On the interaction of promiscuous antigenic peptides with different DR alleles. Identification of common structural motifs. J. Immunol. *147*, 2663–2669.

Okba, N.M.A., Raj, V.S., Widjaja, I., GeurtsvanKessel, C.H., de Bruin, E., Chandler, F.D., Park, W.B., Kim, N.J., Farag, E.A.B.A., Al-Hajri, M., et al. (2019). Sensitive and Specific Detection of Low-Level Antibody Responses in Mild Middle East Respiratory Syndrome Coronavirus Infections. Emerg. Infect. Dis. *25*, 1868–1877.

Paul, S., Lindestam Arlehamn, C.S., Scriba, T.J., Dillon, M.B., Oseroff, C., Hinz, D., McKinney, D.M., Carrasco Pro, S., Sidney, J., Peters, B., and Sette, A. (2015). Development and validation of a broad scheme for prediction of HLA class II restricted T cell epitopes. J. Immunol. Methods *422*, 28–34.

Paul, S., Sidney, J., Sette, A., and Peters, B. (2016). TepiTool: A Pipeline for Computational Prediction of T Cell Epitope Candidates. Curr. Protoc. Immunol. *114*, 18.19.1–18.19.24.

Peeples, L. (2020). News Feature: Avoiding pitfalls in the pursuit of a COVID-19 vaccine. Proc. Natl. Acad. Sci. USA *117*, 8218–8221.

Reiss, S., Baxter, A.E., Cirelli, K.M., Dan, J.M., Morou, A., Daigneault, A., Brassard, N., Silvestri, G., Routy, J.P., Havenar-Daughton, C., et al. (2017). Comparative analysis of activation induced marker (AIM) assays for sensitive identification of antigen-specific CD4 T cells. PLoS ONE *12*, e0186998.

Sallusto, F., Lanzavecchia, A., Araki, K., and Ahmed, R. (2010). From vaccines to memory and back. Immunity *33*, 451–463.

Sette, A., Moutaftsi, M., Moyron-Quiroz, J., McCausland, M.M., Davies, D.H., Johnston, R.J., Peters, B., Rafii-El-Idrissi Benhnia, M., Hoffmann, J., Su, H.P., et al. (2008). Selective CD4+ T cell help for antibody responses to a large viral pathogen: deterministic linkage of specificities. Immunity *28*, 847–858.

Severance, E.G., Bossis, I., Dickerson, F.B., Stallings, C.R., Origoni, A.E., Sullens, A., Yolken, R.H., and Viscidi, R.P. (2008). Development of a nucleocapsid-based human coronavirus immunoassay and estimates of individuals exposed to coronavirus in a U.S. metropolitan population. Clin. Vaccine Immunol. *15*, 1805–1810.

Sidney, J., Peters, B., Frahm, N., Brander, C., and Sette, A. (2008). HLA class I supertypes: a revised and updated classification. BMC Immunol. *9*, 1.

Sidney, J., Steen, A., Moore, C., Ngo, S., Chung, J., Peters, B., and Sette, A. (2010a). Divergent motifs but overlapping binding repertoires of six HLA-DQ molecules frequently expressed in the worldwide human population. J. Immunol. *185*, 4189–4198.

Sidney, J., Steen, A., Moore, C., Ngo, S., Chung, J., Peters, B., and Sette, A. (2010b). Five HLA-DP molecules frequently expressed in the worldwide human population share a common HLA supertypic binding specificity. J. Immunol. *184*, 2492–2503.

Snijder, E.J., Bredenbeek, P.J., Dobbe, J.C., Thiel, V., Ziebuhr, J., Poon, L.L., Guan, Y., Rozanov, M., Spaan, W.J., and Gorbalenya, A.E. (2003). Unique and conserved features of genome and proteome of SARS-coronavirus, an early split-off from the coronavirus group 2 lineage. J. Mol. Biol. *331*, 991–1004.

Southwood, S., Sidney, J., Kondo, A., del Guercio, M.F., Appella, E., Hoffman, S., Kubo, R.T., Chesnut, R.W., Grey, H.M., and Sette, A. (1998). Several common HLA-DR types share largely overlapping peptide binding repertoires. J. Immunol. *160*, 3363–3373.

Sridhar, S., Begom, S., Bermingham, A., Hoschler, K., Adamson, W., Carman, W., Bean, T., Barclay, W., Deeks, J.J., and Lalvani, A. (2013). Cellular immune correlates of protection against symptomatic pandemic influenza. Nat. Med. *19*, 1305–1312.

Stadlbauer, D., Amanat, F., Chromikova, V., Jiang, K., Strohmeier, S., Arunkumar, G.A., Tan, J., Bhavsar, D., Capuano, C., Kirkpatrick, E., et al. (2020). SARS-CoV-2 Seroconversion in Humans: A Detailed Protocol for a Serological Assay, Antigen Production, and Test Setup. Curr. Protoc. Microbiol. *57*, e100.

Takano, T., Kawakami, C., Yamada, S., Satoh, R., and Hohdatsu, T. (2008). Antibody-dependent enhancement occurs upon re-infection with the identical serotype virus in feline infectious peritonitis virus infection. J. Vet. Med. Sci. *70*, 1315–1321.

Thanh Le, T., Andreadakis, Z., Kumar, A., Gómez Román, R., Tollefsen, S., Saville, M., and Mayhew, S. (2020). The COVID-19 vaccine development landscape. Nat. Rev. Drug Discov. *19*, 305–306.

Tian, Y., Grifoni, A., Sette, A., and Weiskopf, D. (2019). Human T Cell Response to Dengue Virus Infection. Front. Immunol. *10*, 2125.

Vennema, H., de Groot, R.J., Harbour, D.A., Dalderup, M., Gruffydd-Jones, T., Horzinek, M.C., and Spaan, W.J. (1990). Early death after feline infectious peritonitis virus challenge due to recombinant vaccinia virus immunization. J. Virol. *64*, 1407–1409.

Vita, R., Mahajan, S., Overton, J.A., Dhanda, S.K., Martini, S., Cantrell, J.R., Wheeler, D.K., Sette, A., and Peters, B. (2019). The Immune Epitope Database (IEDB): 2018 update. Nucleic Acids Res. *47* (D1), D339–D343.

Weiskopf, D., Angelo, M.A., de Azeredo, E.L., Sidney, J., Greenbaum, J.A., Fernando, A.N., Broadwater, A., Kolla, R.V., De Silva, A.D., de Silva, A.M., et al. (2013). Comprehensive analysis of dengue virus-specific responses supports an HLA-linked protective role for CD8+ T cells. Proc. Natl. Acad. Sci. USA *110*, E2046–E2053.

Weiskopf, D., Cerpas, C., Angelo, M.A., Bangs, D.J., Sidney, J., Paul, S., Peters, B., Sanches, F.P., Silvera, C.G., Costa, P.R., et al. (2015). Human CD8+ T-Cell Responses Against the 4 Dengue Virus Serotypes Are Associated With Distinct Patterns of Protein Targets. J. Infect. Dis. *212*, 1743–1751.

Wilkinson, T.M., Li, C.K.F., Chui, C.S.C., Huang, A.K.Y., Perkins, M., Liebner, J.C., Lambkin-Williams, R., Gilbert, A., Oxford, J., Nicholas, B., et al. (2012). Preexisting influenza-specific CD4+ T cells correlate with disease protection against influenza challenge in humans. Nat. Med. *18*, 274–280.

Wong, C.K., Lam, C.W., Wu, A.K., Ip, W.K., Lee, N.L., Chan, I.H., Lit, L.C., Hui, D.S., Chan, M.H., Chung, S.S., and Sung, J.J. (2004). Plasma inflammatory cytokines and chemokines in severe acute respiratory syndrome. Clin. Exp. Immunol. *136*, 95–103.

Wu, Z., and McGoogan, J.M. (2020). Characteristics of and Important Lessons From the Coronavirus Disease 2019 (COVID-19) Outbreak in China: Summary of a Report of 72 314 Cases From the Chinese Center for Disease Control and Prevention. JAMA *323*, 1239–1242.

Xie, X., Muruato, A., Lokugamage, K.G., Narayanan, K., Zhang, X., Zou, J., Liu, J., Schindewolf, C., Bopp, N.E., Aguilar, P.V., et al. (2020). An Infectious cDNA Clone of SARS-CoV-2. Cell Host Microbe *27*, 841–848.

Zhao, J., Zhao, J., Van Rooijen, N., and Perlman, S. (2009). Evasion by stealth: inefficient immune activation underlies poor T cell response and severe disease in SARS-CoV-infected mice. PLoS Pathog. *5*, e1000636.

Zhao, J., Zhao, J., and Perlman, S. (2010). T cell responses are required for protection from clinical disease and for virus clearance in severe acute respiratory syndrome coronavirus-infected mice. J. Virol. *84*, 9318–9325.

Zhao, J., Zhao, J., Legge, K., and Perlman, S. (2011). Age-related increases in PGD(2) expression impair respiratory DC migration, resulting in diminished T cell responses upon respiratory virus infection in mice. J. Clin. Invest. *121*, 4921–4930.

Zhao, J., Zhao, J., Mangalam, A.K., Channappanavar, R., Fett, C., Meyerholz, D.K., Agnihothram, S., Baric, R.S., David, C.S., and Perlman, S. (2016). Airway Memory CD4(+) T Cells Mediate Protective Immunity against Emerging Respiratory Coronaviruses. Immunity *44*, 1379–1391.

Zhao, J., Alshukairi, A.N., Baharoon, S.A., Ahmed, W.A., Bokhari, A.A., Nehdi, A.M., Layqah, L.A., Alghamdi, M.G., Al Gethamy, M.M., Dada, A.M., et al. (2017). Recovery from the Middle East respiratory syndrome is associated with antibody and T-cell responses. Sci. Immunol. *2*, eaan5393.

Zhou, F., Yu, T., Du, R., Fan, G., Liu, Y., Liu, Z., Xiang, J., Wang, Y., Song, B., Gu, X., et al. (2020). Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. Lancet *395*, 1054–1062.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
M5E2 (V500) [anti-CD14]	Becton Dickinson	561391 (RRID:AB_10611856)
HIB19 (V500) [anti-CD19]	Becton Dickinson	561121 (RRID:AB_10562391)
FN50 (BV605) [anti-CD4]	Becton Dickinson	562989 (RRID:AB_2737935)
RPA-T8 (BV650) [anti-CD8]	BioLegend	301042 (RRID:AB_2563505)
FN50 (PE-CF594) [anti-CD69]	Becton Dickinson	562617 (RRID:AB_2737680)
Ber-ACT35 (PE-Cy7) [anti-OX40]	Biolegend	350012 (RRID:AB_10901161)a
4B4-1 (APC) [anti-CD137]	BioLegend	309810 (RRID:AB_830672)
OKT3 (AF700) [anti-CD3]	Biolegend	317340 (RRID:AB_2563408)
G043H7 (BV421) [anti-CD45RA]	BioLegend	353207 (RRID:AB_10915137)
4S.B3 (FITC) [anti-IFNγ]	Thermo Fisher Scientific	11-7319-82 (RRID: AB_465415)
GB11 (PE) [anti-Granzyme B]	Thermo Fisher Scientific	12-8899-41 (RRID: AB_1659718)
Mab11 (PeCy7) [anti-TNFα]	ebioscience	25-7349-82 (RRID:AB_469686)
JES3-19F1 (APC) [anti-IL-10]	BioLegend	506807 (RRID:AB_315457)
3D12 (APC ef780) [anti-CCR7]	eBioscience	47-1979-42 (RRID:AB_1518794)
B56 (FITC) [anti-KI67]	Becton Dickinson	556026 (RRID:AB_396302)
SK3 (percp efluor710) [anti-CD4]	Invitrogen	46-0047-42 (RRID:AB_1834401)
GB11 (af647) [anti-GzmB]	Biolegend	515406 (RRID:AB_2566333)
MHM-88 (af700) [anti-IgM]	Biolegend	314538 (RRID:AB_2566615)
O323 (APC cy7) [anti-CD27]	Biolegend	302816 (RRID:AB_571977)
IA6-2 (PE) [anti-IgD]	Becton Dickinson	555779 (RRID:AB_396114)
HCD56 (PE Dazzle) [anti-CD56]	Biolegend	318348 (RRID:AB_2563564)
HIB19 (Cy5) [anti-CD19PE]	Biolegend	302210 (RRID:AB_314240)
HIT2 (PECy7) [anti-CD38]	Invitrogen	25-0389-42 (RRID:AB_1724057)
J252D4 (bv421) [anti-CXCR5]	Biolegend	356920 (RRID:AB_2562303)
63D3 (bv510) [anti-CD14]	Biolegend	367123 (RRID:AB_2716228)
HI100 (bv570) [anti-CD45RA]	Biolegend	304132 (RRID:AB_2563813)
G025H7 (bv605) [anti-CXCR3]	Biolegend	353728 (RRID:AB_2563157)
2H7 (bv650) [anti-CD20]	Biolegend	302336 (RRID:AB_2563806)
G043H7 (bv711) [anti-CCR7]	Biolegend	353228 (RRID:AB_2563865)
EH12.2H7 (bv786) [anti-PD-1]	Biolegend	329930 (RRID:AB_2563443)
UCHT1 (buv395) [anti-CD3]	Becton Dickinson	563546 (RRID:AB_2744387)
Live/dead (UV) [Zombie]	Biolegend	423108
11A9 (buv496) [anti-CCR6]	Becton Dickinson	612948 (RRID:AB_2833076)
3G8 (buv737) [anti-CD16]	Becton Dickinson	612786 (RRID:AB_2833077)
SK1 (buv805) [anti-CD8]	Becton Dickinson	612889 (RRID:AB_2833078)
LEGENDplex 13-plex kit	Biolegend	740809
Biological Samples		
Healthy donor blood samples	Carter BloodCare	http://www.carterbloodcare.org/
Healthy donor blood samples	LJI Clinical Core	https://www.iedb.org/
Convalescent donor blood samples	UC San Diego Health	http://www.health.ucsd.edu/
Chemicals, Peptides, and Recombinant Proteins		
Synthetic peptides	Synthetic Biomolecules (aka A&A)	http://www.syntheticbiomolecules.com/
SARS-CoV-2 Receptor Binding Domain (RBD) protein	Stadlbauer et al., 2020	N/A

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Critical Commercial Assays		
CoronaCheck COVID-19 Rapid Antibody Test Kit	20/20 BioResponse	https://coronachecktest.com/
Deposited Data		
Wuhan-Hu-1 RNA isolate	NCBI nuccore database	GenBank: MN_908947
ORF10 protein	NCBI protein database	NCBI: YP_009725255.1
Nucleocapsid phosphoprotein	NCBI protein database	NCBI: YP_009724397.2
ORF8 protein	NCBI protein database	NCBI: YP_009724396.1
ORF7a protein	NCBI protein database	NCBI: YP_009724395.1
ORF6 protein	NCBI protein database	NCBI: YP_009724394.1
membrane glycoprotein	NCBI protein database	NCBI: YP_009724393.1
envelope protein	NCBI protein database	NCBI: YP_009724392.1
ORF3a protein	NCBI protein database	NCBI: YP_009724391.1
surface glycoprotein	NCBI protein database	NCBI: YP_009724390.1
orf1ab polyprotein	NCBI protein database	NCBI: YP_009724389.1
Software and Algorithms		
IEDB	Vita et al., 2019	https://www.iedb.org
IEDB-AR (analysis resource)	Dhanda et al., 2019	http://tools.iedb.org
NetMHCpan EL 4.0	Jurtz et al., 2017	http://tools.iedb.org/mhci/
IEDB	Vita et al., 2019	https://www.iedb.org
Tepitool	Paul et al., 2016; Paul et al., 2015	http://tools.iedb.org/tepitool/
FlowJo 10	FlowJo	https://www.flowjo.com/
GraphPad Prism 8.4	GraphPad	https://www.graphpad.com/
LEGENDplex v8.0	Biolegend	https://www.biolegend.com/

RESOURCE AVAILABILITY

Lead Contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Dr. Alessandro Sette (alex@lji.org).

Materials Availability

Aliquots of synthesized sets of peptides utilized in this study will be made available upon request. There are restrictions to the availability of the peptide reagents due to cost and limited quantity.

Data and Code Availability

The published article includes all data generated or analyzed during this study, and summarized in the accompanying tables, figures and Supplemental materials.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Human Subjects

Healthy Unexposed Donors

Samples from healthy adult donors were obtained by the La Jolla Institute for Immunology (LJI) Clinical Core or provided by a commercial vendor (Carter Blood Care) for prior, unrelated studies between early 2015 and early 2018. These samples were considered to be from unexposed controls, given that SARS-CoV-2 emerged as a novel pathogen in late 2019, more than one year after the collection of any of these samples. These donors were considered healthy in that they had no known history of any significant systemic diseases, including, but not limited to, autoimmune disease, diabetes, kidney or liver disease, congestive heart failure, malignancy, coagulopathy, hepatitis B or C, or HIV. An overview of the characteristics of these unexposed donors is provided in Table 1.

The LJI Institutional Review Board approved the collection of these samples (LJI; VD-112). At the time of enrollment in the initial studies, all individual donors provided informed consent that their samples could be used for future studies, including this study.

Convalescent COVID-19 Donors

The Institutional Review Boards of the University of California, San Diego (UCSD; 200236X) and La Jolla Institute (LJI; VD-214) approved blood draw protocols for convalescent donors. All human subjects were assessed for capacity using a standardized and approved assessment. Subjects deemed to have capacity voluntarily gave informed consent prior to being enrolled in the study. Individuals did not receive compensation for their participation.

Study inclusion criteria included subjects over the age of 18 years, regardless of disease severity, race, ethnicity, gender, pregnancy or nursing status, who were willing and able to provide informed consent, or with a legal guardian or representative willing and able to provide informed consent when the participant could not personally do so. Study exclusion criteria included lack of willingness or ability to provide informed consent, or lack of an appropriate legal guardian or representative to provide informed consent.

Blood from convalescent donors was obtained at a UC San Diego Health clinic. Blood was collected in acid citrate dextrose (ACD) tubes and stored at room temperature prior to processing for PBMC isolation and plasma collection. A separate serum separator tube (SST) was collected from each donor. Samples were de-identified prior to analysis. Other efforts to maintain the confidentiality of participants included referring to specimens and other records via an assigned, coded identification number.

Prior to enrollment in the study, donors were asked to provide proof of positive testing for SARS-CoV-2, and screened for clinical history and/or epidemiological risk factors consistent with the World Health Organization (WHO) or Centers for Disease Control and Prevention (CDC) case definitions of COVID-19 or Persons Under Investigation (PUI) (https://www.who.int/emergencies/diseases/novel-coronavirus-2019/technical-guidance/surveillance-and-case-definitions, https://www.cdc.gov/coronavirus/2019-nCoV/hcp/clinical-criteria.html). Per CDC and WHO guidance, clinical features consistent with COVID-19 included subjective or measured fever, signs or symptoms of lower respiratory tract illness (e.g., cough or dyspnea). Epidemiologic risk factors included close contact with a laboratory-confirmed case of SARS-CoV-2 within 14 days of symptom onset or a history of travel to an area with a high rate of COVID-19 cases within 14 days of symptom onset.

Disease severity was defined as mild, moderate, severe or critical based on a modified version of the WHO interim guidance, ''Clinical management of severe acute respiratory infection when COVID-19 is suspected'' (WHO Reference Number: WHO/2019-nCoV/clinical/2020.4). Mild disease was defined as an uncomplicated upper respiratory tract infection (URI) with potential non-specific symptoms (e.g., fatigue, fever, cough with or without sputum production, anorexia, malaise, myalgia, sore throat, dyspnea, nasal congestion, headache; rarely diarrhea, nausea and vomiting) that did not require hospitalization. Moderate disease was defined as the presence of lower respiratory tract disease or pneumonia without the need for supplemental oxygen, without signs of severe pneumonia, or a URI requiring hospitalization (including observation admission status). Severe disease was defined as severe lower respiratory tract infection or pneumonia with fever plus any one of the following: tachypnea (respiratory rate $>$ 30 breaths per minute), respiratory distress, or oxygen saturation less than 93% on room air. Critical disease was defined as the need for ICU admission or the presence of acute respiratory distress syndrome (ARDS), sepsis, or septic shock, as defined in the WHO guidance document.

Convalescent donors were screened for symptoms prior to scheduling blood draws, and had to be symptom-free and approximately 3 weeks out from symptom onset at the time of the initial blood draw. Following enrollment, whole blood from convalescent donors was run on a colloidal-gold immunochromatographic 'lateral flow' assay to evaluate for prior exposure to SARS-CoV-2. This assay detects IgM or IgG antibodies directed against recombinant SARS-CoV-2 antigen labeled with a colloidal gold tracer (20/20 BioResponse CoronaCheck). Ninety percent of convalescent donors tested positive for IgM or IgG to SARS-CoV-2 by this assay (Table 1).

Convalescent donors were California residents, who were either referred to the study by a health care provider or self-referred. The majority (75%) of donors had a known sick contact with COVID-19 or suspected exposure to SARS-CoV-2 (Table 1). The most common symptoms reported were cough, fatigue, fever, anosmia, and dyspnea. Seventy percent of donors experienced mild illness. Donors were asked to self-report any known medical illnesses. Of note, 65% of these individuals had no known underlying medical illnesses.

METHOD DETAILS

Peptide Pools

Epitope MegaPool (MP) design and preparation

SARS-CoV-2 virus-specific CD4 and CD8 peptides were synthesized as crude material (A&A, San Diego, CA), resuspended in DMSO, pooled and sequentially lyophilized as previously reported (Carrasco Pro et al., 2015). SARS-CoV-2 epitopes were predicted using the protein sequences derived from the SARS-CoV-2 reference (GenBank: MN908947) and IEDB analysis-resource as previously described (Dhanda et al., 2019; Grifoni et al., 2020). Specifically, CD4 SARS-CoV-2 epitope prediction was carried out using a previously described approach in Tepitool resource in IEDB (Paul et al., 2015; Paul et al., 2016), to select peptides with median consensus percentile $\leq$ 20, similar to what was previously described, but removing the resulting spike glycoprotein epitopes

from this prediction (CD4-R (remainder) "Non-spike" MP, n = 221). This approach takes advantage of the extensive cross-reactivity and repertoire overlap between different HLA class II loci and allelic variants to predict promiscuous epitopes, capable of binding across the most common HLA class II prototypic specificities (Greenbaum et al., 2011; O'Sullivan et al., 1991; Sidney et al., 2010a, b; Southwood et al., 1998). The algorithm utilizes predictions for seven common HLA-DR alleles (DRB1*03:01, DRB1*07:01, DRB1*15:01, DRB3*01:01, DRB3*02:02, DRB4*01:01 and DRB5*01:01) empirically determined to allow coverage of diverse populations and for different pathogens and antigen systems (Dhanda et al., 2018; Paul et al., 2015).

To investigate in-depth spike-specific CD4 T cells, 15-mer peptides (overlapping by 10 amino acids) spanning the entire antigen have been synthesized and pooled separately (CD-4 S (spike) MP, n = 253).

In the case of CD8 epitopes, since the overlap between different HLA class I allelic variants and loci is more limited to specific groups of alleles, or supertypes (Sidney et al., 2008), we targeted a set of the 12 most prominent HLA class I A and B alleles (A*01:01, A*02:01, A*03:01, A*11:01, A*23:01, A*24:02, B*07:02, B*08:01, B*35:01, B*40:01, B*44:02, B*44:03), which have been shown to allow broad coverage of the general population. CD8 SARS-CoV-2 epitope prediction was performed as previously reported, using NetMHC pan EL 4.0 algorithm (Jurtz et al., 2017) for the top 12 more frequent HLA alleles and selecting the top 1 percentile predicted epitope per HLA allele clustered with nested/overlap reduction (Grifoni et al., 2020). The 628 predicted CD8 epitopes were split in two CD8 MPs containing 314 peptides each (CD8-A and CD8-B). The CMV MP is a pool of previously reported class I and class II epitopes (Carrasco Pro et al., 2015).

Protein peptide pools

In the case of the protein pools, peptides of 15 amino acid length overlapping by 10 spanning each entire protein sequence were tested in a single MP (6-253 peptides per pool). Table S1 lists the number of peptides pooled for each of the viral proteins. Upon request we are prepared to make these MP available to the scientific community for use in a diverse set of investigations.

PBMC isolation

For all samples whole blood was collected in ACD tubes (COVID-19 donors) or heparin coated blood bag (healthy unexposed donors). Whole blood was then centrifuged for 15 min at 1850 rpm to separate the cellular fraction and plasma. The plasma was then carefully removed from the cell pellet and stored at −20C.

Peripheral blood mononuclear cells (PBMC) were isolated by density-gradient sedimentation using Ficoll-Paque (Lymphoprep, Nycomed Pharma, Oslo, Norway) as previously described (Weiskopf et al., 2013). Isolated PBMC were cryopreserved in cell recovery media containing 10% DMSO (GIBCO), supplemented with 10% heat inactivated fetal bovine serum, depending on the processing laboratory, (FBS; Hyclone Laboratories, Logan UT) and stored in liquid nitrogen until used in the assays.

SARS-CoV-2 RBD ELISA

SARS-CoV-2 Receptor Binding Domain (RBD) protein was obtained courtesy of Florian Krammer and Peter Kim (Stadlbauer et al., 2020). Corning 96-well half-area plates (ThermoFisher 3690) were coated with 1μg/mL SARS-CoV-2 RBD overnight at 4°C. ELISA protocol generally followed that of the Krammer lab, which previously demonstrated specificity (Stadlbauer et al., 2020). Plates were blocked the next day with 3% milk (Skim Milk Powder ThermoFisher LP0031 by weight/volume) in Phosphate Buffered Saline (PBS) containing 0.05% Tween-20 (ThermoScientific J260605-AP) for 2 hours at room temperature. Plasma was then added to the plates and incubated for 1.5 hours at room temperature. Prior to plasma addition to the plates, plasma was heat inactivated at 56°C for 30-60 minutes. Plasma was diluted in 1% milk in 0.05% PBS-Tween 20 starting at a 1:3 dilution and diluting each sample at by 1:3. Plates were then washed 5 times with 0.05% PBS-Tween 20. Secondary antibodies were diluted in 1% milk in 0.05% Tween-20 and incubated for 1 hour. For IgG, anti-human IgG peroxidase antibody produced in goat (Sigma A6029) was used at a 1:5000 dilution. For IgM, anti-human IgM peroxidase antibody produced in goat (Sigma A6907) was used at a 1:10,000 dilution. For IgA, anti-human IgA horseradish peroxidase antibody (Hybridoma Reagent Laboratory HP6123-HRP) was used at a 1:1,000 dilution. Plates were washed 5 times with 0.05% PBS-Tween 20. Plates were developed with TMB Substrate Kit (ThermoScientific 34021) for 15 minutes at room temperature. The reaction was stopped with 2M sulfuric acid. Plates were read on a Spectramax Plate Reader at 450 nm using SoftMax Pro, and ODs were background subtracted. A positive control standard was created by pooling plasma from six convalescing COVID-19 patients. Positive control standard was run on each plate and was used to calculate titers (relative units) for all samples using non-linear regression interpolations, done to quantify the amount of anti-RBD IgG, anti-RBD IgM, and anti-RBD IgA present in each specimen. Titers were plotted for each specimen and compared to COVID-19 negative specimens. As a second analytical approach, Area under the curve was also calculated for each specimen to compare COVID-19 to negative specimens, using a baseline of 0.05 for peak calculations.

OC43 and NL63 coronavirus RBD ELISA

An in-house ELISA at UNC was performed by coating with recombinant S RBD antigens (SARS-CoV-2, SARS-CoV, OC43-CoV and NL63-CoV) in TBS for 1 h at 37 C. After blocking, we added 1:20 diluted serum and incubated at 37°C for 1 h. Antigen-specific antibodies (Ig) were measured at 405 nm by using alkaline phosphatase conjugated goat anti-human IgG, IgA and IgM Abs and 4-Nitrophenyl phosphate.

Flow Cytometry

Direct ex vivo PBMC immune cell phenotyping

For the surface stain, 1x10^6 PBMCs were resuspended in 100 μl PBS with 2% FBS (FACS buffer) and stained with antibody cocktail for 1 hour at 4°C in the dark. Following surface staining, cells were washed twice with FACS buffer. Cells were then fixed/permeabilized for 40min at 4C in the dark using the eBioscience FoxP3 transcription factor buffer kit (ThermoFisher Scientific, Waltham, MA). Following fixation/permeabilization, cells were washed twice with 1x permeabilization buffer, resuspended in 100 μl permeabilization buffer and stained with intracellular/intranuclear antibodies for 1 hour at 4°C in the dark. Samples were washed twice with 1x permeabilization buffer following staining. After the final wash, cells were resuspended in 200μl FACS buffer. All samples were acquired on a BD FACSymphony cell sorter (BD Biosciences, San Diego, CA). A list of antibodies used in this panel can be found in Table S2.

T cell stimulations

For all flow cytometry assays of stimulated T cells, cryopreserved cells were thawed by diluting them in 10 mL complete RPMI 1640 with 5% human AB serum (Gemini Bioproducts) in the presence of benzonase [20ul/10mL]. All samples were acquired on a ZE5 Cell analyzer (Bio-rad laboratories), and analyzed with FlowJo software (Tree Star, San Carlos, CA).

Activation induced cell marker assay

Cells were cultured for 24 hours in the presence of SARS-CoV-2 specific MPs [1 μg/ml] or 10 μg/mL PHA in 96-wells U bottom plates at 1x10^6 PBMC per well. A stimulation with an equimolar amount of DMSO was performed as negative control, phytohemagglutinin (PHA, Roche, 1 μg/ml) and stimulation with a combined CD4 and CD8 cytomegalovirus MP (CMV, 1 μg/ml) were included as positive controls. Supernatants were harvested at 24 hours post-stimulation for multiplex detection of cytokines. Antibodies used in the AIM assay are listed in Table S4**.** AIM assays shown in Figures 2 and 3 and AIM assays shown in Figure 6 had five COVID-19 donors in common and nine Unexposed donors. Full raw data is listed in Table S6**.**

Intracellular cytokine staining assay

For the intracellular cytokine staining, PBMC were cultured in the presence of SARS-CoV-2 specific MPs [1 μg/ml] for 9 hours. Golgi-Plug containing brefeldin A (BD Biosciences, San Diego, CA) and monensin (Biolegend, San Diego, CA) were added 3 hours into the culture. Cells were then washed and surface stained for 30 minutes on ice, fixed with 1% of paraformaldehyde (Sigma-Aldrich, St. Louis, MO) and kept at 4°C overnight. Antibodies used in the ICS assay are listed in Table S5**.** The gates applied for the identification of IFNγ, GzB, TNFα, or IL-10 production on the total population of CD8$^+$ T cells were defined according to the cells cultured with DMSO for each individual.

Cytokine bead assays

Supernatants were collected from 24-hour stimulation cultures of the AIM assays and stored in 96 well plates at −20°C. Cytokines in cell culture supernatants of the same samples used for AIM were quantified using a human Th cytokine panel (13-plex) kit (LEGENDplex, Biolegend) according to the manufacturer's instruction. Supernatants were mixed with beads coated with capture antibodies specific for IL-5, IL-13, IL-2, IL-6, IL-9, IL-10, IFNγ, TNFα, IL-17a, IL-17F, IL-4, IL-21 and IL-22 and incubated on a 96 well filter plate for 2 hours. Beads were washed and incubated with biotin-labeled detection antibodies for 1 hour, followed by a final incubation with streptavidin-PE. Beads were analyzed by flow cytometry using a FACS Canto cytometer. Analysis was performed using the LEGENDplex analysis software v8.0, which distinguishes between the 13 different analytes on basis of bead size and internal dye.

Identification of coronavirus epitopes and associated literature references

To identify coronavirus epitopes and associated references, the IEDB was searched (on April 16, 2020) utilizing the following queries. A first query was run to identify references associated with class I restricted CD8 epitopes, which utilized the criteria settings ''Antigen'': Organism = Coronavirus (taxonomy ID 11118); ''Assay'': Positive assays only; ''Assay'': T cell assay; ''MHC restriction'' = MHC Class II; no parameters were defined for ''Host'' or ''Disease.'' This query identified 57 references, which are listed and displayed under the ''References'' tab on the results page.

A second query was run to identify references associated with class II restricted CD4 epitopes which utilized the criteria settings ''Antigen'': Organism = Coronavirus (taxonomy ID 11118); ''Assay'': Positive assays only; ''Assay'': T cell assay; ''MHC restriction'' = MHC Class II; no parameters were defined for ''Host'' or ''Disease.'' This query identified 27 references, which are listed and displayed under the ''References'' tab on the results page.

A third query was run to specifically capture epitopes and map them back to the antigen of origin using the setting; ''Antigen'': Organism = Coronavirus (taxonomy ID 11118); ''Assay'': Positive assays only; ''Assay'': T cell assay; no parameters were defined for ''MHC restriction,'' ''Host'' or ''Disease.'' Results were exported as csv files, and then examined in Excel to tabulate the number of CD4 and CD8 epitopes recognized in humans, mice, transgenic mice and other hosts associated with each respective antigen.

QUANTIFICATION AND STATISTICAL ANALYSIS

Data and statistical analyses were done in FlowJo 10 and GraphPad Prism 8.4, unless otherwise stated. The statistical details of the experiments are provided in the respective figure legends. Data plotted in linear scale were expressed as Mean + Standard Deviation (SD). Data plotted in logarithmic scales were expressed as Geometric Mean + Geometric Standard Deviation (SD). Correlation analyses were performed using Spearman, while Mann-Whitney or Wilcoxon tests were applied for unpaired or paired comparisons, respectively. Details pertaining to significance are also noted in the respective legends. T cell data have been calculated as

background subtracted data or stimulation index. Background subtracted data were derived by subtracting the percentage of AIM^+ cells after SARS-CoV-2 stimulation from the DMSO stimulation. Stimulation Index was calculated instead by dividing the percentage of AIM+ cells after SARS-CoV-2 stimulation with the percentage of AIM^+ cells derived from DMSO stimulation. If the AIM^+ cells percentage after DMSO stimulation was equal to 0, the minimum value across each cohort was used. When two stimuli were combined together, the percentage of AIM^+ cells after SARS-CoV-2 stimulation was combined and either subtracted twice or divided by twice the value of the percentage of AIM+ cells derived from DMSO stimulation. Additional data analysis techniques are described in the STAR Methods sections above.

Supplemental Figures

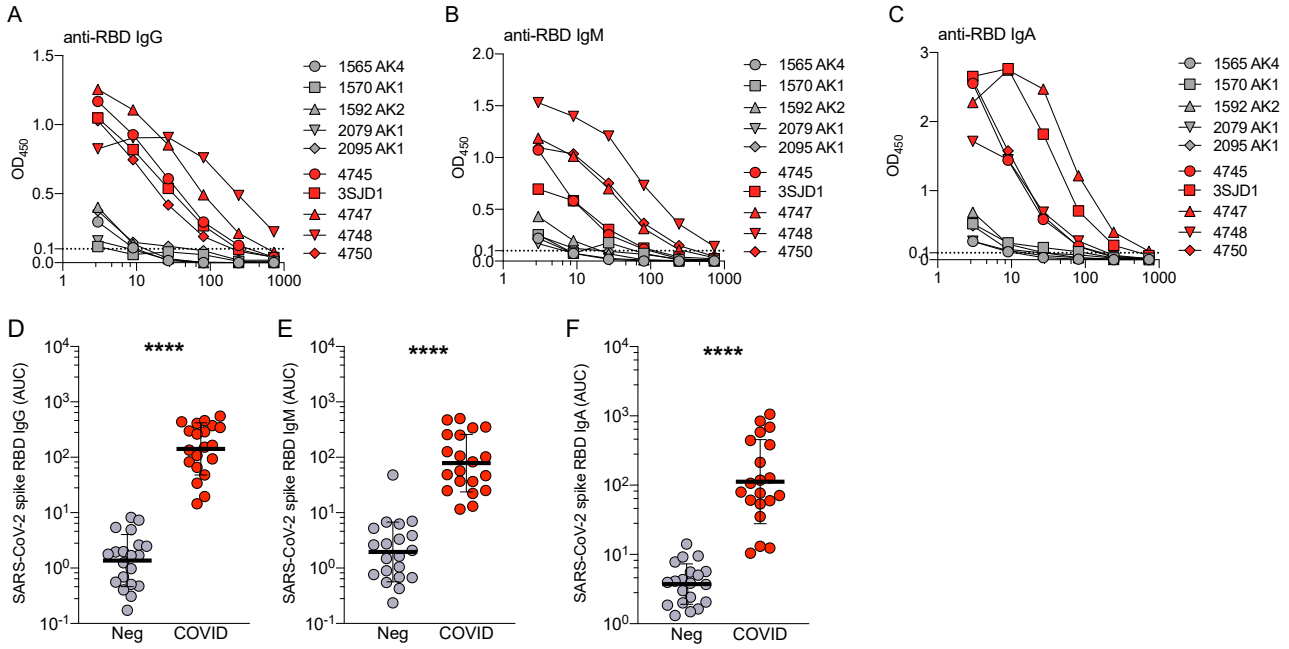


Figure S1. SARS-CoV-2 Spike Protein RBD Serology, Related to Figure 1

(A-C) ELISA curves for (**A**) IgG, (**B**) IgM, and (**C**) IgA from 10 representative donors. Five COVID-19 cases and (**D-F**) Area under the curve (AUC) SARS-CoV-2 spike protein RBD (**D**) IgG (**E**) IgM, and (**F**) IgA, ELISA quantitation, from the same donors and experiments shown in Figure 1. Geometric mean titers with geometric SDs are indicated. P values are two-tailed Mann-Whitney tests.

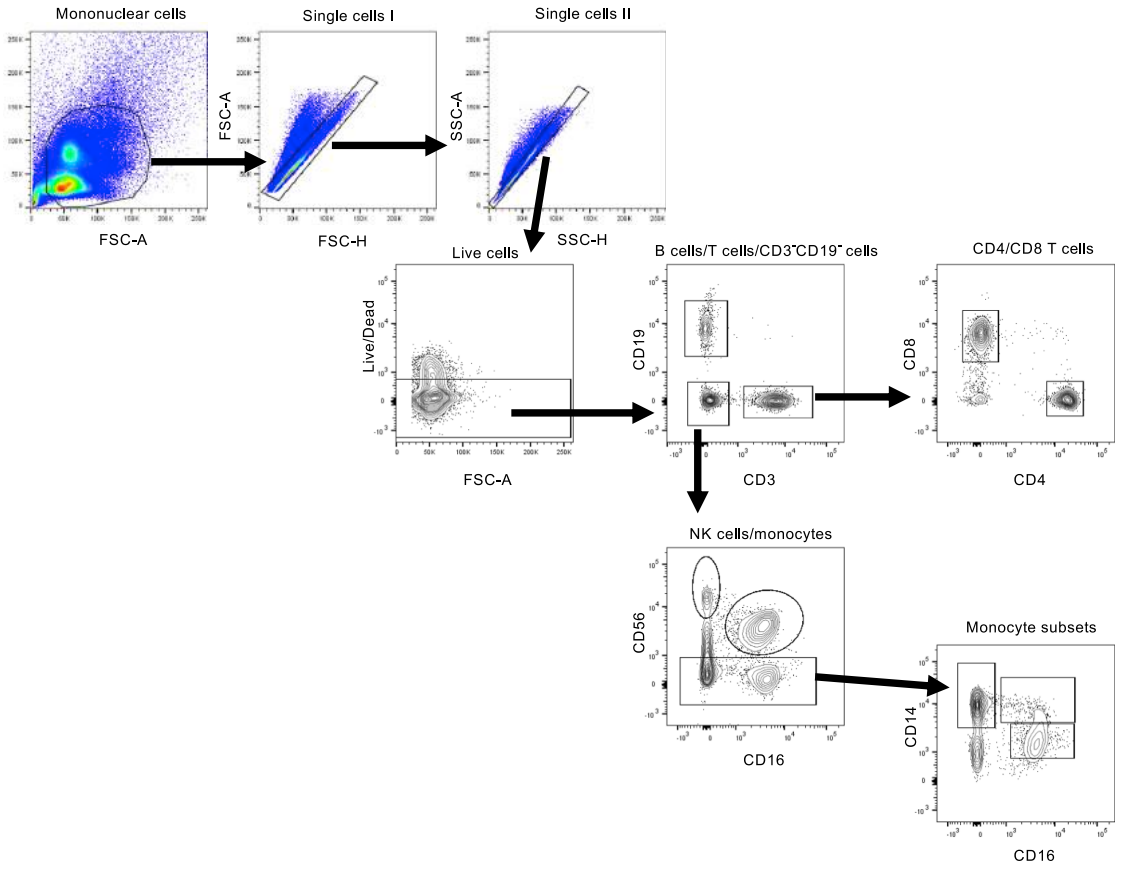


Figure S2. Phenotyping Flow Cytometry, Related to Figure 1

Representative gating of CD3$^+$ T cells, CD19$^+$ B cells, CD3$^-$CD19$^-$ cells, CD4$^+$ T cells, CD8$^+$ T cells and CD14$^+$ monocytes from donor PBMCs is shown. Briefly, mononuclear cells were gated out of all events followed by subsequent singlet gating. Live cells are gated as Zombie UV$^-$. Cells were then gated as CD19-PE-Cy5$^+$, CD3-buv395$^+$ or CD19$^-$CD3$^-$ cells. T cells were further subdivided into either CD8-buv805$^+$ or CD4-PerCPefluor710$^+$ populations. CD3$^-$CD19$^-$ cells were defined as CD56-PE-Dazzle^{bright} NK cells, CD56^{dim}CD-16buv737$^+$ NK cells or CD56$^-$ monocytes. Monocytes were further classified on differential expression of CD14-bv510 and CD16.

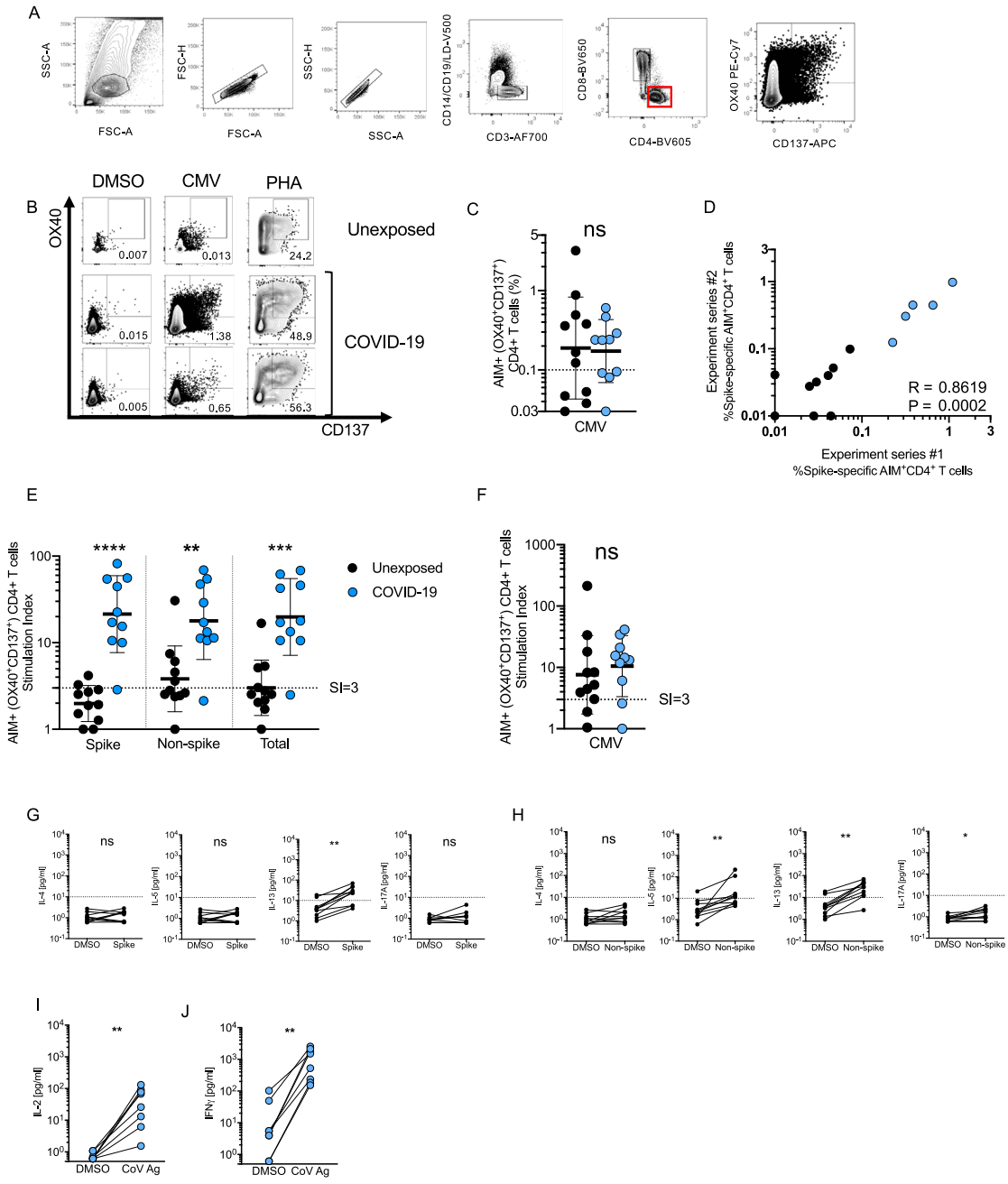


Figure S3. SARS-CoV-2-Specific CD4⁺ T Cell Responses of Recovered COVID-19 Patients, Related to Figure 2

(**A**) Example flow cytometry gating strategy.

(**B**) FACS plot examples for controls. DMSO negative control, CMV positive control, PHA positive control.

(**C**) CMV-specific CD4⁺ T cells as percentage of AIM⁺ (OX40⁺CD137⁺) CD4⁺ T cells after stimulation of PBMCs with CMV peptide pool. Data were background subtracted against DMSO negative control and are shown with geometric mean and geometric standard deviation. Samples were from unexposed donors ("Unexposed," n = 11) and recovered COVID-19 patients ("COVID-19," n = 10).

(**D**) Spearman correlation of SARS-CoV-2 spike–specific CD4⁺ T cells (AIM⁺ (OX40⁺CD137⁺) CD4⁺ T cells, background subtracted) after stimulation with spike pool run on the same donors in two independent experiment series run on different dates. COVID-19 patient samples shown in blue. Unexposed donor samples shown in black.

(**E-F**) Stimulation index quantitation of AIM⁺ (OX40⁺CD137⁺) CD4⁺ T cells; the same samples as in Figure 2 and Figure S3C were analyzed.

(**G-H**) Cytokine levels in the supernatants of AIM assays after stimulation with (**G**) Spike MP (MP_S), or (**H**) CD4-R ("Non-spike"). Data are shown in comparison to the negative control (DMSO), per donor.

(legend continued on next page)

(**I-J**) Cytokine production by $CD4^+$ T cells in response to Non-spike (CD4-R MP) or Spike (MP_S) peptide pools ("CoV antigen (Ag)") was confirmed by analyzing cytokine secretion from the subset of COVID-19 donors determined to have low or negative $CD8^+$ T cell responses (< 0.1% by AIM) to the same peptide pool determined positive for SARS-CoV-2–specific $CD4^+$ T cells by AIM. (**I**) IL-2. (**J**) IFNγ.
Statistical comparisons across cohorts were performed with the Mann-Whitney test, while paired sample comparisons were performed with the Wilcoxon test. **p < 0.01; ***p < 0.001. ns not significant.

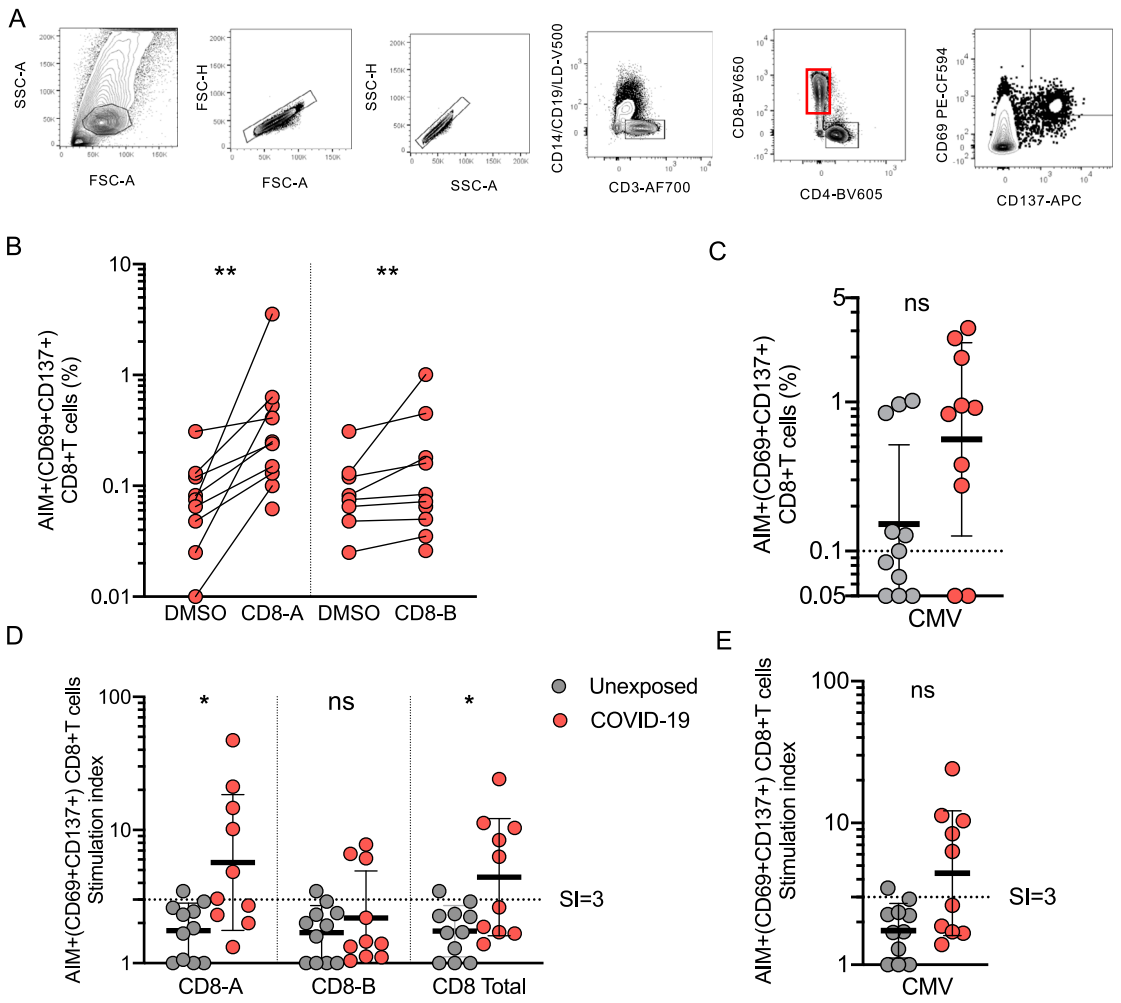


Figure S4. SARS-CoV-2-Specific CD8⁺ T Cell Responses of Recovered COVID-19 Patients, Related to Figure 3

(**A**) Flow cytometry gating strategy.

(**B**) SARS-CoV-2–specific CD8⁺ T cells as determined by AIM⁺ (CD69⁺CD137⁺) CD8⁺ T cells. Response of PBMCs from COVID-19 cases between the negative control (DMSO) and antigen specific stimulation.

(**C**) CMV-specific CD8⁺ T cells as percentage of AIM⁺ (CD69⁺CD137⁺) CD8⁺ T cells after stimulation of PBMCs with CMV peptide pool. Data were background subtracted against DMSO negative control and are shown with geometric mean and geometric standard deviation. Samples were from unexposed donors ("Unexposed," n = 11) and recovered COVID-19 patients ("COVID-19," n = 10).

(**D-E**) Stimulation index quantitation of AIM⁺ (CD69⁺CD137⁺) CD8⁺ T cells; the same samples as in Figure 2 and Figure S4C were analyzed.

Statistical comparisons across cohorts were performed with the Mann-Whitney test, while paired sample comparisons were performed with the Wilcoxon test. **p < 0.01; ***p < 0.001. ns not significant.

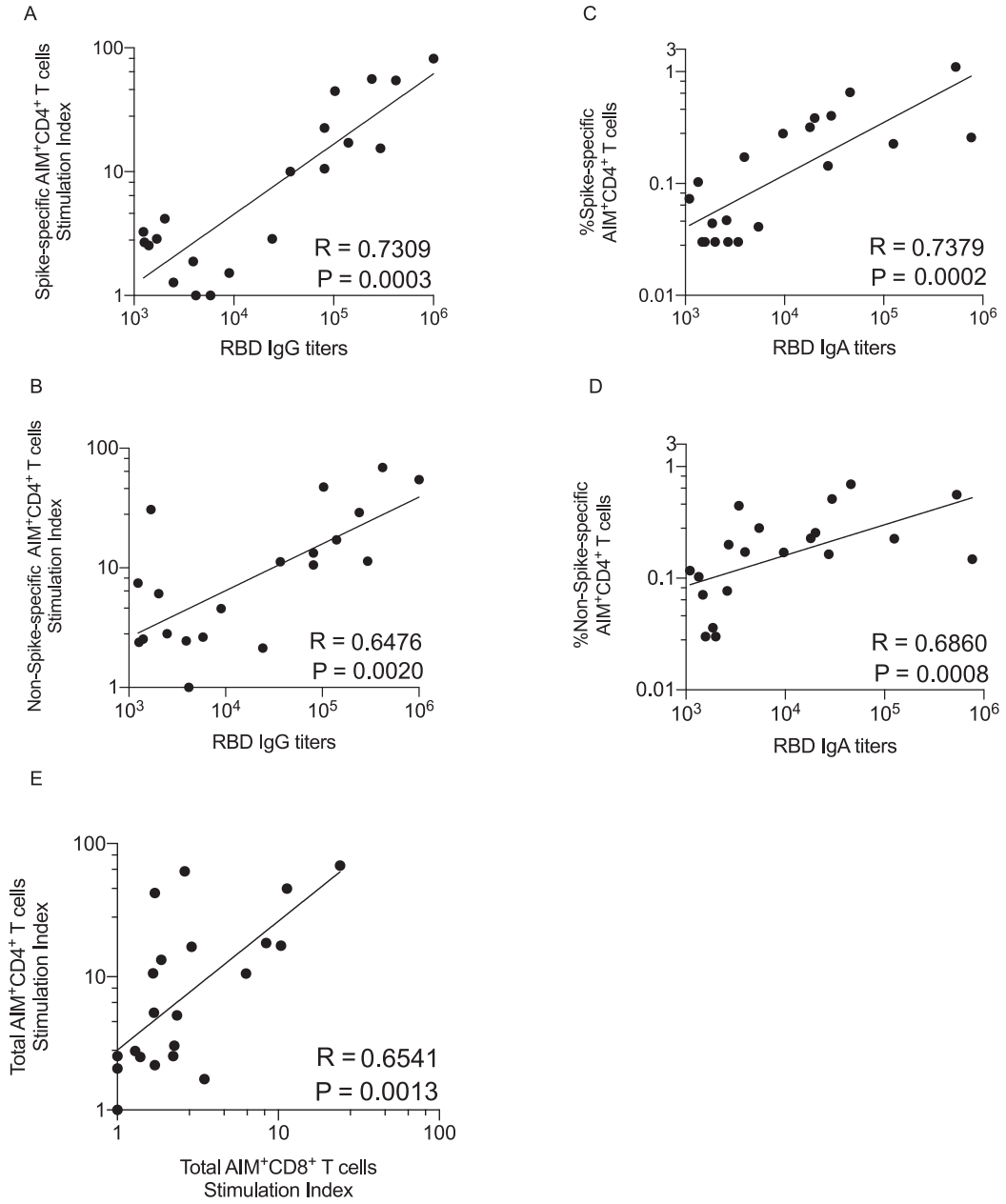


Figure S5. Correlations between SARS-CoV-2-Specific CD4⁺ T Cells, Antibodies, and CD8⁺ T Cells, Related to Figure 4

(**A**) Correlation between SARS-CoV-2 spike–specific CD4⁺ T cells and anti-spike RBD IgG, using CD4⁺ T cell stimulation index.

(**B**) Correlation between SARS-CoV-2 non-spike–specific CD4⁺ T cells and anti-spike RBD IgG, using CD4⁺ T cell stimulation index.

(**C**) Correlation between SARS-CoV-2 spike–specific CD4⁺ T cells (%) and anti-spike RBD IgA.

(**D**) Correlation between SARS-CoV-2 spike–specific CD4⁺ T cells (%) and anti-spike RBD IgA.

(**E**) Correlation between SARS-CoV-2–specific CD4⁺ T cells and SARS-CoV-2–specific CD8⁺ T cells, using stimulation index. Total MP responses per donor were used in each case ("Non-spike" + "spike" (CD4_R + MP_S) for CD4⁺ T cells, CD8-A + CD8-B for CD8⁺ T cells).

Statistical comparisons were performed using Spearman correlation.

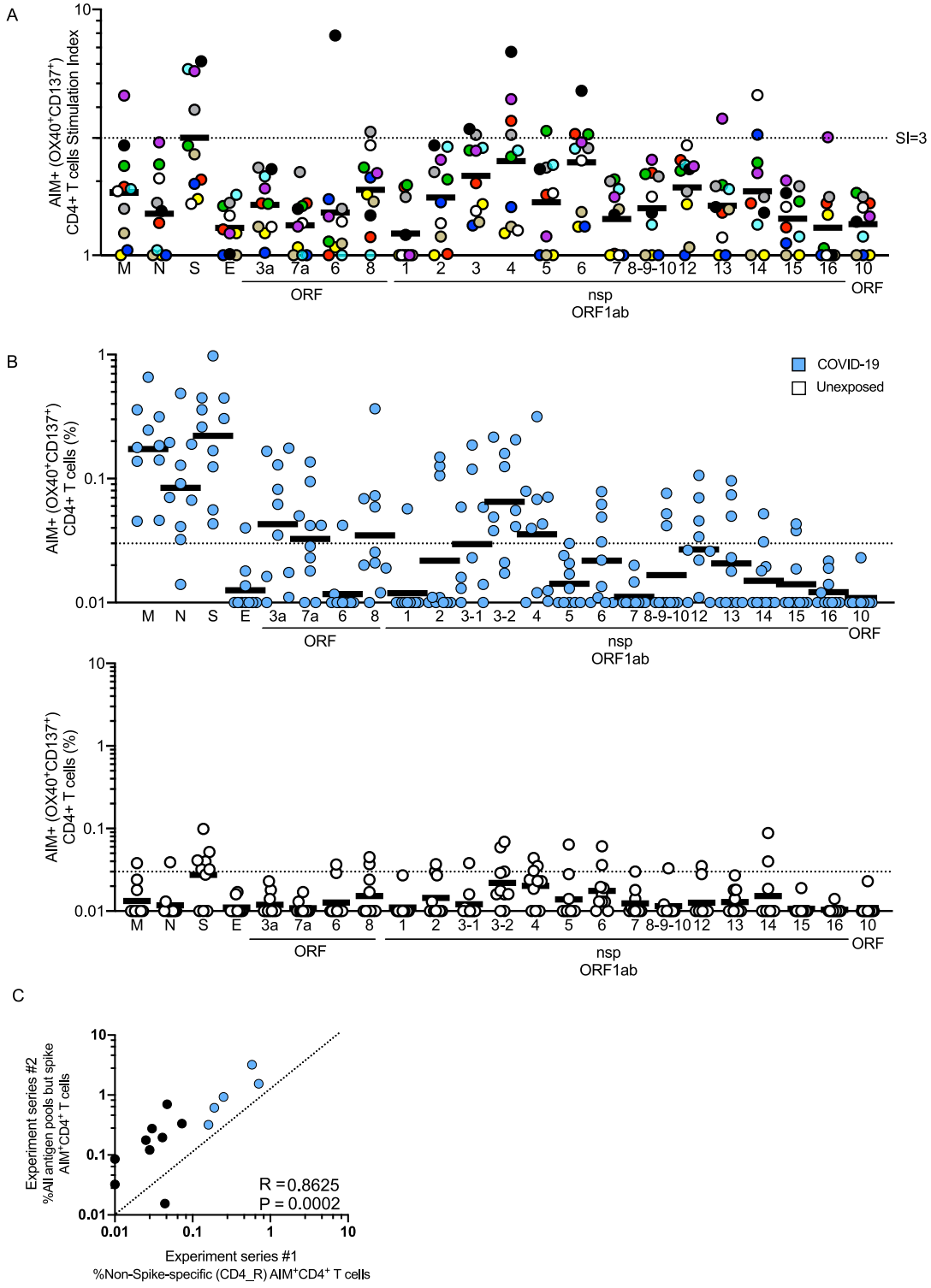


Figure S6. Protein Immunodominance of SARS-CoV-2 Specific CD4⁺ T Cells in Recovered COVID-19 Patients and Unexposed Donors, Related to Figure 6

(**A**) The same data as Figure 6B, but with each unexposed donor color coded.

(**B**) The same experiment as Figure 6B, but with SARS-CoV-2–specific CD4⁺ T cells measured as percentage of AIM⁺ (OX40⁺CD137⁺) CD4⁺ T cells, after background subtraction. COVID-19 cases (top, in blue. n = 10) and unexposed donors (bottom, in white. n = 10).

(legend continued on next page)

(**C**) Correlation of SARS-CoV-2–specific $CD4^+$ T cells detected using the epitope prediction approach (CD4_R MP) compared against the sum total of all antigen pools of overlapping peptides (excluding spike), run with samples from the same donors in two different experiment series. Dotted line indicates 1:1 concordance. Statistical comparison was performed using Spearman correlation.

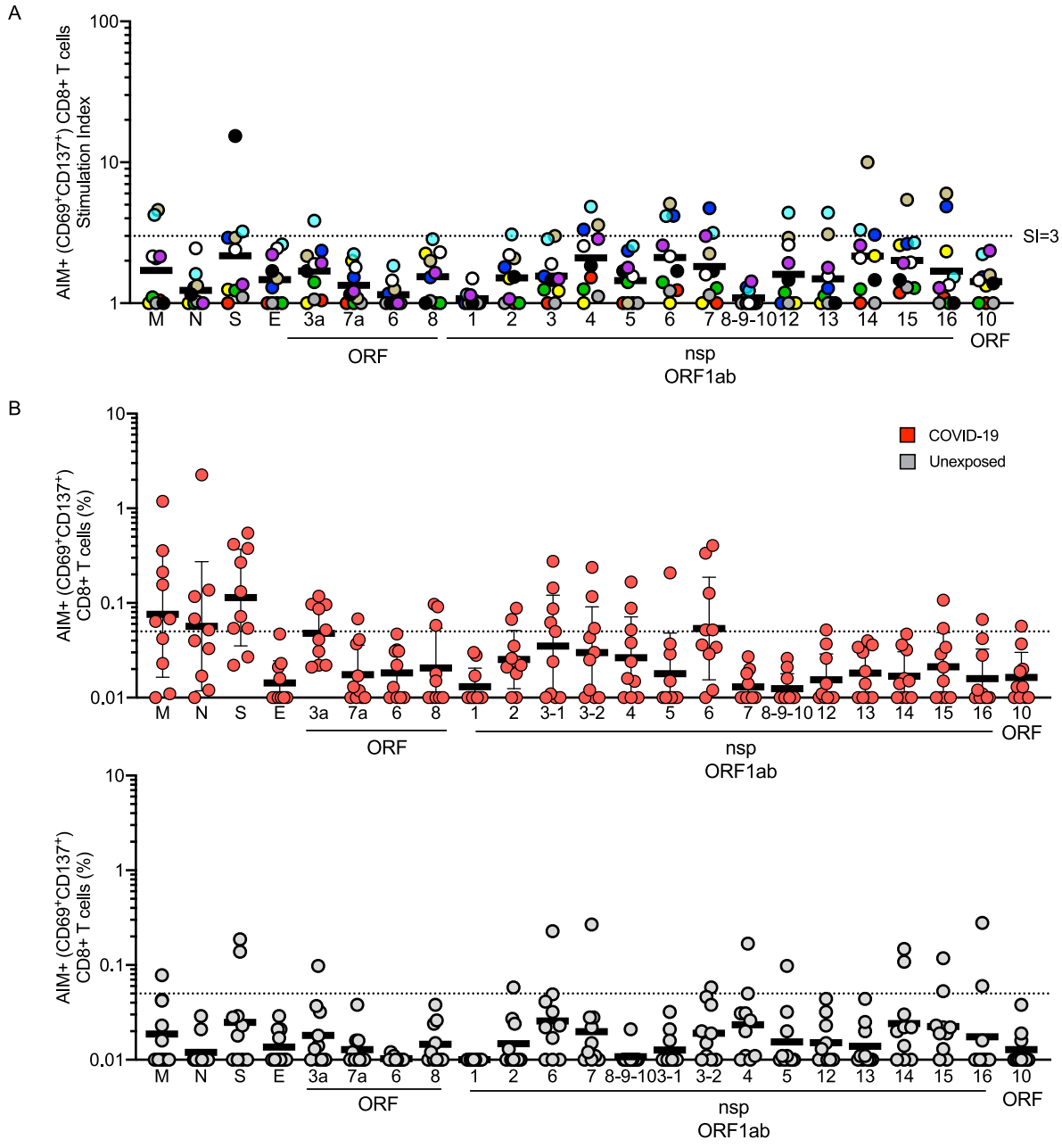


Figure S7. Protein Immunodominance of SARS-CoV-2-Specific CD8⁺ T Cells in Recovered COVID-19 Patients and Unexposed Donors, Related to Figure 6

(**A**) The same data as Figure 6D, but with each unexposed donor color coded.

(**B**) The same experiment as Figure 6D, but with SARS-CoV-2–specific CD8⁺ T cells measured as percentage of AIM⁺ (CD69⁺CD137⁺) CD8⁺ T cells, after background subtraction. COVID-19 cases (top, in red. n = 10) and unexposed donors (bottom, in gray. n = 10).

The Journal of Immunology

Thymus-Derived $CD4^+CD8^+$ Cells Reside in Mediastinal Adipose Tissue and the Aortic Arch

Holger Winkels, Yanal Ghosheh, Kouji Kobiyama, William B. Kiosses, Marco Orecchioni, Erik Ehinger, Vasantika Suryawanshi, Sara Herrera-De La Mata, Paola Marchovecchio, Thomas Riffelmacher, Nicolas Thiault, Mitchell Kronenberg, Dennis Wolf, Gregory Seumois, Pandurangan Vijayanand and Klaus Ley

J Immunol published online 5 November 2021
http://www.jimmunol.org/content/early/2021/11/05/jimmunol.2100208

Supplementary Material **http://www.jimmunol.org/content/suppl/2021/11/05/jimmunol.2100208.DCSupplemental**

The Journal of Immunology is published twice each month by
The American Association of Immunologists, Inc.,
1451 Rockville Pike, Suite 650, Rockville, MD 20852

Print ISSN: 0022-1767 Online ISSN: 1550-6606.

Published November 5, 2021, doi:10.4049/jimmunol.2100208

Thymus-Derived $CD4^+CD8^+$ Cells Reside in Mediastinal Adipose Tissue and the Aortic Arch

Holger Winkels,*,† Yanal Ghosheh,* Kouji Kobiyama,* William B. Kiosses,* Marco Orecchioni,* Erik Ehinger,* Vasantika Suryawanshi,* Sara Herrera-De La Mata,* Paola Marchovecchio,* Thomas Riffelmacher,* Nicolas Thiault,* Mitchell Kronenberg,* Dennis Wolf,‡ Gregory Seumois,* Pandurangan Vijayanand,* and Klaus Ley*,§

Double-positive $CD4^+CD8\alpha\beta^+$ (DP) cells are thought to reside as T cell progenitors exclusively within the thymus. We recently discovered an unexpected $CD4^+$ and $CD8\alpha\beta^+$ immune cell population in healthy and atherosclerotic mice by single-cell RNA sequencing. Transcriptomically, these cells resembled thymic DPs. Flow cytometry and three-dimensional whole-mount imaging confirmed DPs in thymus, mediastinal adipose tissue, and aortic adventitia, but nowhere else. Deep transcriptional profiling revealed differences between DP cells isolated from the three locations. All DPs were dependent on RAG2 expression and the presence of the thymus. Mediastinal adipose tissue DPs resided in close vicinity to invariant NKT cells, which they could activate in vitro. Thymus transplantation failed to reconstitute extrathymic DPs, and frequencies of extrathymic DPs were unaltered by pharmacologic inhibition of S1P1, suggesting that their migration may be locally confined. Our results define two new, transcriptionally distinct subsets of extrathymic DPs that may play a role in aortic vascular homeostasis. ***The Journal of Immunology*****, 2021, 207: 1–13.**

Pro- and anti-inflammatory immune cells modulate the progression of atherosclerosis locally and systemically (1). We and others have recently used high-parametric methodologies, 10× Genomics based single-cell RNA sequencing (scRNA-seq) and mass cytometry (cytometry by time of flight) to define aortic leukocytes in nonatherosclerotic and atherosclerotic mice (2, 3). With both approaches, we identified the unexpected presence of a cell population, which did not express surface TCR but was $CD4^+$, $CD8\alpha^+$, and $CD8\beta^+$.

The only known cell population in the murine and human body harboring these features are $CD4^+CD8^+TCR^-$ (double positives [DPs]), a progenitor stage of T cells found exclusively in the thymus, where the development of T cells takes place. In this process, bone marrow–derived progenitors migrate into the thymus, and a portion of them differentiate into naive single $CD4^+$ or $CD8^+$ TCRαβ-expressing T cells (4). Recombinase activating gene (RAG) 1 and 2 catalyze the rearrangement of the V, D, and J TCRβ gene segments into a unique combination of a TCR β-chain. This chain will pair with an invariant preliminary α-chain (*pTa*) to form a pre-TCR in conjunction with the CD3 complex. Pre-TCR signaling is necessary for thymocyte survival and proliferation and blocks further β-chain loci rearrangement. Simultaneously, the TCR coreceptors CD4 and CD8αβ coemerge on the thymocyte surface to form DP thymocytes. DPs start TCR α-chain rearrangement and subsequently undergo positive selection, based on affinity for unknown self-peptides presented by MHC. Following successful selection, one of the coreceptors (CD4 or CD8) will be downregulated based on the MHC restriction of the TCR, and naive, mature MHC class II–restricted $CD4^+$ or MHC class I–restricted $CD8\alpha\beta^+$ T cells will emigrate into the blood, where they reach secondary lymphoid organs like lymph nodes and spleen.

The present study was undertaken to test whether extrathymic DPs were specifically induced in atherosclerosis. We further aimed to fully characterize their transcriptomes, identify their relationship with immature DP thymocytes, and assess their functionality and their mode of migration.

Materials and Methods

Mice

All experiments followed guidelines of the La Jolla Institute for Immunology Animal Care and Use Committee. Approval for use of rodents was obtained from the La Jolla Institute for Immunology according to criteria outlined in

*La Jolla Institute for Immunology, La Jolla, CA; †Department of Cardiology, Clinic III for Internal Medicine, University of Cologne, Cologne, Germany; ‡University Hospital Freiburg, Freiburg, Germany; and §Department of Bioengineering, University of California, San Diego, La Jolla, CA

ORCIDs: 0000-0003-1333-2031 (H.W.); 0000-0002-2823-439X (E.E.); 0000-0002-9519-9520 (S.H.-D.L.M.); 0000-0002-0372-5272 (T.R.); 0000-0003-1525-4348 (D.W.); 0000-0002-8164-6852 (G.S.); 0000-0001-9339-3672 (K.L.).

Received for publication March 2, 2021. Accepted for publication September 4, 2021.

This work was supported by National Institutes of Health/National Heart, Lung, and Blood Institute Grants HL115232, HL88093, and HL121697 (to K.L.). H.W. was supported by the Deutsche Forschungsgemeinschaft (GZ WI 4811/1-1 and 397484323) and a Tullie and Rickey Families SPARK Award for Innovations in Immunology. The FACSAria II Cell Sorter, the Zeiss LSM 880, and Illumina HiSeq2500 were acquired through National Institutes of Health Office of Research Infrastructure Program Shared Instrumentation Grant Programs S10 RR027366, S10 OD021831, and S10 OD16262.

The sequences presented in this article have been submitted to National Center for Biotechnology Information Gene Expression Omnibus repository (https://www.ncbi.nlm.nih.gov/geo) under accession number GSE180582.

Address correspondence and reprint requests to Prof. Holger Winkels, Department of Cardiology, Clinic III for Internal Medicine, University Hospital Cologne, LFI, 4.042d, Kerpener Strasse 62, 50937, Cologne, Germany. E-mail address: holger.winkels@uk-koeln.de

The online version of this article contains supplemental material.

Abbreviations used in this article: AF647, Alexa Fluor 647; DP, double positive; GaAsP, gallium arsenide phosphide; α-GalCer, α-galactosylceramide; iNKT, invariant NKT; MAT, mediastinal adipose tissue; PFA, paraformaldehyde; RAG, recombinase activating gene; scRNA-seq, single-cell RNA sequencing; S1P1, sphingosine-1-phosphate receptor; Treg, regulatory T cell.

the Guide for the Care and Use of Laboratory Animals from the National Institutes of Health. All mice were maintained in-house on a 12-h light/dark cycle and had access to chow diet ad libitum. Female *C57BL/6J* (catalog no. 000664), *CD45.1*$^+$ (*B6.SJL-PtprcaPep3b/BoyJ*, catalog no. 002014), *BALB/cJ* mice (catalog no. 000651), *Cd1d*$^{-/-}$ (C57BL/6-Cd1d1tm1.1Aben/J, catalog no. 016929), and *Rag2*$^{-/-}$ [*B6(Cg)-Rag2*$^{tm1.1Cgn}$*/J*, catalog no. 008449] were purchased from The Jackson Laboratory. If indicated, sham-operated or thymectomized *C57BL/6J* mice were also commercially obtained from The Jackson Laboratory. *Rag2*^{GFP} and *Cd1d*$^{-/-}$ mice were kindly provided by Dr. Kronenberg. *C57BL/6J* and *CD45.1*$^+$ were crossed to obtain *CD45.1/2* mice.

Mice at indicated ages were euthanized by CO_2 inhalation, followed by blood withdrawal via cardiac puncture and perfusion with 10 ml ice-cold PBS containing 100 U/ml heparin (Fresenius) before further organ and tissue dissection.

Aortic arch and mediastinal adipose tissue collection and processing

After perfusion, the chest of the mouse was opened, and the thymus was removed without damaging the capsule. Aortic arches and mediastinal adipose tissue (MAT) were microsurgically dissected and digested as previously described (5). Briefly, aortic arches and MAT pads were collected in 1 ml RPMI 1640 medium (Thermo Fisher Scientific) containing 10% FBS (Gemini) on ice. The collected tissue was cut into small pieces and individually digested for 1 h at 37°C in HBSS containing 450 U/ml Collagenase I (MilliporeSigma), 250 U/ml Collagenase XI (MilliporeSigma), 120 U/ml Hyaluronidase (MilliporeSigma), and 120 U/ml DNase I (MilliporeSigma).

The digested tissue suspensions were subsequently filtered through a 50-μm cell strainer (Partec) and washed with 37°C prewarmed RPMI 1640 with 10% FBS at room temperature for 5 min at 400 × *g*. Cells were resuspended in prewarmed RPMI 1640 with 10% FBS and incubated for 30 min at 37°C to allow recovery of cell surface molecule expression affected during digestion. Cells were then washed and kept on ice either in PBS (Thermo Fisher Scientific) with 2% FBS for staining with fluorochrome-conjugated Abs and flow cytometric analysis or in RPMI 1640 with 10% FBS for downstream processes.

Cell isolation from lungs

Both lobes of a lung were excised after perfusion and cut into small pieces, followed by the same digestion procedure applied to aortic arches and MAT. After digestion and rescue incubation, lung cells were kept on ice in PBS with 2% FBS for staining with fluorochrome-conjugated Abs and flow cytometric analysis.

Splenocyte isolation

Spleens were homogenized through a 70-μm cell strainer (BD Biosciences), followed by washing with PBS and RBC lysis for 3 min at room temperature using 1× RBC lysis buffer (BioLegend). Splenocytes were washed with PBS and kept in PBS with 2% FCS and kept on ice for staining with fluorochrome-conjugated Abs and flow cytometric analysis.

Thymocyte and lymphocyte isolation

Single-cell suspension of thymocytes and lymph node–residing cells were obtained by homogenizing the tissue through a 70-μm cell strainer. The cell suspension was washed with ice-cold PBS and kept in PBS with 2% FBS on ice for staining with fluorochrome-conjugated Abs and flow cytometric analysis.

Cell isolation from blood

Blood was withdrawn via cardiac puncture and collected in EDTA-coated tubes (Sarstedt). The plasma was removed after centrifugation for 5 min at 400 × *g* and 4°C. Erythrocytes were lysed using 1× RBC lysis buffer (BioLegend) for 10 min at room temperature, and the cell suspension was washed twice with PBS. Cells were kept in PBS with 2% FBS on ice for staining with fluorochrome-conjugated Abs and flow cytometric analysis.

Cell isolation from small and large intestine

The intestine was dissected, and surrounding fat pads and Peyer's patches were surgically removed. Small and large intestine were separated, and both pieces were opened longitudinally and washed in RPMI 1640 containing 10% FBS and HEPES (2×; Thermo Fisher Scientific). The intestinal compartments were cut into small pieces and collected in 30 ml RPMI 1640 with 10% FBS and HEPES (2×; Thermo Fisher Scientific). Subsequently, EDTA (MilliporeSigma) and DTT (MilliporeSigma) were added to a final concentration of 5 and 1 mM, respectively. The suspension was incubated for 25 min at 37°C while shaking at 250 rpm and subsequently filtered through a 70-μm strainer. After washing in PBS, the intestinal cell pellet was resuspended in 4 ml of a 40% Percoll gradient (GE Healthcare). The 40% Percoll suspension was layered on top of 5 ml of a 60% Percoll gradient (GE Healthcare) and centrifuged at room temperature at 900 × *g* for 25 min without acceleration and brake. The cell layer at the interphase was collected, washed with PBS, and kept on ice for staining with fluorochrome-conjugated Abs and flow cytometric analysis.

Cell isolation from liver

The liver was dissected and homogenized through a 100-μm cell strainer (BD Biosciences). After washing in PBS, the liver cell pellet was resuspended in 4 ml of 40% Percoll. The liver Percoll suspension was carefully layered on top of 5 ml 60% Percoll solution and centrifuged at 900 × *g* for 25 min without acceleration and brake. The cell layer at the interphase was collected, washed with PBS, and kept on ice for staining with fluorochrome-conjugated Abs and flow cytometric analysis.

Flow cytometry

Single-cell suspensions were incubated with 50 μl PBS containing 2% FBS, 10% rat serum (MilliporeSigma), and 10 μg/ml anti-CD16/CD32 Ab (FC shield, clone 2.4G2, catalog no. 70-0161-M001; Tonbo Biosciences) and live/dead staining (1:1000; either Thermo Fisher Scientific or Tonbo Biosciences) for 5 min. Subsequently, 50 μl PBS containing 2% FBS, 10% rat serum, 10 μg/ml FC shield, and live/dead staining (1:1000) and fluorochrome-conjugated Abs (2×) were added for 20 min on ice in the dark. Clones and final concentrations used are listed in Supplemental Table I. After incubation, cells were washed with ice-cold PBS containing 2% FBS. For determining cell proliferation, cells were fixed and permeabilized using the Foxp3 Transcription Factor Staining Buffer Set (Thermo Fisher Scientific) and stained with anti-Ki67–AF700 for 30 min on ice and washed. If needed, samples were fixed with 2% paraformaldehyde (PFA; Electron Microscopy Sciences) for 10 min at room temperature. Cells were washed, resuspended in PBS containing 2% FBS, and acquired with an LSR II flow cytometer (BD Biosciences), LSRFortessa (BD Biosciences), or Amnis Imagestream (MilliporeSigma).

FACS sorting

For FACS sorting, cells were stained as previously described. Before sorting, cells were kept in PBS containing 2% FBS and 2 mM EDTA. Viable CD4$^+$CD8$^+$TCR$^-$DUMP$^-$ (DP) cells were sorted with a 70-μm nozzle using the BD FACSAria Fusion and the FACSAria II cell sorter (both BD Biosciences). For imaging analysis, cells were sorted directly on poly-L-lysine glass (MilliporeSigma). For low-input sequencing, 400 DPs were sorted into low-absorbent microcentrifuge tubes (0.2 ml; Eppendorf) containing 8 μl low-input lysis buffer containing rRNAse inhibitor (40 U/μl; Clonetech Laboratories), 2′-deoxynucleoside 5′-triphosphates (Thermo Fisher Scientific), and 0.1% Triton X-100 (Thermo Fisher Scientific). For fetal thymic organ cultures, cells were sorted into low absorbent microcentrifuge tubes (1.5 ml; Eppendorf) containing RPMI 1640 supplemented with 10% FCS.

Clearing and staining of whole-mount tissue

Whole-mount aortic arch, MAT, and thymus were fixed in Cytofix (BD Biosciences) 1:3 diluted in PBS overnight. The tissue block was washed in PBS for 1 d and blocked with 1% mouse serum (MilliporeSigma), 1% BSA (MilliporeSigma), and 0.3% Triton X-100 in PBS for 24 h at room temperature while agitating. Subsequently, primary fluorescently conjugated Abs were added 1:100 for 72 h at 37°C in the dark while agitating. Hoechst (Thermo Fisher Scientific) was added at a concentration of 1:1000 for the final 2 h. After staining, the whole-mount tissue block was washed in PBS with 0.2% Triton X-100 and 0.5% 1-Thioglycerol (MilliporeSigma) overnight at room temperature. The buffer was replaced twice. C_e3D clearing solution (6) was freshly prepared as follows.

One milliliter contained 400 μl (40%) *N*-methylacetamide, 1 μl of Triton X-100 (0.1%), and 5 μl 1-Thioglycerol (0.5%) (all obtained from MilliporeSigma). For optimal solvation of all reagents, the buffer was shaken at 37°C for several hours. The whole-mount tissue was removed from the washing buffer, carefully dried, and transferred into the clearing solution overnight at room temperature in the dark. The following day, the cleared whole-mount tissue was mounted between two 1.5 borosilicate glass coverslips in clearing medium.

Staining of MAT for DPs and invariant NKT cells

Two pads of MAT were kept in 500 μl PBS with 2% FCS and 1:100 of 10 μg/ml anti-CD16/CD32 Ab (FC shield, clone 2.4G2; Tonbo Biosciences) at room temperature. The tissues were washed twice for 5 min in PBS with 2% FCS. The MAT pads were incubated in 500 μl PBS with 2% FCS and

1:100 CD4 A488 (clone RM4-5), 1:100 CD8α PE-Dazzle594 (clone 53-6.7), and either 10 μl α-galactosylceramide (α-GalCer):CD1d-tetramer conjugated with A647 (40683, PBS-57, 1.6 mg/ml) or empty control CD1d tetramer conjugated with A647 (40684, 1.5 mg/ml). Samples were kept shaking overnight at 4°C and washed twice in PBS with 2% FCS. Nuclei were stained with 1:1000 Hoechst PBS with 2% FCS for 1 h at room temperature. Samples were washed twice for 5 min in PBS with 2% FCS and fixed with 4% PFA for 1 h. Samples were washed twice, mounted in Prolong Gold Mounting Medium (Thermo Fisher Scientific), and cured overnight with clamps and sealed with nail polish.

Imaging of whole-mount tissue

Images were acquired at room temperature with a Zeiss LSM880 confocal microscope equipped with a 20× air objective (numerical aperture, 0.8) and Zen browser black v. 2019. Images used for three-dimensional reconstructions were acquired at 405, 488, 561, and 633 nm excitation wavelengths to visualize $CD31^+$ stained vessels single $CD4^+$, single $CD8^+$, and $CD4^+CD8^+$ cells. Spectrally tuned detectors were used to detect fluorescent signals and to minimize the overlap of fluorescent signals (>660, 600–650, 499–553, and 412–475 nm). Main beam splitters were set at 405, 488, 561, or 633 nm and a secondary beam splitter at 660 nm. Signal from Alexa Fluor 647 (AF647) was directed to the AiryScan detector operated in confocal mode, and the signals from the other fluorochromes were read out using internal detectors (405 nm Ch1 bialkali photomultiplier tube [725 V]; 488 nm ChS1 gallium arsenide phosphide [GaAsP; 800 V]; 561 nm far red detector Ch2 [720 V]; and 633 nm ChA [700 V]). The pinhole width was set at 24.6 μm for all channels (=0.49, 0.52, 0.61, and 0.73 Airy units, respectively.). Images were acquired at 5 μm Z-step size and a pixel size of 0.42 μm. Thirty-six tiles were stitched together for a total image size of 5.13 × 5.53 mm and a depth of 100 μm (total pixel size: 12365 × 13320). Acquisition was performed in line-switching mode and unidirectional at a pixel dwell time of 0.77 μs. Image analysis was performed in Imaris v.9.6 (Bitplane). Contrast stretching was used to increase visibility of image features ($\gamma = 1.0$). Isosurfaces were generated based of nuclear signal. Fluorescence signal thresholds were obtained from control tissue to determine a filtering cutoff. This particularly was used to filter out autofluorescence from collagen fibers in the aortic arch. Cells were further filtered for sphericity with a cutoff of 0.7. A fluorescent image and an image generated after isosurface generation and filtering are displayed to demonstrate extrathymic occurrence of CD4CD8 DP cells.

Imaging of MAT and image reconstruction

MAT images for this tissue were acquired using a C-Apo 40 W (1.2 nA) objective using as described a Zeiss LSM880 confocal microscope equipped with GaAsP detectors and Zen browser black v. 2019. In this study, on average, 45 stacks of images per sample area were acquired at 0.5-μm Z-step size and a pixel size of 0.129 μm.

Imaging of sorted DPs

Viable DPs cells were isolated by FACS sorting from the thymus or MAT with aortic arch. Cells were directly sorted on a poly-L-lysine slide (MilliporeSigma). Up to 10000 cells were sorted in a spot placed on the glass slide. After sedimentation, sheath flow was removed and replaced with 4% PFA for 3 min. Subsequently, cells were washed twice with distilled water before mounting with fluorescent mounting gold medium (Thermo Fisher Scientific). After curing overnight, images were acquired at room temperature with the Zeiss LSM780 confocal microscope with a 63× oil objective (numerical aperture, 1.4) and Zeiss Zen browser black v. 2019.

Images were acquired at 405-, 488-, and 633-nm excitation wavelengths. Spectrally tuned GaAsP detectors were used to detect fluorescent signals and to minimize the overlap of fluorescent signals (419–472, 508–543, 668–694, and 650–694 nm). Main beam splitters were set at 405, 488, 561, or 633 nm and the signals were read out using internal detectors (ChS1 GaAsP for all fluorochromes). For thymic cells, detector gains were set at 915 (Hoechst), 1069 (AF488), 1068 (PerCP), and 1007 (AF647), whereas extrathymic-isolated DPs were acquired at a gain of 782 (Hoechst), 1015 (AF488), 1091 (PerCP), and 930 (AF647). The pinhole width was set 1 Airy unit for all channels (= 43, 50, 58, and 64 μm, respectively). Images were acquired at a pixel size of 0.055 μm (total pixel size: 1024 × 1024) and a Z-step depth of 0.374 μm (total depth 4.12 μm). Acquisition was performed in frame-switching mode and unidirectional at a pixel dwell time of 6.3 μs.

EdU injection

Eight-week-old female *C57BL/6J* mice were injected i.p. with 1 mg EdU (Thermo Fisher Scientific) in PBS. Mice were sacrificed at days 1, 3, 5, and 7 postinjection. EdU incorporation into DPs was assessed by the Click-IT EdU plus flow cytometry kit (Thermo Fisher Scientific) according to the manufacturer's instruction.

FTY720

FTY720 (Sellcheck) was dissolved in DMSO (MilliporeSigma) at 2 mg/ml and stored at −20°C. The stock solution was diluted in sterile PBS supplemented with 3% BSA. Eight-week-old female *C57BL/6J* were i.p. injected at 1 mg/kg body weight every day for 7 d. Alternatively, mice were injected with DMSO only.

Thymectomies

Female *C57BL/6J* mice were thymectomized 3 and 21 d after birth at The Jackson Laboratory before shipment to the La Jolla Institute for Immunology animal facility.

Kidney capsule transplantation

Female *C57BL/6J* mice were thymectomized or sham operated 21 d after birth at The Jackson Laboratory before delivery to the La Jolla Institute for Immunology. At the age of 10 wk, mice were anesthetized with isoflurane and kept in deep anesthesia. A sterile field was prepared, and the mouse kept on a heating pad to prevent hypothermia. Fur was removed from the left flank, and the skin was prepared antiseptically. A small incision was performed into the skin and the window was opened to gain access to the peritoneum. The peritoneal cavity was opened close to the left kidney. The kidney was exposed and kept moist in saline. Using a microdissecting microscope and microsurgery tools (Fine Science Tools), a small incision was made into the kidney capsule into which one third of a thymic lobe of a congenically marked female CD45.1 mouse was transplanted. After successful implementation of the thymic piece, the incision in the kidney capsule was sealed with tissue adhesive glue (Vetbond), and the kidney was returned into the peritoneal cavity. The peritoneal cavity and subsequently the skin were closed with 4-0 sterile and absorbable Vicryl sutures (Ethicon). Transplanted mice received an i.p. injection of buprenorphine (0.1 mg/kg body weight) and recovered in a temperature-controlled chamber at 37°C before being transferred into a sterile new cage. Mice were injected with a second dose of buprenorphine 4 h later and monitored for their behavior daily during the first 4 d. Sutures were removed 10 d postsurgery, and mice were euthanized for organ and tissue collection 21 d postsurgery.

Fetal thymic organ culture

DPs were FACS sorted from MAT and aortic arches of 8-wk-old female *C57BL6/J* mice. Cells from three to four donor mice were pooled. The cell count was adjusted to 1×10^6/ml with prewarmed DMEM with GlutaMAX (Thermo Fisher Scientific) and 10% FBS. Individual wells of a Terasaki plate (MilliporeSigma) were filled with 20 μl cell suspension.

CD45.1 and CD45.2 mice were bred, and plug formation of pregnant females was monitored. At embryonic day 14, the uterus of pregnant mice was aseptically removed and transferred into a 90-mm petri dish (MilliporeSigma) containing DMEM with GlutaMAX and 10% FBS, and embryos were removed. Single embryos were washed additionally in DMEM with GlutaMAX and 10% FBS to remove blood. Thymic lobes of E14 embryos were carefully removed from the chest cavity under a dissecting microscope and kept on ice in DMEM with GlutaMAX and 10% FCS. A single thymic lobe was added to a well in the Terasaki plate containing the cell suspension. Subsequently, the Terasaki plate was carefully inverted to allow migration of extrathymic DPs into the fetal thymus for 48 h at 37°C and 5% CO_2. Subsequently, thymic lobes were collected and individually transferred on top of a Transwell filter in submersion culture for another 7 d at 37°C and 5% CO_2. At the end of the culture, thymic lobes were homogenized through a 50-μm filter to obtain a single-cell suspension that was prepared for flow cytometric analysis as previously described.

Glycolipid Ag presentation assay

Viable DPs cells were isolated by FACS sorting from the thymus or MAT with aortic arch from 8-wk-old female *C57BL/6J* mice. Duplicate of cells (10,000–250,000) were cultured with 100 ng/ml α-GalCer (BioVision, Milpitas, CA) or left untreated in DMEM containing 10% FCS, 100 U/ml penicillin, 100 U/ml streptomycin, 2 mM L-glutamine, 1 mM sodium pyruvate, and 10 mM HEPES (all Thermo Fisher Scientific) with 10 ng/ml IL-7 (R&D Systems, Minneapolis, MN) for 3.5 h at 37°C and 5% CO_2. Cells were washed twice with T cell medium and cocultured with 50,000 invariant NKT (iNKT) DN3A4-1.2 hybridoma cells (7) for 12 h in presence of IL-7. The supernatant was analyzed for IL-2 by ELISA while the cultured cells were prepared for flow cytometric analysis of CD69 surface expression.

Single sample gene set enrichment analysis

Single-cell transcriptomes of aortic leukocytes isolated from atherosclerotic and nonatherosclerotic mice (GSM2882367: 8-wk-old male *Ldlr*$^{-/-}$ mice fed 11 wk Western-type diet; GSM2882368: 8-wk-old male *Ldlr*$^{-/-}$ mice fed chow diet for 11 wk, 8-wk-old male *Apoe*$^{-/-}$ mice fed chow diet, 8-wk-old male *Apoe*$^{-/-}$ mice fed chow diet for 12 wk, and 8-wk-old male *Apoe*$^{-/-}$ mice fed Western-type diet for 12 wk) were integratively analyzed with Seurat. Batch correction was performed. The top 100 genes of the newly identified cell subset were extracted and subjected to a single sample gene set enrichment analysis (8, 9) of immune cell transcriptomes (GSE109125) isolated from bone marrow, spleen, and thymus. The enrichment score is reported as spider plot.

Low-input bulk RNA sequencing

For full-length transcriptome using Smart-Seq2, we followed a previously described protocol (10). In brief, 400 DPs were isolated by FACS sorting from aortic arch, MAT, and thymus directly into 8 μl low-input lysis buffer containing 0.1% Triton X-100, 10 mM 2′-deoxynucleoside-5′-triphosphates, and 40 U/μl rRNAse inhibitor in molecular grade water. We sorted at least three technical duplicates from pooled arch, MAT, or thymus cell suspensions of five female *C57BL/6J* mice. Samples were vortexed, centrifuged at 3000 × *g* for 2 min, and stored at −80°C until further processing. To generate cDNA, we PCR amplified 4 μl of lysed cells (equaling 200 cells) with 18 PCR cycles to obtain sufficient material for downstream procedures. The amplified cDNA was size selected and leftover primers eliminated through a magnetic beads dsDNA capture using a 0.8:1 vol/vol AMPure XP beads (Agencourt) procedure. Size-selected cDNA was quantified with a fluorometric DNA quantification assay (PicoGreen dsDNA; Thermo Fisher Scientific). One nanogram of cDNA was used to generate barcoded Illumina sequencing libraries using a standard tagmentation protocol (Nextera XT library preparation kit, Illumina; nine cycles of PCR-based amplification using unique dual-index Illumina adaptors). As previously described, libraries were purified with a double size-selection protocol (0.5 followed by 0.8 volume bead/initial sample volumes), quantified (Picogreen quantification assay; Thermo Fisher Scientific), and fragment size measured by capillary electrophoresis (fragment analyzer; Advance Analytical Labs).

Samples failing our stringent quality control checks (quantity and size) were eliminated from further downstream steps (10). Libraries were pooled at equimolar ratios, quantified by quantitative PCR assay (KAPA SYBR FAST qPCR Kit; Roche Diagnostics), and sequenced using the Illumina sequencing platform and reagents. HiSeq2500 high output v4 (HiSeq SBS Kit v4; Illumina) libraries were sequenced for 50-bp single-end reads. Postsequencing, stringent quality controls were applied, and samples that failed quality control checks were eliminated from further analysis (10). Samples were sequenced to obtain at least 8 million uniquely mapped reads.

Low-input RNA sequencing mapping

All samples passed basic quality check through FASTQC v0.11.5 (11). Then, FASTQ files were mapped using STAR v2.6.0c (12) based on ENSEMBL reference GRCm38 and ENSEMBL mouse annotation release 92 (13). Gene counts were directly obtained from STAR using the parameter quantMode GeneCounts. Postmapping quality check was done through QoRTs v1.3.0 (14) with additional parameter genomeBufferSize 50000. All commands were run using the UNIX command parallel v20161222 (15).

Processed sequencing data are available at the National Center for Biotechnology Information Gene Expression Omnibus repository (accession number GSE180582; https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE180582).

Differentially expressed and coregulated genes

Differential expression was done using DESeq2 v1.24 (16) according to the recommended workflow. Aside from five technical replicates, which were clearly outliers according to their Euclidean distance to other technical replicates, all other technical replicates were collapsed using the DESeq2 function (collapseReplicates). Only genes that had a total of at least 10 reads across all samples were retained. No log fold change shrinkage was done. Adjustment for p values was based on false discovery rate. Significantly differentially expressed genes were those with an adjusted p value ≤ 0.05. UpSet (17) was used to visualize set intersections between the various differentially expressed gene lists. Coregulated genes were analyzed using WGCNA v1.68 (18) and was based on recommended workflow, with specific parameters (power = 9, TOMType = signed, corType = bicor, and networkType = signed hybrid). TCR clonotypes were determined with MiXCR (19).

Statistical analysis

Data were routinely shown as mean with SD. Unless stated otherwise, statistical significance was determined by Student t test or ANOVA using Prism 8.1.2 (GraphPad). A value of more than three SDs from the mean served as criterium to exclude outliers.

Results

CD4$^+$CD8$^+$ cells in the MAT and aortic adventitia

In silico analysis and integrative assessment of single-cell transcriptomic data from aortic leukocytes sets derived from *apolipoprotein E*–deficient mice (aged 8 wk, aged 20 wk, and fed chow diet; aged 20 wk and fed high-fat diet for 12 wk) and low-density lipoprotein receptor–deficient mice (aged 19 wk and fed chow diet; aged 19 wk and fed high-fat diet for 11 wk) showed CD4$^+$ and CD8αβ$^+$ DP T cells in all five datasets (Fig. 1A). To determine the relationship with other immune cells, we tested enrichment of the top 100 genes of DP cells (Supplemental Table I) among bulk transcriptomes of various stem cell populations, thymic T cells progenitors, and splenic naive CD4$^+$ and CD8αβ$^+$ T cells. The gene signature of the aortic mixed CD4$^+$/CD8αβ$^+$ T cell population showed the highest enrichment score with thymic CD4$^+$CD8αβ$^+$ cells, a progenitor stage of single CD4$^+$ and CD8αβ$^+$ TCRαβ$^+$ T cells (Fig. 1B, Supplemental Table I). To confirm extrathymic DPs with an independent method and assess whether DPs (CD4$^+$CD8αβ$^+$ TCRαβ$^-$) are also present in other organs or tissues, we performed a flow cytometry–based screening of 16 tissues of 8-wk-old female wild-type *C57BL/6J* mice (Supplemental Fig. 1). As expected, DPs represented the majority of leukocytes found in the thymus (Fig. 1C). Only two other tissues contained DPs, 1) the MAT connecting the thymus and the aortic arch (mean ± SD: 9 ± 12% of living leukocytes), and 2) aortic arches (mean ± SD: 13 ± 11% of living leukocytes) (Fig. 1C). DPs were not detected in the abdominal aorta, which suggests their local confinement to the arch. Although the percentage of DPs among leukocytes found in the aortic arch was higher compared with the ones in MAT, absolute numbers were higher in MAT (mean ± SD: 6765 ± 9448 versus 315 ± 543 cells per mouse) (Fig. 1D). Of note, all DPs did not express surface CD3e (Supplemental Fig. 2A). Few thymic DPs expressed CD69, a marker of positive selection, which was significantly reduced on MAT and aortic arch DPs (Supplemental Fig. 2B). All three DP populations expressed TCRβ intracellularly, albeit at lower levels compared with lymph node CD4 T cells. The expression of intracellular TCR was significantly reduced in MAT and arch DPs compared with thymic DPs (Supplemental Fig. 2C). To test whether the numbers of MAT and aortic arch DPs are specific for atherosclerosis-prone *C57BL/6J* mice, we also analyzed tissues from 2-y-old *BALB/C* mice, which are resistant to atherosclerosis (20). We found similar numbers in both strains (Fig. 1D). The abundance of aortic DPs was not altered in the course of atherosclerosis (Supplemental Fig. 3). These findings point to homeostatic presence of these cells.

To address the location of extrathymic DPs, we performed three-dimensional whole-mount imaging of cleared thymus with MAT and the aortic arch (Fig. 2A–C). DP cells were present in the MAT pads attached to the thymus capsule and in the adventitial space of the lower curvature of the aortic arch, where they were organized in clusters (Fig. 2B, 2C). Like thymic DPs, extrathymic DPs were small, round cells, which was further established by sorting these cells directly onto glass slides, followed by confocal microscopy (Fig. 2D, Supplemental Fig. 4). Extrathymic DPs were slightly smaller compared with thymic DPs.

Extrathymic DPs have a distinct and region-specific transcriptional profile

To further establish if and how MAT and aortic arch DPs would differ from thymic DPs, we sorted DPs from all three regions by flow cytometry and subjected the cells to a low-input deep transcriptional profiling assay (10). Principal component analysis revealed

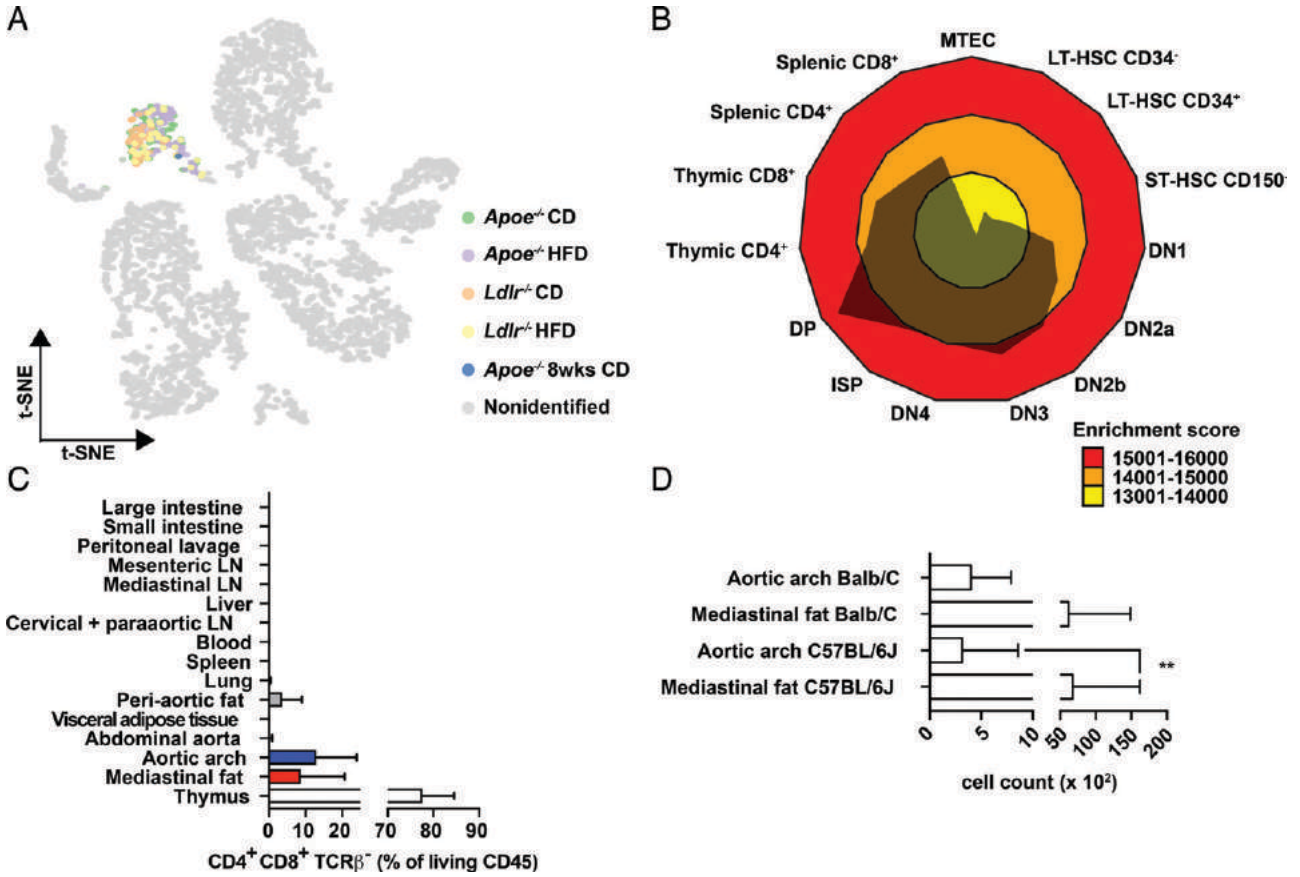


FIGURE 1. A new class of aortic immune cells. (**A**) Integrated scRNA-seq analysis of living leukocytes isolated from atherosclerotic and nonatherosclerotic mice. (GSM2882367: 8-wk-old male *Ldlr*$^{-/-}$ mice fed 11 wk Western-type diet [yellow]; GSM2882368: 8-wk-old male *Ldlr*$^{-/-}$ mice fed chow diet for 11 wk [orange], 8-wk-old male *Apoe*$^{-/-}$ mice fed chow diet [blue], 8-wk-old male *Apoe*$^{-/-}$ mice fed chow diet for 12 wk [green], and 8-wk-old male *Apoe*$^{-/-}$ mice fed Western-type diet for 12 wk [purple]). All cells of the new immune cell subset cluster together and are found across all analyzed scRNA-seq datasets as indicated by the different colors. All other aortic immune cell subsets were colored in gray. (**B**) The top 100 genes of the highlighted cluster in (A) were used for single sample gene set enrichment analysis of immune cell transcriptomes (GSE109125) isolated from bone marrow, spleen, and thymus. The enrichment score is displayed as spider plot. (**C**) Flow cytometric analysis of mentioned tissues for the presence of $CD4^+CD8\alpha^+TCR^-$ cells in 8-wk-old female *C57BL/6J* mice. For thymus $n = 49$, MAT $n = 52$, aortic arch $n = 45$, and all other tissues $n = 5$. (**D**) Absolute numbers of $CD4^+CD8\alpha^+TCR^-$ cells in MAT ($n = 28$) and aortic arches ($n = 29$) of 8-wk-old female *C57BL/6J* mice and 2-y-old female *BALB/C* mice ($n = 4$ per tissue). (C and D) Data are presented as mean ± SD. $^{**}p = 0.0012$.

the close relationship among DPs isolated from the same locations, as the transcriptomes of the individual biological replicates clustered closely together (Fig. 3A) However, the transcriptome of DPs differed in a location-specific and gradual manner in which thymic and aortic arch DPs represent the two ends of the spectrum bridged by MAT DPs. These results are also reflected in a heatmap comparing the collated transcriptomes (Fig. 3A, 3B, Supplemental Fig. 5).

A total of 2147 genes were differentially regulated between the different DP populations (Fig. 3B, Supplemental Table I). Extrathymic DPs had 858 genes upregulated and 1284 genes downregulated compared with thymic DPs. Ninety-four genes were commonly regulated in the two extrathymic DP subsets (Fig. 3B, Supplemental Table I), whereas 23 genes were uniquely expressed in MAT DPs. Thirty-one different genes were exclusively expressed in aortic arch DPs.

Only expression of *Vps37b* increased in MAT in thymic DPs and was even higher expressed in aortic arch DPs compared with MAT DPs (Supplemental Table I). VPS37B is a subunit of the ESCRT-I complex, which regulates vesicle trafficking and was recently identified by scRNA-seq of CD4 memory T cells residing in nonlymphoid tissue (21). *Hist3h2a* encodes for the histone H2A type 3. This gene is expressed higher by extrathymic DPs compared with thymic DPs, whereas its expression is higher in MAT DPs compared with aortic arch DPs. Virus-specific effector $CD8\alpha\beta^+$ T cells expressed higher *Hist3h2a* compared with virus-specific naive T cells (22). MAT DPs also uniquely expressed *Rtel1*, whose protein product regulates telomere lengths. T cells deficient for *Rtel1* have reduced survival in long-term cultures (23). Several other genes were highly expressed among MAT DPs, which are involved in regulating cell growth (*Slc25a10*, *Ftsj2*, and *Ercc16*), apoptosis (*Sulf1* and *Dstyk*), or splicing (*Esrp2*). However, their functions in immune cells have yet to be determined.

Similarly, the top 10 genes uniquely expressed in aortic arch DPs have unknown functions in T cells (*Abcd7*, *1500004A13Rik*, *Metrnl*, *Gm12216*, *P4htm*, *Myl9*, *Rab44*, and *E530011L22Rik*). Two genes were expressed in aortic arch DPs with known functions in T cells. The gene *Slc1a4* encodes for the alanine, serine, cysteine, and threonine amino acid transporter ASCT1, which is strongly upregulated in activated T cells, providing sufficient metabolites for proliferation (24). The other gene was *Prf1* encoding for perforin 1, a well-known effector molecule of cytotoxic T cells. Perforin 1 is also expressed during positive selection of CD8 T cells and by thymocytes receiving MHC class II signals (25). Among the genes expressed by MAT and aortic arch DPs, but not thymic DPs, were a transcription factor (*Klf9*), signaling (*Fosl2*, *Gpr3*, and *Lipt2*), and metabolism-relevant genes (*Dgat2* and *Klhl23*). Both cell populations also expressed *Hspa1a* encoding for the heat shock protein 72, which is upregulated

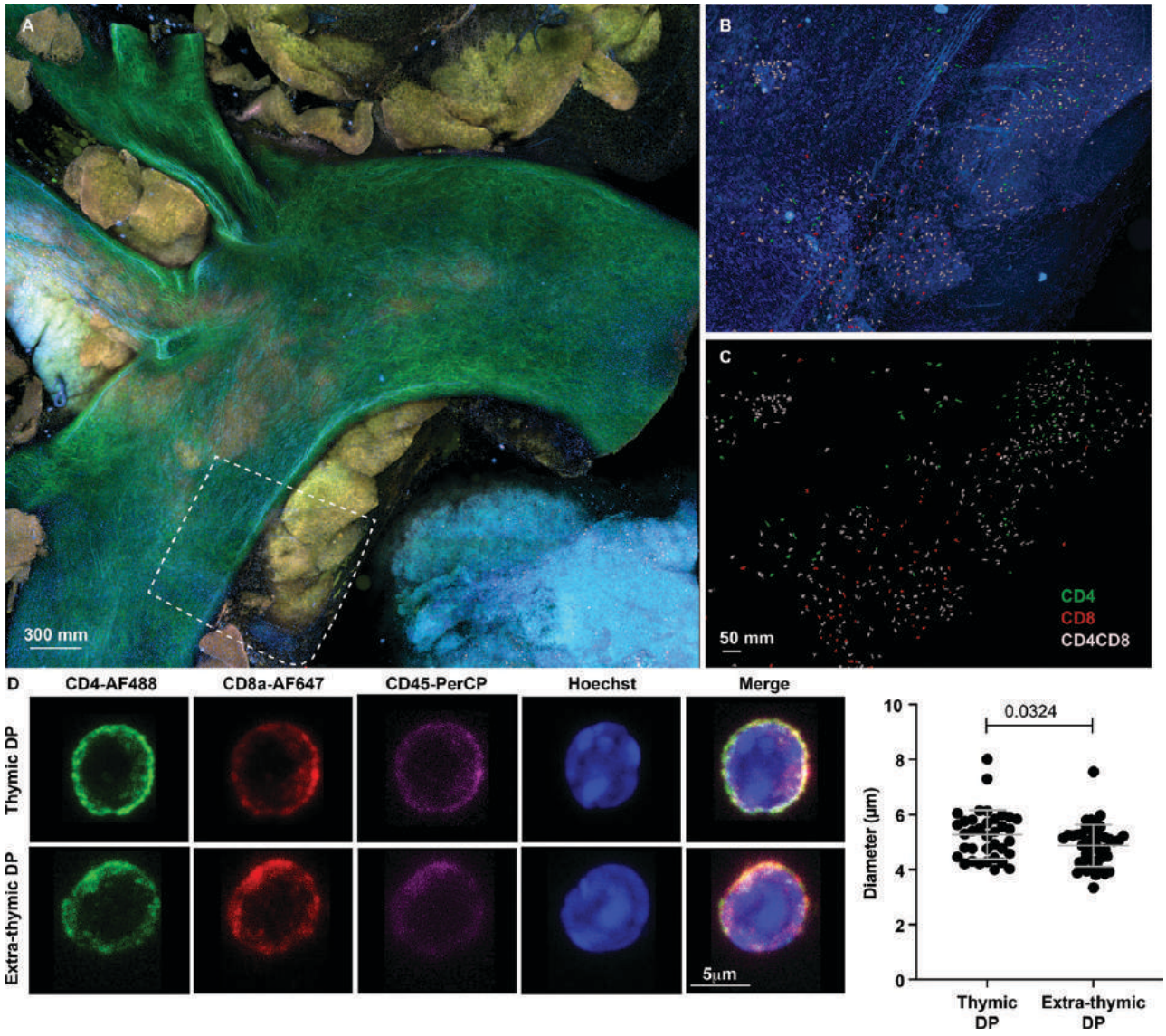


FIGURE 2. CD4CD8 cells are organized in clusters in the aortic arch. (**A**) Whole-mount thymus, MAT, and aortic arch tissue of an 8-wk-old female *C57BL/6J* mouse was cleared according to the clearing-enhanced three-dimensional protocol. The tissue preparation was stained for CD4 (green), CD8α (red), CD31 (pink), and nuclei, (Hoechst) and imaged by confocal microscopy. (**B**) Magnified region outlined by the dashed box in (A). (**C**) Spherical cells within the magnified area (B) were rendered by isosurfaces constructed in Imaris. Isosurfaces of $CD4^+$ were colored in green, $CD8\alpha^+$ positive cells in red, and DP cells are colored in white. The vasculature and nuclei are displayed in different shades of blue. Representative of three independent experiments. (**D**) Thymic and extrathymic $CD4^+CD8\alpha^+TCR^-$ cells of 8-wk-old female *C57BL/6J* mice ($n = 10$) were sorted by flow cytometry onto a poly-L-lysine–coated glass slide and imaged by confocal microscopy. The distribution of the diameter is displayed.

by cellular stress and has anti-inflammatory features (26). *Nr4a2* encodes for a member of the steroid-thyroid hormone-retinoid receptor superfamily. The deletion of this gene in $CD4^+$ T cells prevents FOXP3 expression and hence regulatory T cell (Treg) formation (27). *Rora* is expressed in Th17 cells, but the function in thymocytes is unknown. The transcription factor *Irf8* was described to facilitate TCR signal integration and differentiation of $CD8\alpha\beta^+$ T cells (28) and Th1-like Treg function (29), its function in DPs is, however, unknown. The gene *Il12rb2* encodes for the beta 3 subunit of the IL-12R complex. The cytokine IL-12 prevents thymic involution (30) and is important for differentiation of Th1, T follicular helper cell 1 (31), and TGF-β–induced Tregs (32). Extrathymic DPs also highly expressed *Slfn1*, a member of the Schlafen family. Ectopic *Slfn1* expression in thymocytes blocked DP development and caused cell cycle arrest by inhibiting cyclin D1 (33).

To test whether functional sets of genes were commonly up- or downregulated between thymic and extrathymic DPs, we applied weighted correlation network analysis. A total of 33 gene modules were identified, of which 15 were significantly correlated with either thymic or extrathymic DPs (Fig. 3D). The most significant gene modules of the respective trait (thymic or extrathymic DP) also contained most genes. The turquoise gene module was the most significant module associating with the trait thymic. Pathway analysis demonstrated enrichment for signaling pathways involved in T cell development such as WNT signaling, TCR signaling, and RAP1 signaling (Fig. 3D), suggesting that thymic DPs are more mature in their developmental program. The most significant gene modules found in extrathymic DPs were enriched for cell cycle pathways and stress responses (Fig. 3D). Enrichment for cell cycle pathways suggests that these cells may recently have passed β-selection, advancing into the last proliferative burst, during which their thymic counterparts would initiate CD4 and CD8 expression and TCR α-chain rearrangement. Although not among the differentially expressed genes between aortic arch DPs and thymic DPs, *pTa*, the

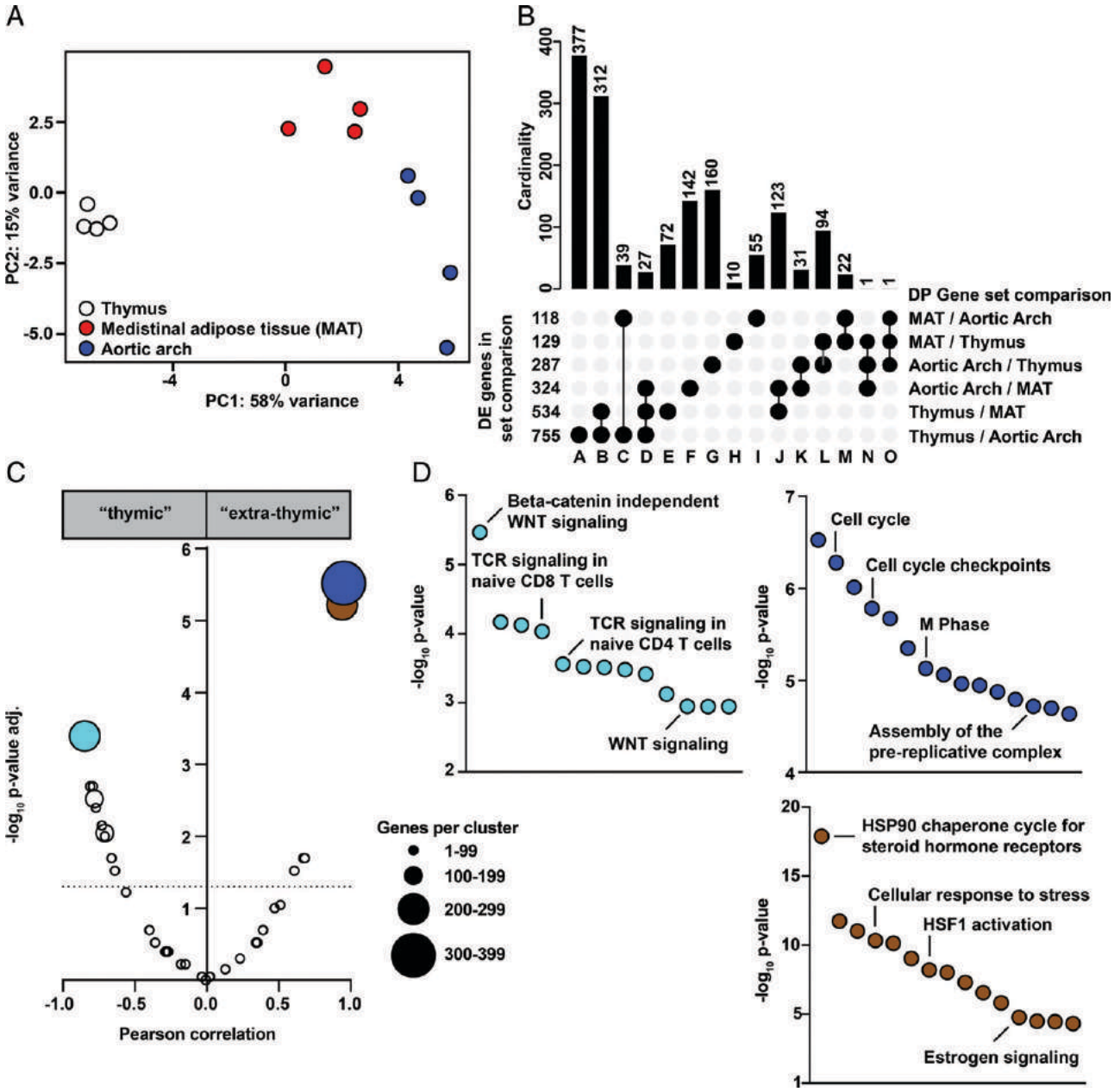


FIGURE 3. Extrathymic DPs differ transcriptomically from thymic DPs. (**A**) Living $CD4^+CD8\alpha^+TCR^-$ (DP) cells were isolated by FACS from thymus (white circles), MAT (red circles), and aortic arches (blue circles) of 8-wk-old female *C57BL/6J* mice ($n = 4$). A low-input SMART-Seq2 protocol was performed on multiple technical replicates (two to three per tissue) containing 200 cells each. The highly variable genes were obtained by principal component analysis. (**B**) UpSet analysis of differentially expressed genes of all possible comparisons between $CD4^+CD8\alpha^+TCR^-$ transcriptomes from the three tissues. Only differentially expressed genes with a log2 fold change above 2 or below −2 and an adjusted p value <0.05 were considered. (**C**) Weighted correlation network analysis (WGCNA) resulting in 15 differentially regulated gene networks between thymic, MAT, and aortic arch $CD4^+CD8\alpha^+TCR^-$ cells. The size of the dot represents the number of genes within the cluster. (**D**) The most significant gene networks of the top three correlating clusters (turquoise, blue, and brown) were subjected to ConsensusPathDB (CPDB) pathway analysis and displayed with their respective p values.

gene encoding for the pre–TCR α-chain, was significantly lower expressed in MAT DPs compared aortic arch and thymic DPs (Supplemental Table I). The appearance of stress response pathways might indicate the adaption of extrathymic DPs to the new and potentially nonfavorable environment outside of the thymus.

Extrathymic and thymic DPs have a similar proliferative capacity and turnover

The increased abundance of cell cycle pathways among extrathymic DPs led us to investigate their proliferative capacity in more detail (Fig. 3D). *Mki67*, encoding for the proliferation marker Ki67, was equally highly expressed among DPs from all three tissues. On the protein level, Ki67 was significantly higher expressed among MAT and aortic arch DPs, suggesting that these cells have a higher proliferative capacity (Fig. 4A). However, Ki67 marks actively proliferating cells and cells that just underwent a cell cycle, thus masking the cells that are currently undergoing proliferation. To further discriminate whether DPs from the three different locations exhibit altered proliferative capacity and turnover, we performed a pulse-chase experiment by injecting EdU i.p. into mice. EdU will be incorporated into the DNA of dividing cells and can be monitored by flow cytometry. One day after EdU injection, 20–30% of DPs across all

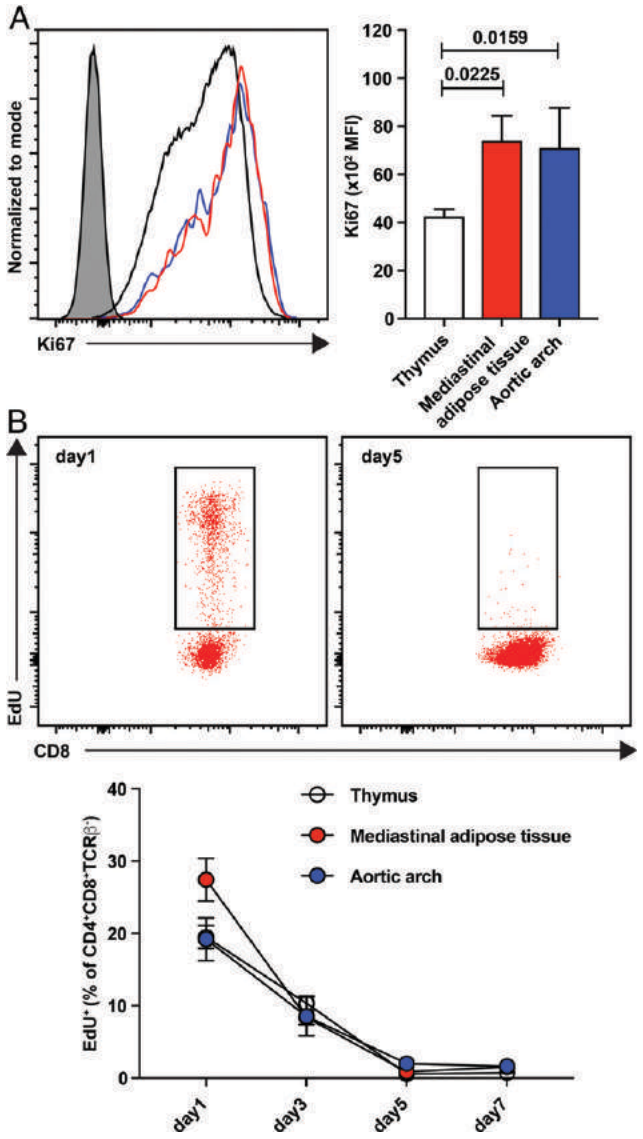


FIGURE 4. Extrathymic DPs are highly proliferative and have a fast turnover. (**A**) Quantification of the proliferation marker Ki67 among $CD4^+CD8\alpha^+TCR^-$ cells (DPs) from thymi (white bars and circles), MAT (red bars and circles), and aortic arches (blue bars and circles) of 8-wk-old female *C57BL/6J* mice ($n = 5$). Representative histograms are displayed on the left and the mean fluorescence intensity is quantified on the right. (**B**) The base analogue EdU was injected i.p. into 8-wk-old female *C57BL/6J* mice (1 mg/mouse) and DPs from thymus, MAT, and aortic arch were analyzed for EdU incorporation by flow cytometry day 1, 3, 5, and 7 after administration ($n = 5$ per time point). Representative flow cytometry plots are displayed on the top part, and presence of EdU-positive DPs is quantified below.

three tissues incorporated EdU, pointing to similar proliferative capacity. The frequency of DPs that were positive for EdU was reduced to 10% after 3 d, whereas no EdU-positive cells were detectable after 5 and 7 d. DPs from all tissues followed the same pattern, indicating that the turnover between thymic and extrathymic DP populations is similar (Fig. 4B).

Extrathymic DPs originate from the thymus

Developing T cells rearrange the V, D, and J gene segments to form a functional TCR chain. This process is catalyzed by *Rag2* and *Rag2*. Transgenic $Rag2^{GFP}$ mice, which express GFP under the control of the *Rag2* promotor, demonstrated that GFP and RAG2 were expressed at the late double-negative stage in the thymus. GFP expression decreased from the double-negative stage to the single-positive stage of thymic T cells but remained detectable, whereas RAG2 expression disappears at the single-positive stage (34, 35). Interestingly, DPs from MAT and aortic arches displayed a significantly lower GFP signal than the two GFP-expressing DP populations—called in this study GFP^{High} and GFP^{Low}—detected in the thymus. The reduced Rag2-GFP expression could be based on several mechanisms: either extrathymic DPs have a constantly low RAG2-GFP expression or extrathymic DPs received a signal for TCR rearrangement before they had left the thymus (Fig. 5A). Bioinformatic TCR α- and TCR β-chain repertoire analysis of thymic and extrathymic DPs revealed that TCR α-chain clonotypes were lower compared with β-chains in all three DP populations, whereas the ratio between both chains did not differ between DPs isolated from the three tissues (Supplemental Fig. 6). This may suggest that the β-chain rearrangement has been finalized in all DPs and the α-chain rearrangement has been initiated. The diversity of clonality defined by the Simpson index is high and indicates a diverse TCR repertoire in all three DP populations. However, extrathymic DP development might have not further progressed past this stage, as no rearranged surface TCR was detectable by flow cytometry (Supplemental Fig. 1).

RAG2 expression is a prerequisite for T cell, B cell, and pre-TCR–expressing T cell progenitor presence in the thymus. To test whether RAG2 is also necessary for extrathymic DP presence, we studied MAT and aortic arches of $Rag2^{-/-}$ mice. *Rag2* deficiency coincides with absence of MAT and aortic arch DPs (Fig. 5B). We further established the relationship of extrathymic with thymic DPs by thymectomizing 3- and 21-d-old *C57BL/6J* mice. After a 6-wk recovery period, we could not detect any MAT DPs by flow cytometry, and the abundance of aortic arch DPs was drastically reduced (Fig. 5C).

Extrathymic DPs migrate locally

The drastically reduced presence of extrathymic DPs after thymectomy suggests that the thymus is the source of these cells. To further test this, we transplanted a third of a thymic lobe of a female congenically marked CD45.1 mouse under the kidney capsule of a female CD45.2 mice, which were previously sham operated or thymectomized. Presence of MAT and aortic arch DPs and their chimerism was assessed by flow cytometry 21 d after thymus transplantation (Fig. 6A). Thymectomy reduced the presence of DPs among living leukocytes in MAT (sham [$n = 7$] versus thymectomy [$n = 5$]: 10 ± 12% versus 0.05 ± 0.03% of living CD45) and aortic arch (sham [$n = 6$] versus thymectomy [$n = 5$]: 2.4 ± 3.1% versus 0.2 ± 0.1% of living $CD45^+$ cells). Transplantation of thymi under the kidney capsule did not lead to the accumulation of transplant-derived DPs in MAT or aortic arches (Fig. 6B) of sham-operated mice. Similar results were obtained in thymectomized recipient mice. As a positive control, we confirmed accumulation of transplant-derived $CD4^+$ T cells in inguinal lymph nodes (Supplemental Fig. 7). The failed population of transplant-derived DPs in MAT and aortic arches and their absence in the circulation (Fig. 1C) suggests that DPs do not enter the circulation. Thus, we hypothesized that extrathymic DPs migrate from the thymus and seed the MAT and aortic arch locally. To test this hypothesis, we blocked lymphocyte migration by pharmacologically inhibiting the sphingosine-1-phosphate receptor (S1P1) by injecting FTY720 every day for 7 d. As expected, the frequency of thymic DPs decreased, whereas frequencies of thymic single $CD4^+$ T cells and $CD8^+$ T cells increased when they were trapped in the thymus (Fig. 6C–E). EdU injection demonstrated turnover of all DPs within 5 d (Fig. 4C). Thus, if S1P1 was required for extrathymic DPs, we would expect reduced

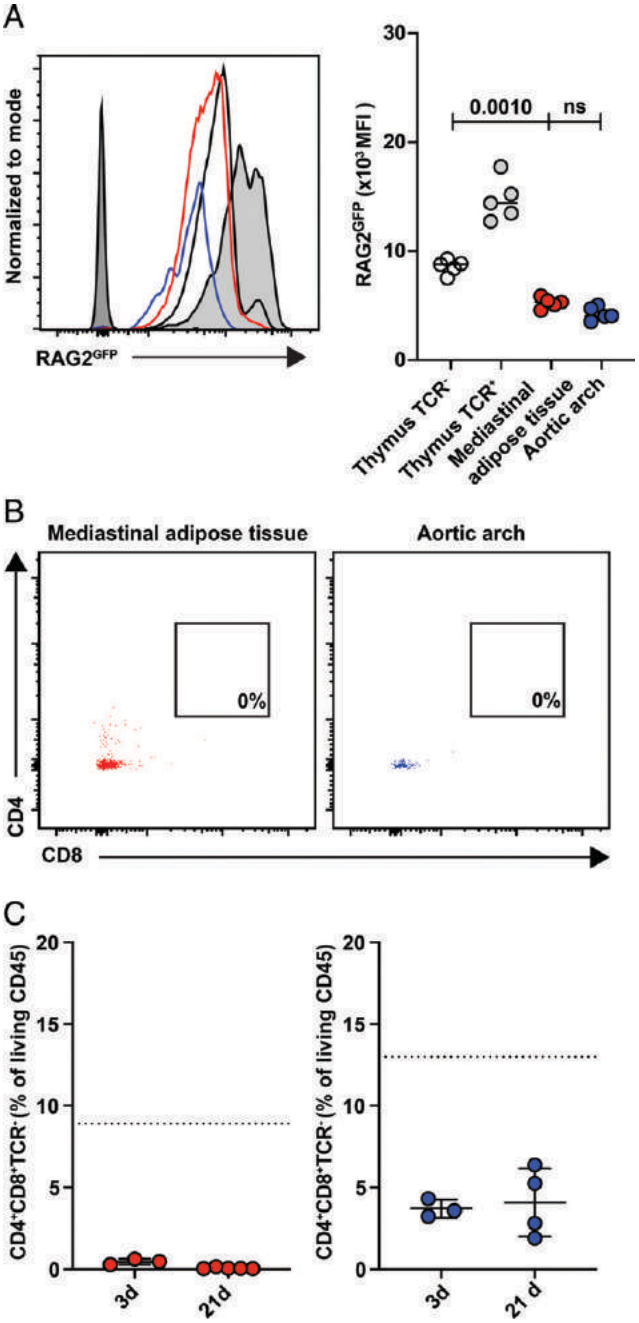


FIGURE 5. Extrathymic DP presence is thymus dependent. (**A**) $CD4^+CD8\alpha^+TCR\beta^-$ cells DPs from thymi (white circles and solid black line), MAT (red circles and red line), and aortic arches (blue circles and blue line) and $CD4^+CD8\alpha^+TCR\beta^+$ cells in the thymus (gray circles and gray shaded line) were analyzed for RAG2 expression using 8-wk-old female *Rag2*^{GFP} mice. Representative histograms are displayed on the left and the mean fluorescence intensity is quantified on the right. The dark gray line in the histogram indicates GFP background fluorescence in a *C57BL/6J* mouse. (**B**) MAT and aortic arches of 8-wk-old female *Rag2*$^{-/-}$ mice were analyzed for presence of DPs by flow cytometry. Representative flow cytometry plots of three independent experiments are displayed. (**C**) Female *C57BL/6J* mice were thymectomized 3 d ($n = 3$) and 21 d after birth ($n = 4$). After a 6-wk recovery period, MAT and aortic arches were analyzed for presence of DPs by flow cytometry. Dotted lines, levels in intact mice (not thymectomized).

DP frequencies in MAT and aortic arches. However, neither MAT nor aortic arch abundance was altered, pointing to a locally confined mode of DP migration that does not require S1P1.

Extrathymic DPs differentiate into single-positive T cells with a CD4 bias

To test whether extrathymic DPs are capable of developing into single $CD4^+$ and $CD8\alpha\beta^+$ T cells, we flow sorted DPs from MAT and aortic arches of female 8-wk-old *C57BL6/J* mice and subjected them to a hanging-drop fetal thymic organ culture (Fig. 7). Thymi were isolated from E14 *CD45.1/CD45.2* embryos. At day 10, of eight individual fetal thymic organ cultures, 20% of extrathymic $CD4^+CD8\alpha\beta^+$ CD45.2 cells expressed surface TCRβ. Half of these cells were expressing both CD4 and CD8, whereas the majority of single-positive T cells were CD4 positive (Fig. 7). TCRβ-negative extrathymic DPs were mostly CD4 and CD8 DP.

Next, we tested whether extrathymic DPs interacted with other cells in the tissue. iNKT cells need interactions with thymic DPs for their development and differentiation (36). iNKT cells are present in adipose tissue and play a prominent role in adipose tissue inflammation (37), although the exact nature of this role is controversial. First, we tested whether MAT, like other adipose tissues, contains iNKT cells. Confocal imaging of whole-mount MAT fat pads stained with a fluorochrome-conjugated CD1d tetramer loaded with α-GalCer showed the presence of iNKT cells (Fig. 8A). The detection threshold was determined by a sample stained with empty fluorochrome-conjugated CD1d tetramer. In 10 independent regions of interest, we found 76 CD1d $tetramer^+$ cells. We simultaneously stained for CD4 and CD8α in the staining and detected 37 $CD4^+CD8\alpha^+$ cells, indicating that iNKT cells outnumber DPs in MAT. Isosurface rendering revealed that DPs clustered around and in close proximity to iNKT cells (Fig. 8B–D). Quantification of 10 independent regions of interest in MAT showed that the distance from DPs to iNKT cells was 23 μm, whereas iNKT cells have a distance of 56 μm to $CD4^+CD8\alpha^+$ cells (Fig. 8E). This supports the notion that $CD4^+CD8\alpha^+$ cells in MAT reside close to iNKT cells, whereas some iNKT cells are distant from DPs. We further confirmed expression of CD1d by extrathymic DPs (Supplemental Fig. 8). To test whether MAT DPs might interact with iNKT cells, we designed a glycolipid Ag restimulation experiment. In brief, DPs were isolated from MAT and as a control from the thymus of *C57BL/6J* mice and pulsed with α-GalCer. Varying numbers of MAT and thymic DPs were cocultured with the DN3A4-1.2 hybridoma iNKT cell line for 12 h (7). This cell line specifically responds to antigenic stimulation by CD69 expression and IL-2 secretion. Because of limited DP numbers in MAT and aortic arches, we decided to coculture 50,000 MAT DPs with 50,000 DN3A4-1.2 hybridoma iNKT cells. The coculture of DPs from either origin with the hybridoma cells induced a similar expression of surface CD69 (Fig. 8F) on the iNKT cells and similar amount of IL-2 (Fig. 8G) secreted into the culture supernatant. Both responses were detectable over the included coculture controls without α-GalCer, and there was no difference whether thymic or MAT DPs were used at 50,000 cells. This experiment demonstrates that MAT DPs can activate iNKT cells in vitro.

Discussion

In the current study, we demonstrate the presence of DPs in MAT and aortic arches of wild-type mice. DP thymocytes are a heterogenous population. Blastic DPs are highly proliferative and do not express a surface TCR or CD3e, similar to extrathymic DPs. The process of β-selection induces a proliferative burst dependent on c-Myc activity (38) and pre-TCR (39) and Notch signaling (40). We show extrathymic DP proliferation by Ki67 staining and turnover by a BrdU pulse experiment. Signaling via the pre-TCR suppresses Notch signaling dependent pathways, which halts the proliferative

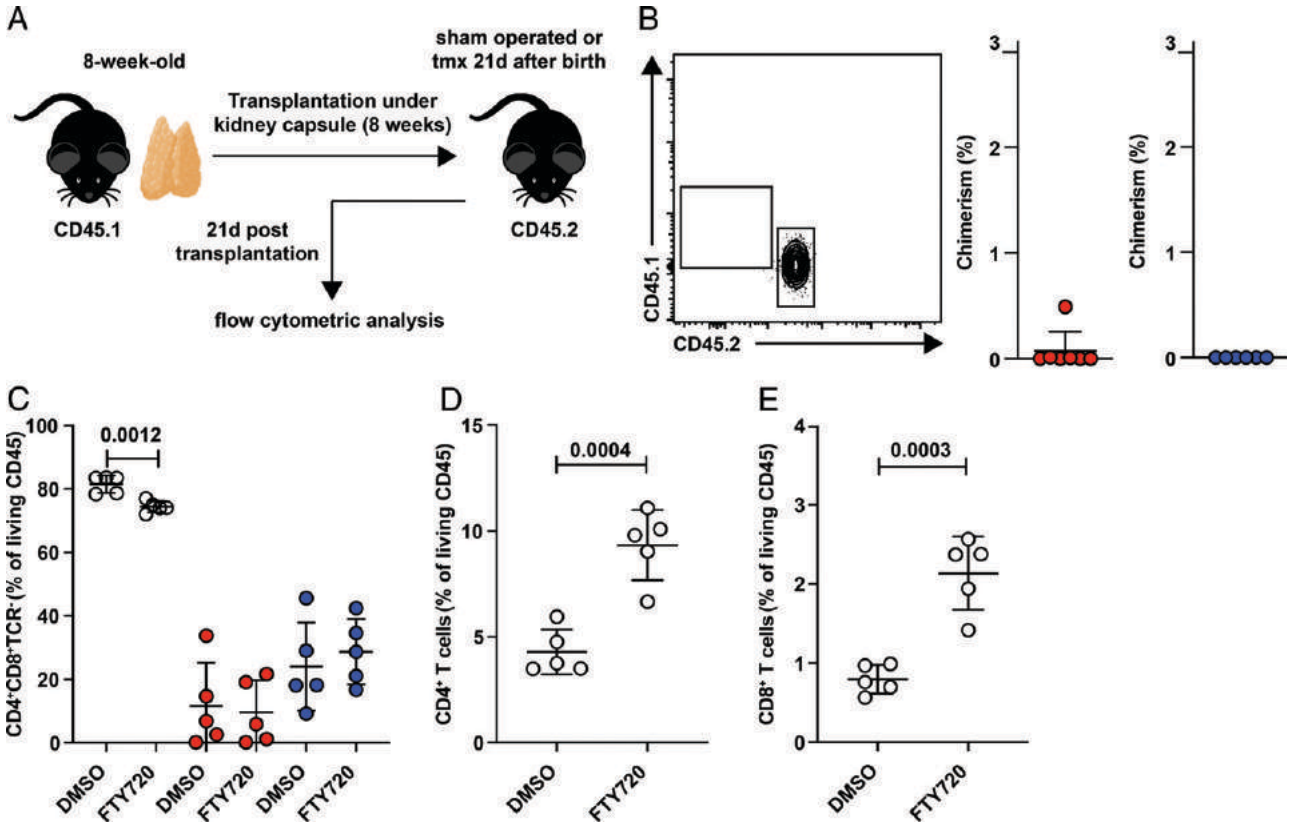


FIGURE 6. Extrathymic DP migration is locally restrained. (**A**) Experimental layout. Three-week-old female *C57BL/6J* mice were thymectomized or sham operated. At the age of 8 wk, mice were transplanted with one third of a thymic lobe of same-age female *CD45.1 C57BL/6J* mice under the kidney capsule. Mice were analyzed for DP presence in MAT and aortic arches 21 d after transplantation. (**B**) Representative flow cytometry plot of DPs in MAT of sham-operated mice. The percentage of chimerism in MAT (red dots) and aortic arch (blue dots) was quantified on the right. (**C**–**E**) Eight-week-old female *C57BL/6J* mice were injected i.p. with DMSO alone or with 1 mg/kg FTY720 in DMSO daily for 7 d ($n = 5$ per group). (C) Thymus, MAT, and aortic arch were analyzed for presence of DPs by flow cytometry. Thymic $CD4^+$ (D) and $CD8^+$ T cell content (E) in DMSO- and FTY720-treated mice.

burst of blastic DPs and starts their transition into quiescent small DPs (41). The shift from blastic DPs to small DPs was associated with a strong change in gene expression and downregulated many genes, 100 of which were involved in metabolism (41). Small DPs engage in positive selection to test their newly rearranged and surface expressed TCRαβ for binding to self-peptides presented in MHC complexes (42). Like blastic thymic DPs, extrathymic lack surface expression of CD3e and TCRβ but are considerably smaller. Successful rearrangement and signaling via the αβ-TCR upregulated surface expression of the activation marker CD69 on the small DPs,

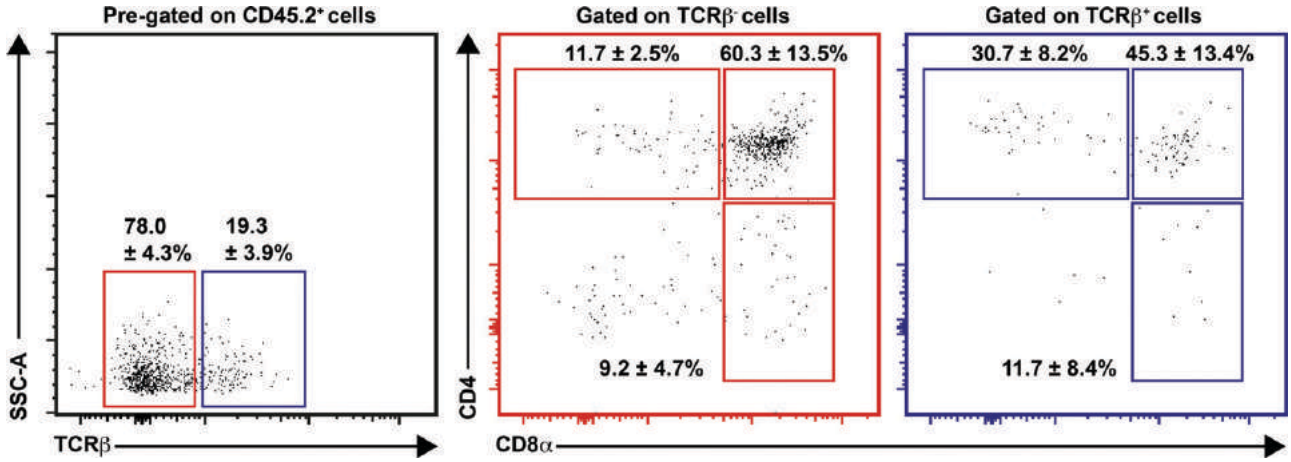


FIGURE 7. Extrathymic DPs differentiate into $CD4^+$ and $CD8\alpha\beta^+$ T cells. $CD4^+CD8\alpha\beta^+TCR^-$ cells (DPs) were isolated by flow cytometry from MAT and aortic arches of 10-wk-old female *C57BL/6J* mice ($n = 25$ mice, $n = 8$ pools of cells from three to four mice). E14 thymic lobes were dissected from *CD45.1/2 C57BL/6J* mice and cultured with 20,000 extrathymic DPs in a hanging-drop culture. After 2 d, the thymic lobes were transferred into a submersion culture for 8 d and subsequently analyzed by flow cytometry. Representative flow cytometric plots of living extrathymic DPs in the fetal thymic organ culture. Frequencies of $CD4^+$ and $CD8\alpha^+$ positive cells among TCRβ-positive (blue) and TCRβ-negative cells (red) are reported.

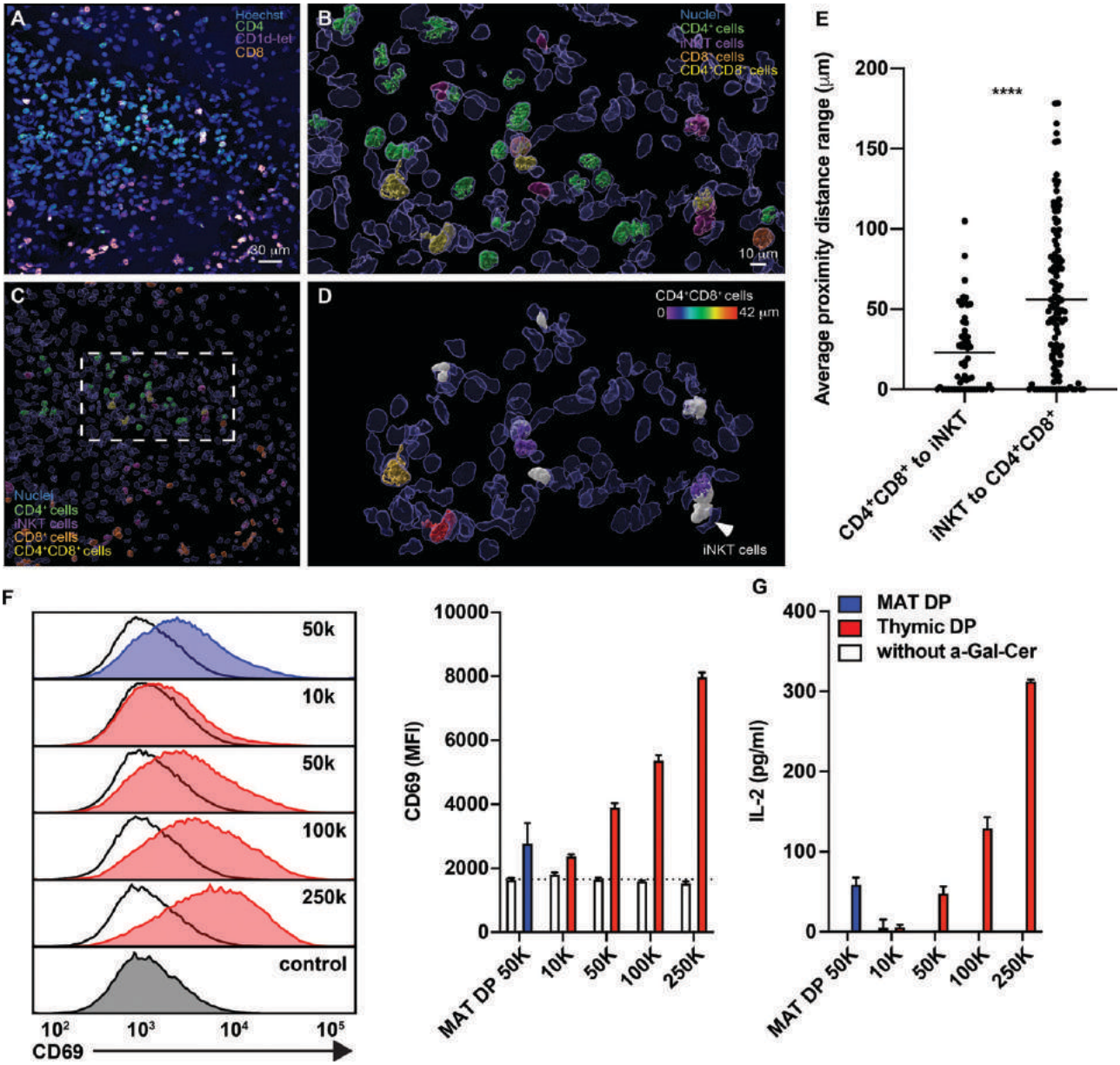


FIGURE 8. MAT DPs reside close to iNKT cells and can activate them. (**A**–**D**) Whole-mount MAT of an 8-wk-old female *C57BL/6J* mouse was stained for CD4 (green), CD8α (orange), α-GalCer/CD1d-tetramer (iNKT cells; purple/magenta), and nuclei (Hoechst; blue) and imaged using a C-Apochromat 40×/1.2 objective by confocal microscopy. (B) Rendered isosurfaces of labeled cells were constructed in Imaris. $CD4^+$ cells (green), $CD8\alpha^+$ cells (orange), iNKT cells (purple/magenta), nuclei (blue), and $CD4^+CD8\alpha^+$ cells (yellow) are displayed. (C) Magnified area highlighted in (D). iNKT cells are highlighted in white. The distance of $CD4^+CD8\alpha^+$ cells to the closest iNKT cells is encoded by the color scale. (**E**) Ten independent regions of interest were analyzed for the distance between iNKT cells and $CD4^+CD8\alpha^+$ cells. Each with a fixed location of one population to compare variable distance of the other from the cell center. (**F**) Thymic (red line) and MAT (blue line) $CD4^+CD8\alpha^+TCR^-$ cells were FACS sorted, pulsed with α-GalCer for 3.5 h or left untreated (black line), and cocultured for 12 h in duplicate with 50,000 iNKT DN3A4-1.2 hybridoma cells. CD69 expression on the iNKT DN3A4-1.2 hybridoma cells was measured by flow cytometry ($n = 3$ per condition). A hybridoma control only was included (control). Quantification of CD69 expression is reported; dashed line represents expression of CD69 in the iNKT DN3A4-1.2 hybridoma only control. (**G**) The culture supernatant was analyzed for IL-2 levels by ELISA. ****$p < 0.0001$.

labeling these cells $CD69^+$ DPs (43). Extrathymic DPs do not express CD69, and they have lower $Rag2^{GFP}$ expression, which suggests that they are similar to blastic DPs in the thymus and are in a proliferative phase.

Deep transcriptional profiling revealed differences between DP cells isolated from the three locations. Thymic DPs were enriched for pathways involved in TCR signaling and extrathymic DPs demonstrated stress responses and elevated cell cycle pathways. Only 94 genes are significantly overexpressed by aortic arch and MAT DPs compared with thymic DPs. Both populations of extrathymic DPs expressed *Grn* encoding for granulin. Granulin is a secreted protein with numerous functions including tissue growth, wound healing, and inflammation (44) and has been implicated in myelopoiesis (C. Campbell, O. Fursova, X. Cheng, E. Snella, A. McCune, L. Li, B. Solchenberger, B. Schmid, D. Sahoo, M. Morton, et al., manuscript posted on bioRxiv, DOI: 10.1101/2020.07.23.217067). The cognate receptor is unknown, but granulin can bind many receptors including TNFRs, Notch, TLR9, low-density lipoprotein receptor related protein 1 (LRP1), and sortilin1 and induce their downstream signaling cascades (45). It is thus tempting to speculate that extrathymic

DPs could contribute to local tissue homeostasis and wound repair. Although the three DP populations are transcriptomically different from each other, we would like to acknowledge that the current analysis is limited, and the magnitude of this difference in comparison with other thymocyte populations is unknown.

Although both extrathymic DP populations express genes that are involved in the development and effector function of CD4 and CD8 T cell subsets [*Irf8* (28, 29) and *Il12rb2* (31, 32], neither DP population differentially expressed genes that would determine commitment to the CD4 [*Thpok* (46)] or CD8 [*Runx3* (47)] T cell lineage. Aortic arch DPs uniquely expressed *Prf1* encoding for perforin, which is expressed during positive selection of CD8 T cells and by thymocytes receiving MHC class II signals (25). We demonstrate that extrathymic DPs are capable of developing into single-positive T cells in an in vitro hanging-drop culture. The MAT isolated DP cells outnumbered the aortic arch DPs by 10:1 in our cell culture system, as determined by flow cytometry. Because of the low cell numbers, we were unable to test the aortic arch DPs in the hanging-drop culture system.

Our data that extrathymic DP cells reside near iNKT cells and that DP cells express the MHC-like molecule CD1d suggest that they may be involved in activating iNKT cells. It should be noted that many cell types express CD1d and can activate iNKT cells (48). iNKT cells exist in different flavors and exert tissue-specific roles (49). They are prominent mediators of adipose tissue inflammation (37) or atherosclerosis (50) but also can protect from obesity through regulatory cytokine production (51). We provide evidence that MAT DPs can present lipid Ag and activate iNKT hybridoma cells in vitro and cluster around them in vivo. However, further research will have to clarify which Ags MAT DPs would present and if they indeed can activate a polyclonal iNKT response and how this would be shaped.

The occurrence of a cervical thymus is common in humans (52). The function of these structures is unknown, although they are considered to be origin sites of cervical thymomas in humans (53, 54). In 2006, Terszowski and colleagues (55) reported presence of a cervical thymus in multiple mouse strains including C57BL/6J mice. This structure, however, is located at a different anatomical side in the neck in distance to the aortic arch or MAT. The close vicinity of the aortic arch and MAT to the thymus and the rather large distance to cervical thymi suggest that the DPs in these to extrathymic tissues migrate locally from the thymus, which we further confirmed experimentally. The pharmacological inhibition of S1P1R trapped $CD4^+$ and $CD8^+$ T cells in the thymus as expected but did not alter the presence of DP numbers in MAT and the aortic arch, suggesting that they do not enter the circulation and migrate locally. Transplantation of congenically marked thymi under the kidney capsule failed to reconstitute DPs in the MAT or aortic arch.

Recent studies employing mass cytometry, scRNA-seq, or cellular indexing of transcriptomes and epitopes by sequencing to uncover cellular heterogeneity in human and mouse atherosclerosis identified the existence of aortic and plaque-resident $CD4^+CD8\alpha\beta^+$ T cells (3, 56, 57). In a recent meta-analysis integrating single-cell transcriptomes from aortic leukocytes in atherosclerosis, we identified a cluster with mixed phenotype of single-positive *Cd4*$^+$ and *Cd8*$^+$ T cells (58). This cluster accounted for ∼13–26% of aortic T cells in scRNA-seq, a frequency we also observed by flow cytometry in young and old wild-type mice in the current study. Although the frequency of DPs was higher in atherosclerotic mouse aortas as assayed by flow cytometry, their relative abundance did not change with disease progression. Thus, we have no evidence that aortic and MAT DPs are related to cardiovascular disease. Mass cytometric analysis of human atherosclerotic plaques also uncovered $CD4^+CD8\alpha\beta^+$ T cells. Unlike mouse DPs, these cells had relevant cell surface expression of CD3 (56). Indeed, mature human $CD3^+CD4^+CD8\alpha\beta^+$ T cells have been identified in tumors (59, 60) and the circulation of patients with rheumatoid arthritis (61). These cells lacked expression of relevant thymocyte markers and are likely a mature independent subset different from the immature DP cells uncovered in this study.

Taken together, our findings define two new, transcriptionally distinct subsets of extrathymic DPs that escape from the thymus independent of S1P1.

Acknowledgments

We thank Cheryl Kim, Denise Hinz, Christopher Dillingham, and Matthew Haynes for cell sorting and Zbigniew Mikulski and Sara McArdle for expert help with imaging analysis. We also thank Dr. Ellen Rothenberg for kind advice and critical feedback throughout the preparation of the manuscript.

Disclosures

The authors have no financial conflicts of interest.

References

1. Wolf, D., and K. Ley. 2019. Immunity and inflammation in atherosclerosis. *Circ. Res.* 124: 315–327.
2. Winkels, H., E. Ehinger, M. Vassallo, K. Buscher, H. Dinh, K. Kobiyama, A. Hamers, C. Cochain, E. Vafadarnejad, A. E. Saliba, et al. 2018. Atlas of the immune cell repertoire in mouse atherosclerosis defined by single-cell RNA-sequencing and mass cytometry. *Circ Res.* 122: 1675–1688.
3. Cochain, C., E. Vafadarnejad, P. Arampatzi, J. Pelisek, H. Winkels, K. Ley, D. Wolf, A. E. Saliba, and A. Zernecke. 2018. Single-cell RNA-seq reveals the transcriptional landscape and heterogeneity of aortic macrophages in murine atherosclerosis. *Circ. Res.* 122: 1661–1674.
4. Hosokawa, H., and E. V. Rothenberg. 2018. Cytokines, transcription factors, and the initiation of T-cell development. *Cold Spring Harb. Perspect. Biol.* 10: a028621.
5. Galkina, E., A. Kadl, J. Sanders, D. Varughese, I. J. Sarembock, and K. Ley. 2006. Lymphocyte recruitment into the aortic wall before and during development of atherosclerosis is partially L-selectin dependent. *J. Exp. Med.* 203: 1273–1282.
6. Li, W., R. N. Germain, and M. Y. Gerner. 2017. Multiplex, quantitative cellular analysis in large tissue volumes with clearing-enhanced 3D microscopy (C_e3D). *Proc. Natl. Acad. Sci. USA* 114: E7321–E7330.
7. Burdin, N., L. Brossay, Y. Koezuka, S. T. Smiley, M. J. Grusby, M. Gui, M. Taniguchi, K. Hayakawa, and M. Kronenberg. 1998. Selective ability of mouse CD1 to present glycolipids: alpha-galactosylceramide specifically stimulates V alpha 14+ NK T lymphocytes. *J. Immunol.* 161: 3271–3281.
8. Subramanian, A., P. Tamayo, V. K. Mootha, S. Mukherjee, B. L. Ebert, M. A. Gillette, A. Paulovich, S. L. Pomeroy, T. R. Golub, E. S. Lander, and J. P. Mesirov. 2005. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. USA* 102: 15545–15550.
9. Barbie, D. A., P. Tamayo, J. S. Boehm, S. Y. Kim, S. E. Moody, I. F. Dunn, A. C. Schinzel, P. Sandy, E. Meylan, C. Scholl, et al. 2009. Systematic RNA interference reveals that oncogenic KRAS-driven cancers require TBK1. *Nature* 462: 108–112.
10. Rosales, S. L., S. Liang, I. Engel, B. J. Schmiedel, M. Kronenberg, P. Vijayanand, and G. Seumois. 2018. A sensitive and integrated approach to profile messenger RNA from samples with low cell numbers. [Published erratum appears in 2018 *Methods Mol. Biol.* 1799: C3]. *Methods Mol. Biol.* 1799: 275–302.
11. Andrews, S. 2010. FastQC: A Quality Control Tool for High Throughput Sequence Data [Online]. Available at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc/. Accessed: August 13, 2019.
12. Dobin, A., C. A. Davis, F. Schlesinger, J. Drenkow, C. Zaleski, S. Jha, P. Batut, M. Chaisson, and T. R. Gingeras. 2013. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 29: 15–21.
13. Yates, A. D., P. Achuthan, W. Akanni, J. Allen, J. Allen, J. Alvarez-Jarreta, M. R. Amode, I. M. Armean, A. G. Azov, R. Bennett, et al. 2020. Ensembl 2020. *Nucleic Acids Res.* 48(D1): D682–D688.
14. Hartley, S. W., and J. C. Mullikin. 2015. QoRTs: a comprehensive toolset for quality control and data processing of RNA-Seq experiments. *BMC Bioinformatics* 16: 224.
15. Tange, O. 2011. GNU Parallel - the command-line power tool. *The USENIX Magazine* 36: 42–47.
16. Love, M. I., W. Huber, and S. Anders. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15: 550.
17. Lex, A., N. Gehlenborg, H. Strobelt, R. Vuillemot, and H. Pfister. 2014. UpSet: visualization of intersecting sets. *IEEE Trans. Vis. Comput. Graph.* 20: 1983–1992.
18. Langfelder, P., and S. Horvath. 2008. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* 9: 559.

19. Bolotin, D. A., S. Poslavsky, A. N. Davydov, F. E. Frenkel, L. Fanchi, O. I. Zolotareva, S. Hemmers, E. V. Putintseva, A. S. Obraztsova, M. Shugay, et al. 2017. Antigen receptor repertoire profiling from RNA-seq data. *Nat. Biotechnol.* 35: 908–911.
20. Lusis, A. J., M. M. Seldin, H. Allayee, B. J. Bennett, M. Civelek, R. C. Davis, E. Eskin, C. R. Farber, S. Hui, M. Mehrabian, et al. 2016. The hybrid mouse diversity panel: a resource for systems genetics analyses of metabolic and cardiovascular traits. *J. Lipid Res.* 57: 925–942.
21. Miragaia, R. J., T. Gomes, A. Chomka, L. Jardine, A. Riedel, A. N. Hegazy, N. Whibley, A. Tucci, X. Chen, I. Lindeman, et al. 2019. Single-cell transcriptomics of regulatory T cells reveals trajectories of tissue adaptation. *Immunity* 50: 493–504.e7.
22. Wherry, E. J., S. J. Ha, S. M. Kaech, W. N. Haining, S. Sarkar, V. Kalia, S. Subramaniam, J. N. Blattman, D. L. Barber, and R. Ahmed. 2007. Molecular signature of CD8+ T cell exhaustion during chronic viral infection. *Immunity* 27: 670–684.
23. Speckmann, C., S. S. Sahoo, M. Rizzi, S. Hirabayashi, A. Karow, N. K. Serwas, M. Hoemberg, N. Damatova, D. Schindler, J. B. Vannier, et al. 2017. Corrigendum: clinical and molecular heterogeneity of rtel1 deficiency. [Published erratum appears in 8 Front Immunol. 8: 449]. *Front. Immunol.* 8: 1250.
24. Levring, T. B., A. K. Hansen, B. L. Nielsen, M. Kongsbak, M. R. von Essen, A. Woetmann, N. Odum, C. M. Bonefeld, and C. Geisler. 2012. Activated human CD4+ T cells express transporters for both cysteine and cystine. *Sci. Rep.* 2: 266.
25. Liu, X., B. J. Taylor, G. Sun, and R. Bosselut. 2005. Analyzing expression of perforin, Runx3, and Thpok genes during positive selection reveals activation of CD8-differentiation programs by MHC II-signaled thymocytes. *J. Immunol.* 175: 4465–4474.
26. Borges, T. J., L. Wieten, M. J. van Herwijnen, F. Broere, R. van der Zee, C. Bonorino, and W. van Eden. 2012. The anti-inflammatory mechanisms of Hsp70. *Front. Immunol.* 3: 95.
27. Sekiya, T., I. Kashiwagi, N. Inoue, R. Morita, S. Hori, H. Waldmann, A. Y. Rudensky, H. Ichinose, D. Metzger, P. Chambon, and A. Yoshimura. 2011. The nuclear orphan receptor Nr4a2 induces Foxp3 and regulates differentiation of CD4+ T cells. *Nat. Commun.* 2: 269.
28. Miyagawa, F., H. Zhang, A. Terunuma, K. Ozato, Y. Tagaya, and S. I. Katz. 2012. Interferon regulatory factor 8 integrates T-cell receptor and cytokine-signaling pathways and drives effector differentiation of CD8 T cells. *Proc. Natl. Acad. Sci. USA* 109: 12123–12128.
29. Lee, W., H. S. Kim, S. Y. Baek, and G. R. Lee. 2016. Transcription factor IRF8 controls Th1-like regulatory T-cell function. *Cell. Mol. Immunol.* 13: 785–794.
30. Li, L., H. C. Hsu, C. R. Stockard, P. Yang, J. Zhou, Q. Wu, W. E. Grizzle, and J. D. Mountz. 2004. IL-12 inhibits thymic involution by enhancing IL-7- and IL-2-induced thymocyte proliferation. *J. Immunol.* 172: 2909–2916.
31. Powell, M. D., K. A. Read, B. K. Sreekumar, D. M. Jones, and K. J. Oestreich. 2019. IL-12 signaling drives the differentiation and function of a T_H1-derived T_{FH1}-like cell population. *Sci. Rep.* 9: 13991.
32. Zhao, Z., S. Yu, D. C. Fitzgerald, M. Elbehi, B. Ciric, A. M. Rostami, and G. X. Zhang. 2008. IL-12R beta 2 promotes the development of CD4+CD25+ regulatory T cells. *J. Immunol.* 181: 3870–3876.
33. Schwarz, D. A., C. D. Katayama, and S. M. Hedrick. 1998. Schlafen, a new family of growth regulatory genes that affect thymocyte development. *Immunity* 9: 657–668.
34. Boursalian, T. E., J. Golob, D. M. Soper, C. J. Cooper, and P. J. Fink. 2004. Continued maturation of thymic emigrants in the periphery. *Nat. Immunol.* 5: 418–425.
35. Yu, W., H. Nagaoka, M. Jankovic, Z. Misulovin, H. Suh, A. Rolink, F. Melchers, E. Meffre, and M. C. Nussenzweig. 1999. Continued RAG expression in late stages of B cell development and no apparent re-induction after immunization. *Nature* 400: 682–687.
36. Hogquist, K., and H. Georgiev. 2020. Recent advances in iNKT cell development. *F1000 Res.* 9: F1000:127.
37. Park, Y. J., J. Park, J. Y. Huh, I. Hwang, S. S. Choe, and J. B. Kim. 2018. Regulatory roles of invariant natural killer T cells in adipose tissue inflammation: defenders against obesity-induced metabolic complications. *Front. Immunol.* 9: 1311.
38. Dose, M., I. Khan, Z. Guo, D. Kovalovsky, A. Krueger, H. von Boehmer, K. Khazaie, and F. Gounari. 2006. c-Myc mediates pre-TCR-induced proliferation but not developmental progression. *Blood* 108: 2669–2677.
39. Ciofani, M., T. M. Schmitt, A. Ciofani, A. M. Michie, N. Cuburu, A. Aublin, J. L. Maryanski, and J. C. Zúñiga-Pflücker. 2004. Obligatory role for cooperative signaling by pre-TCR and Notch during thymocyte differentiation. *J. Immunol.* 172: 5230–5239.
40. Maillard, I., L. Tu, A. Sambandam, Y. Yashiro-Ohtani, J. Millholland, K. Keeshan, O. Shestova, L. Xu, A. Bhandoola, and W. S. Pear. 2006. The requirement for Notch signaling at the beta-selection checkpoint in vivo is absolute and independent of the pre-T cell receptor. *J. Exp. Med.* 203: 2239–2245.
41. Immunological Genome Consortium. 2013. The transcriptional landscape of αβ T cell differentiation. *Nat. Immunol.* 14: 619–632.
42. Starr, T. K., S. C. Jameson, and K. A. Hogquist. 2003. Positive and negative selection of T cells. *Annu. Rev. Immunol.* 21: 139–176.
43. Swat, W., M. Dessing, H. von Boehmer, and P. Kisielow. 1993. CD69 expression during selection and maturation of CD4 + 8+ thymocytes. *Eur. J. Immunol.* 23: 739–746.
44. Bateman, A., S. T. Cheung, and H. P. J. Bennett. 2018. A brief overview of progranulin in health and disease. *Methods Mol. Biol.* 1806: 3–15.
45. Chitramuthu, B. P., H. P. J. Bennett, and A. Bateman. 2017. Progranulin: a new avenue towards the understanding and treatment of neurodegenerative disease. *Brain* 140: 3081–3104.
46. He, X., X. He, V. P. Dave, Y. Zhang, X. Hua, E. Nicolas, W. Xu, B. A. Roe, and D. J. Kappes. 2005. The zinc finger transcription factor Th-POK regulates CD4 versus CD8 T-cell lineage commitment. *Nature* 433: 826–833.
47. Woolf, E., C. Xiao, O. Fainaru, J. Lotem, D. Rosen, V. Negreanu, Y. Bernstein, D. Goldenberg, O. Brenner, G. Berke, et al. 2003. Runx3 and Runx1 are required for CD8 T cell development during thymopoiesis. *Proc. Natl. Acad. Sci. USA* 100: 7731–7736.
48. Brigl, M., and M. B. Brenner. 2004. CD1: antigen presentation and T cell function. *Annu. Rev. Immunol.* 22: 817–890.
49. Crosby, C. M., and M. Kronenberg. 2018. Tissue-specific functions of invariant natural killer T cells. *Nat. Rev. Immunol.* 18: 559–574.
50. Getz, G. S., and C. A. Reardon. 2017. Natural killer T cells in atherosclerosis. *Nat. Rev. Cardiol.* 14: 304–314.
51. Lynch, L., M. Nowak, B. Varghese, J. Clark, A. E. Hogan, V. Toxavidis, S. P. Balk, D. O'Shea, C. O'Farrelly, and M. A. Exley. 2012. Adipose tissue invariant NKT cells protect against diet-induced obesity and metabolic disorder through regulatory cytokine production. *Immunity* 37: 574–587.
52. Prabhu, A. V., H. A. Kale, and B. F. Branstetter IV. 2015. Residual cervical thymus: a normal CT finding that may be present throughout patients' lives. *AJNR Am. J. Neuroradiol.* 36: 1525–1528.
53. Tovi, F., and A. J. Mares. 1978. The aberrant cervical thymus. Embryology, pathology, and clinical implications. *Am. J. Surg.* 136: 631–637.
54. Yamashita, H., N. Murakami, S. Noguchi, A. Noguchi, S. Yokoyama, A. Moriuchi, and I. Nakayama. 1983. Cervical thymoma and incidence of cervical thymus. *Acta Pathol. Jpn.* 33: 189–194.
55. Terszowski, G., S. M. Müller, C. C. Bleul, C. Blum, R. Schirmbeck, J. Reimann, L. D. Pasquier, T. Amagai, T. Boehm, and H. R. Rodewald. 2006. Evidence for a functional second thymus in mice. *Science* 312: 284–287.
56. Fernandez, D. M., A. H. Rahman, N. F. Fernandez, A. Chudnovskiy, E. D. Amir, L. Amadori, N. S. Khan, C. K. Wong, R. Shamailova, C. A. Hill, et al. 2019. Single-cell immune landscape of human atherosclerotic plaques. *Nat. Med.* 25: 1576–1588.
57. Winkels, H., E. Ehinger, M. Vassallo, K. Buscher, H. Q. Dinh, K. Kobiyama, A. A. J. Hamers, C. Cochain, E. Vafadarnejad, A. E. Saliba, et al. 2018. Atlas of the immune cell repertoire in mouse atherosclerosis defined by single-cell RNA-sequencing and mass cytometry. *Circ. Res.* 122: 1675–1688.
58. Zernecke, A., W. H., C. Cochain, J. Williams, D. Wolf, O. Soehnlein, C. Robbins, M. Monaco, I. Park, C. McNamara, C. Binder et al. 2020. Meta-analysis of leukocyte diversity in atherosclerotic mouse aortas. *Circ. Res.* 127: 402–426.
59. Overgaard, N. H., J. W. Jung, R. J. Steptoe, and J. W. Wells. 2015. CD4+/CD8+ double-positive T cells: more than just a developmental stage? *J. Leukoc. Biol.* 97: 31–38.
60. Bohner, P., M. F. Chevalier, V. Cesson, S. C. Rodrigues-Dias, F. Dartiguenave, R. Burruni, T. Tawadros, M. Valerio, I. Lucca, D. Nardelli-Haefliger, et al. 2019. Double positive $CD4^+CD8^+$ T cells are enriched in urological cancers and favor T helper-2 polarization. *Front. Immunol.* 10: 622.
61. Quandt, D., K. Rothe, R. Scholz, C. W. Baerwald, and U. Wagner. 2014. Peripheral CD4CD8 double positive T cells with a distinct helper cytokine profile are increased in rheumatoid arthritis. *PLoS One* 9: e93293.