Shots to prevent cancer show early promise p. 126

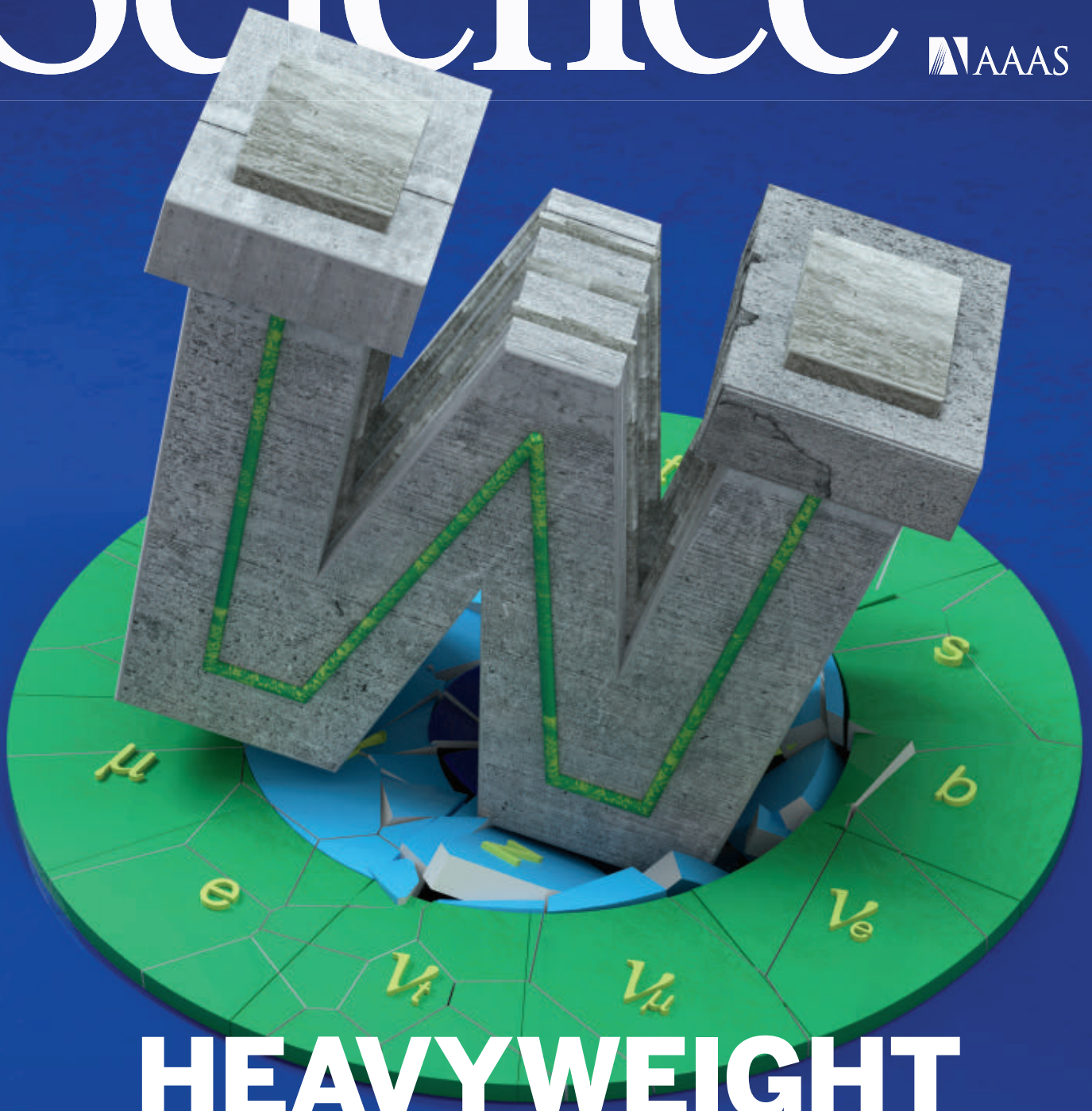
Visualizing a key step in cytokine signaling pp. 139 & 163

Silk-wrapped food wins BII & *Science* Prize p. 146

Science

$15
8 APRIL 2022
science.org

AAAS



HEAVYWEIGHT

W boson mass measures higher than expected pp. 125, 136, & 170

It's prime time!

Introducing LyoPrime™ Luna Probe One-Step RT-qPCR Mix with UDG – Lyophilized for increased flexibility, same optimized performance

Supplied as a lyophilized cake, the LyoPrime Luna Probe One-Step RT-qPCR Mix with UDG enables sensitive detection of targeted RNA sequences in a room temperature-stable format, and contains the same versatile features and strong performance as the liquid version. This is the first in a series of lyophilized products developed jointly by New England Biolabs and Fluorogenics™ Limited, which is now a subsidiary of New England Biolabs, Inc.

- Simply add nuclease-free water for rapid rehydration
- Store at room temperature for up to 2 years prior to rehydration
- Eliminate cold chain shipping requirements
- Multiplex up to 5 targets to increase throughput
- Increase reaction specificity and robustness with our unique pairing of Luna WarmStart® RT and Hot Start *Taq*
- Prevent carryover contamination with Thermolabile UDG and dUTP included in optimized mix
- Maintain RNA integrity with Murine RNase Inhibitor included in optimized mix
- Eliminate pipetting errors with non-interfering, visible blue tracking dye

By simply adding water, the cake can be rehydrated and is ready for use.

Learn more at
www.neb.com/L4001

NEW ENGLAND BioLabs® Inc.

be INSPIRED
drive DISCOVERY
stay GENUINE

CONTENTS

8 APRIL 2022 • VOLUME 376 • ISSUE 6589

130

PHOTO: ERIC GREGORY/THE JOURNAL-STAR/AP PHOTO

NEWS

IN BRIEF

IN DEPTH

FEATURES

INSIGHTS

POLICY FORUM

PERSPECTIVES

WE'RE CELEBRATING.

A NEW ANTIBODY FAILED TO FAIL.

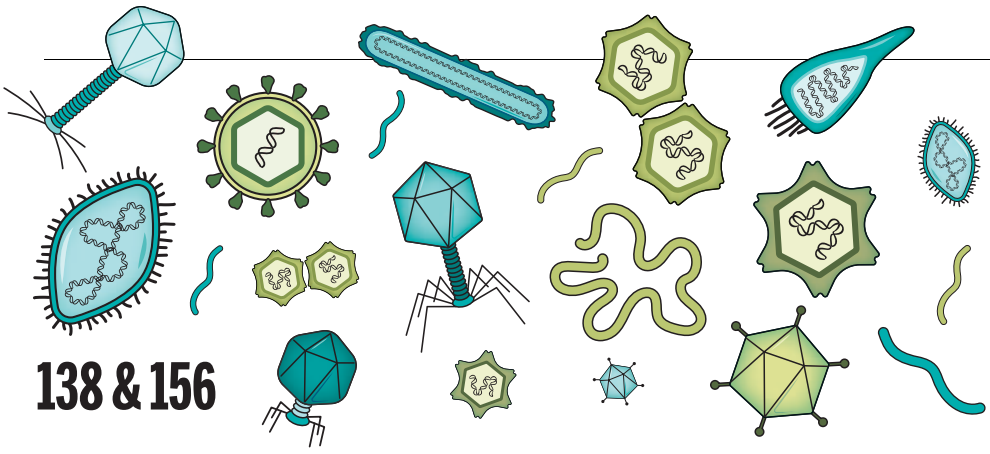
The US FDA has granted a Breakthrough Therapy designation to the first anti-TIGIT therapy for use in NSCLC in combination with PD-L1 blockade. TIGIT expression in paraffin-embedded non-small cell lung cancer is observed using TIGIT (E5Y1W) XP® Rabbit mAb #99567.

OUR PROCESS — YOUR SUCCESS.

As a privately held company founded and led by active research scientists, we understand your needs as a researcher. Like you, we only want to use products that are specific and deliver reproducible results. We validate every one of our thousands of antibodies and research products in-house. If any product fails to meet our stringent standards, we won't sell it. But when one passes, it's a big deal. For over 20 years, researchers have trusted CST to help them achieve experimental success. You can, too.

www.cellsignal.com

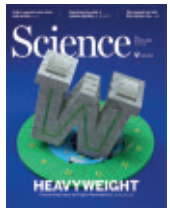
Cell Signaling TECHNOLOGY®



138 & 156

ON THE COVER

Symbolic illustration of the *W* boson weighing down on the standard model of particle physics. The mass of the *W* boson, a mediator of the weak nuclear force, is tightly constrained by the theory. A new, very-high-precision measurement of the *W* boson mass is in significant tension with the standard model expectation and suggests that improvements to calculations or extensions to the standard model might be needed. See pages 125, 136, and 170.
Illustration: Carlo Cadenas



GRAPHIC: A. MASTIN/*SCIENCE*

SCIENCE (ISSN 0036-8075) is published weekly on Friday, except last week in December, by the American Association for the Advancement of Science, 1200 New York Avenue, NW, Washington, DC 20005. Periodicals mail postage (publication No. 484460) paid at Washington, DC, and additional mailing offices. The title SCIENCE is a registered trademark of the AAAS. Domestic individual membership, including subscription (12 months): $165 ($74 allocated to subscription). Domestic institutional subscription (51 issues): $2212; Foreign postage extra: Air assist delivery: $98. First class, airmail, student, and emeritus rates on request. Canadian rates with GST available upon request, GST #125488122. Publications Mail Agreement Number 1069624. **Printed in the U.S.A.**
Change of address: Allow 4 weeks, giving old and new addresses and 8-digit account number. **Postmaster:** Send change of address to AAAS, P.O. Box 96178, Washington, DC 20090–6178. **Single-copy sales:** $15 each plus shipping and handling available from backissues.science.org; bulk rate on request. **Authorization to reproduce** material for internal or personal use under circumstances not falling within the fair use provisions of the Copyright Act can be obtained through the Copyright Clearance Center (CCC), www.copyright.com. The identification code for *Science* is 0036-8075. *Science* is indexed in the *Reader's Guide to Periodical Literature* and in several specialized indexes.

NOMIS & Science
YOUNG EXPLORER AWARD 2022

Research at the intersection of the social and life sciences

Unconventional. Interdisciplinary. Bold.

The **NOMIS & *Science* Young Explorer Award** recognizes and rewards early-career M.D., Ph.D., or M.D./Ph.D. scientists that perform research at the intersection of the social and life sciences. Essays written by these bold researchers on their recent work are judged for clarity, scientific quality, creativity, and demonstration of cross-disciplinary approaches to address fundamental questions.

A cash prize of up to 15 000 USD will be awarded to essay winners, and their engaging essays will be published in ***Science***. Winners will also be invited to share their work and forward-looking perspective with leading scientists in their respective fields at an award ceremony as well as a meeting of the NOMIS Board of Directors to consider future funding.

Apply by May 15, 2022

at www.science.org/nomis

NOMIS FOUNDATION
Creating the Spark

Science
AAAS

EDITORIAL

New goals for science philanthropy

Science philanthropy is experiencing a growth spurt, propelled by the newly acquired wealth of individuals and foundations, as well as a desire to address challenges such as infectious disease, fire, drought, and food and water security. Especially in the United States, this is altering the dynamics of the research ecosystem, which has been dominated by government funding since the end of World War II. This change comes with new perspectives and approaches to solving the world's problems. And it comes with a commitment to increase equity in funding.

Current philanthropy supports basic research in the United States with about $5 billion annually. When legacy philanthropic endowments spent by research institutions are taken into account, that number is about $25 billion per year. These estimates, based on US National Science Foundation (NSF) data, indicate that philanthropy accounts for 42% of support for basic science at US research institutions.

> "This change comes with... a commitment to increase equity in funding."

Entrepreneurs are deploying newfound wealth to form foundations and philanthropic organizations, joining the ranks of more established foundations, some with a century-old history. Their origin story is not so different from that of the agricultural, oil, gas, and railroad barons of yore—they have become wealthy through private enterprise. What is new is their willingness to confront confounding issues of the day, such as how to identify unexplored areas of research and apply new technologies for discovery, how to leverage funding through creative partnerships, how to redress societal inequities, and how to involve the public in research design.

Philanthropies are now partnering with public entities such as government agencies to extend their impact. "We're being partners when we identify areas where the federal government cannot easily invest and we can make those investments," said David Spergel, president of the Simons Foundation, "Sometimes philanthropic funding can be about 'de-risking' projects." An example is the Vera C. Rubin Observatory in Chile, where philanthropists assumed the risk of funding the development of a new mirror technology before the NSF stepped in with support. Philanthropy can provide flexibility that government agencies may lack. With the NSF-Simons Research Centers for Mathematics of Complex Biological Systems, Spergel says, "We were able to provide funding for the centers in ways that were more difficult for NSF to fund, through fewer rules on things like supporting visitors, conferences, [and] postdocs." NSF in turn brought the benefits of the new center to a broader community. "The whole was greater than the sum of the parts," says Spergel.

The new collaborations are working to overcome past limitations in which some philanthropies followed too narrowly the predilections of their founders or tended to direct money to high-profile universities and already established scientists. The new philanthropy is placing more emphasis on positioning equity among its goals. Some members of the Science Philanthropy Alliance, composed of 35 of the largest science funders, expressly seek out underrepresented scientists. For example, the Sloan Foundation widens education pathways for students at minority-serving institutions. Lyda Hill Philanthropies envisions a culture shift among young girls, opening their eyes to careers in science by involving media, sports, fashion, and female science innovators as role models. At the same time, philanthropies are focusing more on efficiency and effectiveness in their grant making. The Research Corporation for Science Advancement, with partner foundations and federal agencies, sponsors interdisciplinary dialogs among early-career researchers to develop innovative, collaborative proposals born "on the spot" during meetings that are reviewed rapidly for seed funding.

Many foundations are building communities that extend beyond researchers, collapsing silos and encouraging interactions across groups and disciplines. The Chan Zuckerberg Initiative funds patient communities to build research networks and partner in research project design. The goals of civic science are also a priority. Foundations, including the Rita Allen, Kavli, Gordon and Betty Moore, Heising-Simons, and Packard, along with the Burroughs Wellcome Fund, support the Civic Science Fellows program to catalyze interactions between science and society. And the Kavli Foundation recently funded university centers to engage the public in ethical issues in fields such as artificial intelligence, neuroscience, and genomics.

Societally responsible philanthropy recognizes the need to improve the world through funding science. Foundation leaders are taking bolder actions. The result will be a more responsive science that pushes the frontiers of knowledge in service of humanity.

–France A. Córdova

PHOTO: STEPHEN VOSS

France A. Córdova is president of the Science Philanthropy Alliance, Palo Alto, CA, USA. president@sciphil.org

10.1126/science.abq2259

Underrepresented STEM students can visit and conduct research in numerous locations in South America, Africa, and Europe, including Bourdeaux, France.

LSAMP-NICE expands international exchanges between underrepresented STEM students and their host laboratories

In 2019, as a U.S. scientific ambassador on a trip to South Africa, then-graduate-student Joshua Ames was looking forward to practicing his talk on his virology Ph.D. work. At the time, he didn't realize he would also bring back more than the souvenirs packed in his suitcase.

When presenting his work to faculty members at the Stellenbosch University medical school, Ames highlighted how he had deleted a protein from both genetically engineered mice and human cells in culture to show the protein's importance in protecting against viral infections. "The cancer biologists and immunologists there suggested several other cell culture lines to test," he recalls. "It was a small suggestion from a group that wouldn't have encountered my work otherwise."

This brief moment of scientific exchange led Ames to go back and add human corneal, neuronal, and skin cells to his project. "It expanded our understanding of how this protein protects certain tissues in the body," says Ames, now a postdoctoral fellow at University of Washington in Seattle. It also earned him a first-author *Nature Communications* paper in September 2021.

This free-flowing exchange of knowledge between science, technology, engineering, and mathematics (STEM) students and leading international researchers is exactly why the Louis Stokes Regional NSF International Center of Excellence (LSAMP-NICE) program was created. Now, the program is expanding its horizons to include more international partnerships so that more underrepresented graduate students can benefit from global research exchanges and mentorship [see sidebar: Louis Stokes Alliances for Minority Participation].

These partnerships are two-way, mutually beneficial endeavors, says Romilla Maharaj, executive director of Human and Infrastructure Capacity Development for South Africa's National Research Foundation (NRF) in Pretoria.

"There are a lot of ways that we as a country punch above our weight. We have internationally competitive, world-leading researchers and facilities like the Square Kilometre Array radio telescope and the iThemba LABS particle accelerators" from which students can learn, develop, and grow as researchers, she says.

Expanding who gets to do global research

Ames and another doctoral student at the time, Jason Garcia, represented the Illinois LSAMP alliance on the South Africa visit. "We went to show the level of research that LSAMP Ph.D. students are performing and to start building relationships with top universities in South Africa," says Ames. During the trip, the two attended Science Forum South Africa in Pretoria, toured the iThemba LABS particle accelerator facility near Faure, visited the Cradle of Humankind museum, hiked up Table Mountain, and presented their work at Stellenbosch University and Cape Town University.

The trip was culturally and scientifically eye-opening for the two students, and in some ways, helped shape their career directions [see sidebar: LSAMP Alumni Success Stories]. The LSAMP-NICE program broadens the participation of underrepresented students in STEM fields by facilitating international research experiences through conferences, internships, or months-long research exchanges between laboratories.

PHOTO: COURTESY OF LSAMP

It's essential for students from underrepresented groups in STEM to see themselves as international, global researchers if they want to be competitive in today's scientific marketplace, says Bill McHenry, an LSAMP-NICE advisory board member and organic chemist at Jackson State University in Jackson, Mississippi. "Where else would you find a program that supports the participation of Native Americans, Hispanic Americans, and African Americans to do an international research endeavor? We are opening doors that were closed to these groups of students for too long."

Students should absorb the "think globally, act locally" mantra, because every modern scientific puzzle requires international collaboration among diverse scholars.

"Problems and challenges today, such as COVID-19 and climate change, are not just local; they reverberate around the globe," says Mary Benjamin, LSAMP-NICE advisory board member and vice chancellor emeritus of research, innovation and economic development at University of Arkansas at Pine Bluff. "Students can fully appreciate that global connectivity sooner through an international research experience where they interact with people who have different mindsets."

These research experiences also mean that students on their way to becoming Ph.D. holders get hands-on experience, meet the people behind specific concepts, experience culture in different parts of the world, and develop a professional relationship with researchers who will be in their networks for years to come.

"It's a great opportunity to land in a safe environment to conduct research and to explore their host country," says Benjamin Flores, advisory board member and electrical engineer at the University of Texas at El Paso. He says the exposure to different cultures, academic systems, and ways of doing science helps spur students' development and reinforces their desire to pursue research careers.

LSAMP-NICE students attend a poster competition at the King Abdullah University of Science and Technology (KAUST) in Saudi Arabia.

PHOTO: COURTESY OF LSAMP

Louis Stokes Alliances for Minority Participation (LSAMP)

The Louis Stokes Alliances for Minority Participation (LSAMP), currently represented by 57 member alliances across the United States, was authorized by the U.S. Congress in 1991. Its mission is to help diversify the nation's science, technology, engineering, and mathematics (STEM) workforce (including academia) by funding institutions of higher learning to implement evidence-based strategies for recruiting and retaining students historically underrepresented in STEM disciplines, and imbuing them with academic integration, social integration, and professionalism.

In 2018 the Louis Stokes Regional National Science Foundation (NSF) International Center of Excellence (LSAMP-NICE) received an NSF award—as one of six NSF regional centers to do so—to broaden underrepresented students' participation in international collaborative research through mentorship, research experience, internships, conferences, and other programming opportunities with an aim toward social professionalization. Funded through the NSF Broadening Participation initiative, it is a collaboration of four institutions:

- Salish Kootenai College, a tribal college in Pablo, Montana (Steve Dupuis, LSAMP-NICE PI)
- Jackson State University, a historically black university in Jackson, Mississippi (Glake Hill and Martha Mondoa-Tchounwou, co-PIs)
- Louisiana State University, a research-intensive (R1) institution in Baton Rouge, Louisiana (Zakiya S. Wilson-Kennedy, co-PI)
- The University of Texas at El Paso, a Hispanic-serving R1 institution (Denise Yates, co-PI).

LSAMP-NICE is overseen by an advisory board whose members include:

- Dario Bassani (France)
- William McHenry (Jackson State University)
- Mary Benjamin (The University of Arkansas at Pine Bluff)
- Benjamin Flores (The University of Texas at El Paso)
- John Goeltz (California State University, Monterey Bay)

LSAMP Alumni Success Stories

Three researchers who participated in the Illinois LSAMP Bridge to the Doctorate program while getting their Ph.D.s at the University of Illinois Chicago found that their experiences with international collaborators shaped their career paths and decisions.

Chemical and materials engineer Deisy Carvalho Fernandes is now a Presidential Diversity Postdoctoral Fellow at Brown University in Providence, Rhode Island. During her doctoral studies, Carvalho Fernandes spent two 3-month sessions at the University of Bordeaux in France in the laboratory of Philippe Poulin, learning how to work with gel polymers for 3D printing.

"The techniques I learned in France helped me create a diverse research portfolio to be competitive in the job market," says Carvalho Fernandes, who is currently applying for assistant professor positions. The ability to make high-tech materials out of a 3D-printed gel opened up new directions for her research into membranes, filtration, electronic devices, and biomedicine.

Carvalho Fernandes, who is Brazilian, notes that experiencing two new national and academic cultures—those of the United States and France—as a Ph.D. student has also strengthened her leadership capabilities. "My group will have people from different backgrounds with different ways of working. It will make me a more adaptable PI."

Cancer biologist and first-generation scientist Jason Garcia now works at Tempus Labs in Chicago in their biological modeling lab, a facility that tests new therapies on patient-derived organoids. He says the LSAMP program helped him in every aspect of his graduate education. "It helped me financially with classes and tutoring, boosted my confidence, and allowed me to begin research sooner," he says.

Garcia's visit to South Africa came at a time when he was deciding whether to continue in a research career or to go into teaching or policy. While on tours of labs and universities there, he saw many people of color doing research successfully, which reaffirmed his desire to continue on a research path. "It made me realize how much I enjoy contributing to cancer research and it solidified my decision to go into industry."

Postdoctoral fellow Joshua Ames says his visit to South Africa was a much-needed "maturity boost" for him as a senior doctoral student. Just as it did for Garcia, his exposure to a different scientific setting instilled confidence in him to pursue a research career.

"I didn't think I was competitive for grants and awards, but having structured mentoring and mentors on my side was transformative," says Ames, who won an NSF Graduate Research Fellowship as a graduate student. Now as a postdoctoral researcher in immunology at the University of Washington in Seattle, Ames's goal is to eventually start his own research lab at a research-intensive institution. "I may not have continued on to do a postdoc if not for the LSAMP program—it helped me maintain my love for academic science without feeling like an imposter."

The cultural and academic exchanges go both ways, too. Students are often educating their host mentors and labmates about American traditions and education systems, including how community college students can work their way up to doctoral studies. And those international colleagues are keen for collaborations with U.S. research groups. "It is really crucial for their success and ours," says Flores.

Faculty committed to developing world leaders

International researchers seek out student talent to develop; they also hope such talent will enrich the culture of their laboratory and their institution. In addition, researchers want to diversify their portfolio of research resources. "It's a small universe," says Flores. And to put it simply, he adds, "we can do better science together."

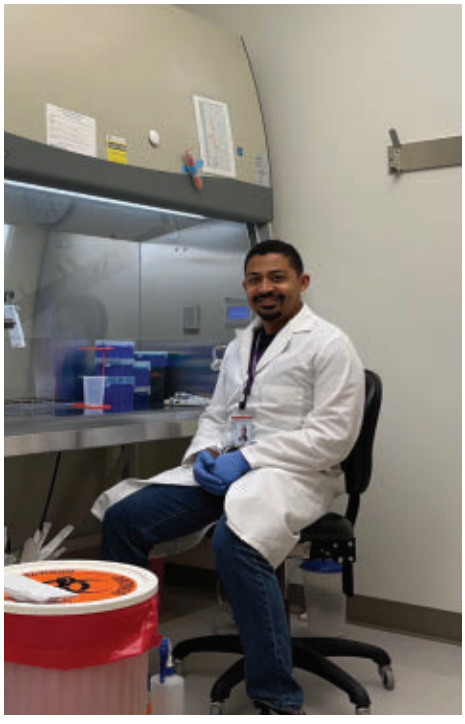
Georges Zissis says the LSAMP-NICE international exchanges help his university increase its visibility to U.S. Ph.D. students. Even though the University of Toulouse III Paul Sabatier, named for a chemistry Nobel laureate and specializing in STEM fields, is the second-largest research university in France, it is not well known internationally, says Zissis, a professor of electrical engineering and vice rector for international and European projects. And, he says, the university's values align nicely with the program's.

"As a university that values diversity and inclusion, we are always looking to attract students with different horizons, cultures, and ways of thinking," says Zissis. Hosting U.S. students in Toulouse's laboratories opens doors to future collaborations and connections. "Those students may come back for Ph.D.s or postdocs, or they will move somewhere else in the world, but we'll still have them in our network." Building expanded research networks effectively is something that French universities want to import from the United States, too, he says.

Zissis is eager to host students in his own research group in the future. He leads an artificial-lighting technology group of about 20 members with five other professors. LSAMP-NICE students have a quality that he seeks in students, called *gnac* in French slang (pronounced "nyak"), which means "driven."

"People coming from marginalized areas of society want to become someone; they have this determination and motivation, this *gnac*," he says. "If these students have the necessary knowledge and resourcefulness, then it's up to me to build excellence within them."

It's that commitment to mentoring and nurturing student excellence that defines LSAMP-NICE research faculty. "Science should be open to everyone," says Zissis, and he knows that attracting more women and students of color to electrical engineering research will improve his entire field. In his own research, he notes that the way people use color in lighting is highly dependent on their global culture and environment. "Researchers from diverse backgrounds bring societal knowledge and ideas that you will not find elsewhere," he says.



Joshua Ames **(left)** and Jason Garcia **(right)** presented their graduate work at Stellenbosch University, South Africa; Deisy Carvalho Fernandes **(center)** spent 6 months doing her doctoral disseration research at the University of Bordeaux, France.

Benjamin notes that by hosting students from diverse cultures, LSAMP-NICE mentors will end up training people who can translate and articulate the impact of their work to broader populations. Anecdotally, the public has more trust in scientific messages delivered by someone who looks like them, speaks like them, or has a shared background with them.

Expanding horizons with new partners/directions

With its established partners in France, Saudi Arabia, and South Africa, LSAMP-NICE is moving toward the next phase of its development, exploring the possibility of having joint graduate programs in which students would spend roughly half their time abroad and receive either a Master's degree or a Ph.D. from both their home U.S. university and their international host university.

LSAMP-NICE is also adding more international partners that will sponsor student exchanges, including the Brazilian and Panamanian Embassies in Washington, D.C., the Brazilian Research Foundation FAPESP (Fundação de Amparo à Pesquisa do Estado de São Paulo), and the Panamanian Research Foundation SENACYT (La Secretaría Nacional de Ciencia, Tecnología e Innovación de la República de Panamá).

"It's crucial to keep expanding our network to make sure we have additional opportunities for our students—one country may not fit everyone's needs," says Flores. There are rich opportunities for research, education, and cultural exchanges in neighboring regions to the south of the United States, he says. "Reaching out to our partners in Central and South America is really serving all of the Americas' needs."

Maharaj says that the partnership between South Africa's NRF and LSAMP-NICE grew out of their aligned aspiration for graduate students to have international research exposure.

"Research and innovation are disciplines where excellence is defined by what your global peers perceive excellence to be," she says. NRF, which is akin to the U.S. NSF, aligns with the LSAMP philosophy in other key ways, too. "Given our history in South Africa of [racial segregation and transformation], we resonate with the U.S. program—bringing in previously disadvantaged individuals is paramount in transforming both of our research cohorts."

Her colleague Sepo Hachigonta, director of strategic partnerships at NRF, says partnerships like those with LSAMP-NICE also fulfill NRF's mandate to train students. NRF encourages the Ph.D. students they fund to spend 3–18 months outside of South Africa. He says they are eager to tap into the vast LSAMP network of universities to find placements for their students in the United States.

Again he emphasizes that these are values where the two partners are perfectly aligned: "If you want to fund the best students or become the best in innovation, you cannot do it in isolation."

PHOTOS: COURTESY OF LSAMP

NEWS

> “In many ways, justice has not been served.”
>
> **Anthropologist Eben Kirksey,** in *MIT Technology Review*, noting that although He Jiankui has just been released from a Chinese prison after a 3-year term for creating gene-edited babies, his U.S. collaborators faced no punishment. Kirksey wrote a book about the case.

IN BRIEF

Edited by **Jeffrey Brainard**

CLIMATE POLICY

IPCC makes new pitch for renewables, fast climate action

The world's governments must immediately make a wholesale switch to carbon-free energy to have a shot at preventing catastrophic effects of climate change. That's the conclusion of the final section of the Intergovernmental Panel on Climate Change's *Sixth Assessment Report*, released this week, which assesses policies needed to effectively restrain global warming. The report updates previous calls by the U.N.-sanctioned panel and climate scientists for rapid action to avoid warming above 1.5°C, the threshold for catastrophic effects such as flooding and crop failures. Planned fossil fuel plants must be canceled, most existing plants must be decommissioned, spending on renewable energy must increase three- to sixfold by 2030, and politicians must back incentives for these technologies, it says. Installing new solar and wind farms and other renewables is already cheaper in many cases than building new fossil fuel power plants, the report adds. U.N. Secretary-General António Guterres called the report “a file of shame” and “litany of broken climate promises” by government and business leaders “that put us firmly on track toward an unlivable world.”

To reduce emissions, many more houses need rooftop solar panels, like these in Dallas, a report says.

PHOTO: ART WAGER/GETTY IMAGES

WHO pauses India vaccine supply

COVID-19 | The World Health Organization (WHO) last week suspended shipments through U.N. channels of a COVID-19 vaccine made in India after an inspection revealed manufacturing deficiencies. WHO said Bharat Biotech, maker of the Covaxin vaccine, which uses an inactivated virus, promised to stop exporting it to any customer until the firm addresses the problems. But the company said it will continue to sell doses from the plant for use in India. The country is the largest consumer of Covaxin, with 308 million doses administered so far. India's drug regulatory body, the Central Drugs Standard Control Organization, has not taken regulatory action or commented on WHO's move. WHO's action is significant because it authorized Covaxin's use in November 2021,

and several low-income countries have also authorized it; the vaccine is easier for them to distribute than messenger RNA vaccines because it does not need to be stored at low temperatures.

Biomedical agency lands at NIH

POLICY | President Joe Biden's new biomedical research agency for high-risk, cutting-edge research won't have the full autonomy many backers had sought. Instead, the Advanced Research Projects Agency for Health (ARPA-H) will sit within the National Institutes of Health's organizational chart—but, to promote its independence, its director will report to the NIH director's boss, Secretary of Health and Human Services Xavier Becerra, who announced that compromise in a letter to Congress last week. Many ARPA-H supporters argued it needed to be independent of NIH and its grantmaking culture, which they see as insufficiently innovative. Becerra testified at a House of Representatives hearing last week that NIH's role will be to provide administrative support, such as human resources and payroll. Becerra also said ARPA-H will not be housed on NIH's main campus in Bethesda, Maryland.

Max Planck director fired, again

LEADERSHIP | For a second time, archaeologist Nicole Boivin has been removed as director of the Max Planck Institute for the Science of Human History (MPI-SHH), following a vote last month by a governing board of the Max Planck Society (MPG). Its president first removed her in October 2021, citing evidence of bullying and scientific misconduct. Boivin, who has denied the allegations, sued, and a Berlin court reinstated her, saying the removal violated procedures. But on 25 March, MPI-SHH's Senate voted overwhelmingly to dismiss her as director, pointing to a confidential report whose summary supported the allegations. She remains a researcher at MPG. The case has drawn wide attention in Germany and from women scientists elsewhere, who noted that recent demotions at MPG have disproportionately affected women. Others said Boivin created an abusive work environment that harmed young women scholars.

Kyoto shutters primate institute

PRIMATOLOGY | One of the world's leading groups that studies primate behavior, Kyoto University's Primate Research Institute (PRI), closed last week following a scandal. A new Human Behavior Evolution Research Center is taking over the institute's facilities and animals; researchers within and outside the institute fear the scientific focus will shift from the lab-based cognitive studies and field observations that earned PRI international recognition to genetics, neuroscience, and biomedicine. The university made the move after investigations in 2020 uncovered mishandling of $9.7 million provided to build a chimpanzee enclosure at the institute's campus in Inuyama, near Nagoya. As a result, the university dismissed then-Director Tetsuro Matsuzawa, known for work documenting the cognitive abilities of captive chimps.

PHOTO: RYAN DICKIE

Workers wait for a sedated female caribou to wake up in a maternal enclosure that keeps out predators.

CONSERVATION

Protections give woodland caribou a boost

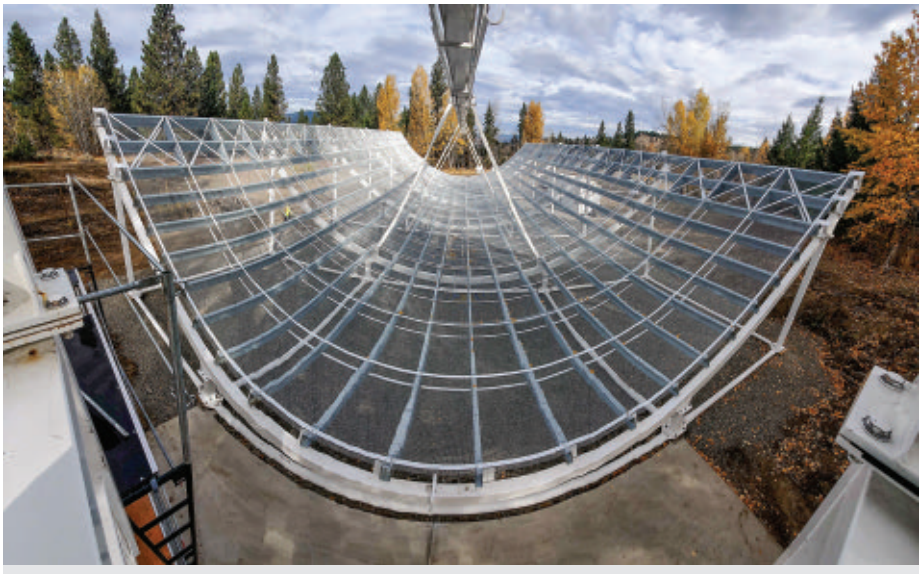
A rescue effort led by Indigenous First Nations has roughly tripled the size of a British Columbia caribou herd in less than 10 years, one of the few examples of success at reversing declines of this species. The Klinse-Za herd grew from 38 animals in 2013 to more than 100 by 2021, after wildlife officials authorized the killing of hundreds of wolves that prey on the caribou, and housed female caribou in fenced enclosures while they give birth, scientists reported last month in *Ecological Applications*. Researchers highlight the leadership of the West Moberly and Saulteau First Nations in the work and for pressuring Canadian federal and provincial governments to agree to protect 8000 square kilometers of forestland. Since 2000, nearly one-third of 38 caribou herds in southwestern Canada have disappeared, largely because logging and oil and gas exploration drove predators into caribou habitat.

Law aims at research dog breeder

RESEARCH ANIMALS | Virginia Governor Glenn Youngkin (R) this week signed into law a first-of-its-kind statute that would shut down research animal breeders that commit a single serious violation of the U.S. Animal Welfare Act (AWA). The law was prompted by complaints about mistreatment at Envigo, a contract research company that has housed more than 4000 beagles at a facility in Cumberland, Virginia; between July 2021 and last month, U.S. Department of Agriculture (USDA) inspectors documented 73 AWA violations there. Thirty-five were in the most serious categories, including a finding of more than 300 uninvestigated puppy deaths. In recent years, Envigo has supplied beagles to labs at the U.S. National Institutes of Health, the Food and Drug Administration, research universities, and pharmaceutical companies. The new law, which takes effect on 1 July 2023, prohibits organizations that breed dogs and cats for research from selling them for 2 years if they sustain a single USDA citation in the most serious categories.

Mexico harassment cases dropped

WORKPLACE | A government office in Mexico has drawn criticism for dismissing allegations of sexual harassment against a leading plant geneticist. The complaints



A new CHIME outrigger telescope has been built in British Columbia with funds from a $10 million gift.

ASTRONOMY

With upgrade, telescope to pinpoint radio bursts

A groundbreaking radio telescope is being expanded to sharpen its eyesight to find the mysterious, milliseconds-long flashes known as fast radio bursts (FRBs). They were first detected in 2007, but only a few dozen were known before the 2018 debut of the Canadian Hydrogen Intensity Mapping Experiment (CHIME). It has glimpsed more than 500 since. Although the telescope has a wide field of view that is ideal for detecting FRBs, it cannot locate them very well. Now, the management team is building three mini-CHIMEs, or outriggers, in British Columbia, California, and West Virginia. Their wide separation will allow researchers to pinpoint FRBs to a patch of sky no bigger than a coin viewed from 40 kilometers away, down from the size of the full Moon. Better accuracy helps other telescopes zoom in on the FRBs' home neighborhoods for clues to their origins.

came from four women scientists, three of whom worked with or under the supervision of Jean-Philippe Vielle Calzada of Mexico's National Laboratory of Genomics for Biodiversity (Langebio). They alleged that he touched them without their consent, pressured them to enter a romantic relationship, and retaliated professionally after they rejected him (*Science*, 1 October 2021, p. 17). (Vielle Calzada has denied the allegations.) The office, the Internal Control Organ, found in 2021 that Vielle Calzada had committed "serious" misconduct. But in recent weeks, it dismissed two of the complaints, in one case citing a procedural issue, which the complainant plans to appeal. Critics of the office's move have vowed to press for giving Langebio and its parent institution, the Center for Research and Advanced Studies, authority to sanction its researchers for harassment. From 2016 to 2018, only 1% of 399 cases of sexual harassment reported in Mexico's federal institutions led to a sanction for the accused harasser.

Trump plug increases vaccination

COVID-19 | An online advertisement created by political scientists and economists that featured former President Donald Trump recommending COVID-19 shots led to increased uptake of the vaccines in U.S. counties that had low vaccination rates, an analysis has concluded. COVID-19 vaccine hesitancy is higher in U.S. regions that voted heavily for Trump in the 2020 election, so the research team targeted them by creating a 30-second YouTube ad that featured a Fox News TV interview in which Trump recommends the vaccine. The team spent nearly $100,000 on Google Ads to place it online in 1083 U.S. counties in which fewer than 50% of adults were vaccinated; an additional 1085 similar counties that did not receive the ads served as a control group. Compared with control counties, the study found an increase of 104,036 people receiving first vaccinations in areas that observed the ad, a statistically significant difference. The intervention's cost was just under $1 per vaccinated person. In contrast, U.S. locales that used lottery tickets as a reward spent $60 to $80 per vaccination, according to the preprint study posted at the National Bureau of Economic Research.

Biologist quits over #MeToo ruling

WORKPLACE | David Sabatini, the high-profile biologist forced out of the Whitehead Institute in 2021 after a probe found he violated its sexual harassment policies, has resigned his professorship at the Massachusetts Institute of Technology after three senior MIT officials recommended revoking his tenure. They found that Sabatini violated MIT policy by engaging in a consensual sexual relationship with a person over whom he held a career-influencing role. In an emailed statement, Sabatini, who codiscovered a key mammalian signaling pathway, called the outcome "out of all proportion to the actual, underlying facts. I look forward to setting the record straight and standing up for my integrity." Nancy Hopkins, an emeritus professor of biology who helped lead a landmark push for gender equality on the MIT faculty in the 1990s, called his resignation "a milestone," noting in an email, "A young woman had the courage to demand that the rules be enforced. And she was heard."

Cancer institute head steps down

BIOMEDICINE | Norman "Ned" Sharpless, director of the U.S. National Cancer Institute (NCI) since 2017, will step down at the end of April. He said he will spend time with his family in North Carolina before deciding what's next, but could return to academia. Sharpless, who studies aging and cancer, was appointed to lead NCI, which has a $6.9 billion budget, by former President Donald Trump after directing the University of North Carolina's Lineberger Comprehensive Cancer Center. In 2019, Sharpless served a 7-month stint as acting chief of the Food and Drug Administration, then returned to NCI. There he highlighted the harm caused by missed cancer screenings during the pandemic and launched efforts to improve immunotherapies, share data on pediatric tumors, and improve diversity in the cancer research workforce. He also worked to find funding that has enabled NCI to raise its grant success rates, which have been the lowest of the National Institutes of Health's largest institutes because of soaring applications.

SCIENCE.ORG/NEWS
Read more news from *Science* online.

PHOTO: JANE KACZMAREK

eppendorf & Science

PRIZE FOR NEURO BIOLOGY

2021 Winner
Amber L. Alhadeff, Ph.D.
Monell Chemical Senses Center, USA
For research on the gut-brain control of hunger circuits

Tell the World About Your Work!

Application Deadline
June 15, 2022

Eppendorf & Science Prize for Neurobiology
The annual Eppendorf & Science Prize for Neurobiology is an international prize which honors young scientists for their outstanding contributions to neurobiological research based on methods of molecular and cell biology. The winner and finalists are selected by a committee of independent scientists, chaired by *Science's* Senior Editor, Dr. Peter Stern. If you are 35 years of age or younger and doing great research, now is the time to apply for this prize.

As the Grand Prize Winner, you could be next to receive

> Prize money of US$25,000
> Publication of your work in *Science*
> Full support to attend the Prize Ceremony held in conjunction with the Annual Meeting of the Society for Neuroscience in the USA
> 10-year AAAS membership and online subscription to *Science*
> Complimentary products worth US$1,000 from Eppendorf
> An invitation to visit Eppendorf in Hamburg, Germany

It's easy to apply! Write a 1,000-word essay and tell the world about your work. Learn more at:

eppendorf.com/prize

eppendorf Science AAAS

IN DEPTH

A COVID-19 vaccine made by BioNet-Asia in Ayutthaya, Thailand, should be cheaper than the two messenger RNA vaccines used in richer countries.

COVID-19

New crop of mRNA vaccines aim for accessibility

If approved, they could bring the pandemic's star vaccine technology to more of the world

By **Jon Cohen**

The two COVID-19 vaccines based on messenger RNA (mRNA) have been the breakout stars of the pandemic. Both trigger impressive immune responses with minimal side effects, and both did exceptionally well in efficacy trials. But the vaccines, produced by the Pfizer-BioNTech partnership and Moderna, have also split the world. Because of their high prices and their need to be stored at extremely low temperatures, few people in lower and middle-income countries have had access to them.

That might soon change. More than a dozen new mRNA vaccines from nine countries are now advancing in clinical studies, including one from China that's already in a phase 3 trial. Some are easier to store, and many would be cheaper. Showing they work won't be easy: The number of people who don't already have some immunity to COVID-19 because of vaccination or infection is dwindling. But if one or more of the candidates gets the green light, the mRNA revolution could reach many more people.

The Pfizer-BioNTech and Moderna shots rely on mRNA to direct cells to produce spike, a protein on SARS-CoV-2's surface. Although 23 COVID-19 vaccines are in use around the world, based on technologies including inactivated SARS-CoV-2 and cold viruses engineered to carry the spike gene, the two mRNA vaccines account for about 30% of the 13.2 billion doses produced so far, according to health care data company Airfinity. But the companies have been reluctant to share their intellectual property (IP) and know-how, which would allow manufacturers in poorer countries to produce the shots.

Instead, BioNTech and Moderna each recently announced plans to build their own plants in African countries. In a separate effort, the World Health Organization has created a training hub for mRNA vaccines that will teach scientists from low- and middle-income countries how to build and run their own plants. But it may take years before these efforts bear fruit.

The candidates already under development could reach the marketplace much faster. IP protections are still a challenge, says Melanie Saville, who heads vaccine R&D at the Coalition for Epidemic Preparedness Innovations: "Who can do what and where is going to be a critical question." But the new mRNA developers have managed to dodge some of the showstoppers.

Furthest along is a vaccine made by Walvax Biotechnology in Kunming, China, together with Suzhou Abogen Biosciences and the Chinese Academy of Military Science. Details are hard to come by and Walvax did not respond to detailed questions from *Science*, but a paper about a phase 1 trial, published in *The Lancet Microbe* in January, offers some information. Instead of using mRNA that encodes the entire spike protein, the Walvax team only included the sequence of a key portion known as the receptor binding domain. In July 2021, the company launched a placebo-controlled phase 3 trial in 28,000 people in Mexico, Indonesia, Nepal, and China.

A key advantage is that Walvax's product can be kept in a standard refrigerator, says Víctor Bohórquez López, a clinician who leads trials at five sites in Mexico for Red OSMO, a network based in Oaxaca. A

PHOTO: ADAM DEAN

company official told Reuters in January that Walvax can produce 400 million doses a year.

In Thailand, a team lead by Kiat Ruxrungtham at Chulalongkorn University has developed an mRNA vaccine, produced by the French-Thai company BioNet-Asia, that has completed phase 1/2 studies. The team followed a key step in the playbook used by the Pfizer-BioNTech collaboration and Moderna: replacing uridine—one of the four basic building blocks of RNA—with methylpseudouridine, a substitution that reduces the toxicity of mRNA and increases the amount of spike protein cells produce. The substitution is "the most important thing that people have done with mRNA vaccines," says Philip Krause, a former top vaccine official at the U.S. Food and Drug Administration (FDA). BioNet-Asia can use the replacement for free because the company that licensed the technology from the University of Pennsylvania, where it was invented, has not sought protection in Southeast Asia.

The vaccine differs from the marketed ones in other ways, however. Kiat's team did not introduce two mutations in spike that stabilize the protein, which would have required an expensive IP license. They avoided another licensing issue by having the code direct cells to secrete the spike protein, rather than leaving it bound to the membrane. Some comparative studies have found this leads to a weaker immune response, but Kiat's mouse studies saw no difference, and human data show the vaccine triggers robust levels of antibodies that can neutralize the virus, he says.

BioNet-Asia can make up to 100 million doses a year, Kiat says, at a lower price than the Pfizer-BioNTech collaboration and Moderna. Japan's Daiichi Sankyo and Canada's Providence Therapeutics have mRNA vaccines at similar stages of development.

About half of the new candidates are "self-amplifying": They include harmless genes from an alphavirus that code for an enzyme used in RNA replication, enabling the spike mRNA to make additional copies of itself. Each dose can get by with less mRNA, which could make it easier to vaccinate more people. A downside is that self-amplifying mRNA vaccines can't use the methylpseudouridine substitution—they need the natural uridine to replicate.

A phase 1 study of a self-amplifying vaccine developed at Imperial College London triggered such mediocre immune responses that the researchers went back to the drawing board. But a similar candidate from GlaxoSmithKline solidly protected hamsters against SARS-CoV-2 infection, a January paper in *Molecular Therapy* showed. That vaccine is now being tested in a 10-person phase 1 trial.

Showing that the new vaccines work in humans presents formidable challenges. "I'm in trouble because I can't find the population right now for the phase 3 trial," Kiat says. Not only is it becoming more difficult to find people who have no immunity at all against SARS-CoV-2, but enrolling participants in a placebo-controlled study is increasingly ethically fraught, because proven COVID-19 vaccines are now widely available. Producers of self-amplifying vaccines in India and Vietnam instead plan to compare the vaccines with others already in use.

> **"We can make the price lower than the Big Pharmas."**
> **Kiat Ruxrungtham**, Chulalongkorn University

Kiat hopes to judge his candidate based on a proxy measure: how well it boosts antibody levels in people who are fully vaccinated. Past studies of the marketed mRNA vaccines have shown that specific levels of neutralizing antibodies are correlated with protection from disease, and BioNet-Asia and other manufacturers hope regulators will accept similar data to authorize use of their vaccines. The European Medicines Agency and regulators from several countries have indicated they will accept such "immunobridging" data in some circumstances, Krause says. FDA has yet to issue guidelines. "I know from talking to people at FDA that they are reluctant" to rely on antibody data, says Stanley Plotkin, a veteran vaccine researcher who consults with Moderna and many other companies.

One problem is that antibodies are only part of the immune response triggered by mRNA vaccines. T cells—which are more difficult to measure—play a role in preventing severe disease by eliminating infected cells. They also offer better protection against new virus variants than antibodies and help ensure the durability of immunity. Still, Plotkin and others say, antibody levels are good enough surrogates to issue emergency use authorizations. For full approval, they say, vaccines will have to prove effective in real-world studies.

"We know that there are a lot of hurdles ahead," Kiat says. But even if their COVID-19 vaccine fails, his team is building capacity for the future, he says. "We can now manufacture new mRNA vaccines very quickly, so that's a way to solve the next pandemic—and we can make the price lower than the Big Pharmas." ■

U.S. FUNDING

Congress restores earmarking—but with limits

One legislator simulates peer review to make her selection process more rigorous

By **Jeffrey Mervis**

The $1.5 trillion spending bill enacted last month did more than fund U.S. government operations for the next 6 months. It also revived congressional earmarking—the controversial practice of allowing legislators, often at the behest of powerful constituents, to allocate money for specific projects in their district or state that federal agencies did not request.

Earmarks, such as a new bridge or refurbished airport, have traditionally given lawmakers a reason to vote for legislation they otherwise might not support, making the wheels of Congress turn more easily. But the U.S. higher education community is deeply divided over the practice.

Many academic institutions have sought—and won—earmarks, seeing them as a quick and easy route to growing their research capacity. At the same time, the higher education organizations to which they belong have long argued that scarce federal dollars should be allocated based on peer review rather than the whims of a single powerful legislator.

A crescendo of costly projects of dubious merit led Congress to ban earmarking in 2010. But the itch never went away. And last year, the Democratic majority in both chambers of Congress adopted new rules that require earmark requests to be posted online, limit eligibility to nonprofit organizations and projects in which the legislator has no personal or financial interest, and cap the total spending on earmarks at 1% of overall discretionary spending.

Lawmakers welcomed their return, inserting more than 4000 projects totaling $9 billion into this year's spending bill. Research-related activities make up about 10% of both totals, according to an analysis by AAAS (which publishes *Science*). Re-

tiring Senator Richard Shelby (R–AL), a master earmarker who topped this year's list with some $548 million in home-state projects, added a novel twist to the traditional funding for academic bricks and mortar with a $50 million endowment at the University of Alabama to attract and retain world-class faculty in the sciences.

The new rules haven't won over opponents. A spokesperson for the Association of American Universities, for example, says it stands behind a 2018 statement that declares "should Congress restore earmarks, AAU respectfully urges that competitive peer-review continue to be the primary method for allocating federal research funding."

Representative Chrissy Houlahan (D–PA, center) won an earmark for a program to draw students into science.

But AAU and other higher education organizations that oppose earmarks acknowledge their appeal. Jeff Lieberson of the Association of Public and Land-grant Universities says, "APLU's focus is on programmatic requests," referring to its traditional advocacy for more federal spending on certain activities or for an entire agency rather than for a specific project. "But we understand member institutions [also] may seek congressionally directed spending consistent with the rules of Congress."

The new rules prompted at least one legislator to choose her earmarks in a way meant to address some of the flaws in the old system. Earmarks should represent the "highest and best use" of federal dollars, says Representative Chrissy Houlahan (D–PA), who won a seat in Congress in 2018 touting her expertise as a scientist—she's an industrial engineer with a master's degree from the Massachusetts Institute of Technology—educator, and serial entrepreneur. To meet that goal, Houlahan created a process that parallels how government agencies like the National Institutes of Health (NIH) and the National Science Foundation assess the merit of grant proposals.

Traditionally, those seeking earmarks might hire a lobbyist to make their case—or go directly to a lawmaker. But Houlahan required any group seeking an earmark to submit a written proposal, complete with a budget justification and outside letters of recommendation. She chose nine community leaders—given anonymity to ensure they would not be the target of lobbying—to score each request using criteria such as the project's potential benefits for the regional economy and public health and safety, as well as whether it would improve equity.

The panel met several times to discuss proposals in the top half of the rankings, much as an NIH study section would, and agreed that roughly one-third of the 53 requests were worthy of funding. (The losers were given tips on how to improve their proposals and encouraged to reapply next year, again mirroring the federal process.) Houlahan then chose 10—the maximum number allowed each House of Representatives member—to be considered by congressional appropriators.

Houlahan won approval for all but one project, totaling $6.2 million. The biggest payout, at $1.5 million, went to Albright College, a small liberal arts school in Reading, Pennsylvania, to expand an afterhours and summer program that draws middle and high school students into science by encouraging them to find real-world applications for what they're learning.

"It checked all of her boxes," says a Houlahan staffer about the program, called the Science Research Institute (SRI). Houlahan was especially impressed by the program's track record of serving low-income students, minorities, and those with disabilities, as well as the fact that several older students have developed technologies they are hoping to patent.

Adelle Schade, a high school biology teacher, began SRI in 2014 to supplement classroom science instruction at her school. Operating on a shoestring budget, Schade secured donations from area hospitals and medical supply companies to outfit labs with professional-grade equipment suitable for student research projects.

In 2020, Albright College acquired SRI, which has served 6000 students since its inception, and hired Schade as dean of precollege and summer programs. Albright's goal is to further expand the program and perhaps export the model to other localities.

The institute's emphasis on tackling real-world problems appealed to Albright's president, Jacquelyn Fetrow, a biochemist who founded a bioinformatics company early in her academic career. Besides getting middle and high school students excited about science, Fetrow believes SRI can help the college produce graduates with the technical skills and business savvy to revitalize the local economy, which has been shedding manufacturing and retail jobs for decades.

Seeking an earmark was the only way a small college that emphasizes teaching over research could attract federal dollars to realize SRI's potential, she notes. "We can't follow the traditional route of bringing in superstar faculty who win hundreds of millions in federal grants," says Fetrow, who built her career at large research universities before coming to Albright in 2017.

Science doesn't know of other lawmakers who followed Houlahan's path in selecting earmarks this year. And those who decry the practice are still assessing its impact on agency budgets. The 1% cap removes some of the foul odor emanating from earmarks, says one higher education lobbyist, before adding, "But we're going to watch closely to see if they start to get out of control again." ■

PHOTO: BELL MEDIA GROUP

PHOTO: REUTERS/GONZALO FUENTES

The reelection of French President Emmanuel Macron could see universities gain more control over national research laboratories.

EUROPE

French election could buoy president's R&D overhaul

Macron vows reforms of national research bodies

By **Elisabeth Pain**

In a repeat of 2017, centrist French President Emmanuel Macron and far-right nationalist Marine Le Pen are the leading contenders in presidential elections on 10 April. Academics, who are generally left leaning, dislike Le Pen for her anti-immigration and isolationist views. But many scientists are also uneasy with Macron, because a second term would let him pursue controversial efforts to strengthen universities at the expense of national research organizations like CNRS and INSERM.

Macron views universities as more nimble and innovative than the national bodies, which are still the backbone of research in France. The "potential danger" is that the organizations will lose autonomy and become subservient to universities, says Patrick Monfort, a microbial ecologist at the University of Montpellier and member of a researcher trade union. He worries about Macron's vision that "to improve the efficiency of universities, we must give them all the resources of the research organizations."

The latest opinion polls put Macron at 27% of the vote, versus 21% for Le Pen. Far-left candidate Jean-Luc Mélenchon is polling at 15%. If no one wins an absolute majority on 10 April, the two leaders would face each other in a runoff 2 weeks later. The traditional parties are lagging behind, as in 2017. Conservative candidate Valérie Pécresse, a former research minister, is polling at about 10%, whereas socialist candidate Anne Hidalgo has struggled to get any notice.

Although Le Pen's popularity is rising, many pundits expect Macron to win. That would allow him to make good on plans laid out in a 2020 science bill, which promises to raise public research spending from about €15 billion per year to €20 billion by 2030, aiming for 3% of gross domestic product. Researchers welcome that goal but point out that even then, R&D spending would fall short of that of competitors such as Germany. "There has been a catch-up effort, but it is not sufficient," says Manuel Tunon de Lara, president of France Universités, an association of 74 universities.

The law also launched a battery of measures to make French science more competitive. In its first year of implementation, the law increased funding at the National Research Agency, allowing it to raise the success rate for competitive grant applications to 23%, compared with 17% in 2020. University professors and permanent researchers are now guaranteed a salary of at least €3200 per month, twice the minimum wage. The law also created nearly 100 "junior professor" positions resembling tenure-track posts elsewhere. Trade unions criticized the new positions as an attack on France's tradition of offering permanent jobs even at entry level.

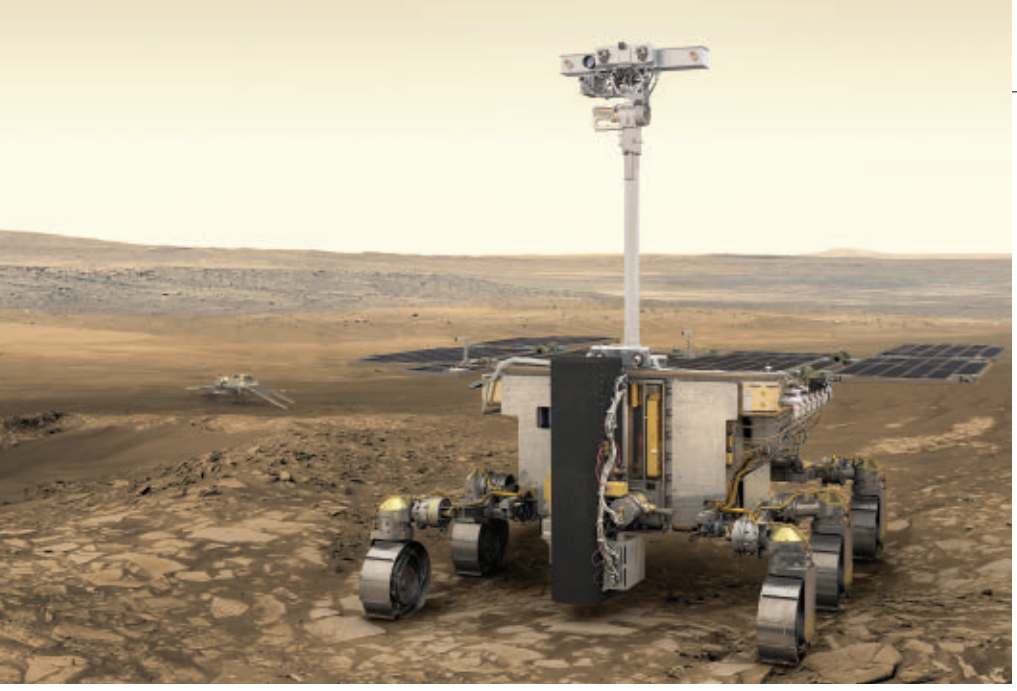
Macron's efforts to reorganize French science around elite universities have sparked more unease. In March, Macron earmarked €300 million per year to support education, research, and innovation at 17 university-led alliances, the first of which were launched in 2011. Macron credits them with raising the international profile of select French universities and boosting France's success at winning grants from the European Research Council. But Bruno Andreotti, a physicist at Paris City University and member of the radical researcher collective RogueESR, says the initiative has led to a culture of the have and the have-nots. "There is this fantasy … to have 10 cutting-edge research universities and the … others, to abandon them."

Even more controversial changes could come with Macron's reelection. At his first press conference as a presidential candidate on 17 March, Macron declared he would "make [universities] fully fledged research performers." This would require giving them "full autonomy and go all the way through on reforms initiated a decade ago."

He was alluding to a long-standing effort to reform a peculiarity of the French research system: the so-called mixed research units, which bring together researchers from both universities and the national research organizations. At a university congress in January, Macron indicated the research organizations should merely provide support for universities. Tunon de Lara, who supports Macron's proposal, says the duplicated and ill-defined roles between universities and research organizations are hampering efficiency. "This is an orchestra with different instruments—it cannot be a cacophony," he says. If "we have an ambition of international competition … then everyone must be able to play their role."

But Sabrina Speich, a climate scientist at the École Normale Supérieure who works in a mixed research unit, says they are "an incredible strength" of the French system, one that allows for a richness of both research and training. Previous attempts by Pécresse to reform the mixed research units failed. In campaign pledges, Mélenchon said he would maintain the current roles of universities and research organizations. Le Pen has not offered a position.

Macron's announcements have already met some backlash. Patrick Flandrin, a CNRS physicist and president of the French Academy of Sciences, says universities and research organizations must continue to "cohabitate." The real problem is chronic underfunding, Monfort says. Universities "are deluding themselves into believing that they would have more resources" if they had more control over the mixed research units, Monfort says. They would just have "more power … with a constant budget." ■



A European rover was set to go to Mars with a Russian landing platform—until the invasion of Ukraine.

PLANETARY SCIENCE

Europe tries to save Mars rover after split with Russia

Team leader says delay to 2028 or 2030 is likely

By **Daniel Clery**

After a technical review, the European Space Agency (ESA) confirmed last week that its Mars rover was ready for launch. The only problem: The rover has neither a ride to the Red Planet nor a landing craft to get it safely to the surface. Russia was supposed to provide both, but ESA suspended ties and canceled a planned launch in September after the country invaded Ukraine. "There was no real alternative," says ExoMars team leader Thierry Blancquaert of ESA's technology center in the Netherlands.

Now, ESA is studying options to keep the €1 billion mission alive. Even if the agency can replace the Russian technologies—and pay for them—a delay to 2028 or even 2030 is likely, Blancquaert says. Planetary scientists say the rover will be worth the wait. "The mission will still be cutting edge even for potentially later launch windows in the next decade," says Andrew Coates of University College London, principal investigator of the rover's panoramic camera.

ExoMars has not had an easy gestation. It was originally an ESA-NASA collaboration, but the United States pulled out in 2012 for budgetary reasons. Russia stepped in to provide a Proton rocket for its launch, and a landing vehicle called Kazachok. Russian scientists also provided most of the instruments on Kazachok, which will form a sensor station after landing. And Russia contributed two of the nine instruments for the mission's centerpiece: a rover the size of a golf cart, named Rosalind Franklin after the British DNA pioneer. It can drill 2 meters below the surface in search of pristine samples that may yield evidence of past life.

The change in parentage caused the launch date to slip from 2018 to 2020. (Launch opportunities for Mars occur roughly every 2 years when the planets align.) But in early 2020, difficulties with the parachutes designed to slow descent into the martian atmosphere led to another 2-year delay.

This time around, "All the hardware was ready to start the launch campaign," Blancquaert says. The rover and Kazachok were ready to be shipped to Russia's spaceport in Baikonur, Kazakhstan, when Russia began its war with Ukraine. "As soon as we saw this Iron Curtain descend again, we thought: What can we do to save ExoMars?" Blancquaert says.

ESA is embarking on a 3-month study to assess what's possible. If relations with Russia are restored swiftly, a 2024 launch is still possible, Blancquaert says, but, "If we need to change hardware, there's no way we would be ready." As the war drags on in Ukraine, a quick rapprochement looks increasingly unlikely.

In that case, ESA will have to replace several critical technologies. One is radioisotope heating units (RHUs), small capsules of radioactive plutonium-238 that keep the rover warm during the frigid martian night. NASA has provided RHUs in the past, and the United Kingdom could develop some later this decade, but there are no other European providers. ESA has RHUs leftover from the development of its Huygens Probe, which dropped to the surface of Titan in 2005. But because they put out less power than the Russian ones, more would have to be packed aboard, possibly leading to the ouster of an instrument, Blancquaert says.

Another technology Russia was to provide is retrorockets, used to take over from parachutes in the final stages of descent. ESA tested some simple retrorockets in 2016 with its Schiaparelli lander, which crashed in its final approach to Mars because of a software error. "In Europe we have not yet matured this technology," Blancquaert says.

A 2026 launch date might be possible if ESA had help. NASA has said it is in discussions with ESA to see what it could provide. "NASA doesn't have a billion dollars to build ESA a lander," says aerospace engineer Zachary Putnam of the University of Illinois, Urbana-Champaign, who has studied Mars landing systems for NASA. But if ESA was just looking for RHUs and retrorockets, U.S. suppliers could provide them, he says. If ESA must go it alone, however, 2028 looks like the earliest possible launch date, Blancquaert says: "It gives us more time to finalize European technology."

ESA is relatively comfortable with other lander elements trialed by Schiaparelli, such as the heat shield, parachutes, and navigation system. And the agency also has a ready launch solution: its new Ariane 6 rocket, which may fly before the end of this year.

But such rockets aren't cheap, and adapting ExoMars for a new launcher will cost money. Keeping the spacecraft fit to fly and employing research teams during the delay also add to the expense. When the launch of NASA's InSight Mars lander was delayed, it cost the agency about $150 million per year to keep it on ice, Putnam says. On top of that, developing a new lander for ExoMars "could run to as much money as the rover itself," he says.

Meanwhile, ESA is planning another rover and an orbiter, due for launch in 2027, that are part of a $7 billion joint effort with NASA to bring martian rocks back to Earth. ExoMars team members will be eyeing a November meeting of the ESA council, when budgets are set and the troubled mission will be weighed against other Mars plans. ■

IMAGE: ESA/ATG MEDIALAB

PARTICLE PHYSICS

Particle's mass may tip the scales to new physics

New estimate of *W* boson mass conflicts with prediction from "standard model"

By **Adrian Cho**

Particle physicists may have finally poked a hole in their understanding of the subatomic realm—which they would relish. A new look at old data suggests an ephemeral particle called the *W* boson is heavier than predicted by physicists' "standard model" of particles and forces. The discrepancy could hint at particles not included in the 40-year-old theory, says Doreen Wackeroth, a theorist at the University at Buffalo who was not involved in the work. "I'm very excited about the result!"

But the finding, reported today in *Science* (p. 170), also clashes with previous measurements, giving some physicists pause. "All these measurements claim to measure the same quantity," says Martin Grünewald, an experimental physicist at University College Dublin. "Somebody must be, I will not say wrong, but maybe made a mistake or pushed the error evaluation too aggressively."

Vexingly successful, the standard model was completed in 2012, when the world's largest atom smasher, the Large Hadron Collider (LHC) at the European particle physics laboratory CERN, discovered its last missing piece, the long-predicted Higgs boson. The theory accounts for every particle interaction seen so far, but it suffers obvious deficiencies. It includes three forces—electromagnetic, strong, and weak—but leaves out gravity. It also contains no dark matter, the invisible stuff that makes up 85% of the universe's matter.

Now that all the standard model particles are known, physicists can test the theory's internal consistency, because each particle's properties depend on those of others. For example, the mass of the *W* boson—which conveys the weak nuclear force just as the photon conveys the electromagnetic force—depends on those of the Higgs and a heavy but fleeting subatomic particle called the top quark. So, from those input measurements, physicists can predict the W's mass and look for a discrepancy with the measured value.

The measurement is tricky. Created in a high energy particle collision, a W quickly decays into either an electron or its heavier cousin, a particle called a muon, and an antineutrino. The antineutrino cannot be detected, so physicists must deduce its presence by summing up the momenta and energies of all the other particles spewing from each collision and looking for events in which something unseen seems to fly out the side of the cylindrical detector. From the energy and momentum of the decay particles, analyzed statistically over many events, they can estimate the W's mass.

Now, one team says its reading conflicts with the standard model prediction. The data come from the Collider Detector at Fermi National Accelerator Laboratory (CDF), a particle detector fed by the Tevatron collider, which ran at Fermilab from 1984 until 2011. After a decade of work, Ashutosh Kotwal, a particle physicist at Duke University, and his 397 CDF collaborators find the *W* boson has a mass of 80,443.5 megaelectron volts—86 times that of a proton. The measurement differs from the predicted mass by seven times the experimental uncertainty.

"What does it mean? That's the next big question," Wackeroth says. Physicists have spotted a couple of other small anomalies that suggest the standard model may finally be cracking, she says. For example, she notes that the muon appears to be slightly more magnetic than predicted (*Science*, 9 April 2021, p. 113).

However, earlier measurements of the W's mass generally agreed with the standard model (see chart, below). The new result even contradicts the CDF's previous result, published in 2012, which was based on the first quarter of the current data set, notes Dmitri Denisov, a physicist at Brookhaven National Laboratory who worked on D0, a rival Tevatron detector. "That's my first concern," he says.

But CDF researchers made several improvements in the analysis that account for the difference, Kotwal says. "We are confident in the techniques we have used," he says. "It is a distinct possibility that there is something new in nature that the standard model does not capture."

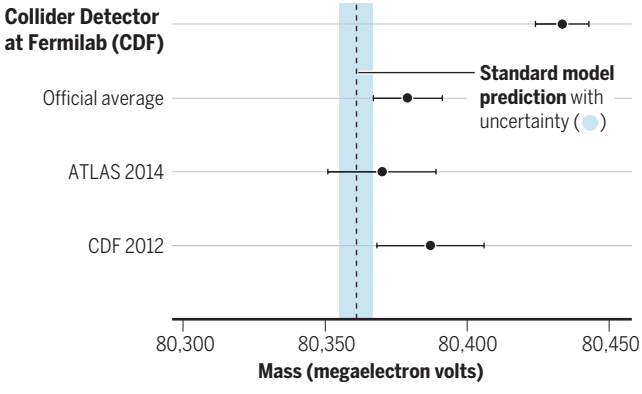
Physicists should soon get yet another *W* boson mass measurement. Scientists with the Compact Muon Solenoid, a detector at the LHC, hope to publish one early next year, says Guillelmo Gomez-Ceballos, a CMS physicist at the Massachusetts Institute of Technology. He is also a CDF member, and although he didn't work on the new study, he says, "I don't remember any analysis that has been done with so much care."

It may take years to reconcile the measurements. But physicists won't be left rudderless in the meantime. Since 1957, the Particle Data Group (PDG) at Lawrence Berkeley National Laboratory (LBNL) has maintained a compendium of particles and arbitrated disputes over their measured properties. The new *W* boson mass value comes as the PDG is preparing its latest annual update, says Michael Barnett, a retired LBNL physicist who led the PDG from 1990 to 2015 and still works on it. "We're going to have to stop the presses, just like we did when the Higgs was discovered," he says.

For a parameter like the *W* boson's mass, the PDG averages the most current and reliable measurements. If they disagree far beyond their uncertainties, the group applies a specific mathematical algorithm that effectively widens the error bars to encompass the discordant individual results, Barnett says. Ironically, even though the CDF has now reported the single most precise measurement of the W mass, the official value will likely become even less certain than before. ■

Weighty issue

A new measurement of the mass of a particle called the *W* boson disagrees strongly with the theoretical prediction—and with previous measurements, including one from the same group.



CREDITS: (GRAPHIC) C. BICKEL/*SCIENCE*; (DATA) P.A. ZYLA *ET AL.*, *PROG. THEOR. EXP. PHYS.* 2020, 083C01 (2020); T. AALTONEN *ET AL.*, *PHY. REV. LETT.* 108, 151803 (2012); M. AABOUD *ET AL.*, *EUR. PHYS. J.* C78, 100 (2018); T. AALTONEN *ET AL.*, *SCIENCE*, 376, 6589 (2022)

FEATURES

TAKING A SHOT AT CANCER

Preventive vaccines that wipe out nascent tumors are now being tested in healthy people

By **Jocelyn Kaiser**

When Dave Dubin learned at age 29 that he had colon cancer, it wasn't a big surprise. His grandfather and father had both survived the disease. "It was almost the Dubin way, and we just went on," Dubin says. He had surgery and chemotherapy, but his cancer came back 10 years later. Genetic testing finally found an explanation for his family's trials: a mutation in a DNA repair gene that lets genetic errors pile up in dividing cells. The disease, Lynch syndrome, comes with up to a 70% lifetime risk of cancer.

Dubin, 55, gets annual colonoscopies, endoscopies, and imaging scans, which caught a third cancer, in his kidney. His eldest son, Zach Dubin, 26, inherited the DNA repair mutation and also regularly gets checked for cancer. "It's no fun. Nobody enjoys it," Dave Dubin says—not the 2-day colonoscopy prep and procedure, nor the worrying about possible tumors. The disease also turned him into an activist. He and his family in Haworth, New Jersey,

PHOTO: KRISTON JAE BETHEL

Zach Dubin (left) and Dave Dubin hope for vaccines that could prevent cancer in families like theirs.

launched a nonprofit, AliveAndKickn, to promote research and awareness of Lynch syndrome, which affects an estimated 1.1 million people in the United States.

"There is a lot of anxiety in this patient population," says oncologist and geneticist Eduardo Vilar-Sanchez of the MD Anderson Cancer Center. "It is a big psychological burden." In hopes of easing that strain, Vilar-Sanchez will soon lead a clinical trial of a vaccine to prevent or at least delay Lynch-related cancers. If it works, Dave Dubin says, "it could be huge."

Vaccines to prevent certain types of cancer already exist. They target viruses: hepatitis B virus, which can trigger liver cancer, and human papillomavirus, which causes cervical and some other cancers. But most cancers are not caused by viruses. The Lynch vaccine trial will be one of the first clinical tests of a vaccine to prevent nonviral cancers.

The idea is to deliver into the body bits of proteins, or antigens, from cancer cells to stimulate the immune system to attack any incipient tumors. The concept isn't new, and it has faced skepticism. A decade ago, a *Nature* editorial dismissed a prominent breast cancer advocacy group's goal of developing a preventive vaccine by 2020 as "misguided," in part because of the genetic complexity of tumors. The editorial called the goal an "objective that science cannot yet deliver." But now, a few teams—including one funded by the same advocacy group, the National Breast Cancer Coalition (NBCC)—are poised to test preventive vaccines, in some cases in healthy people at high genetic risk for breast and other cancers. Their efforts have been propelled by new insights into the genetic changes in early cancers, along with the recognition that because even nascent tumors can suppress the immune system, the vaccines should work best in healthy people who have never had cancer.

> **"We're inspired because the impact will be massive."**
>
> **Robert Vonderheide**, Penn Medicine

Researchers are trying out several vaccine strategies. Some use so-called tumor antigens, molecular markers that are scarce on healthy cells but plentiful on cancer cells. The Lynch vaccine instead targets "neoantigens," a potent type of antigen only found on tumor cells. Some deploy just a single antigen whereas others use a large number, in a bid to broadly shield against cancer. The best approach is unclear, and developers also face the difficult challenge of measuring success without waiting decades for healthy people to develop cancers.

Early trials are yielding glimmers of promise. If the idea works to prevent one or a few cancers, it could be extended to meet an ambitious goal suggested by President Joe Biden: developing a vaccine that could prevent many types of cancer, modeled on the messenger RNA (mRNA) vaccines that have helped fight the COVID-19 pandemic. "We are a long way from a general vaccine" to prevent cancer, says medical oncologist Shizuko Sei of the National Cancer Institute's Division of Cancer Prevention. "But it could be in the distant future. It's a stepwise approach."

EFFORTS TO HARNESS the immune system to fight cancer have a long history. In the 1890s, physician William Coley reported that injections of bacterial toxins—a vaccine of sorts—sometimes shrank patients' tumors, apparently by stimulating the immune system. Decades later, researchers discovered that immune cells called T cells could recognize tumor antigens as foreign and attack cancers. This finding led to two classes of approved therapies: drugs that lift molecular brakes on T cells so they can intensify their anticancer attack, and T cells engineered to home in on cancer cells. Both kinds of treatment have had striking success against certain cancers.

A third type of immunotherapy, vaccines to treat cancer, has lagged. Efforts took off in the early 1990s, when researchers began to tally dozens of tumor antigens that might rouse a patient's immune defenses. Often these antigens are proteins that cancer cells use to grow or spread, so the antigens are good markers of cancer cells.

But despite promising data from animal experiments, most treatment vaccines failed to halt tumor growth in people. Because tumor-associated antigens can also be present in scant quantities on normal cells, the immune system tends to ignore them. The chemotherapy or other harsh treatments cancer patients receive also weaken their immune response, and tumors are protected by their "microenvironment"—surrounding cells and molecules that suppress killer T cells and block them from entering tumors. The only approved treatment vaccine, for advanced prostate cancer, extends life by just 4 months.

Some scientists thought cancer vaccines might work better to prevent rather than treat the disease. One proponent was University of Pittsburgh cancer immunologist Olivera Finn, whose team in 1989 discovered the first tumor-associated antigen: a version of MUC1, a sugar-laden cell-surface protein. The altered version dots many types of cancer cells.

Finn developed a vaccine consisting of short stretches of MUC1. In the first study of a preventive vaccine in healthy people, she tested safety in 39 people who had previously had precancerous colon polyps, which put them at elevated risk for colon cancer. In 2013, her team reported 17 had a strong immune response, with much higher levels of antibodies to the tumor version of MUC1 than previously seen in cancer patients who got the vaccine as treatment. The other 22 people, who didn't make antibodies, had immune-suppressing cells in their blood, apparently lingering from their removed polyps, Finn says.

The trial's modest success led to a larger, placebo-controlled trial to see whether the vaccine prevented new polyps in people who had had them removed. This time, just 11 of 53 participants who received

the vaccine produced plentiful antibodies, possibly because the patients' immune-suppressing polyps had been removed only recently. But among the 11 responders, only three had polyps recur within 1 year of receiving the vaccine, compared with 31 of 47 participants in a placebo group, Finn's team reports in a paper submitted to a journal.

"It was very encouraging," Finn says. "When you have no recurrence in responders, you know the vaccine is working." Adding a treatment that blocks immune-suppressing cells may boost response rates, she says. Her team now plans MUC1 vaccine trials for several precancerous conditions.

ONE DRAWBACK of Finn's vaccine strategy is that the short proteins, or peptides, it contains mainly trigger one arm of the immune system: the B cells that make antibodies. "For immunity against cancer we really need to mobilize T cells," says cancer immunologist Robert Vonderheide, director of Penn Medicine's Abramson Cancer Center. That's best done by injecting the genetic instructions for the antigen rather than the antigen itself. Special immune cells then take up the DNA or RNA, manufacture the antigen, chop it up, and display bits tailored to that person's immune system on their cell surfaces. These antigen-presenting cells then teach T cells to recognize and kill tumor cells.

Vonderheide's team is testing a DNA-based vaccine targeting a different antigen that marks many tumors: hTERT, a small chunk of telomerase, an enzyme that protects chromosomes as cancer cells proliferate.

Results of a trial testing the vaccine's safety in 93 patients in remission after treatment for various cancers were encouraging. All but four people made T cells that home in on hTERT, the team reported in the *Journal for ImmunoTherapy of Cancer* in July 2021. And there was a hint the vaccine was warding off cancer. Among the 34 people who had had pancreatic cancer, 41% were still cancer free after 18 months. In other pancreatic cancer patients in remission, their tumor reappears within an average of 12 months.

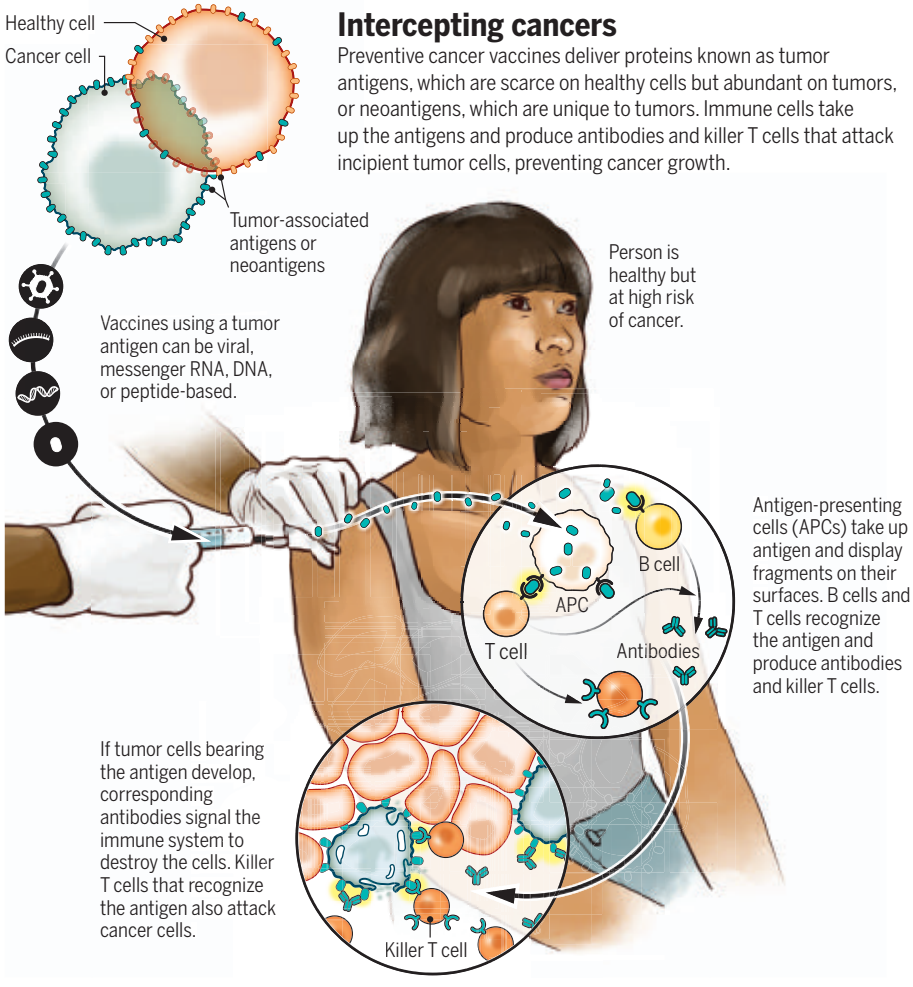
The Penn team is now studying safety and immune responses to the vaccine in 16 people in remission from previous cancers who have inherited mutations in *BRCA1* or *BRCA2*, relatively common cancer genes that raise risk for breast and some other cancers. Next year, the researchers expect to give the vaccine to 28 people with *BRCA* mutations who have never had cancer.

But because hTERT is found on some normal cells as well as cancerous ones, a vaccine could trigger an autoimmune attack on healthy cells, suggests immunologist Vincent Tuohy of the Cleveland Clinic. He has devised a breast cancer prevention vaccine that may be safer because it contains a breast cell protein called alpha-lactalbumin that people only make during late pregnancy and breastfeeding. Production of the protein also occurs in triple negative breast cancer, an aggressive form of the disease. Tuohy's team is testing whether his protein vaccine can stimulate an immune response in 24 women who have been treated for triple negative breast cancer and have no plans to get pregnant. The next step, he says, will be a trial in healthy women with *BRCA1* mutations, who are prone to this cancer type.

Other teams hope to offer broader protection against breast cancer. Undeterred by being called "misguided" in 2012, NBCC is close to testing a breast cancer vaccine, initially in healthy breast cancer survivors. The advocacy group's president, Fran Visco, says it set the ambitious goal because it was "frustrated with the lack of innovation in breast cancer." With scientist partners, it has settled on a vaccine that combines six tumor antigens, including hTERT and MUC1. "We don't know what type of breast cancer a woman is going to get," explains trial leader Keith Knutson, an immunologist at the Mayo Clinic. Multipronged vaccines "are probably going to be more effective than vaccines targeting one individual protein," says cancer immunologist Nora Disis of the University of Washington, Seattle, who is developing such a vaccine to prevent colon cancer.

Intercepting cancers

Preventive cancer vaccines deliver proteins known as tumor antigens, which are scarce on healthy cells but abundant on tumors, or neoantigens, which are unique to tumors. Immune cells take up the antigens and produce antibodies and killer T cells that attack incipient tumor cells, preventing cancer growth.



GRAPHIC: V. ALTOUNIAN/*SCIENCE*

AS SOME TEAMS are trying to broaden the immune response triggered by cancer vaccines, others want to make it safer and more precise by targeting neoantigens, only found on cancer cells. Those efforts have accelerated over the past decade thanks to a surge in tumor genome sequencing, which has revealed a flood of neoantigens. Some drive cancer growth, whereas others have no apparent function. Most are unique to an individual cancer—an obstacle for developing preventive vaccines, which have to target markers that can be predicted in advance.

Some neoantigens reliably appear on many people's tumors, however. For instance, pancreatic cancer is almost always triggered by mutations in a growth protein called KRAS, which give rise to a predictable set of neoantigens. This spring, Johns Hopkins University immunologist Elizabeth Jaffee and colleague Neeha Zaidi will begin to safety test a vaccine containing mutated KRAS peptides in 25 men and women who haven't had cancer but are at high risk because of an inherited mutation or family history. KRAS is like pancreatic cancer's Achilles' heel, Jaffee says: It's the first of several genes to get mutated. As a result, the team hopes early tumor cells won't be able to evade the vaccine by ditching KRAS and finding another way to grow.

Lynch syndrome cancers also sport a predictable set of neoantigens. That's because patients' DNA repair problem leads to "frameshift" mutations, which shift how a cell's proteinmaking machinery reads a gene, scrambling the resulting protein in a consistent way. A peptide vaccine containing a few of these neoantigens, which was developed by a German team, caused no serious side effects when tested in people with cancer. A similar vaccine designed for mice with Lynch syndrome reduced tumor growth, researchers reported in July 2021 in *Gastroenterology*.

The vaccine Vilar-Sanchez's team will test is more ambitious: It consists of viruses modified to carry DNA for a whopping 209 frameshift neoantigens found in Lynch tumors. People's immune systems vary in how they respond to specific neoantigens, and different individuals' tumors won't all make the same set. "Therefore, the best [approach] is to have many," says Elisa Scarselli, chief scientific officer of Nouscom, an Italian company developing the vaccine.

The vaccine is also being developed as treatment, and in an early test Nouscom is giving it along with an immunotherapy drug to patients who have metastatic cancers with frameshift mutations like those in Lynch syndrome. At a meeting in fall 2021, the company reported the treatment shrank tumors in seven of the first 12 patients. "We really believe we will see even more immunogenicity in healthy carriers of Lynch disease" because they should have stronger immune systems, Scarselli says.

Vilar-Sanchez's trial, beginning within a few months, will give the vaccine to 45 volunteers with Lynch syndrome—both people in remission after cancer treatment and others who have never had tumors. Investigators will assess whether the vaccine stimulates an immune response and has any apparent effect on polyps or tumor formation.

If the results look good, the next step will be a randomized study of hundreds of patients over perhaps 5 to 10 years. "There's a lot to be gained" if the vaccine works, Vilar-Sanchez says. "A cancer vaccine is not going to reduce the risk to zero, but it could impact how often we perform screening." It could also help patients decide whether to have a hysterectomy to prevent endometrial cancers, which are common in people with Lynch syndrome.

All prevention vaccines would face a long road to regulatory approval if researchers must wait for tumors to appear to judge the vaccine's efficacy. So they will also look for surrogate measures of protection, such as reduced growth of polyps in people prone to colon cancer. For breast cancer, researchers don't have biomarkers yet but hope to find them, perhaps a change in blood-borne immune cells or breast tissue, Vonderheide says.

Eduardo Vilar-Sanchez is testing a vaccine to prevent Lynch syndrome cancers.

"We have to be smart enough to present to the FDA [U.S. Food and Drug Administration] a biomarker of success," Vonderheide says. "This is formidable. But we're inspired because the impact will be massive."

WHATEVER THEIR PREFERRED antigens, many scientists expect to model their next preventive vaccines on the leading COVID-19 vaccines, which use a lipid particle to ferry mRNA for antigens into cells. mRNA vaccines are easier to make and deliver than DNA or viral vaccines, and the pandemic has shown they're generally safe and stimulate a strong response. "The fact that mRNA vaccines have shown safety in billions of healthy people of all ages makes [mRNA] a very good platform" for preventive cancer vaccines, Jaffee says.

The White House is gunning for mRNA vaccines to prevent cancer, too. They are on the list of potential projects for a reignited Cancer Moonshot and the new high-risk, high-reward research agency, the Advanced Research Projects Agency for Health (ARPA-H). A concept paper for ARPA-H puts the goal this way: "Use mRNA vaccines to teach the immune system to recognize 50 common genetic mutations that drive cancers, so that the body will wipe out cancer cells when they first arise."

That description raises some eyebrows. "That would be heroic," Finn says, because the vaccine antigens would have to cover not only a huge number of cancer mutations, but also "the incredible genetic diversity" in individuals' immune responses. "Not impossible but not simple," she says.

Clinical geneticist Steven Lipkin of Weill Cornell Medicine, who works on Lynch syndrome vaccines, is cautiously optimistic, noting that a vaccine that cut the rates of the most common cancers "by say one-third or one-half in a large number of people would be a tremendous benefit."

One team is already testing a multicancer prevention vaccine—not yet in people, but in dogs. In a 5-year trial, a team is giving 400 middle-age dogs a vaccine that contains 31 antigens from eight common dog cancers. (Another 400 dogs are getting a placebo vaccine.) It relies on RNA neoantigens, little-studied molecules that result from RNA processing errors rather than mutations in DNA. They are far more abundant than DNA neoantigens in dogs and people, and are "highly immunogenic," says developer and biochemist Stephen Johnston of the Biodesign Institute at Arizona State University, Tempe. If they prove effective, they might make it easier to reach the White House's goal of developing a pancancer human vaccine, he says.

Another proponent of a universal cancer prevention vaccine is Johns Hopkins cancer geneticist Bert Vogelstein. He notes that sequencing has shown "a relatively small number of genes are involved in most cancers," suggesting a limited number of antigens could lead to broad protection. Such a vaccine "seems like science fiction," Vogelstein says, but "a concerted effort by many labs" might succeed. Sei agrees: "That's not crazy. That's possible."

For Dave Dubin, even a narrower success—a Lynch syndrome vaccine—"could be game-changing," he says, if it meant fewer cancer screenings and no more major surgeries. "The goal would be almost to live a normal life." ■

PHOTO: MD ANDERSON CANCER CENTER

POLICY FORUM

AGRICULTURE

Policy reforms for antibiotic use claims in livestock

Antibiotics detected in "raised without antibiotics" cattle underscore the need to ensure the integrity of labeling claims

By **Lance B. Price[1], Laura Rogers[1], Kevin Lo[2]**

Public concern regarding antibiotic use in food-animal production has driven a rapidly growing market for meat products from animals that have been raised without antibiotics (RWA) (*1*). RWA is a credence attribute that cannot be easily verified by consumers (*2*). Instead, they must rely on producers, retailers, third-party certifiers, or government regulators to ensure that these claims are valid. Meat labels in the United States, one of the leading beef-producing

PHOTO: ERIC GREGORY/THE JOURNAL-STAR/AP PHOTO

Feedyards face added expenses to feed cattle longer and with less energy-rich diets to reduce the risk of liver abscesses when not using antibiotics.

and -exporting nations, must be approved by the US Department of Agriculture (USDA). Although USDA-approved labels give RWA claims credibility and value in the marketplace (*3*), the agency does not require empirical antibiotic testing to validate them. Absent verification, there are incentives for parties throughout the supply chain to cheat or limit scrutiny (*4*). We present empirical evidence that some beef cattle processed for the RWA market have been administered antibiotics and propose policies to reform the system.

Consumers choose RWA meats for both private and public benefits. Some consumers make this choice because they believe it is safer and healthier for them. Others may choose RWA products to support market-based efforts to reduce antibiotics in food-animal production and preserve the effectiveness of these critical medicines (*5*). However, neither public nor private benefits can be achieved if RWA labels are applied to animals that have been treated with antibiotics.

RWA label claims that are approved by the USDA through the Food Safety Inspection Service (FSIS) include "Raised Without Antibiotics," "No Antibiotics Administered," "No Added Antibiotics," "Raised Antibiotic Free," and "No Antibiotics Ever." Producers wishing to market their products with one of these labels must submit (i) a description of controls to ensure that animals are not given antibiotics; (ii) a protocol for tracing and segregating RWA products; (iii) a protocol for identifying and segregating nonconforming animals (i.e., those treated with antibiotics); and (iv) a signed affidavit describing how the animals were raised to support label claims (*6*). The USDA does not conduct or mandate empirical antibiotic testing for these labels.

Although the USDA occasionally tests for antibiotic residues in meat animals, these tests are not conducted to verify RWA claims. Among the more than 9 billion animals that are slaughtered in the US for meat each year, the USDA tests fewer than 7000 for antibiotics through the US National Residue Program. Technicians from this program conduct tests to determine whether antibiotics in target tissues exceed their maximum residue limits (a threshold defined as safe for public consumption) and effectively blind themselves to antibiotics below these concentrations (*7*). The National Residue Program is not designed to assess RWA claims and is not used to do so.

Other programs, including the Global Animal Partnership's Animal Welfare Certified Program and the USDA's National Organic Program, prohibit the use of antibiotics in beef cattle and lend credence to RWA claims (*8*, *9*). Beef products are often branded with multiple label claims, including RWA, Animal Welfare Certified, and Organic, which implies added layers of scrutiny. However, none of these programs require empirical antibiotic testing.

RWA cattle generate increased price premiums over conventional products at every step along the supply chain (fig. S1). There are also increased costs associated with RWA production, so these premiums should not be interpreted simply as added profit. For example, cow–calf operators—the farmers and ranchers who raise beef cattle—spend more money on supplements and spend more time weaning calves without antibiotics. Feedyards—the companies that fatten cattle for market—pay higher prices for RWA cattle and then take on the added expenses of feeding animals longer with less energy-rich diets to reduce the risk of liver abscesses without antibiotics (*10*).

Cattle producers use antibiotics to treat, control, and prevent infections. In the absence of robust verification, the prospect of sick animals creates a dilemma: Producers must decide whether to withhold antibiotics and potentially sacrifice animal welfare; openly administer antibiotics and forgo investments and premiums; or covertly administer antibiotics and enjoy the benefits of treatment without the costs. From the perspectives of animal welfare and consumer protection, openly treating animals is the best option; however, this places a financial burden on cow–calf operators and feedyards (*11*). This also creates supply disruptions for processors and retailers, who can experience lost revenue, lower plant utilization, and reduced customer confidence. The stakes are highest for retailers that exclusively sell RWA meats as they cannot substitute with conventionally labeled products when supplies are disrupted. This could mean costly periods with empty shelves and missing menu items. In a system characterized by lax verification and enforcement, these financial incentives may be difficult to overcome.

In a well-functioning market, concern for one's reputation should counterbalance the incentives to cheat. In the case of RWA labels, the USDA grants credence and also confers a degree of liability protection. The law states and courts confirm that the USDA has sole authority to determine whether meat labels are truthful and accurate. Thus, an approved USDA label cannot be deemed false or misleading by any entity other than the USDA, even when the evidence suggests otherwise (*12*). This changes every player's risk calculation. For example, retailers can avoid doing their own quality control by relying on the legal safe harbor granted by an approved USDA label. Indeed, meat companies refer to the USDA's duty to review and approve meat labels as a means of preempting consumer protection laws when challenged in court for mislabeling products (*13*). These incentives further limit scrutiny on a set of claims that are otherwise relatively easy to confirm.

To determine whether antibiotic-treated animals are making their way into the RWA supply chain, we tested for antibiotics in urine from beef cattle being slaughtered for the RWA market. All of the cattle were part of a "No Antibiotics Ever" program, with a subset produced under the third-party–audited Global Animal Partnership program (see supplementary materials). Using a rapid immunoassay that screens for 17 antibiotics commonly administered in feed and water, we sampled animals from every lot of RWA cattle delivered for processing at a single slaughter facility over the course of 7 months (mean lot size = 122 cattle; mean number of animals tested per lot = 2). A total of 699 animals were tested from 312 lots and 33 different RWA-certified feedyards (see the figure). The 312 lots sampled in this study included 38,219 head of cattle, representing ~12% of US RWA beef production for this period.

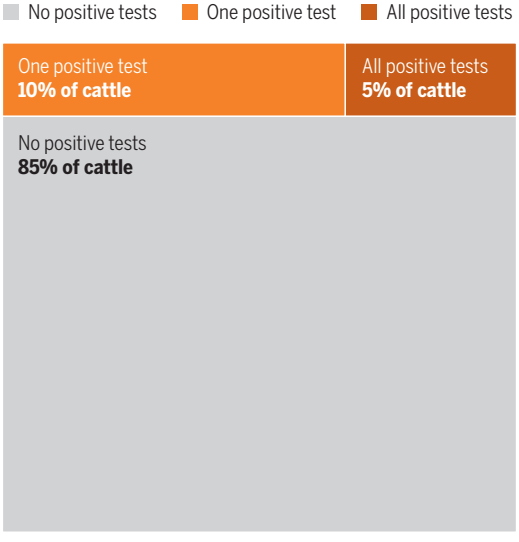
Three feedyards (9%) had multiple lots in which all samples tested positive for antibiotics; 4 feedyards (12%) had all samples test positive in a single lot; 7 (21%) had a positive sample in more than one lot; and 14 (42%) had at least one animal test positive (see the figure). Lots with at least one positive test represented ~15% of the RWA cattle processed at the slaughter facility during the study period (see the figure). These findings provide empirical evidence that a material portion of beef products currently being marketed with RWA labels is from cattle that were treated with antibiotics.

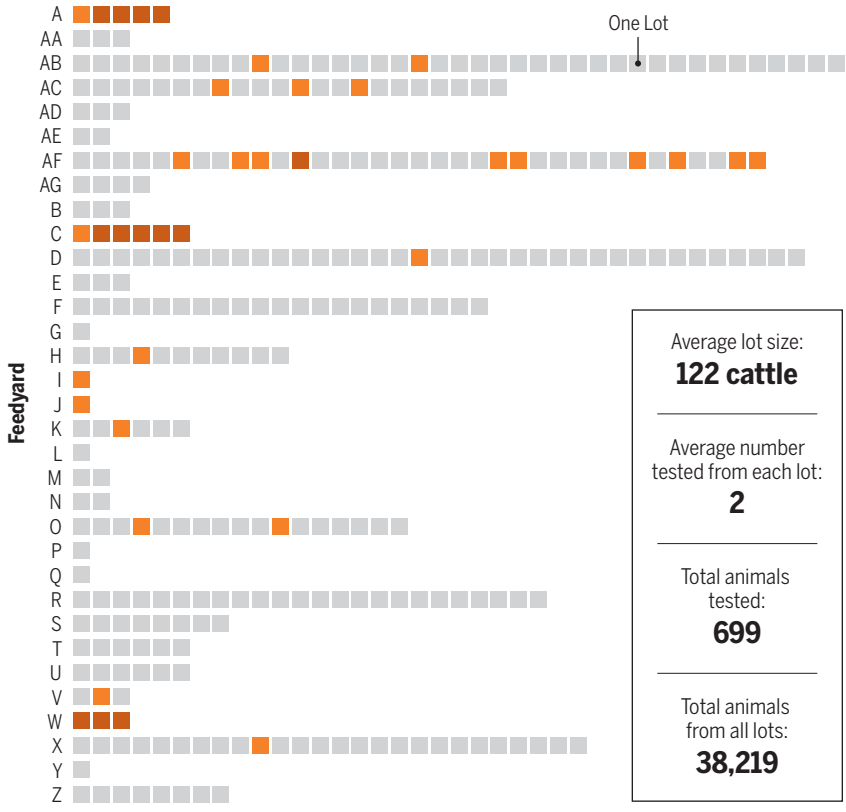
These findings suggest that today's RWA labels lack integrity. Although our testing was limited to beef cattle, other meat and poultry sectors are vulnerable to similar incentives. To protect consumers and restore the integrity of RWA labels, we recommend the following policy reforms grounded in the literature, which has shown that testing programs with robust standards and public disclosure can overcome incentives to cheat (*4*, *14*, *15*).

[1]Antibiotic Resistance Action Center, Environmental and Occupational Health, Milken Institute School of Public Health, George Washington University, Washington, DC, USA. [2]Food In-Depth, San Mateo, CA, USA.
Email: lprice@email.gwu.edu

Raised with antibiotics

(Right) Antibiotic testing results for 312 lots of cattle from 33 feedyards. Each box represents a single lot in sequential testing order. (Below) Percentage of 38,219 cattle coming from lots in which zero tests were positive, one test was positive, or all tests were positive. See supplementary materials for details and data.





The USDA should establish a rigorous verification system to ensure that RWA claims are truthful and accurate, or they should cease approving these labels. For meaningful verification, the USDA should conduct or require continuous, on-site empirical testing for antibiotics on a meaningful number of animals from every lot delivered for processing. For testing to be effective, the USDA must move beyond maximum residue levels and use sensitive, real-time technologies that identify animals that have been treated with antibiotics. Lots testing positive should be rerouted and sold on the conventional market. Positive lots should be tracked and published on a public ledger. Repeat offenders should be excluded from supplying animals for RWA programs until they can demonstrate that they have taken meaningful steps to eliminate undisclosed antibiotic use.

To ensure that animal welfare is not pitted against the financial welfare of producers, the USDA must eliminate the financial disincentives for treating sick animals. We recommend that the USDA create a fund to compensate RWA producers for lost premiums if they are periodically forced to administer antibiotics and segregate animals from the RWA market. To offset expenses of robust verification and animal-welfare compensation, we recommend that the USDA implement a RWA label user fee. This should help ensure that these new costs are passed to RWA producers, retailers, and consumers of RWA products rather than placing the burden on the general public.

Until the USDA acts, additional studies such as ours could weaken confidence in the RWA labels and decrease consumers' willingness to pay for these products. Retailers can limit these undesirable outcomes by taking responsibility for the integrity of the food they sell and implement a robust, industry-wide standard that incorporates empirical testing, strict enforcement, and transparent administration.

Growing demand for RWA meats and poultry has the potential to curb antibiotic use in food-animal production; however, the integrity of the USDA's RWA labels is being undermined by lax verification and enforcement. Until either the USDA acts to rigorously verify RWA claims or retailers eliminate their own safe harbor of ignorance, consumers should not rely on the accuracy of these labels. ■

REFERENCES AND NOTES

1. C. Dubois, A.M. Roerink, "Mid-Year Meat Department Performance Review, Animal Protein Reigns High Amid Pandemic; What's Next?" *IRI*, June 2020; https://www.iriworldwide.com/IRI/media/Library/IRI-Mid-Year-Meat-Performance.pdf.
2. N. D. Fortin, *Credence Claims and Conditional Labeling in Food Regulation: Law, Science, Policy, and Practice* (Wiley, ed. 2, 2017), chap. 6, pp. 107–115.
3. E. T. Page, G. Short, S. Sneeringer, M. Bowman, "The Market for Chicken Raised Without Antibiotics," 2012–17, EIB-224, US Department of Agriculture, Economic Research Service (September 2021).
4. B. Holmstrom, *Bell J. Econ.* **10**, 74 (1979).
5. M. Bowman, K. Marshall, F. Kuchler, L. Lynch, *Am. J. Agric. Econ.* **98**, 622 (2016).
6. USDA, "Food Safety and Inspection Service labeling guideline on documentation needed to substantiate animal raising claims for label submissions, December 2019"; https://www.fsis.usda.gov/sites/default/files/media_file/2021-02/RaisingClaims.pdf.
7. USDA, "Residue Sampling Plan: Fiscal Year 2020 Blue Book," December 2019; https://www.fsis.usda.gov/node/1982.
8. Global Animal Partnership, "Our Standards"; https://globalanimalpartnership.org/standards/.
9. USDA, "Organic Livestock Requirements," July 2013; https://www.ams.usda.gov/sites/default/files/media/Organic%20Livestock%20Requirements.pdf.
10. C. D. Reinhardt, M. E. Hubbert, *Appl. Anim. Sci.* **31**, 101 (2013).
11. R. Singer *et al.*, *Front. Vet. Sci.* **6**, 452 (2019).
12. United States Court of Appeals for the Ninth Circuit, "Robert Cohen, a consumer, on behalf of himself and all others similarly situated, Plaintiff-Appellant, v. Conagra Brands, Inc., a Delaware corporation" (2021); https://cdn.ca9.uscourts.gov/datastore/opinions/2021/10/26/20-55969.pdf.
13. United States Court of Appeals for the Ninth Circuit, Robert Cohen, Plaintiff-Appellant, v. Conagra Brands, Inc., Defendant-Appellee, Appellee's Answering Brief (2021); https://www.law360.com/dockets/documents/602ea2afef2b439ca7120489.
14. M. Ollinger, J. Bovay, *Am. J. Agric. Econ.* **102**, 186 (2020).
15. S. A. Starbird, *Am. J. Agric. Econ.* **87**, 15 (2005).

ACKNOWLEDGMENTS

We thank M. Ollinger and T. Roberts for reviewing the manuscript and for providing valuable insights. We also thank S. Levitan for data visualization, manuscript review, and editing. Antibiotic testing was funded by Food In-Depth Inc. All data, except for the identities of the feedyards and processing plant, are available in the supplementary materials.

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abj1823

10.1126/science.abj1823

GRAPHIC: K. FRANKLIN/*SCIENCE*

PLANT BIOLOGY

The quest for optimal plant architecture

Changes in plant architecture can improve cereal crop yield

By **G. Wilma van Esse**

Cereals, such as barley and wheat, are vital crops for both food and feed. Modern cultivars emerge, grow, flower, and mature uniformly, and they carry more and bigger seeds that do not shatter compared with their wild relatives (*1*). Modification of plant architecture is a key driver for further improving crop yield but is challenging because different factors that affect yield are often negatively correlated (*2*). On page 180 of this issue, Zhang *et al.* (*3*) describe the identification in common wheat (*Triticum aestivum*) of CONSTANS-like B5 (TaCOL-B5) that affects plant architecture by increasing both the number of tillers [seed head (spike)–bearing stems] and seeds per spike, which enhances yield potential by ~12%. TaCOL-B5 is a transcriptional regulator that is closely related to the flowering time gene *CONSTANS* (*3*, *4*). The discovery of TaCOL-B5 is a milestone toward enhancing yield in cereals because it improves our understanding of molecular mechanisms that control yield-related architectural traits.

Agriculture is disproportionally affected by climate change, which affects plant growth owing to changing growing temperatures (including heat waves), water availability, and disease and pest pressures (*5*). In combination with a rapidly growing human population (*6*), there is a need for crop varieties that produce high yields with limited inputs of artificial fertilizers and pesticides and that are resilient to unpredictable weather. Optimizing yield in cereals involves modulating the delicate balance between key yield-related traits, such as seed size, seed number, and tiller number, as well as the timing of the transition from the vegetative to generative phase (flowering time). Variation in the gene networks that control these factors leads to different balances in plant development and architecture (*7*). This offers opportunities to use genetic variation to select or engineer cereal plants that can adapt to specific growing environments and be resilient to changing conditions.

Yield potential in wheat and barley is determined by flowering time, number of tillers, number of seeds per spike, and seed weight. The trade-off between individual components, such as seed weight and number, is a major bottleneck for further yield improvement (*2*). The identification of genes that control yield-related architectural traits in wheat is not trivial because common wheat has a large (16 giga–base pairs) and complex hexaploid genome. An added complexity is that the wheat genome contains >80% of repetitive DNA (*8*)—with so many similar genomic pieces, it is hard to assemble the sequence jigsaw. Additionally, transformation efficiencies are genotype dependent, which limits routine genetic modification to only a subset of cultivars (*9*). The release of a fully annotated wheat genome and the use of speed-breeding technology has accelerated research in wheat (*8*–*10*). Zhang *et al.* made shrewd use of these resources to identify *TaCol-B5* as a major regulator of yield.

Wheat yield is a complex trait that is determined by factors such as the number of spike-bearing tillers per unit area, the number of seeds per spike, and seed weight.

PHOTO: NIGEL CATTLIN/MINDEN PICTURES

The authors identified *TaCol-B5* through map-based cloning in a population derived from two common wheat cultivars, Cltr17600 and Yangmai18. They found that the dominant *TaCol-B5* allele from Cltr17600 has a positive yield effect. To fast-track the functional evaluation of *TaCol-B5* in wheat, they expressed the *TaCol-B5* allele from Cltr17600 in Yangmai18, which increased tiller number and seeds per spike under field conditions. Notably, there was no negative effect on seed size, which indicates that breaking negative correlations between yield components is possible. Although world cereal production increased annually by 1.9% between 1961 and 2007, this growth rate is projected to reduce to a 0.9% annual increase between 2007 and 2050 (*6*). Considering this, a potential yield increase of ~12%, as shown by Zhang *et al.*, is a leap forward. The *TaCol-B5* allele from Cltr17600 is not commonly used in cultivated germ plasm. It is therefore important to test the effect of allelic variation in *TaCol-B5* in wheat grown in multiple environments, as well as in other genetic backgrounds, to get a more accurate assessment of the potential yield increases. In addition, these results might be translatable to other key cereal crops, such as rice, barley, and rye.

Zhang *et al.* also provide a detailed analysis of the mechanism by which the TaCOL-B5 transcription factor functions. They found that a single amino acid substitution (Ser^{269}→Gly) in TaCOL-B5 from Cltr1700 resulted in differential protein phosphorylation by TAK4, a serine-threonine protein kinase that they identified as a key protein-protein interac-

Cluster of Plant Developmental Biology, Laboratory of Molecular Biology, Wageningen University and Research, Wageningen, Netherlands. Email: wilma.vanesse@wur.nl

tor of TaCOL-B5. According to functional analysis of the conserved domains, the authors postulate that the TaCOL-B5 transcription factor modulates multiple traits, such as flowering time and plant height, through different conserved domains.

The next challenge is to untangle these different yield-related traits through targeted modification of specific domains or specific amino acids. Further studies into TaCOL-B5, its conserved domains, its role in growth regulation networks, and its responses to diverse environmental cues can help to fine-tune wheat cultivars to the specific needs of growers worldwide. Such research should also include other regulators of flowering time and plant architecture, such as FLOWERING LOCUS-T family members (*7*). Fundamental knowledge of the underlying molecular-genetic networks provides opportunities for generating new variation and increasing yield potential through knowledge-driven breeding as well as by genetic modifications and gene editing (*2*, *3*, *7*, *11*).

Rapid climate change, reduced resources, and biotic and abiotic stress (*5*, *6*) call for a multidisciplinary approach to tackle these challenges. Innovations and technologies, such as genomic selection, gene editing, precision agriculture, and intercropping (growing at least two different crops in a field at the same time), as well as advanced phenotyping technologies are needed to apply scientific knowledge of plant development and plasticity to breeding and growing practices (*7*, *12*–*14*). Nonetheless, the introduction of genes or alleles into new varieties that increase the yield potential of cereals is a major goal for plant breeders and scientists to enable sustainable crop production. The identification of *TaCol-B5* by Zhang *et al.* offers a new route to maximize yield in wheat. ■

REFERENCES AND NOTES

1. M. Haas, M. Schreiber, M. Mascher, *J. Integr. Plant Biol.* **61**, 204 (2019).
2. D. F. Calderini *et al.*, *New Phytol.* **230**, 629 (2021).
3. X. Zhang *et al.*, *Science* **376**, 180 (2022).
4. J. Putterill, F. Robson, K. Lee, R. Simon, G. Coupland, *Cell* **80**, 847 (1995).
5. Food and Agriculture Organization of the United Nations (FAO), "The impact of disasters and crises on agriculture and food security: 2021" (2021); https://doi.org/10.4060/cb3673en.
6. N. Alexandratos, J. Bruinsma, "World Agriculture Towards 2030/2050: The 2012 Revision" (FAO, ESA Working Paper no. 12-03, 2012); www.fao.org/3/ap106e/ap106e.pdf.
7. Y. Eshed, Z. B. Lippman, *Science* **366**, eaax0025 (2019).
8. International Wheat Genome Sequencing Consortium *et al.*, *Science* **361**, eaar7191 (2018).
9. P. Borrill, *New Phytol.* **228**, 1721 (2020).
10. A. Watson *et al.*, *Nat. Plants* **4**, 23 (2018).
11. L. Liu *et al.*, *Nat. Plants* **7**, 287 (2021).
12. C. Li *et al.*, *Nat. Plants* **6**, 653 (2020).
13. D. Sun, K. Robbins, N. Morales, Q. Shu, H. Cen, *Trends Plant Sci.* **27**, 191 (2022).
14. C. Uauy, *Curr. Opin. Plant Biol.* **36**, 142 (2017).

10.1126/science.abo7429

GENOMICS

Population genetics meets single-cell sequencing

Single-cell technology can be used to understand the genetic basis of human diseases

By **Tomokazu S. Sumida**[1,2] *and* **David A. Hafler**[1,2,3]

Using high-throughput genotyping platforms to identify single-nucleotide polymorphisms (SNPs) has allowed genome-wide association studies (GWASs) to generate an unbiased classification of human diseases that are associated with common genetic variation. There are many common allelic variants in noncoding regions that have small effect sizes with complex interactions that are highly cell type and cell state dependent. Moreover, progress toward understanding disease mechanisms has been limited by the challenges of assigning molecular function to most GWAS "hits" that are noncoding sequences associated with disease. On pages 154 and 153 of this issue, Yazar *et al.* (*1*) and Perez *et al.* (*2*), respectively, use multiplexed single-cell RNA sequencing (scRNA-seq) with fine mapping of autoimmune disease–associated genetic variants to provide a resource that allows the large-scale identification of genotype-phenotype interactions. Notably, these two studies provide a comprehensive catalog of immune cell profiles that opens the door to a new era of functional genetics.

Most GWAS variants associated with diseases map to noncoding regions that are highly enriched in regulatory elements, indicating that those variants are likely to exert their effects through the modulation of gene expression (*3*). Expression quantitative trait locus (eQTL) analysis is used to measure the association between genetic variants and gene expression. This requires RNA expression to be averaged across bulk populations of cells (*4*), allowing the characterization of genetic variants that are significantly associated with gene expression in a population sample (*5*–*8*). Previous integration of bulk RNA-seq–based eQTL analysis with tissue- or cell type–specific gene expression profiles provided further evidence that eQTL effects act in a tissue- and cell type–specific manner, although this approach is biased to known cell types and does not allow the identification of cell subsets and transitional states. Although scRNA-seq coupled with GWAS to identify disease-associated variants has the potential of revealing both individual cell types and states where variants exert their effects, the cost and low throughput of scRNA-seq had not allowed this approach at the necessary scale.

By overlaying likely causative SNPs onto maps of histone modifications (which regulate gene expression), such as the histone 3 Lys^{27} (H3K27) acetylation maps of different cell types from the ENCODE project, cells that are likely influenced by disease-associated gene variants can be identified (*3*). Although the effect size of allelic variants can be small, functional analysis of single haplotypes (a group of alleles of different genes that are inherited together) revealed that the biologic effects could be substantial. For example, risk variants associated with the autoimmune disease multiple sclerosis that were proximal to the nuclear factor κB subunit 1 (*NFKB1*) gene were associated with increased pathogenic NFκB signaling with tumor necrosis factor–α stimulation in healthy individuals (*9*). Although studies that investigated specific genotype-phenotype interactions were informative, they did not provide the necessary scalability to broadly examine hundreds of allelic variants.

To overcome those limitations, high-throughput scRNA-seq was previously used to conduct a cell type–specific eQTL study at subpopulation scale as a proof of concept (*10*). Cell type–specific eQTL analysis was performed with six immune cell types identified from ~25,000 peripheral blood mononuclear cells (PBMCs) extracted from 45 individuals. The study confirmed that scRNA-seq–based eQTL analysis can replicate observations of previously known "local" eQTLs that act in cis to modulate gene expression (cis-eQTLs). This study showed the utility of scRNA-seq for integrating multiple sets of immune cell tran-

[1]Department of Neurology, Yale School of Medicine, New Haven, CT, USA. [2]Broad Institute of Massachusetts Institute of Technology and Harvard University, Cambridge, MA, USA. [3]Department of Immunobiology, Yale School of Medicine, New Haven, CT, USA. Email: tomokazu.sumida@yale.edu; david.hafler@yale.edu

scriptomic data to better determine cell type–specific eQTLs and highlighted the potential of the scRNA-seq approach over bulk RNA-seq. However, the numbers of individuals and cells analyzed by this study were relatively small. Therefore, a larger population-scale scRNA-seq study with appreciable cell numbers per individual was warranted. The studies by Yazar *et al.* and Perez *et al.* addressed this need.

Yazar *et al.* characterized the transcriptome and genetic variation across a total of 1,267,758 PBMCs from 982 individuals. Their eQTL mapping at single cell resolution enabled the identification of 117 loci outside of the major histocompatibility complex (MHC) region that exert cell type-specific causal effects that account for autoimmune disease risk. Perez *et al.* applied the same method for patients with the autoimmune disease systemic lupus erythematosus (SLE), profiling 1,263,676 PBMCs from 264 individuals (162 cases, 99 controls). Of note, in the SLE dataset, Perez *et al.* observed that cell type–specific eQTL effects of SLE risk variants were highly enriched in classical monocytes and B cells. Moreover, by using type I interferon (IFN-I) response gene expression as a proxy for IFN-I–induced activation, they found that the cell type–specific cis-eQTL effect was modified by IFN-I responses. Given that an IFN-I signature has been observed in SLE patients, these results highlight the importance of disease-relevant cellular and biological contexts to better understand the disease-associated genetic effects.

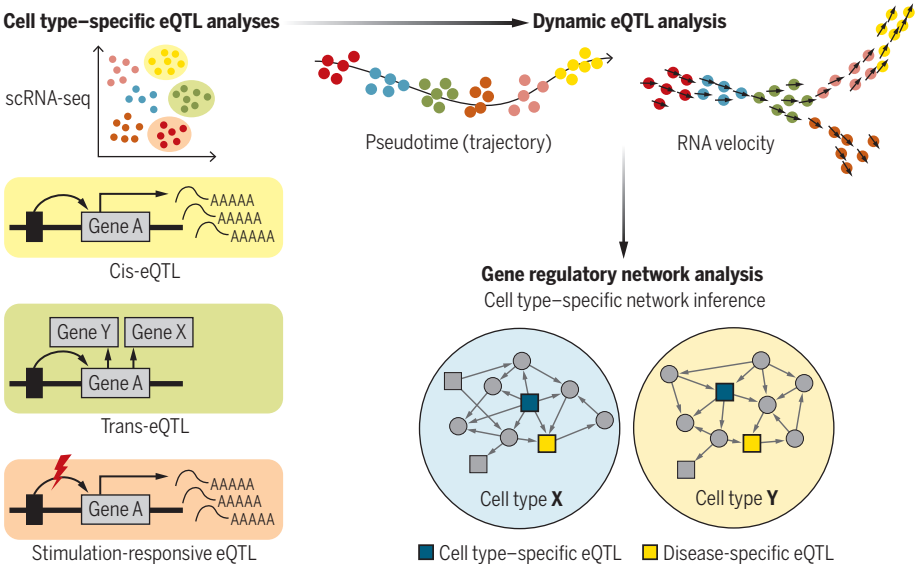
A distinct feature of the scRNA-seq–based approach over bulk RNA-seq is the ability to compute dynamic transcriptional transitions of cellular state and project cells onto an axis called pseudotime that represents the progression trajectory. By inferring the dynamic trajectory of cells, the effects of eQTLs on this cellular dynamism can be investigated. Yazar *et al.* applied this method to identify dynamic eQTLs during B cell maturation that were not detected by cell type–specific cis-eQTL analysis. In addition, RNA velocity, which allows inference of future transcriptional state direction, can be integrated with genetic effects in scRNA-seq data. These new features, which can only be achieved by scRNA-seq–based approaches, could expand our understanding of context-dependent genetic effects beyond the framework of conventional cell type–specific effects (see the figure).

Although scRNA-seq has advantages, several shortcomings are apparent. The low number of cells within the minor subpopulations of PBMCs make it difficult to perform cell type–specific eQTL analysis. Indeed, less than 15 cell types were investigated in these studies for cell type–specific eQTL analysis, which is fewer than a recent study using a bulk RNA-seq–based approach with 28 immune cell populations (*11*). For example, plasmablasts were identified as an immune cell that exhibits substantial overlap with SLE GWAS top hits and cell type–specific eQTLs (*11*); however, this signal was not captured by Yazar *et al.* and Perez *et al.*, likely because of low numbers of plasmablasts for analysis. Moreover, the insufficient resolution of T cell subclustering limited investigation of immunologically meaningful T cell subsets, such as regulatory T cells, which are a small fraction of $CD4^+$ T cells but play critical roles in regulating autoinflammatory diseases. To include those minor populations in the analysis, larger numbers of cells per donor or enrichment of those cell types are required for scRNA-seq.

There are several ways to expand the capacity of single-cell functional genomics. scRNA-seq data can be used to reconstruct gene regulatory networks (GRNs) for cell types or cell lineages by integrating coexpression matrices and RNA velocity (*12*). GRNs can also be inferred by integrating chromatin accessibility data (*13*). Recent advances in single-cell technology make it possible to jointly profile messenger RNA, protein, and chromatin accessibility (*14*). This single-cell multi-omics approach provides further layers of information that can improve causal GRN inference. A promising avenue to understand functional features of genetic susceptibility is to interrogate specific responses to stimulation using eQTLs. Given that immune cells can dynamically change their characteristics in response to external stimuli and that disease-associated eQTL effects can be context specific, genetic effects could be observed in each context. Thus, interrogating immune cell responses under different stimulation conditions may potentiate the detection of eQTL effects that may not be apparent at steady state. The advancement of single-cell technology will further expand the application of functional genetics. Integration of scRNA-seq data with available functional genetic resources could pave the way for our understanding of causal mechanisms of complex diseases. ■

Single-cell technology applied to functional genomics

Population-scale single-cell RNA sequencing (scRNA-seq) analyses have the potential to perform multiple expression quantitative trait locus (eQTL) analyses in a cell type–specific manner. scRNA-seq can be used for pseudotime-trajectory and RNA-velocity analyses, which can reconstruct cell type–specific gene regulatory networks. By integrating cell type–specific genetic eQTL effects and eQTLs in response to stimuli, personalized cell type–specific regulatory networks can be inferred.



GRAPHIC: K. HOLOSKI/*SCIENCE* BASED ON T. S. SUMIDA AND D. A. HAFLER

REFERENCES AND NOTES

1. S. Yazar *et al.*, *Science* **376**, eabf3041 (2022).
2. R. K. Perez *et al.*, *Science* **376**, eabf1970 (2022).
3. K. K. Farh *et al.*, *Nature* **518**, 337 (2015).
4. E. Choy *et al.*, *PLOS Genet.* **4**, e1000287 (2008).
5. E. E. Schadt *et al.*, *Nature* **422**, 297 (2003).
6. T. Lappalainen *et al.*, *Nature* **501**, 506 (2013).
7. GTEx Consortium, *Science* **348**, 648 (2015).
8. L. R. Lloyd-Jones *et al.*, *Am. J. Hum. Genet.* **100**, 371 (2017).
9. W. J. Housley *et al.*, *Sci. Transl. Med.* **7**, 291ra93 (2015).
10. M. G. P. van der Wijst *et al.*, *Nat. Genet.* **50**, 493 (2018).
11. M. Ota *et al.*, *Cell* **184**, 3006 (2021).
12. X. Qiu *et al.*, *Cell Syst.* **10**, 265 (2020).
13. V. K. Kartha *et al.*, bioRxiv 10.1101/2021.07.28.453784 (2021).
14. E. P. Mimitou *et al.*, *Nat. Biotechnol.* **39**, 1246 (2021).

10.1126/science.abq0426

PARTICLE PHYSICS

An upset to the standard model

Latest measurement of the *W* boson digs at the most important theory in particle physics

By **Claudio Campagnari**[1] *and* **Martijn Mulders**[2]

Over the past 60 years, the standard model (SM) has established itself as the most successful theory of matter and fundamental interactions—to date. The 2012 discovery of the Higgs boson only added to the streak of triumphs for the theory (*1*, *2*). However, the SM is known to be incomplete and has noticeable shortcomings, such as its inability to account for dark matter in the universe or to include gravity in a consistent fashion. Physicists have looked for phenomena that directly challenge the SM in the hope of finding hints on what a more complete theory may look like. Although no "new" particle has yet been found, a few fissures have recently been exposed in the SM by precise measurements that are at odds with the model's predictions (*3*, *4*). On page 170 of this issue, the Collider Detector at Fermilab (CDF) Collaboration (*5*) adds further intrigue with its measurement of the *W* boson mass.

The *W* boson, whose existence and detailed properties were first predicted in the 1960s and confirmed at CERN in 1983, is a key building block of the SM. It is a particle that is associated with the weak force, which is responsible for radioactive nuclear β decay, and that plays a similar role as that of the photon in the electromagnetic interaction. Although the photon is massless, the *W* boson is massive; it is about 80 times the mass of a hydrogen nucleus. Within the theoretical framework of the SM, the *W* boson mass is a parameter, with a value that is bounded by other observables such as the electron charge and the masses of other particles, including the top quark and the Higgs boson. A very accurate measurement of the *W* boson mass can therefore provide a stringent test of the self-consistency of the SM.

Over the past 30 years, there have been ever more precise measurements of the *W* boson mass, and the CDF Collaboration now adds to these reports. Based on 10 years of data recorded at the CDF, they report a *W* boson mass with an impressive precision of 117 parts per million (ppm)—twice as precise as the previous most accurate measurement. Their measured *W* boson mass is in direct contention with the SM because it is heavier than the SM prediction by seven standard deviations. This could be a signature for new interactions or new particles that are either too massive to be produced or too hard to detect at existing accelerators. Nonetheless, such yet-to-be-known particles and physical interactions could alter the relationships between the various observables through hidden interactions with the *W* boson and cause the observed deviation from SM predictions.

Effects on the *W* boson mass from previously undetected particles have been observed before. Notably, the observations of these effects were used to probe the masses of the top quark and the Higgs boson before their direct detection. After the observation and precise measurement of each discovered particle, the web of SM predictions was weaved with greater strength and accuracy. With more and more precise measurements of physical quantities—such as cross sections, decay rates, and masses of fundamental particles—fissures between SM predictions and reality may have begun to show. When not in agreement with the theoretical predictions, such measurements can provide a first glimpse of physics beyond the SM.

> **"This could be a signature for new interactions or new particles..."**

Because extraordinary claims require extraordinary evidence, the claim by the CDF Collaboration will require additional experiments to provide an independent confirmation. Scientists at the Large Hadron Collider (LHC) have already collected samples of *W* bosons that are larger than those available at Fermilab and, in principle, could achieve better precision. The Tevatron experiment at Fermilab—DZero—may also get back in the *W* boson mass-measuring race. The result from the CDF Collaboration provides an impetus to improve the measurements of other SM parameters that can help to test and constrain the theory, such as the top quark mass, the strong coupling constant, and the Weinberg angle, named after the late Steven Weinberg (*6*, *7*), a founding father of the electroweak model that is currently being challenged.

The High-Luminosity LHC—an upgraded version of the LHC at CERN that will come online later in this decade—will provide higher beam energy and collision rates with updated and more powerful detectors. The upgraded collider will offer ample opportunity for more precise measurements and for direct searches for new particles. Particle physicists are also looking forward to the next generation of accelerators. Electron-positron colliders are particularly well suited for carrying out precision measurements. Several proposals for electron-positron colliders—such as the International Linear Collider in Japan, the Compact Linear Collider, the Future Circular Collider (FCC-ee) at CERN, and the Circular Electron Positron Collider in China (*8*)—are under consideration in the ongoing discussions for the future of particle physics. Among them, the FCC-ee would offer the best prospects for an improved *W* boson mass measurement, with a projected sensitivity of 7 ppm (*9*), more than 10 times better than the current best measurement.

Among possible theories that could explain the discrepancy with the SM prediction is the theory of supersymmetry (SUSY), which is an old favorite of particle physicists because it provides a plausible explanation for some of the SM's unexplained properties and forms a natural connection to deeper level descriptions of the universe such as string theory. However, none of the many exotic particles predicted by SUSY have been observed, despite extensive searches at particle detectors around the world. The surprisingly high value of the *W* boson mass reported by the CDF Collaboration directly challenges a fundamental element at the heart of the SM, where both experimental observables and theoretical predictions were thought to have been firmly established and well understood. The finding of the CDF Collaboration offers an exciting new perspective on the present understanding of the most basic structures of matter and forces in the universe. ■

[1]Department of Physics, University of California, Santa Barbara, Santa Barbara, CA, USA. [2]CERN, Geneva, Switzerland. Email: claudio@physics.ucsb.edu; martijn.mulders@cern.ch

REFERENCES AND NOTES

1. G. Aad *et al.*, *Phys. Lett. B* **716**, 1 (2012).
2. S. Chatrchyan *et al.*, *Phys. Lett. B* **716**, 30 (2012).
3. B. Abi *et al.*, *Phys. Rev. Lett.* **126**, 141801 (2021).
4. R. Aaji *et al.*, arxiv:2103.11769 [hep-ex] (2021).
5. CDF Collaboration, *Science* **376**, 170 (2022).
6. F. Wilczek, *Nature* **596**, 183 (2021).
7. J. Preskill, *Science* **373**, 1092 (2021).
8. J. de Blas *et al.*, *J. High Energy Phys.* **2020**, 139 (2020).
9. M. Benedikt, A. Blondel, P. Janot, M. Mangano, F. Zimmermann, *Nat. Phys.* **16**, 402 (2020).

10.1126/science.abm0101

OSTEOLOGY

Enhancing strength in mineralized collagen

X-ray data reveal the role of prestress in hierarchical biocomposites at the nanoscale

By **Fabio Nudelman**[1] and **Roland Kröger**[2]

Living organisms build an assortment of mineralized tissues by combining biopolymers and minerals. Mineralization is fundamental to many biological functions, ranging from mechanical shock protection by shells, mastication by teeth, linear acceleration detection by otoconia in the inner ear, and body support by skeletons. Scientists have been investigating the material properties of these biominerals with the focus on the combination of organic and inorganic phases and on the organization of microscopic building blocks across several length scales. Bone, which consists of nanocrystalline calcium phosphate in the form of hydroxyapatite embedded within collagen fibrils (*1*), is one of the most extensively studied biominerals. Fracture resistance of bones is generally attributed to the mineralized collagen fibril (*2*). On page 188 of this issue, Ping *et al.* (*3*) report that mineral growth inside collagen generates a fibril that is under tension, similar to prestressed concrete.

Bones have a hierarchical architecture where mineralized collagen fibrils are assembled into higher-order structures ranging from the submicrometer to the macroscopic scale (*4, 5*). The main advantage of this type of organization is twofold: It provides many interfaces that serve as efficient crack deviation, which enhances the toughness of bone; and it allows the formation of tissues with the mineralized collagen fibrils organized in different motifs, thereby imparting different mechanical properties (*4*). This hierarchical structure of bone is key for understanding both the mechanisms of bone formation and how its mechanical properties arise from its composition and the arrangement of its building blocks.

Using a combination of in-operando x-ray diffraction and Raman microscopy, Ping *et al.* used carbonate-based minerals to observe how mineral growth inside the collagen generates compression on the fibrils, and how this stress is subsequently transferred from the fibrils to the mineral. This tension-transducing process leads to prestressed mineralized collagen fibrils. These are strengthened against external tensile pressures that, when organized into higher-order structures, generate the micro- and macroscopic stress as observed in bone (see the figure) (*6, 7*).

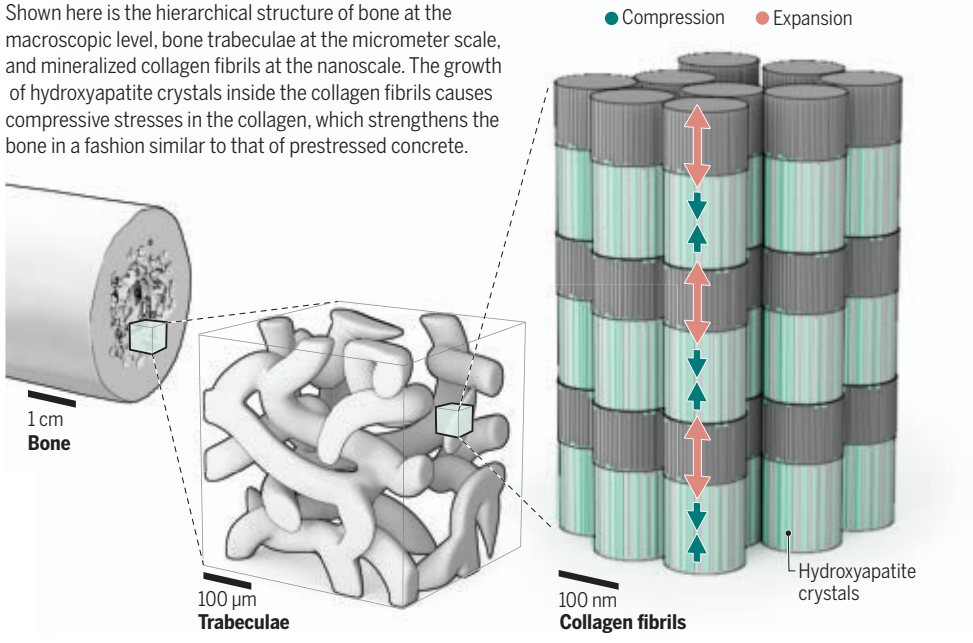
Ping *et al.* highlight prestressing as a widespread strategy to strengthen natural materials with load-bearing functions. A notable example is the trunk of a tree, which is under compression in the central region, whereas the outer layers are under tension (*8*). This combination of forces helps the trunk dissipate stress when a load is applied and allows the tree to sustain bending forces without breaking. It is conceivable that the prestressed mineralized collagen fibrils affect the mechanical properties of bone in a similar way. An important difference between wood and bone, though, is that in the former, prestressing is generated not by reinforcing fibrils with minerals, but through the organization of cells in the interior of the tree and the orientation of the cellulose fibrils. This similarity in material properties—shared by vastly different biological systems—shows that different organisms can evolve similar strategies to achieve the prestressing of their structural tissues.

With regard to experimental techniques, Ping *et al.* provide a neat proof-of-principle demonstration for the in-operando use of advanced x-ray scattering for studying collagen mineralization. Small-angle x-ray scattering facilitates the determination of changes in the overall structure resulting from mineralization, whereas wide-angle x-ray scattering enables the characterization of the mineral at a much smaller crystal-structure scale (*9*). By combining both x-ray scattering techniques, one may investigate molecular-level responses to mineralization not only in biomimetic systems, but also in real bones. Moreover, these measurements can be combined with high-resolution three-dimensional x-ray imaging techniques to reveal the nanostructure, orientation, and organization of the hydroxyapatite crystals (*10, 11*). This would provide information on the relationship between the mechanical properties and bone structure at the nanometer and micrometer scales.

The findings of Ping *et al.* raise several questions about collagen mineralization. Future studies may seek to address the mechanisms behind the molecular contraction and the dehydration of collagen, to explore the impact of size, shape, and

Where bones get their strength

Shown here is the hierarchical structure of bone at the macroscopic level, bone trabeculae at the micrometer scale, and mineralized collagen fibrils at the nanoscale. The growth of hydroxyapatite crystals inside the collagen fibrils causes compressive stresses in the collagen, which strengthens the bone in a fashion similar to that of prestressed concrete.



GRAPHIC: V. ALTOUNIAN/*SCIENCE*

[1]EaStCHEM School of Chemistry, The University of Edinburgh, Edinburgh, UK. [2]Department of Physics, University of York, York, UK. Email: fabio.nudelman@ed.ac.uk

orientation of the hydroxyapatite crystals, and to determine the degree of mineralization on the generation of compression forces inside the collagen. These factors are particularly interesting, given the complexity and multilevel organization of the hydroxyapatite crystals in the collagen fibrils, spanning both the intra- and extrafibrillar spaces (*12*). Quantifying the contribution of prestressing to the overall mechanical properties of bone, and how it scales with the hierarchical organization of the fibrils, will constitute an important step toward understanding how the properties of the tissue arise from its composition and structure across length scales. It will be exciting to determine whether, and how, prestressing varies between bone tissues with different mechanical requirements and across different species.

This work draws attention to a broader perspective—namely, the large variety of biominerals with load-bearing functions found in nature. Enamel and dentin, which compose the vertebrate tooth, are subject to forces during mastication. Shells have to be tough enough to provide protection without fracturing, and in some cases, can withstand large deformations (*13*). This raises interesting questions as to whether prestresses at the submicrometer and micrometer scales constitute a mechanism to strengthen the mechanical properties of other mineralized tissues. Given the diversity of compositions, structures, and functions of biominerals, it is crucial to elucidate how prestressing is enabled in each case. Hence, using advanced correlative and in situ characterization methods, demonstrated in this work, constitutes a step change in addressing these questions for our general understanding of biomineralization as well as the application of this knowledge in biomedicine, environmental protection, materials design, and engineering. ■

REFERENCES AND NOTES

1. S. Weiner, W. Traub, *FASEB J.* **6**, 879 (1992).
2. J. W. C. Dunlop, P. Fratzl, *Annu. Rev. Mater. Res.* **40**, 1 (2010).
3. H. Ping *et al.*, *Science* **376**, 188 (2022).
4. S. Weiner, H. D. Wagner, *Annu. Rev. Mater. Sci.* **28**, 271 (1998).
5. N. Reznikov, R. Shahar, S. Weiner, *Acta Biomater.* **10**, 3815 (2014).
6. M. G. Ascenzi, *J. Biomech.* **32**, 935 (1999).
7. Y. N. Yeni, M. B. Schaffler, G. Gibson, D. P. Fyhrie, *Ann. Biomed. Eng.* **30**, 217 (2002).
8. R. Ennos, *Solid Biomechanics* (Princeton Univ. Press, 2012).
9. Y. Xu *et al.*, *Nat. Commun.* **11**, 5068 (2020).
10. T. A. Grünewald *et al.*, *Sci. Adv.* **6**, eaba4171 (2020).
11. F. Schaff *et al.*, *Nature* **527**, 353 (2015).
12. N. Reznikov, M. Bilton, L. Lari, M. M. Stevens, R. Kröger, *Science* **360**, eaao2189 (2018).
13. J. Ihli *et al.*, *Nat. Commun.* **12**, 5383 (2021).

10.1126/science.abo1264

MICROBIOLOGY

There's more to RNA viruses than diseases

A simple, pervasive biological entity in the ocean sheds light on evolution

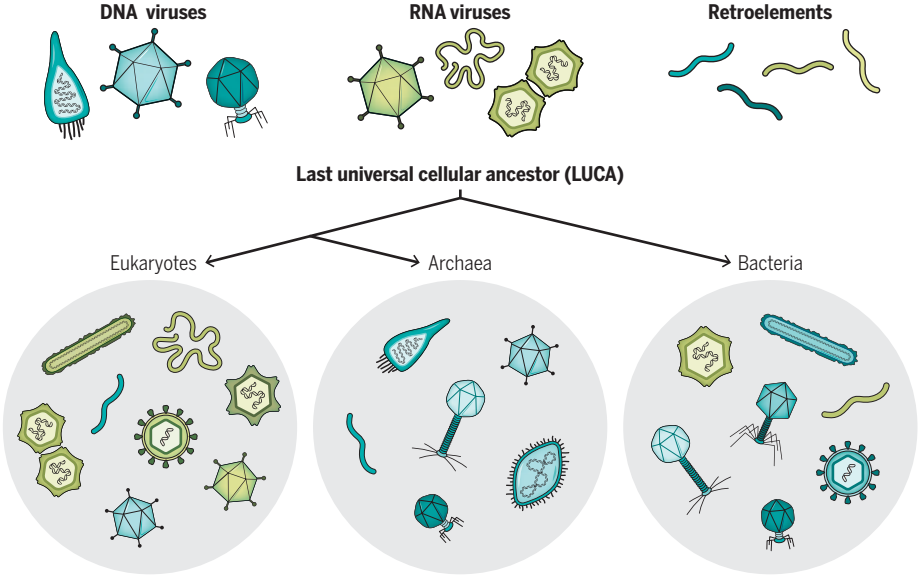
By **Jessica M. Labonté** and **Kathryn L. Campbell**

Viruses infect and affect all domains of life, playing roles as drivers for evolution, diversity, and global geochemical cycling. Because viruses evolve with and depend on their hosts for replication, they are essential in understanding the origins of life. RNA viruses are notorious for being agents of disease in humans and agriculturally important plants and animals. However, because of the focus on studying RNA viruses in pathologies, there is a lack of research on their abundance and diversity in the environment. This paucity of data has challenged evolutionary studies aimed at determining the origin of RNA viruses. On page 156 of this issue, Zayed *et al.* (*1*) report the identification of thousands of RNA viruses in the ocean. These new sequences fill previously missing gaps, enabling the construction of a more robust phylogenetic tree and confirming hypotheses regarding the evolution of RNA viruses.

Every day, viral infections are estimated to kill anywhere between 10 and 20% of all microbial biomass (*2*). These infections have impacts on microbial community composition through population control, on evolution as agents of horizontal gene transfer, and on global geochemical cycles and nutrient recycling (*3*). The development of next-generation sequencing, which facilitates sequencing of total genetic material, provided opportunities for discoveries. Sequencing of the total genetic material of viruses, known as viromics, has demonstrated that viruses are highly diverse (*4*) and globally distributed (*5*) and play important roles in the geochemical cycles (*6*). However, most efforts have focused on the study of DNA viruses. Viral genomes, especially RNA genomes, are smaller and less stable, and difficulties in extracting high-quality viral RNA for sequencing have impaired the exploration of RNA viruses in the environment.

Evolution of life and viruses are intertwined

The divergence of cellular life has led to the majority of RNA viruses infecting eukaryotes, the phylogenetic details of which have been unclear. Zayed *et al.* have found the missing link between retroelements and RNA viruses, which suggests that RNA viruses were present before the LUCA and had multiple points of origin.



GRAPHIC: A. MASTIN/*SCIENCE*

Recently, total viral RNA was extracted from ocean water in a harbor in China. The researchers were able to identify more than 4500 new viruses. Zayed *et al.* cast an even wider net to identify RNA viruses from water samples collected during the *Tara* Oceans expeditions. The *Tara* Oceans expeditions sailed 125,000 km across the global ocean, sampling 210 sites through all oceanic basins at depths down to 1000 m to get a three-dimensional picture of the microbial diversity and ecology (*7*). Zayed *et al.* mined 771 metatranscriptomes—the sequences from total RNA—from different depths at 121 different locations. Using innovative bioinformatics strategies to identify distant homologs of the RNA-directed RNA polymerase (RdRp)—a protein found only in orthornavirans RNA viruses—the authors doubled the number of orthornaviran phyla from 5 to 10. From there, they reconstructed a robust phylogenetic tree that revealed new insights into the evolution of RNA viruses.

Most known RNA viruses infect eukaryotes, with very few infecting bacteria and none infecting archaea (see the figure). However, retroelements have been found in both eukaryotes and prokaryotes, which include bacteria and archaea. Retroelements are RNA genetic elements that can move to new locations in the genome, which they do through an RNA intermediate similar to orthornavirans. This behavior suggests that there is an early origin of the RdRp (*8*). Zayed *et al.* discovered a globally distributed phylum, "*Taraviricota*," that provides the missing link for the evolutionary origins of RNA viruses with regard to retroelements. They propose that retroelements and *Taraviricota* viruses share a common ancestor. This ancestor could be a capsidless RNA replicon, as opposed to viruses, which have an outer shell called a capsid.

The presence of retroelements in all domains of life but the absence of RNA viruses in archaea suggests an important path of evolution resulting from the separation of cellular life from the last universal cellular ancestor (LUCA). All cellular life shares a set of universal genes hypothesized to have been inherited from the LUCA, which was likely a complex community of organisms that shared features of both bacteria and archaea (*9*). It is hypothesized that the LUCA's virome was a complex assemblage of viruses that included both DNA and RNA viruses, indicating that viruses had several points of origin before the LUCA (*9*).

The divergence of cellular life has played a major role in RNA virus evolution. When cellular life evolved to include a nucleus and an endomembrane system—both distinct features of eukaryotes—it created a barrier for DNA virus replication while creating a favorable niche for RNA virus replication (*10*, *11*). Studying viral evolution highlights the importance of the coevolution of viruses with their hosts and the implications it has in understanding the evolution of life. Viral evolution is more complicated than simply tracking the presence and absence of viruses that infect the three distinct domains. In all, there are five major branches of viruses that are distinguished by the nucleic acid they use for their genomes: double-stranded DNA (dsDNA), single-stranded DNA (ssDNA), double-stranded RNA (dsRNA), and positive- and negative-sense single-stranded RNA (+ssRNA and –ssRNA, respectively), which indicate polarity in respect to messenger RNA (*12*). Orthornavirans, the group of RNA viruses investigated by Zayed *et al.*, include dsRNA, +ssRNA, and –ssRNA, suggesting that all three share a common evolutionary origin pre-LUCA.

It is difficult for multiple reasons to track the inheritance of genes between lineages of viruses. Gene acquisition is not linear, and viruses can acquire genes from cellular hosts and other viruses (*11*). Additionally, viral genomes have high mutation rates that lead to rapid evolution (*13*). To further complicate matters, viruses do not share universal genes with highly conserved sequences and functions, such as the ribosomal RNA genes found in cellular life that inform the tree of life (*14*). Even though many groups of viruses share genes, this does not necessarily translate to a common ancestor. There are genes referred to as viral hallmark genes that are shared among two or more branches of viruses, which is the case for the RdRp of orthornavirans. Studies such as that by Zayed *et al.* create connections between viral and cellular worlds, allowing for the possibility of a fully integrated tree of life and a more complete understanding of the origins and evolution of all life. ■

Department of Marine Biology, Texas A&M University at Galveston, Galveston, TX, USA. Email: labontej@tamug.edu

REFERENCES AND NOTES

1. A. A. Zayed *et al.*, *Science* **376**, 156 (2022).
2. C. A. Suttle, *Nat. Rev. Microbiol.* **5**, 801 (2007).
3. L. F. Jover, T. C. Effler, A. Buchan, S. W. Wilhelm, J. S. Weitz, *Nat. Rev. Microbiol.* **12**, 519 (2014).
4. J. R. Brum *et al.*, *Science* **348**, 1261498 (2015).
5. A. C. Gregory *et al.*, *Cell* **177**, 1109 (2019).
6. S. Roux *et al.*, bioRxiv 10.1101/053090 (2016).
7. S. Sunagawa *et al.*, *Nat. Rev. Microbiol.* **18**, 428 (2020).
8. S. T. de Farias *et al.*, *Front. Genet.* **8**, 125 (2017).
9. E. V. Koonin, *Nat. Rev. Microbiol.* **1**, 127 (2003).
10. M. Krupovic, V. V. Dolja, E. V. Koonin, *Nat. Rev. Microbiol.* **18**, 661 (2020).
11. E. V. Koonin *et al.*, *Microbiol. Mol. Biol. Rev.* **84**, e00061 (2020).
12. A. E. Gorbalenya *et al.*, *Nat. Microbiol.* **5**, 668 (2020).
13. S. Duffy, L. A. Shackelton, E. C. Holmes, *Nat. Rev. Genet.* **9**, 267 (2008).
14. L. A. Hug *et al.*, *Nat. Microbiol.* **1**, 16048 (2016).

10.1126/science.abo5590

MEDICINE

Unlocking the secrets to Janus kinase activation

The full-length structure of a Janus kinase provides insights for drug development

By **Ross L. Levine**[1] *and* **Stevan R. Hubbard**[2]

Members of the Janus family of nonreceptor tyrosine kinases (JAK1, JAK2, JAK3, and TYK2) transmit a diversity of ligand-mediated signals, from cytokines and hormones, resulting in activation of downstream signaling pathways and alterations in gene expression. They have important roles in key physiologic functions, including hematopoiesis and immune effector function. Additionally, aberrant activation of JAK signaling plays a critical role in various disease states, including autoimmune disorders and various malignancies. This has led to the development of small-molecule JAK inhibitors, which provide therapeutic benefit to patients with inflammatory diseases, including rheumatoid arthritis (*1*). However, new approaches are needed to inhibit JAK signaling in other diseases. On page 163 of this issue, Glassman *et al.* (*2*) present a full-length JAK structure, which provides a structural roadmap for understanding the regulatory mechanisms that govern JAK activity and the promise of new therapeutic approaches.

Activating point mutations and fusion events that involve JAKs have been identified in different human cancers. The most common oncogenic events target JAK2, most frequently through $Val^{617}\rightarrow$ Phe (V617F) substitution in most patients with myeloproliferative neoplasms (MPNs) (*3*). First-generation JAK inhibitors show JAK2 inhibitory efficacy in the MPNs myelofibrosis and polycythemia vera (*4*). Although JAK2 inhibitors can improve disease parameters in MPN patients,

[1]Human Oncology and Pathogenesis Program, Memorial Sloan Kettering Cancer Center, New York, NY, USA. [2]Department of Biochemistry and Molecular Pharmacology, New York University Grossman School of Medicine, New York, NY, USA. Email: leviner@mskcc.org; stevan.hubbard@nyulangone.org

they do not lead to regression, suggesting a need for new strategies to inhibit JAKs. The major limitation in informing new therapeutic approaches has been sufficient insight into JAK regulatory mechanisms, particularly how the V617F mutation causes ligand-independent JAK2 signaling.

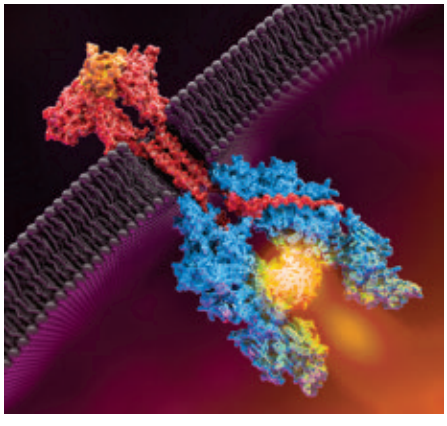
JAK proteins, which bind to the cytoplasmic region of cytokine receptors, comprise four domains: a FERM (band 4.1, ezrin, radixin, moesin) domain, a Src homology-2–like (SH2L) domain, a pseudokinase domain (PKD), and a carboxyl-terminal tyrosine kinase domain (TKD). The first step in JAK signaling is activation of the TKD, which occurs through reciprocal transphosphorylation of two tyrosines in the TKD activation loop, mediated by cytokine-dependent receptor dimerization. Once activated, JAKs phosphorylate the cytokine receptor itself and subsequently the signal transducer and activator of transcription (STAT) proteins that are recruited to the tyrosine-phosphorylated receptors.

There have been considerable efforts to determine the three-dimensional structure of a full-length JAK. As often is the case, this work initially went piece by piece, starting with the crystal structure of the JAK3 TKD in 2005 (*5*), followed by the JAK2 PKD in 2012 (*6*), and then the crystal structure of the integrated FERM-SH2L domains in 2014 (*7*). However, in the absence of a full-length JAK structure, it was unclear how the various domains cooperated to regulate JAK activity.

A crystal structure and molecular dynamics–derived model of the PKD-TKD of TYK2 (*8*) and JAK2 (*9*), respectively, revealed an autoinhibitory interaction that rationalized activating mutations found in human cancers, such as Arg^{683}→Gly (R683G) in the PKD and Asp^{873}→Asn (D873N) in the TKD. However, V617F (in the PKD) was conspicuously absent from the PKD-TKD interface. Several mutagenesis studies established that the hyperactivity of V617F and other mutants was quashed by additional mutations in the PKD, such as Phe^{595}→Ala (F595A) (*10*) or those that destabilized adenosine triphosphate (ATP) binding to the PKD (*11*). These studies suggested that the PKD is directly involved in JAK2 dimerization and that V617F enhances dimerization, a hypothesis confirmed with single-molecule fluorescence studies in cells, which demonstrated that expression of the JAK2-V617F mutant caused a substantial increase in the basal (no cytokine) level of dimerized JAK2 on cytokine receptors (*12*).

What does the cryo–electron microscopy (cryo-EM) structure of full-length JAK1 from Glassman *et al.* tell us? The structure provides the mechanism by which V658F in human JAK1 (equivalent to V617F in JAK2) leads to cytokine-independent signaling. The structure shows a parallel JAK1 homodimer that is mediated solely by a PKD-PKD interaction, with F658 in the heart of the dimer interface. The authors show that for wild-type JAK1 (V658), the PKD-PKD interaction would not be as snug as with F658, which explains why normal signaling by wild-type JAKs requires the assistance of cytokine-mediated receptor dimerization for activation.

The structure also shows how the PKD interacts with the FERM-SH2L domains—experimental information that was completely lacking. Notably, AlphaFold [artificial intelligence–based protein structure prediction (*13*)] predicted the correct FERM-SH2L-PKD interaction. Now that this prediction has been verified by the JAK1 cryo-EM structure, there is a reliable structural model for the autoinhibited form of JAKs, based on the AlphaFold structure prediction for TYK2.



Model of an activated Janus kinase dimer (blue; pseudokinase domain highlighted in yellow) interacting with a dimerized receptor (red) bound to a ligand (orange).

The JAK1 cryo-EM structure also depicts a previously unknown interaction between the TKD and the PKD with a limited interface. It is conceivable that the addition of a nanobody to promote receptor dimerization could have influenced the positions of the TKDs in the structure. The AlphaFold predictions for full-length JAK1, JAK2, and JAK3 feature an elongated structure, similar to the JAK1 cryo-EM structure, but with a different PKD-TKD interaction. Further structural and mutagenesis analyses will be required to determine whether and how the TKD interacts with the PKD to facilitate various phosphorylation events.

The cryo-EM structure of JAK1 provides the crucial missing piece in the JAK activation process: the specific JAK dimer configuration that triggers TKD transphosphorylation and downstream signaling. The model for JAK activation that has emerged is as follows: JAK molecules bound to monomeric cytokine receptors undergo conformational equilibrium between a closed, autoinhibited state and an open state in which the PKD is accessible for homotypic (JAK2-JAK2) or heterotypic (all JAKs) dimerization, and the TKD is available to serve as enzyme or substrate in a transphosphorylation event. For wild-type JAKs in the absence of cytokine, equilibrium favors the closed state, establishing a low basal kinase activity. For cytokine-mediated receptor dimerization, or for activating mutations such as JAK2 V617F in the absence of cytokine, the equilibrium is shifted toward the PKD-mediated dimerized state and concomitant TKD transphosphorylation.

Initial efforts to inhibit JAKs focused on the ATP-binding pocket of the TKD, but attention has recently shifted to the same pocket in the PKD, and compounds that bind to the TYK2 PKD are in phase 2 and phase 3 clinical trials for psoriasis, Crohn's disease, and psoriatic arthritis. How these PKD-targeted compounds actually inhibit JAKs is not understood, but the cryo-EM structure of dimeric JAK1 will provide important clues and enable other therapeutic strategies for mitigating hyperactive signaling by JAKs. More broadly, the study of Glassman *et al.* sets the stage for structural analyses of additional complexes—including cytokine receptors, JAK heterodimers, and STAT proteins—paving the way for a comprehensive structure-based approach to abrogate pathologic JAK-STAT signaling in a spectrum of human diseases. ■

REFERENCES AND NOTES

1. A. Rubbert-Roth *et al.*, *N. Engl. J. Med.* **383**, 1511 (2020).
2. C. R. Glassman *et al.*, *Science* **376**, 163 (2022).
3. C. James *et al.*, *Nature* **434**, 1144 (2005).
4. S. Verstovsek *et al.*, *N. Engl. J. Med.* **366**, 799 (2012).
5. T. J. Boggon, Y. Li, P. W. Manley, M. J. Eck, *Blood* **106**, 996 (2005).
6. R. M. Bandaranayake *et al.*, *Nat. Struct. Mol. Biol.* **19**, 754 (2012).
7. H. J. Wallweber, C. Tam, Y. Franke, M. A. Starovasnik, P. J. Lupardus, *Nat. Struct. Mol. Biol.* **21**, 443 (2014).
8. P. J. Lupardus *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **111**, 8025 (2014).
9. Y. Shan *et al.*, *Nat. Struct. Mol. Biol.* **21**, 579 (2014).
10. A. Dusa, C. Mouton, C. Pecquet, M. Herman, S. N. Constantinescu, *PLOS ONE* **5**, e11157 (2010).
11. H. M. Hammarén *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **112**, 4642 (2015).
12. S. Wilmes *et al.*, *Science* **367**, 643 (2020).
13. J. Jumper *et al.*, *Nature* **596**, 583 (2021).

ACKNOWLEDGMENTS

R.L.L. is on the supervisory board of Qiagen and is a scientific advisor to Imago, Mission Bio, Zentalis, Ajax, Auron, Prelude, C4 Therapeutics, and Isoplexis, for which he has received equity. R.L.L. receives research support from and has consulted for Celgene and Roche and has consulted for Incyte, Janssen, Astellas, Morphosys, and Novartis. He has received honoraria from Roche, Lilly, and Amgen for invited lectures and from Gilead for grant reviews. S.R.H. is a cofounder and scientific advisory board member of Ajax Therapeutics, for which he has received equity.

10.1126/science.abo7788

ILLUSTRATION: CHRIS GARCIA AND ERIC SMITH

BOOKS *et al.*

A diver swims over a coral reef near the Meemu Atoll in the Maldives.

PHOTO: REINHARD DIRSCHERL/ULLSTEIN BILD/GETTY IMAGES

MARINE ECOLOGY

Investing in what matters most

Faced with a family crisis, a marine scientist finds parallels with Earth's imperiled coral reefs

By **Steven Mana'oakamai Johnson**

In the Anthropocene, saving the planet is a business venture. But, like the most ambitious enterprises, it is also a labor of love. In *Life on the Rocks*, Juli Berwald takes readers on a globe-trotting adventure to fight for the future of coral reefs—a world of million-dollar prizes for ecosystem-saving breakthroughs and where a trillion dollars could potentially save one-tenth of Earth's reefs. Berwald weaves into this narrative her own story of coping with a mental health issue afflicting her daughter, drawing parallels with the story of the world's coral reefs.

Life on the Rocks: Building a Future for Coral Reefs
Juli Berwald
Riverhead Books, 2022. 352 pp.

The primary narrative of *Life on the Rocks* centers around the "badass merger" established by the coral animal (*Scleractinia*) and its endosymbiotic microalgae (*Symbiodiniaceae*). Corals and their endosymbionts provide each other with the ecological and evolutionary ingredients for success. This partnership, rooted in cooperation and coordination, is so beneficial to both parties that it is responsible for the most biodiverse ecosystem on the planet. But this symbiosis exists at its physiological limits and begins to break down under various stressors, resulting in coral bleaching (when the microalgae and coral host part ways, leaving only the ghostly white skeleton seen through the translucent body of the coral).

The leading cause of coral bleaching is elevated ocean temperatures. Climate change is increasing the frequency of severe bleaching events, and the long-term projections are bleak: Annual bleaching is expected to occur in nearly every coral reef by 2050. Having witnessed and documented the 2013–2017 global bleaching event from the reefs of my home island, Saipan, I can attest that coral reefs will be in bad shape long before then if something does not change.

Berwald may earn a living as a science writer, but she holds a PhD in marine science. This insider knowledge and expertise give her the ability to deconstruct the "wicked problem" of climate change with a stylistic ease reminiscent of Elizabeth Kolbert and Susan Casey. She exposes the breadth and depth of the challenges facing reefs, introducing readers to the businesspeople and scientists collaborating to save them along the way. We meet a candy bar–making CEO in Bali who creates an army of "spider" frames for coral gardening, for example, and accompany Berwald as she visits Caribbean hotels hosting coral "refugees" that are fleeing a devastating disease outbreak. Here, the author challenges readers to focus on the generosity of business entities in taking up this noble pursuit, eliding (temporarily) their own substantial contributions to climate change.

The menagerie of scientists Berwald talks to highlights the diversity of people committing their life's work to coral reef conservation and their various attitudes and outlooks on coral futures. Misha Matz, a researcher at the University of Texas at Austin, comes off as quite cheery. His models suggest that we "cannot make [coral] go extinct…it's impossible." Meanwhile, Megan Morikawa, a coral geneticist who works for the Spanish hotel chain Iberostar, is pragmatic about our need to feel like we are helping, even when other actions might be more effective. Tourists concerned about using reef-friendly sunscreens, for example, rarely confront the carbon emissions associated with their flights to tropical locales.

Berwald frequently finds ways to show that we are more like our planet than we often appreciate. Weathering, seasons, fires, and floods—these processes can be jarring and violent, but they help reveal Earth's potential, creating the diversity that enriches our planet. The same can be said of our own personal crises. According to Berwald, her daughter's mental health only turned the corner when the family decided that her well-being was priceless and committed to the long process of a systemic reset. Saving reefs will require the same thing from us.

If the book has one weak point, it is that the perspectives and relationships of Indigenous people to coral reefs only make brief appearances. But Berwald is right on the cusp of engaging with these worldviews with her discussion of reticulated evolution. This process, which involves "separation and repackaging, divergence and convergence," just might be the conceptual framework we need in this moment. What would it look like to overcome the legacies of colonialism that persist in the form of capitalism and climate change? Can we pair the technological advancements of the 21st century with Indigenous worldviews that continue to be cast out and marginalized?

Partnerships have trade-offs. Some symbionts help corals succeed under one set of conditions, for example, but not under others. Will the partnerships highlighted in this book be the ones that save coral reefs? We don't have much time, so let's hope so. ■

The reviewer is at the School of Geographical Sciences and Urban Planning, Arizona State University, Tempe, AZ 85287, USA. Email: stevenjohnson@asu.edu

10.1126/science.abo5005

BIOLOGY

Gender, biology, and behavior

There is much to learn from a primatologist's framework for gender diversity

By **Barbara J. King**

Frans de Waal's books, beginning with *Chimpanzee Politics* (1982) up through *Mama's Last Hug* (2019), illuminate for wide audiences the social lives, cognitive abilities, and emotions of animals, with an emphasis on monkeys and apes. Now, in his 13th volume, *Different*, de Waal takes on a fresh and controversial topic: the contribution of biology to gender in humans.

This subject is an "ideological minefield," de Waal notes, and he wonders whether writing this book is one of his "most foolish decisions." But there is much to learn from his perspective. A smart interactionist framework, in which biological and socioenvironmental influences on human behavior are entwined, sits at the book's heart.

De Waal embraces gender variability even as he describes evolutionary influences on gender by comparing humans to our closest living relatives, chimpanzees and bonobos. He misses opportunities, however, to educate his readers accurately when he underestimates ways in which humans vary by both gender identity and sex.

"I sincerely believe that the best way to achieve greater equality will be to learn more about our biology instead of trying to sweep it under the rug," he writes, urging the retirement of theories that see gender as socially constructed, full stop. "The most meaningful expressions of gender have deeper roots, including the generally greater physical combativeness of men or the devotion of many women to children."

Already, readers of this review may be bristling. Aren't these inaccurate, outdated gender stereotypes? Yet de Waal offers solid evidence to show that across primates—including our own species—physical violence is associated with males far more often than with females, and attraction to infants with females far more often than males, in alignment with evolutionary pressures that differ by sex. Along the way, he effectively deploys anecdotes from primate research to emphasize how far biology is from determinist.

Male chimpanzees may become violent, but they can also display cooperation. In the Burgers' Zoo chimp colony in the Netherlands, for example, alpha male Nikkie showed signs of being a potential threat to the infant Roosje, who—along with her adoptive mother—was being reintroduced to the group after a period apart. Nikkie was at first held back, and when he was finally released, the colony's two oldest males did something unexpected: They positioned themselves strategically between Roosje and her mother and the path of Nikkie's approach, with arms wrapped around each other's shoulders. "This was a sight to behold, given that these two had been arch enemies for years," de Waal notes. Nikkie carried out no violence that day and, in fact, acted gently around Roosje.

A bonobo mother looks on as her baby nurses.

Meanwhile, chimpanzee Donna at the Yerkes Field Station in the United States was a physically "robust" female with broad shoulders "who acted more masculine than other females." Donna never exhibited full sexual swelling at the time of ovulation, as is typical, nor did she mate or have offspring. De Waal concludes that Donna was a "largely asexual gender-nonconforming individual," a remarkable inference that extends concepts usually reserved for humans into the nonhuman world.

Different: Gender Through the Eyes of a Primatologist
Frans de Waal
Norton, 2022. 408 pp.

De Waal speaks up repeatedly for the rights of women as well as transgender and gay people (overlapping categories, of course). However, his language is at times troublesome. His definition of gender, for example, involves "culturally encouraged sex roles in society." This is fine, but he also writes that "gender refers to the learned overlays that turn a biological female into a woman and a biological male into a man," a statement in line with his view that sex is all about biology and that we can assert another person's sex in "only a second" just by looking.

To problematize the notion that a person's sex or gender is detectable on sight as an exemplar of a fixed biological sex goes beyond mere political correctness. For one thing, intersex individuals, whom de Waal mentions in passing, have genital tissue and/or chromosomes that depart from what is considered typical presentation. They are biologically neither exclusively male nor exclusively female. As biologist Anne Fausto-Sterling pointed out years ago, a person's sex may be socially constructed just as one's gender is.

Moving into the realm of gender identity, a person who is assigned female at birth may be assumed by others to be a woman on the basis of external cues, whereas in fact they may identify not as a woman but as nonbinary and agender. Similarly, transgender people may not inevitably feel that they "belong to the opposite sex." Such language fails to recognize people who identify, for example, as both man and woman, or neither.

Overall, however, *Different* offers a fascinating and mostly forward-thinking look at the biology and culture of human gender by an esteemed primatologist. Occasionally, it requires correction to reflect the full range of human sex and gender diversity. ■

The reviewer is professor emerita of anthropology at William & Mary, Williamsburg, VA 23185, USA.
Email: bjking@wm.edu

10.1126/science.abo1569

PHOTO: FRANS DE WAAL

gairdner
LES PRIX CANADA GAIRDNER AWARDS

CONGRATULATIONS

TO THE 2022 CANADA GAIRDNER AWARD LAUREATES

2022 CANADA GAIRDNER **INTERNATIONAL AWARD**

Dr. Stuart H. Orkin

Awarded for “the discovery of the molecular mechanism responsible for the switch from fetal to adult hemoglobin gene expression during human development and translating that knowledge into a novel treatment for the hemoglobin disorders – sickle cell disease and beta-thalassemia.”

Dr. John E. Dick

Awarded for “the discovery and characterization of leukemic stem cells, providing insights into the understanding, diagnosis and treatment of acute myeloid leukemia”



Dr. Pieter Cullis, Dr. Katalin Karikó, and Dr. Drew Weissman

Awarded “For their pioneering work developing nucleoside-modified mRNA and lipid nanoparticle (LNP) drug delivery: the foundational technologies for the highly effective COVID-19 mRNA vaccines”

2022 JOHN DIRKS CANADA GAIRDNER **GLOBAL HEALTH AWARD**

Dr. Zulfiqar Bhutta

Awarded “For the development and evaluation of evidence-based interventions in child and maternal health for marginalized populations, focusing on outcomes for the ‘first thousand days’ of life.”

2022 CANADA GAIRDNER **WIGHTMAN AWARD**

Dr. Deborah J. Cook

Awarded for “pioneering research that has developed and defined evidence-based critical care medicine in Canada, informing best practices around the world.”

CELEBRATING EXCELLENCE
CONVENING LEADERS
INSPIRING THE NEXT GENERATION

Water extractions for greenhouse agriculture have drained the aquifer under Doñana National Park in Spain.

Edited by **Jennifer Sills**

Spain's Doñana World Heritage Site in danger

Spain's Doñana National Park, established in 1969, was listed as a UN Educational, Scientific, and Cultural Organization (UNESCO) World Heritage Site in 1994 in recognition of its wide range of habitats, including seasonal ponds, lagoons, and marshlands, and its biodiversity (*1*). A diverse combination of European and African flora and fauna inhabit the park, including many endemic species (*2*). Doñana also supports several migratory waterbird populations (*3*), many of which are globally threatened and show long-lasting declines despite increasing international investments in their conservation (*4*, *5*). However, human activities and environmentally questionable political decisions have put Doñana at risk.

As in most Mediterranean wetlands, the availability of shallow lagoons for waterbird populations critically depends on groundwater discharges from the main aquifer (*6*). For more than two decades, Doñana has been drying out. Although rising temperatures and rainfall shortages contribute to this trend, the World Heritage Committee has determined that Doñana's shrinking aquifer is primarily the result of groundwater pumping and upstream retrieval of river water for intensive agricultural purposes (*7*, *8*), particularly greenhouse-grown blueberries and strawberries.

The threat that agriculture poses to Doñana was recognized by the Court of Justice of the European Union last year (*9*). The court reminded the Spanish government of its obligation to protect Doñana from illegal water extractions. Yet, despite the Spanish government's opposition, the regional parliament of Andalucía has approved a proposal to amnesty and legalize unregulated groundwater pumping (*10*).

Legalizing unregulated groundwater pumping may well be a death sentence for Doñana, known as the "jewel in the crown" of Mediterranean biodiversity hotspots (*6*). The citizens of Andalucía should demand that their government consider the environmental risks of the groundwater pumping proposal before ratifying it. Instead of approving unregulated groundwater pumping, the government should give rights to access surface irrigation water exclusively to farmers who are operating legally. As food retailers have suggested (*11*), the international community, which serves as the market for Doñana's berries, should leverage economic power to ensure that the products they consume come from sustainable agriculture and do not threaten Doñana or other protected areas. Finally, we urge UNESCO to add Doñana to the List of World Heritage in Danger (*12*).

Juan G. Navedo[1,2]*, Theunis Piersma[3,4], Jordi Figuerola[5], Wouter Vansteelant[3,5]
[1]Área de Zoología, Universidad de Extremadura, Badajoz, Spain. [2]Bird Ecology Lab, Instituto de Ciencias Marinas y Limnológicas, Universidad Austral de Chile, Valdivia, Chile. [3]Groningen Institute for Evolutionary Life Sciences, University of Groningen, Groningen, Netherlands. [4]Department of Coastal Systems, NIOZ Royal Netherlands Institute for Sea Research, Den Burg, Texel, Netherlands. [5]Estación Biológica de Doñana, Consejo Superior de Investigaciones Científicas, Sevilla, Spain.
*Corresponding author. Email: jgnavedo@unex.es

REFERENCES AND NOTES

1. UNESCO, World Heritage List, Doñana National Park; https://whc.unesco.org/en/list/685.
2. C. M. Finlayson, G. R. Milton, R. C. Prentice, N. C. Davidson, *The Wetland Book II: Distribution, Description and Conservation* (Springer, 2016).
3. M. A. Rendón, A. J. Green, E. Aguilera, P. Almaraz, *Biol. Conserv.* **141**, 1371 (2008).
4. R. Kentie *et al.*, *Ardea* **114**, 213 (2016).
5. M. van Roomen *et al.*, *Ocean Coast. Manage.* **68**, 79 (2012).
6. C. Martín-Cabrera, F. García-Novo, Eds., *Doñana, Water and Biosphere* (Ministerio de Medio Ambiente, Gobierno de España, 2005).
7. A. J. Green *et al.*, *Front. Ecol. Environ.* **15**, 99 (2017).
8. World Heritage Committee, "WHC/21/44.COM/18" (2021), p. 291; https://whc.unesco.org/archive/2021/whc-21-44com-18-en.pdf.
9. EU Court Sentence ECLI:EU:C:2021:512 (2021); www.stradalex.com/en/sl_src_publ_jur_int/document/cjeu2021_C_559_19_57.
10. Diario de Sesiones del Parlamento de Andalucía, Sevilla, Núm. 124, XI LEGISLATURA (2022); www.parlamentodeandalucia.es/webdinamica/portal-web-parlamento/actividadparlamentaria/tramitacionencurso/legislativas.do?numexp=11-22/PPL-000003& [in Spanish].
11. World Wildlife Fund, "Los principales supermercados europeos piden al Presidente Juan Manuel Moreno Bonilla que abandone el plan para legalizar el regadío ilegal en Doñana" (2022) [in Spanish].
12. UNESCO, World Heritage in Danger (https://whc.unesco.org/en/danger/).

10.1126/science.abo7363

US conservation atlas needs biodiversity data

The US administration has proposed a Conservation and Stewardship Atlas that would facilitate the conservation of 30% of US lands and waters by 2030 (30x30) under its "America the Beautiful" initiative (*1*). To maximize the benefits of the initiative, decisions about which lands to prioritize for conservation and restoration should be based on not only an area's current protection and management status (*2*) but also its potential to safeguard the nation's biodiversity (*3*–*5*). A rigorous system to coordinate the collection and interpretation of spatial biodiversity data would facilitate informed decisions.

The 30x30 target is an element of the Global Biodiversity Framework, which will be finalized at the 2022 meeting of the Convention on Biological Diversity (*6*). The Global Biodiversity Framework emphasizes that the 30x30 target will effectively address the biodiversity crisis only if placement and management of conserved areas are coordinated with efforts to achieve targets for halting loss of ecosystems (*7*), species (*8*), and genetic diversity (*9*). Therefore, where conserved lands are located and what biodiversity they support

PHOTO: EBD-CSIC IMAGE BANK

are as important as how much area is conserved (*3*–*9*). The United States, although not formally a party to the convention, has made high-level commitments to achieve 30x30 through "ecologically representative and well-connected" areas that "deliver the greatest benefits for global biodiversity, ecosystem services and climate protection" (*10*), but the metrics by which such benefits will be assessed remain undefined.

Although the America the Beautiful initiative includes a goal of tracking "fish and wildlife habitats and populations," little detail is provided as to what spatial data will be collected and how it will inform decisions (*1*). The administration should use this opportunity to support and better coordinate efforts to track the status and distribution of the nation's ecosystems and species within a coherent and evidence-based framework. This effort should build on existing data developed by federal, state, nonprofit, and tribal remote sensing and species monitoring programs (*11*). Ideally, the data would be synthesized in a manner similar to the Global Biodiversity Framework. The administration would then have the information required to make effective decisions about lands in need of conservation and restoration. Synthesized data would also help to assess conservation goals moving forward (*5*–*9*), as suggested in recent proposals for development of a National Biodiversity Assessment and Strategy analogous to the existing quadrennial National Climate Assessment (*12*).

C. Carroll[1]*, R. F. Noss[2], Bruce A. Stein[3]
[1]Klamath Center for Conservation Research, Orleans, CA 95556, USA. [2]Florida Institute for Conservation Science, Melrose, FL 32666, USA. [3]National Wildlife Federation, Washington, DC 20005, USA.
*Corresponding author.
Email: carlos@klamathconservation.org

REFERENCES AND NOTES

1. US Department of the Interior, "Request for information to inform interagency efforts to develop the American Conservation and Stewardship Atlas," *Fed. Reg.* **87**, 235 (2022); www.govinfo.gov/content/pkg/FR-2022-01-04/pdf/2021-28548.pdf.
2. J. Yachnin, "Does Biden's '30x30' plan trade science for popularity?" *E&E News* (2021); www.eenews.net/stories/1063734011.
3. C. Carroll, R. F. Noss, *Conserv. Biol*. 10.1111/cobi.13869 (2021).
4. E. Dinerstein *et al*., *Sci. Adv*. **5**, eaaw2869 (2019).
5. W. Jetz *et al*., *Nat. Ecol. Evol*. **6**, 123 (2021).
6. Convention on Biological Diversity (CBD), "First draft of the post-2020 Global Biodiversity Framework, CBD/WG2020/3/3" (CBD, Montreal, 2021).
7. E. Nicholson *et al*., *Nat. Ecol. Evol*. **5**, 1338 (2021).
8. M. D. A. Rounsevell *et al*., *Science* **368**, 1193 (2020).
9. S. Hoban *et al*., *Biol. Conserv*. **248**, 108654 (2020).
10. Group of 7, "G7 Climate and Environment Ministers' Meeting Communiqué" (2021).
11. H. Hamilton *et al*., *Ecol. Appl*. 10.1002/eap.2534 (2022).
12. "Expressing the need for the Federal Government to establish a national biodiversity strategy for protecting biodiversity for current and future generations," H. Res. 69, 117th Cong. (2021).

10.1126/science.abo0526

LIFE IN SCIENCE

Drifting away in the Atlantic

Floating 20 meters below the surface, in the warm Atlantic waters around the desert islands of the Cape Verde archipelago, my dive buddy and I beheld a marine biologist's dream: schools of thousands of fish, sharks, and a beautiful underwater landscape of barely explored habitats. We were 2 weeks into a sailing trip retracing the steps of the second voyage of the HMS *Beagle*, the ship that carried Charles Darwin around the world in the 1830s. Our mission was to explore how biodiversity had thrived in the relative absence of human pressure. We had jumped from the deck of our boat, the *Captain Darwin*, just 1 hour before. The current was stronger than expected, and dusk was approaching, but we lingered as long as we could to enjoy the amazing view. As we ascended, we tried to express to one another how incredible the dive had been using hand signs and screaming and laughing into the water. Then we surfaced and looked around for our boat. It was gone.

We quickly realized that in the little time we had been underwater, we had drifted miles from our drop-off point. The boat's crew could not possibly spot us from such a distance. Although there was land nearby, we had no way to access the rocky shore, pummeled by powerful waves. As the sun set, we waved our long, red surface marker buoy high up in the air and blew our safety whistles, but to no avail. The sky darkened, and stories of regional spearfishermen taken away by the current came to mind. We tried shining our flashlights on the water to illuminate a large area that might be visible to the crew. Still, no help arrived.

Then we had an idea. We placed our flashlights inside the buoy. Bobbing in the now pitch-black ocean, the marker glowed with an intense red color, mimicking the inflatable tube figures that dance in the wind in front of car dealerships. We waited nervously, listening to the waves and gazing toward the dark horizon. Finally, we saw it: A masthead light, heading straight toward our bright red beacon.

Eduardo Sampaio
Department of Collective Behavior, Max Planck Institute of Animal Behavior, University of Konstanz, Konstanz, Germany; Centre for the Advanced Study of Collective Behaviour, University of Konstanz, Konstanz, Germany; Department of Biology, University of Konstanz, Konstanz, Germany. Email: esampaio@ab.mpg.de

10.1126/science.abo5607

Call for submissions
Life in Science is an occasional feature highlighting some of the humorous or unusual day-to-day realities that face our readers. Can you top this? Submit your story to www.submit2science.org.

A diver views the *Captain Darwin* from a distance.

PHOTO: VICTOR RAULT/*CAPTAIN DARWIN*

PRIZE ESSAY

GRAND PRIZE WINNER
Benedetto Marelli

Benedetto Marelli received undergraduate degrees from Politecnico di Milano and a PhD from McGill University. After completing his postdoctoral fellowship at Tufts University, he started his laboratory in the Department of Civil and Environmental Engineering at the Massachusetts Institute of Technology in late 2015. His research focuses on nanomanufacturing of structural biopolymers to engineer a new generation of advanced materials that can be interfaced with food and plants. **www.science.org/doi/10.1126/science.abo4233**

BII Science | Prize for Innovation

INNOVATION

Biomaterials for boosting food security

Renewable silk-protein technologies promote plant growth and reduce food waste

By **Benedetto Marelli**

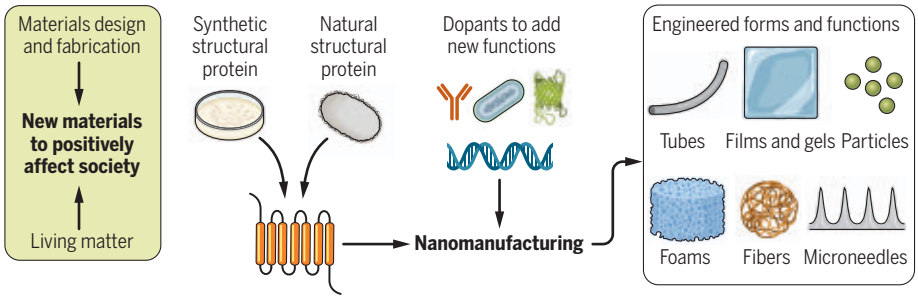
In the 20th century, new material-based technologies have positively affected many aspects of human life—including health management, communication, education, and transport—as well as improved our access to energy, water, and food. Continued technological advancements to improve quality of life must now consider sustainability alongside mitigation of and adaptation to climate change (*1*). Scientists and engineers are looking to living systems to learn how to translate sustainability principles into material design. Soft matter and structural biopolymers (e.g., polysaccharides, proteins, and DNA) are being used to design technologies that address unmet challenges in the health, energy, food, and education sectors. These natural polymers are biomaterials that can be extracted in high volumes and at low cost from by-products of food and textile industries and upscaled into advanced materials (see the figure).

There is wide interest in the development of biomaterials, but their application in agro-food systems (i.e., all actors and activities involved in food production, distribution, regulation, and consumption) has lagged. The infrastructure of agro-food systems is responsible for more than 25% of anthropogenic greenhouse gas (GHG) emissions. These systems face pressure to support an increasing world population and to simultaneously minimize inputs (e.g., water, fertilizers, pesticides) and mitigate environmental impact. For the first time in history, the availability of arable land has plateaued, and crop yields are

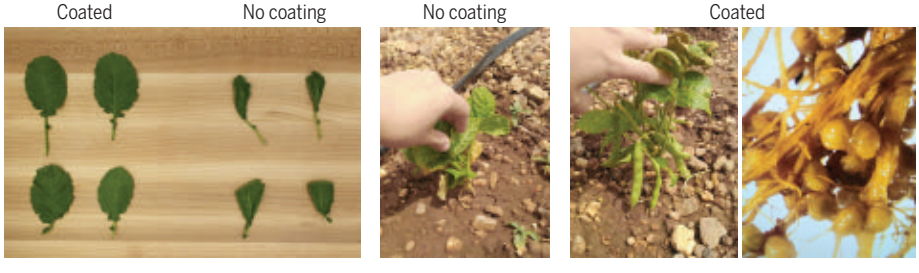
From mission to materialization

The Marelli Laboratory's long-term research mission flows into its general process to engineer structural proteins in advanced biomaterials (top). New discoveries in biomaterials science spur the translation of new solutions in agro-food systems. Examples of these are the use of silk fibroin–based technology as an edible food coating (bottom left) to extend perishable produce shelf life in kale and as a seed coating (bottom right) to deliver biofertilizers that boost germination and mitigate soil salinity.

Mission and research



Translation and impact



CREDITS: (PHOTO) KATHY BRIANA; (GRAPHIC) KELLIE HOLOSKI/*SCIENCE* BASED ON B. MARELLI

threatened by soil salinity and water scarcity—stressors that are exacerbated by climate change. Food security and food waste are twin crises; more than 800 million people are undernourished, and 30% of food is lost or wasted from farm to fork. Food waste could potentially feed 1.6 billion people; instead, it is responsible for 25% of global freshwater consumption and, when considered en masse, is the third largest producer of GHGs after China and the United States (*2*, *3*). New technologies that are economically sustainable, scalable, and rapidly deployable to market are needed to address these challenges. These innovations must also meet stringent requirements for safety and biodegradability. The environmental responsibility of consumers is increasing, and new laws that limit the environmental impact of materials are on the horizon (e.g., the European Union's ban on microplastics that begins in 2025).

An opportunity lies for biomaterials to address these challenges in the agro-food industry. Our laboratory strives to reinvent silk as an advanced material to extend food shelf life, boost crop production, and precisely deliver payloads in plants (*4*–*6*). Silk is an abundant, natural fiber produced by *Bombyx mori* caterpillars when making their cocoons. Silk fibroin, which is an edible, nontoxic protein, can be extracted at low cost from by-products of the textile industry (*7*). This protein is well known for its mechanical strength, but its structural polymorphism (i.e., the ability to fold in stable configurations ranging from a random coil to a β sheet) is ideal for applications as a technical material. The polymorphism of silk fibroin enables its low-energy, water-based regeneration in water-soluble or water-insoluble materials, dependent on molecular structure, and allows for nanomanufacturing in numerous material formats (*4*, *7*, *8*).

Our laboratory has investigated the self-assembly of regenerated silk fibroin in transparent coatings that can adhere to three-dimensional substrates through spray drying or dip coating, which are retrofitting tools commonly used in the agro-food industry (*9*–*11*). Modulation of polymorphism in silk coatings provides extraordinary barrier properties to water and oxygen as well as resistance to microbial spoilage and contamination. Payloads such as bacteria can be encapsulated and preserved in these silk coatings, and composite materials can be easily manufactured to further tailor coating properties. In the past few years, research I have contributed to has led to the spinout of technologies that use silk-based materials to enhance food security (see the figure).

We developed safe-to-eat food coatings using this protein that extend the shelf-life of perishable foods (*10*, *12*). This edible silk coating can be applied to numerous types of foods, including produce, meats, fish, and consumer packaged goods. The coating decreases evaporation and oxidative stress and may contribute to a reduction in natural microbial spoilage. In 2018, the technology spun out to a company called Mori (formerly Cambridge Crops, Inc.), which uses intellectual property (IP) developed with my research. Mori was also recognized as a 2021 World Economic Forum Technology Pioneer. Since its founding, Mori has raised upward of $88 million and currently employs more than 55 people in offices across the United States and in Mexico. The food coating is designated as "generally recognized as safe" in the United States and has obtained "non-novel" food status from Health Canada. Silk fibroin is also considered safe to eat in other countries based on the historic consumption of *B. mori* in those countries. Mori can scale up this technology and intends to continue using it to extend shelf life and build a more resilient food supply.

In addition, our laboratory has also developed a silk-based seed-coating technology for the delivery of plant growth–promoting rhizobacteria (PGPRs). PGPRs boost plant health and crop yield by increasing the availability of macronutrients, decreasing the use of synthetic fertilizers and pesticides, and mitigating abiotic stressors (*13*). The use of PGPRs is frequently hindered by limited viability outside the soil and during desiccation. By using a combination of rational design and bioinspiration, silk and polysaccharides are combined to adhere to a seed surface, encapsulate and preserve bacteria in a dry state, and modulate their delivery and growth in the spermosphere (*14*). In field tests conducted at an experimental farm in Ben Guerir, Morocco, in collaboration with Mohammed VI Polytechnic University, the delivery of PGPRs through seed coatings boosted growth when plants were grown in saline soil and under water-stress conditions (see the figure) (*15*). A technology spinout effort is underway to commercialize these coatings and have a positive impact on our society by mitigating climate and food crises.

Together, these technologies open the door to the application of biomaterials to boost food security and enhance agro-food resilience. We are bringing innovation to a field that needs creative solutions to enhance food production while minimizing inputs and mitigating environmental impacts. ■

Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, Cambridge, MA, USA. Email: bmarelli@mit.edu

FINALIST

Philipp Mews

Philipp Mews received an undergraduate degree from the Free University of Berlin and a BSc and PhD from the University of Pennsylvania. In 2017 he cofounded EpiVario, Inc., and is currently an instructor in the Neuroscience Department at the Icahn School of Medicine at Mount Sinai. His research explores the interplay between metabolism and epigenetics in the adult brain, with an emphasis on brain circuits involved in the formation and maintenance of memory. **www.science.org/doi/10.1126/science.abo4234**

FINALIST

Jacqueline Douglass

Jacqueline Douglass received her undergraduate degree from the Massachusetts Institute of Technology and an MD-PhD from Johns Hopkins University School of Medicine. She is currently an internal medicine resident at Johns Hopkins and plans to pursue medical oncology training. **www.science.org/doi/10.1126/science.abo4237**

PHOTOS: (TOP TO BOTTOM) VERONICA SZAREJKO; JOHNS HOPKINS PATHOLOGY PHOTOGRAPHY LAB

REFERENCES AND NOTES

1. National Academy of Engineering (NAE), NAE grand challenges for engineering; www.engineeringchallenges.org.
2. H. C. J. Godfray *et al.*, *Science* **327**, 812 (2010).
3. Food and Agriculture Organization of the United Nations (FAO), "The state of food security and nutrition in the world" (FAO, 2019); www.fao.org/3/ca5162en/ca5162en.pdf.
4. H. Sun, B. Marelli, *MRS Commun.* **11**, 31 (2021).
5. D. Kim *et al.*, *Adv. Funct. Mater.* **31**, 2005370 (2021).
6. Y. Cao, E. Lim, M. Xu, J.-K. Weng, B. Marelli, *Adv. Sci.* **7**, 1903551 (2020).
7. B. Marelli *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **114**, 451 (2017).
8. Z. Zhou *et al.*, *Adv. Mater.* **30**, 1706983 (2018).
9. H. Sun, B. Marelli, *Nat. Commun.* **11**, 351 (2020).
10. B. Marelli, M. A. Brenckle, D. L. Kaplan, F. G. Omenetto, *Sci. Rep.* **6**, 25263 (2016).
11. E. Ruggeri *et al.*, *ACS Sustain. Chem. Eng.* **8**, 14312 (2020).
12. S. Yigit *et al.*, *Food Chem. Toxicol.* **151**, 112117 (2021).
13. A. T. Zvinavashe, E. Lim, H. Sun, B. Marelli, *Proc. Natl. Acad. Sci. U.S.A.* **116**, 25555 (2019).
14. A. T. Zvinavashe, I. Mardad, M. Mhada, L. Kouisni, B. Marelli, *J. Agric. Food Chem.* **69**, 13270 (2021).
15. A. T. Zvinavashe *et al.*, *Nat. Food* **2**, 485 (2021).

10.1126/science.abo4233

PRIZE ESSAY

FINALIST

Philipp Mews

Philipp Mews received an undergraduate degree from the Free University of Berlin and a BSc and PhD from the University of Pennsylvania. In 2017 he cofounded EpiVario, Inc., and is currently an instructor in the Neuroscience Department at the Icahn School of Medicine at Mount Sinai. His research explores the interplay between metabolism and epigenetics in the adult brain, with an emphasis on brain circuits involved in the formation and maintenance of memory.

www.science.org/doi/10.1126/science.abo4234

BII | Prize for Science | Innovation

INNOVATION

Targeting memories to treat trauma

Blocking a metabolic enzyme controls the encoding of memories

By **Philipp Mews**

Are we really what we eat? EpiVario, cofounded by Shelley L. Berger and myself, was born of this fundamental question rooted in an age-old adage. We discovered that we—or more accurately, our brains—are, in some respects, the product of what we eat and drink. Brain functions, including memory formation, can be affected by metabolism. We found that metabolic enzymes fed by what we consume can alter gene expression in the brain's learning centers. Our company, EpiVario, was established to help bring the benefits of our discoveries to the fields of psychology and addiction studies—two areas that can have an outsized effect on societal health at large.

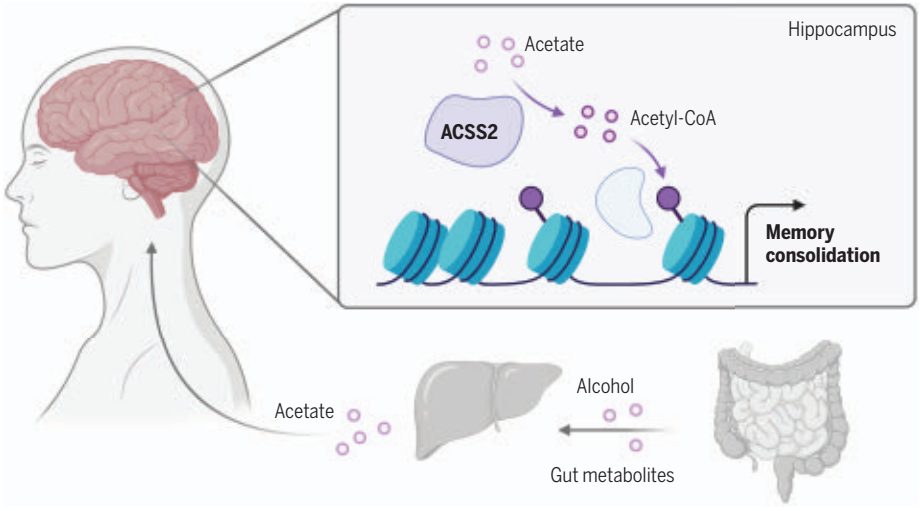
Through our metabolism, the body turns what we eat and drink into energy and molecular building blocks. Neurobiologists have traditionally viewed these metabolic processes in the body as wholly separate from the cognitive functions of the mind. Our research has helped shift this paradigm by demonstrating that metabolism directly affects learning and memory (*1*). Metabolic enzymes are emerging as key players in the nuclei of neurons, where, fueled by food metabolites, they work as engines to drive gene expression. It is this enzymatic process that activates neuronal genes whenever we learn or create a new memory.

Memories are stored in the connections between neurons, and forming a memory requires new proteins at the synapse. These proteins are encoded by neuronal genes in the cell nucleus, where the genetic material is tightly wrapped around histone proteins to form a complex called chromatin (*2*). When the compact chromatin structure is unwound, the histones are modified with chemicals called acetyl groups, leading to an increase in the production of synapse proteins (*3*).

In relation to those acetyl groups, our research in mouse brains has found a metabolic enzyme responsible for turning on the production of acetyl coenzyme A (acetyl-CoA) in memory. This enzyme, acetyl-CoA synthetase 2 (ACSS2), binds to chromatin in neurons, fueling the acetylation of histones throughout the hippocampus, the brain's memory center (*1*, *4*). This process activates genes that reshape synaptic con-

Alcohol metabolism affects memory

Alcohol is metabolized in the liver to acetate, which is released into circulation and enters the brain. In the brain, acetate is used to generate acetyl coenzyme A (acetyl-CoA) by acetyl-CoA synthetase 2 (ACSS2), boosting histone acetylation and gene expression involved in memory.



CREDITS: (PHOTO) VERONICA SZAREJKO; (GRAPHIC) CREATED WITH BIORENDER.COM

nections, sculpting new brain circuits that encode the memory. This finding led to the idea that ACSS2 could be involved in unwanted memory formation. We confirmed this idea by showing that blocking ACSS2 affected the ability of mice to encode fear memories. Animals without the active enzyme displayed reduced aversion to an environment in which they had previously experienced an electric shock. The mice lacking ACSS2 are also completely healthy beyond their impaired memory. Because this work had great clinical promise, we formed our company to bring new treatments to people with disorders associated with traumatic or burdensome memories.

We launched EpiVario to target ACSS2 in people with posttraumatic stress disorder (PTSD) to diminish their traumatic memories. PTSD can develop after events such as interpersonal violence, combat, or even the physical and emotional stress linked to severe cases of COVID-19. Symptoms include intrusive retrieval of traumatic memories, insomnia, and irritability. Unfortunately, current treatments are desperately insufficient. EpiVario aims to treat patients during psychotherapy sessions, by administering the drug as a clinician works with the patient to elicit the stress-inducing memories. Recall of the traumatic event opens a window during which recollection can either reinforce or weaken the memory. Our product is a short-lived drug that readily crosses the blood-brain barrier to transiently block the ACSS2 pathway. EpiVario's goal is to reduce the stress associated with the unwanted trauma memory through treatment that provides a lasting effect. The company's intellectual property is well protected. The pathway and mode of action were previously unknown, and patents were exclusively licensed worldwide from the University of Pennsylvania, where our spin-off was incubated.

Our recognition that acetate metabolism is closely linked to learning and memory inspired us to explore a nutrient well known for altering memory: alcohol. When consumed, alcohol is metabolized in the liver and causes a surge in circulating acetate (*5*). We hypothesized that this acetate spike might fuel ACSS2-driven acetylation of neuronal genes. In mice injected with isotopically labeled alcohol, we used mass spectrometry to track alcohol molecules as they traveled throughout the body. Minutes after alcohol consumption, we detected labeled acetate on histones within neurons, indicating a link between liver alcohol metabolism and gene regulation in the brain (*5*).

Icahn School of Medicine at Mount Sinai, New York, NY, USA.
Email: philipp.mews@gmail.com

Bringing the two research areas together synergistically, we hope to target memories linked to substance use disorders (*5*, *6*). When ACSS2 is blocked in the mouse brain, alcohol was prevented from contributing to gene activation, confirming that its metabolites play an important role in controlling how alcohol affects our memory (see the figure). Mice with lowered ACSS2 do not form a preference for environments where they have been given alcohol. These findings are notable because the memory of alcohol-associated cues is a primary driver of craving and relapse in people with alcohol use disorder (*7*).

Memory-related diseases are insidious and often share a common ability to silently erode joy from our lives. In the past, PTSD and substance use disorders were frequently written off as moral failures of character, but today they are recognized as complex disorders with biological underpinnings.

Our research shows that trauma and stress memory are influenced by a link between nutrient metabolism and histone acetylation in the brain. EpiVario is developing future therapeutics that target this link to treat memory-related disorders, including PTSD; alcohol addiction; and, most recently, smoking cessation. In the future, our company aims to build a broader drug-development platform for screening epigenetic enzymes that could be suitable for modulating neuronal processes in a host of anxiety and addiction disorders.

Our research further suggests that other external sources of acetate (e.g., sour foods and various gut microbiota) may similarly affect histone acetylation to modulate memory. Exploring the ways that our metabolism can shape genes and neural circuits promises to improve our understanding of mental health and disease. As we continue to raise money to fund our forthcoming clinical trials, we are driven by our passion for having a positive impact on the many individuals affected by traumatic memories and addiction. ■

REFERENCES AND NOTES

1. P. Mews *et al.*, *Nature* **546**, 381 (2017).
2. P. Mews *et al.*, *J. Neurosci.* **41**, 873 (2021).
3. E. J. Nestler, C. Lüscher, *Neuron* **102**, 48 (2019).
4. J. Gräff, L.-H. Tsai, *Nat. Rev. Neurosci.* **14**, 97 (2013).
5. P. Mews *et al.*, *Nature* **574**, 717 (2019).
6. S. C. Pandey, E. J. Kyzar, H. Zhang, *Neuropharmacology* **122**, 74 (2017).
7. G. F. Koob, N. D. Volkow, *Lancet Psychiatry* **3**, 760 (2016).

10.1126/science.abo4234

PRIZE ESSAY

FINALIST

Jacqueline Douglass

Jacqueline Douglass received her undergraduate degree from the Massachusetts Institute of Technology and an MD-PhD from Johns Hopkins University School of Medicine. She is currently an internal medicine resident at Johns Hopkins and plans to pursue medical oncology training. **www.science.org/doi/10.1126/science.abo4237**

BII Science | Prize for Innovation

INNOVATION

Mutation-guided therapeutics

Development of bispecific antibodies to target mutant peptides in cancer

By **Jacqueline Douglass**

At the start of my PhD, my advisers tasked me with the challenge of using antibodies to target driver mutation-derived mutant peptides that are presented on the surface of cancer cells. I was new to the field of tumor immunology and did not appreciate how wild a concept this was or all the hurdles that would await. My naïvety allowed me to believe that this approach might work against cancer cells in a Petri dish or even in a mouse model. Yet I never imagined that this challenge would blossom into a project encompassing numerous team members and multiple laboratories, resulting within a few short years in several licensed patents and a start-up company that develops clinical-grade products.

Cancer development is driven by driver mutations in critical oncogenes and tumor suppressor genes. These mutations are specific to cancer cells; required for initiation and maintenance of the cancerous state, thus unlikely to be lost during the evolution of cancer cells; and often shared among many patients. These characteristics make driver mutations preferred targets for cancer therapy.

The general concept of using antibodies to target mutant peptides seemed straightforward enough. Mutations in the DNA of cancer cells translate to mutant proteins. When mutant proteins are degraded, some of the resulting peptides that contain the mutated amino acids can be presented on the cancer cell surface by a type of protein called human leukocyte antigen (HLA), resembling a hot dog (the 8– to 11–amino acid peptide) in a hot dog bun (the HLA protein) (*1*–*3*). A mutant peptide HLA (pHLA) on the surface of a cancer cell would be an ideal therapeutic target; its mutant amino acid sets it apart from its wild-type counterpart. The location of the mutant pHLA on the cell surface would make these protein complexes amenable to targeting through a variety of therapeutic modalities, including antibody-based therapies.

We began by developing a method to identify antibodies specific to these mutant pHLAs. Traditional hybridoma technology failed to generate antibodies able to discriminate between wild-type and mutant peptides, which differ by a single amino acid. We then designed and built two diverse phage display libraries to screen different bacteriophages—viruses that only infect bacteria—that each express a distinct antibody fragment for antibody clones specific to the target protein of interest (*4*, *5*).

We then needed to decide which mutant pHLA complexes to target. Our primary criterion was mutant peptides derived from common driver mutations in cancers. However, we needed to assess which peptides are presented on the cell surface by HLA. We initially relied on in silico algorithms to predict binding of peptides to HLA proteins but soon found that the predictions could be misleading. Therefore, we developed a mass spectrometry–based method that was quantitative to determine which peptides are presented as pHLA, which revealed that these pHLA were typically present at very low numbers on the cell surface (*6*).

After identifying mutant pHLAs to target, we had to convert our antibodies into a format capable of killing cancer cells presenting these complexes. However, there was no precedent for an antibody-based therapy that targets such a low number of pHLAs on a cancer cell. We assessed a range of modalities, settling on bispecific antibodies: small bivalent molecules that redirect T cells to kill cancer cells. We tested dozens of bispecific antibody formats before identifying one with the sensitivity required to target the very low-density mutant pHLAs (*5*, *7*).

We then had to demonstrate that the bispecific antibodies were effective and specific therapeutic agents. We used gene-editing techniques to produce cancer cell lines that differed by only a single mutation of interest. These tailored cell lines allowed us to show that our bispecific antibodies could redirect T cells to kill only cancer cells that harbor the mutant protein in in vitro and in vivo mouse models (*5, 7*). To better understand the determinants of the antibodies' specificities, we collaborated with biophysicists to solve the structures of our antibodies in complex with the mutant pHLAs (*7*–*9*).

Johns Hopkins School of Medicine, Johns Hopkins University, Baltimore, MD, USA. Email: jdouglass86@gmail.com

PHOTO: JOHNS HOPKINS PATHOLOGY PHOTOGRAPHY LAB

In retrospect, our project was successful because we operated as a small biotechnology company would. We combined the efforts of scientists and clinicians across several laboratories, allowing everyone to contribute their own expertise. We had weekly meetings as a group to discuss data, troubleshoot problems, and brainstorm, where everyone's opinions and ideas were equally valued. We were fortunate to operate in an academic climate that encouraged us to patent our intellectual property by working closely with legal experts. As a graduate student, I had the opportunity to pitch our approach to biotechnology companies and potential investors, who were interested in translating our technology to large-scale clinical adoption. More recently, we established a start-up company to advance the technologies our group has developed into the clinical stage.

I am amazed how a simple concept has blossomed into a detailed and optimized pipeline for developing an innovative approach to treating cancer. Our environment fostered teamwork and creativity, allowing us to overcome seemingly impossible challenges. As I look to the future, I am excited to see our work advance beyond the academic realm—one step closer to the end goal of providing new treatment options for cancer patients and ultimately hope for those patients and their families. ■

REFERENCES AND NOTES

1. T. N. Schumacher, R. D. Schreiber, *Science* **348**, 69 (2015).
2. A. H. Pearlman *et al.*, *Nat. Can.* **2**, 487 (2021).
3. E. Blass, P. A. Ott, *Nat. Rev. Clin. Oncol.* **18**, 215 (2021).
4. A. D. Skora *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **112**, 9967 (2015).
5. J. Douglass *et al.*, *Sci. Immunol.* **6**, eabd5515 (2021).
6. Q. Wang *et al.*, *Cancer Immunol. Res.* **7**, 1748 (2019).
7. E. H. Hsiue *et al.*, *Science* **371**, eabc8697 (2021).
8. M. S. Miller *et al.*, *J. Biol. Chem.* **294**, 19322 (2019).
9. M. S. Hwang *et al.*, *Nat. Commun.* **12**, 5271 (2021).

10.1126/science.abo4237

RESEARCH

IN *SCIENCE* JOURNALS

Edited by **Michael Funk**

CLIMATOLOGY

Climate history of the central Sahara

Our understanding of the climate history of the Sahara, Earth's largest warm desert, is limited by a paucity of local records, because all of the previously available records are from the fringes of the desert. Van der Meeren *et al.* describe a new record from the central Sahara that provides important constraints on the developmental history of the desert over the past several thousand years. The authors found evidence that before 4200 years ago, the Sahara was even more arid than today, and that the central Saharan climate over the past 3000 years is closely linked with the intensity of the tropical West African monsoon. —KVH

Sci. Adv. 10.1126/sciadv.abk1261 (2022).

Sediment from the Ounianga Serir oasis reveals changes in aridity in the central Sahara over the past several thousand years.

PLANT SCIENCE

Increasing wheat grain yield

In wheat, the numbers of tillers, spikes, and spikelets determine how much grain is produced. Beginning with a cross between two common wheat cultivars, Zhang *et al*. cloned a gene that affects wheat plant architecture and, consequently, grain yield (see the Perspective by van Esse). Exon capture analyses identified the same gene in wild emmer wheat. The gene nonetheless remains rare among contemporary US wheat cultivars. In field trials in Jiangsu, China, overexpression of the dominant allele in transgenic wheat increased grain production by about 12%. —PJH

Science, abm0717, this issue p. 180; see also abo7429, p. 133

PHYSICS

Fresh test of quantum electrodynamics

One of the best ways to advance our understanding of nature is to challenge the fundamental theories developed to describe its laws mathematically. Quantum electrodynamics (QED) theory of the interaction of matter with light is currently one of the most accurate fundamental theories, and the search for QED deviations is of considerable interest. Henson *et al*. measured and theoretically calculated the helium 2^3S_1–$2^3P/3^3P$ tune-out frequency with an accuracy that made it possible to discern its QED contributions and previously omitted components. The tune-out frequency is sensitive to a different part of QED compared with other, more common atomic structure probes, and the present work is an important step in expanding the horizon of possible QED tests. —YS

Science, abk2502, this issue p. 199

NANOMATERIALS

Imaging particles and patterns

A key feature of nanoscale materials is the ability to tune their properties through small changes in particle size or chemical composition. Assembly into three-dimensional superstructures provides a platform for building complex, multifunctional materials, but it also makes it harder to understand

PHOTO: GUENTERGUNI/ISTOCKPHOTO

the structure at a particle level. Michelson *et al*. present the nondestructive three-dimensional imaging of a superstructure made of thousands of particles at 7-nanometer resolution, from which they were able to map both position and composition. The authors were also able to view defects in the crystalline lattice of the superstructures. —MSL

Science, abk0463, this issue p. 203

OPTICS

Electrical control of topological light

Most closed physical systems are described as Hermitian in that they can have a single or a set of distinct resonant modes. Open systems, however, are non-Hermitian, and engineering the gain and loss of such systems can produce exceptional points where the resonant modes coalesce. Ergoktas *et al*. demonstrate an electrically tunable system that allows for reconstruction of the complex energy landscape and provides topological control of light by tuning the loss-imbalance and frequency detuning of the interacting modes. Electrical tuneability provides a route for exploiting the sensitivity of exceptional point singularities for device applications. —ISO

Science, abn6528, this issue p. 184

GRAPHENE

Zooming into trilayer graphene

Stacking and twisting graphene layers with respect to each other can lead to exotic transport effects. Recently, superconductivity was observed in graphene trilayers in which the top and bottom layers were twisted with respect to the middle layer by the same, "magic" angle. Turkel *et al*. used scanning tunneling microscopy to take a closer look into the stacking structure. They found that a small misalignment between the top and bottom layers caused the lattice to rearrange itself into a pattern of triangular domains. The domains had a magic-angle twisted trilayer structure and were separated by a network of line and point defects. —JS

Science, abk1895, this issue p. 193

DEVELOPMENTAL BIOLOGY

Generating functional rat gametes

In the past decade, methods have been developed to generate germ cells from pluripotent stem cells for studies of development and in vitro gametogenesis. However, offspring from in vitro–derived germ cells has only been achieved in mice. Oikawa *et al*. extend this work beyond mice to a second rodent species, the rat, a leading animal model for biomedical research with many physiological similarities to humans. A stepwise protocol allows for the production of fetal stage rat germ cells that can produce viable offspring upon maturation in the testis and injection of the sperm into unfertilized oocytes. This system will allow comparative studies and enable broader execution and analysis of in vitro gametogenesis. —BAP

Science, abl4412, this issue p. 176

TUMOR IMMUNOLOGY

Monitoring immune cells in tumors

To predict the effects of immunotherapy in cancer, better spatial understanding of the tumor microenvironment is needed. Hoch *et al.* used multiplex imaging mass cytometry of protein (immune cell markers) and RNA (chemokine ligands) targets to define immune cell interactions in the tumor microenvironment of melanoma samples. The authors found that the chemokines *CXCL9* and *CXCL10* were coexpressed in patches with *CXCL13*-expressing exhausted T cells, suggesting that they recruited B cells and aided in the formation of tertiary lymphoid structures in melanoma tumors. These structures had a spatial enrichment of naïve and naive-like T cells, which are involved in antitumor responses. —DAE

Sci. Immunol. **7**, eabk1692 (2022).

IN OTHER JOURNALS

Edited by **Caroline Ash** *and* **Jesse Smith**

Oomycete spores spread infection by swarming using two polar flagella that coordinate speed and direction of movement.

IMAGE: FRANÇOIS ORANGE/CCMA, UNIVERSITÉ CÔTE D'AZUR

PLANT PATHOLOGY

Two flagella cooperate to steer

Several of the oomycete *Phytophthora* species are plant pathogens that cause diseases such as potato late blight, sudden oak death, and cocoa black pod disease. Affecting food crops as diverse as tomato, onion, soybean, and cucumber, *Phytophthora* pathogens are a worldwide threat to food security. Spores of *Phytophthora* swarm through thin layers of moisture across soil and leaves, some initiating infection as they go, steering in response to environmental signals in their search for a new victim. Tran *et al.* investigated how the two flagella on each spore coordinate for speed and steering. The flagellum at the front is the source of most of the straightforward action. Not unlike how two canoeists alter strokes when turning, the spore puts the rear flagellum on pause while the anterior flagellum changes its action from sinusoidal waves to full-power stroke. —PJH

eLife **11**, e71227 (2022).

LUNG DISEASE

Neural signals control lung fluid

Fluid buildup in the lungs after injury or infection interferes with the transfer of oxygen into the bloodstream. This occurs in acute respiratory distress syndrome (ARDS) patients and causes breathing difficulties. Prior work has shown that neuroendocrine substances originating from lung cells increase during ARDS. By developing a mouse model for neuroendocrine cell hyperplasia of infancy (NEHI), Xu *et al.* found that elevated neuropeptides

EXTREME POLITICS

Ebbs and flows of political polarization

Across 12 advanced democracies, affective polarization, the degree to which people feel more negatively toward other political parties than toward their own, has increased the most since the 1980s in the United States and to a lesser extent in Canada, Denmark, France, New Zealand, and Switzerland, and has decreased in Australia, Britain, (West) Germany, Japan, Norway, and Sweden. Boxell *et al.* harmonized results from 149 surveys and assembled data on economic, media, demographic, and political trends. Trends in the nonwhite share of the population and in the polarization of political elites were most strongly associated with trends in polarization of the general public. —BW

Rev. Econ. Stat. 10.1162/rest_a_01160 (2022).

signal to endothelial cells, alter junctional proteins, and result in a compromised (leaky) lung with excess fluid. If neuropeptides are reduced in NHEI mutant animals, then excess fluid buildup is reversed. This model for NEHI disease could offer useful therapeutic targets for pulmonary edema. —BAP

Dev. Cell 10.1016/j.devcel.2022.02.023 (2022).

SEX

Fruit fly's symmetrical serenade

A symmetrical mate is more attractive for many bilateral species, possibly because such balance signals health and quality. Appearances are not everything and can fluctuate, so the symmetrical mate idea has been controversial. Instead, Vijendravarma *et al.* decided to test the symmetry of a nonvisual signal. Fruit fly males serenade females by wing vibrations, and symmetrical males will have symmetrical songs. By manipulating wing symmetry during fly development in laboratory culture, the authors were able to change the fly's song. The asymmetric songs from the asymmetric males were rejected by females in mate choice experiments. However, if females were bred in the absence of mate choice, they were just as happy with males singing an off-kilter song as a symmetrical one. —CA

Proc. Natl. Acad. Sci. U.S.A. **119**, e2116136119 (2022).

CELL BIOLOGY

Pericytes protect the kidneys

Kidney damage caused by renal ischemia is a common complication after surgery. Freitas and Attwell examined the cause of the sustained reduction in renal blood flow known as "no-reflow," which can exacerbate kidney injury. Working in rats and mice, the authors found that 60 minutes of kidney ischemia followed by 30 to 60 minutes of reperfusion resulted in prolonged reduction in renal blood flow. A type of cells known as pericytes were identified as the culprits. Pericytes enwrap capillaries and constrict and occlude blood flow. Treatments that blocked pericyte constriction, including rho kinase inhibitors, helped to prevent capillary obstruction and reduced kidney damage. Thus, therapies targeting pericytes might prove fruitful in the treatment of acute kidney injury. —SMH

eLife **11**, e74211 (2022).

GEOPHYSICS

Weak under pressure

Calcium silicate is one of the major crystal phases in Earth's mantle, but its mechanical properties are poorly constrained. Immoor *et al.* discovered that this phase is surprisingly weak when in the cubic crystal structure. The strength and viscosity are much lower than the other major mantle phases, which makes it critical in affecting how subducting slabs sink into the mantle. The properties also may dictate the fate of the accompanying oceanic crust, which may either founder in the midmantle or sink all the way to the core. —BG

Nature **603**, 276 (2022).

QUANTUM OPTICS

Modulating spontaneous emission

When an excited quantum two-level system relaxes back to the lower level, it can do so through a process of spontaneous emission whereby a single photon is emitted. It is well known that the interaction between the emitter and an optical cavity in which it is placed can alter the emission process. Tian *et al.* show that hybridizing two-level solid-state emitters with a tunable optomechanical cavity can provide a dynamical aspect to the spontaneous emission process. Optical stimulation of the optomechanical cavity shifts its resonance frequency to couple with that of the emitter and thereby alter its emission rate. The ability to modulate the spontaneous emission rate of solid-state emitters should be useful for applications in photonic integrated circuits and photonic quantum technologies. —ISO

Optica **9**, 309 (2022).

ILLUSTRATION: DAVIDE BONAZZI/SALZMANART

ALSO IN *SCIENCE* JOURNALS

Edited by **Michael Funk**

IMMUNOGENOMICS

Analyzing immune system gene expression

Diseases involving the immune system are heritable, but it is unknown how genetic variation contributes to different diseases. To identify how implicated loci affect gene expression in immune cells from individuals from different populations, two groups performed single-cell RNA sequencing of immune cells, with each study investigating hundreds of individuals and more than 1 million immune cells (see the Perspective by Sumida and Hafler). These studies examined both proximal (cis) and distal (trans) genetic variants affecting gene expression in 14 different immune cell types. Perez *et al*. studied healthy individuals of both European and Asian descent, as well as individuals diagnosed with systemic lupus erythematosus. Yazar *et al*. performed a population-based study investigating how segregating alleles contribute to variation in immune function. Integrating these data with autoimmune disease cohorts identifies causal effects for more than 160 loci. Both studies observed how gene expression patterns are cell-type and context specific and can explain observed variation in immune cell function among individuals. Both studies also identified causal links between genome-wide analyses and expression quantitative trait loci, identifying potential mechanisms underlying autoimmune diseases. —LMZ

Science, abf1970, abf3041, this issue p. 153, p. 154; see also abq0426, p. 134

PROTEIN ENGINEERING

Increasing potency but not toxicity

Cancer immunotherapy takes advantage of natural immune responses. In one approach, T cells are engineered to be activated in response to a tumor-specific antigen. A challenge is that increasing the affinity of a T cell receptor (TCR) for the tumor-specific antigen to increase cancer cell killing can lead to off-target toxicities. Zhao *et al*. took advantage of the fact that an extended bond lifetime characteristic of so-called catch bonds is associated with agonist potency. The authors screened for TCR mutants that acquired catch bonds by selecting those that showed high activation paired with low antigen-binding affinity. They engineered a tumor antigen–specific TCR that had killing potency at least equal to a previously described high-affinity TCR but without the associated adverse cross-reactivity. —VV

Science, abl5282, this issue p. 155

NANOMATERIALS

Diversifying nanoparticles

Multielement nanoparticles are attractive for a variety of applications in catalysis, energy, and other fields. A more diverse range and larger number of elements can be mixed together because of high-entropy mixing states accessed by a number of recently developed techniques. Yao *et al*. review these techniques along with characterization methods, high-throughput screening, and data-driven discovery for targeted applications. The wide range of different elements that can be mixed together presents a large number of opportunities and challenges. —BG

Science, abn3103, this issue p. 151

CANCER GENOMICS

Noncoding mutations decoded

Numerous large-scale efforts have been undertaken to catalog and understand the biology of cancer-associated mutations in regions that directly code for proteins. Much of the genome, however, consists of noncoding regions that do not directly encode specific proteins, but instead perform other functions such as regulating protein expression. These genome regions can also play key roles in cancer. Dietlein *et al*. developed a computational approach to systematically detect cancer-associated mutations in noncoding regions of different cancer types and directly examined the biological function of one such region involved in breast cancer. Using this genome-wide approach, researchers should be able to comprehensively examine the contributions of noncoding regions to cancer development. —YN

Science, abg5601, this issue p. 152

STRUCTURAL BIOLOGY

How JAKs are activated

Janus kinases (JAKs) are essential to the many biological outcomes of cytokine signaling. They bind to cytokine receptors and are activated when cytokine binding leads to receptor dimerization. Dimerization leads to activation of signal transducer and activator of transcription (STAT) transcription factors, which translocate to the nucleus and initiate the transcription of cytokine-responsive genes. Mutations in JAKs and STATs lead to immunodeficiency and myeloproliferative disorders. Glassman *et al*. report the structure of a full-length JAK bound to an engineered construct that displays a dimer of the intracellular domains of a cytokine receptor (see the Perspective by Levine and Hubbard). The structure provides insight into how cytokine receptor dimerization drives JAK activation and how a range of disease mutations affect function. —VV

Science, abn8933, this issue p. 163; see also abo7788, p. 139

PARTICLE PHYSICS

Weighing the *W* boson

W bosons mediate the weak interaction, one of the fundamental forces in physics. Because the Standard Model (SM) of particle physics places tight constraints on the mass of the *W* boson, measuring the mass puts the SM to the test. The Collider Detector at Fermilab (CDF) Collaboration now reports a precise measurement of the *W* boson mass extracted from data taken at the Tevatron particle accelerator (see the Perspective by Campagnari and Mulders). Surprisingly, the researchers found that the mass of the boson was significantly higher than the SM predicts, with a discrepancy of 7 standard deviations. —JS

Science, abk1781, this issue p. 170; see also abm0101, p. 136

BIOMATERIALS

Key aspects of bone mineralization

Bone is a hierarchical material consisting of organic fibers, mainly in the form of collagen, which are mineralized with inorganic crystals, primarily hydroxyapatite. It is this structure that gives bone its remarkable combination of strength and toughness. Ping *et al*. examined the deposition of minerals on both the outside and inside of the fibers over time (see the Perspective by Nudelman and Kröger). They found that large contractile forces occur within the collagen during intrafibrillar mineralization regardless of the mineral type, thus giving bone its unusual combination of mechanical properties. This feature is analogous to the reinforcement of concrete using prestressed steel rods. —MSL

Science, abm2664, this issue p. 188; see also abo1264, p. 137

RESPIRATORY DISEASE

Characterizing cardiopulmonary disease

Bronchopulmonary dysplasia (BPD) and BPD-associated pulmonary hypertension (BPD-PH)

cause substantial morbidity and mortality in preterm infants. However, the drivers of BPD and BPD-PH are not well understood. Lao *et al.* characterized immune polarization in blood samples from preterm infants and in a mouse model of BPD and BPD-PH. In both cases, BPD was associated with type 2 immune polarization, characterized by interleukin (IL)-4, IL-5, and IL-13, as well as activation of signal transducer and activator of transcription 6 (STAT6). STAT6 deficiency or blockade of type 2 immune mediators partially reversed evidence of alveolar and pulmonary vascular disease in mice, suggesting that this signaling axis could be targeted for preterm infants with BPD. —CSM

Sci. Transl. Med. **14**, eaaz8454 (2022).

CANCER

A PAX8-SOX17 duo in tumor angiogenesis

The transcription factor PAX8 is essential for the development of the female reproductive tract but is frequently amplified in and supports the growth of ovarian cancers. By comparing ovarian cancer and nonmalignant fallopian tube cells and tissues, Chaves-Moreira *et al*. found that PAX8 interacted with another transcription factor, SOX17, and that the complex in cancer cells transcriptionally promoted a pro-angiogenic secretome. Repressing the complex inhibited tumor cell–induced angiogenesis in both cell culture and in vivo models. —LKF

Sci. Signal. **15**, eabm2496 (2022).

VIROME

Expanding the RNA catalog

Apart from their roles in human infectious diseases, we understand relatively little about RNA viruses in the wider world. Recently, the discovery curve has been spectacular and has revealed unexpected diversity. Zayed *et al*. optimized discovery and classification methods on Tara Oceans RNA sequence data to double the roster of known RNA virus phyla (see the Perspective by Labonté and Campbell). This is not just a numbers game; the authors also found a missing link in RNA virus evolution and discovered new phyla that dominate in the oceans and might infect mitochondria. These viruses require an ancient enzyme, RNA-directed RNA polymerase (RdRp) for replication, which is thus used as a marker of deep evolutionary relationships. In addition to the primary sequence data, information on the three-dimensional structures of the RdRp, network-based clusters, other genomic domains, and whole-genome characteristics help to reshape the outlines of the evolutionary history of RNA viruses. —CA

Science, abm5847, this issue p. 156; see also abo5590, p. 138

REVIEW SUMMARY

NANOMATERIALS

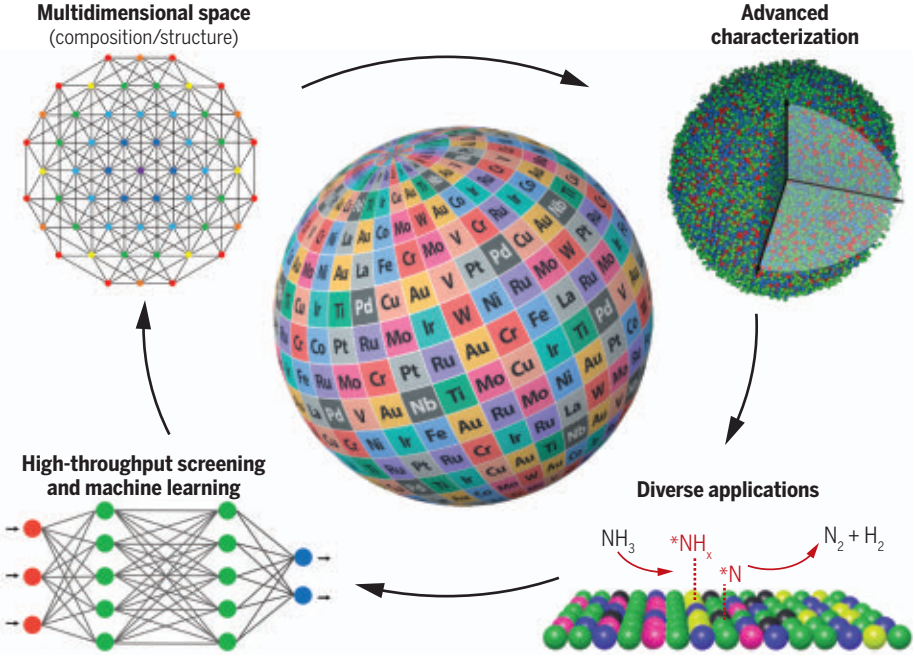
High-entropy nanoparticles: Synthesis-structure-property relationships and data-driven discovery

Yonggang Yao†, Qi Dong†, Alexandra Brozena, Jian Luo, Jianwei Miao, Miaofang Chi, Chao Wang, Ioannis G. Kevrekidis, Zhiyong Jason Ren, Jeffrey Greeley, Guofeng Wang, Abraham Anapolsky, Liangbing Hu*

BACKGROUND: High-entropy nanoparticles contain more than four elements uniformly mixed into a solid-solution structure, offering opportunities for materials discovery, property optimization, and advanced applications. For example, the compositional flexibility of high-entropy nanoparticles enables fine-tuning of the catalytic activity and selectivity, and high-entropy mixing offers structural stability under harsh operating conditions. In addition, the multielemental synergy in high-entropy nanoparticles provides a diverse range of adsorption sites, which is ideal for multistep tandem reactions or reactions that require multifunctional catalysts. However, the wide range of possible compositions and complex atomic arrangements also create grand challenges in synthesizing, characterizing, understanding, and applying high-entropy nanoparticles. For example, controllable synthesis is challenging given the different physicochemical properties within the multielemental compositions combined with the small size and large surface area. Moreover, random multielemental mixing can make it difficult to precisely characterize the individual nanoparticles and their statistical variations. Without rational understanding and guidance, efficient compositional design and performance optimization within the huge multielemental space is nearly impossible.

ADVANCES: The comprehensive study of high-entropy nanoparticles has become feasible because of the rapid development of synthetic approaches, high-resolution characterization, high-throughput experimentation, and data-driven discovery. A diverse range of compositions and material libraries have been developed, many by using nonequilibrium "shock"–based methods designed to induce single-phase mixing even for traditionally immiscible elemental combinations. The nanomaterial types have also rapidly evolved from crystalline metallic alloys to metallic glasses, oxides, sulfides, phosphates, and others. Advanced characterization tools have been used to uncover the structural complexities of high-entropy nanoparticles. For example, atomic electron tomography has been used for single-atom-level resolution of the three-dimensional positions of the elements and their chemical environments. Finally, high-entropy nanoparticles have already shown promise in a wide range of catalysis and energy technologies because of their atomic structure and tunable electronic states. The development of high-throughput computational and experimental methods can accelerate the material exploration rate and enable machine-learning tools that are ideal for performance prediction and guided optimization. Materials discovery platforms, such as high-throughput exploration and data mining, may disruptively supplant conventional trial-and-error approaches for developing next-generation catalysts based on high-entropy nanoparticles.

OUTLOOK: High-entropy nanoparticles provide an enticing material platform for different applications. Being at an initial stage, enormous opportunities and grand challenges exist for these intrinsically complex materials. For the next stage of research and applications, we need (i) the controlled synthesis of high-entropy nanoparticles with targeted surface compositions and atomic arrangements; (ii) fundamental studies of surfaces, ordering, defects, and the dynamic evolution of high-entropy nanoparticles under catalytic conditions through precise structural characterization; (iii) identification and understanding of the active sites and performance origin (especially the enhanced stability) of high-entropy nanoparticles; and (iv) high-throughput computational and experimental techniques for rapid screening and data mining toward accelerated exploration of high-entropy nanoparticles in a multielemental space. We expect that discoveries about the synthesis-structure-property relationships of high-entropy nanoparticles and their guided discovery will greatly benefit a range of applications for catalysis, energy, and sustainability. ■



High-entropy nanoparticles and data-driven discovery. Emerging high-entropy nanoparticles feature multielemental mixing within a large compositional space and can be used for diverse applications, particularly for catalysis. High-throughput and machine-learning tools, coupled with advanced characterization techniques, can substantially accelerate the optimization of these high-entropy nanoparticles, forming a closed-loop paradigm toward data-driven discovery.

CREDITS: TOP RIGHT: YANG *ET AL.*, *NATURE* **592**, 60–64 (2021); CENTER: JIAQI DAI; BOTTOM RIGHT: XIE *ET AL.*, *NAT. COMMUN.***10**, 4011 (2019)

The list of author affiliations is available in the full article online.
*Corresponding author. Email: binghu@umd.edu
†These authors contributed equally to this work.
Cite this article as Y. Yao *et al.*, *Science* **376**, eabn3103 (2022). DOI: 10.1126/science.abn3103

READ THE FULL ARTICLE AT
https://doi.org/10.1126/science.abn3103

REVIEW

NANOMATERIALS

High-entropy nanoparticles: Synthesis-structure-property relationships and data-driven discovery

Yonggang Yao[1]†, Qi Dong[1]†, Alexandra Brozena[1], Jian Luo[2], Jianwei Miao[3], Miaofang Chi[4], Chao Wang[5], Ioannis G. Kevrekidis[5], Zhiyong Jason Ren[6], Jeffrey Greeley[7], Guofeng Wang[8], Abraham Anapolsky[9], Liangbing Hu[1,10]*

High-entropy nanoparticles have become a rapidly growing area of research in recent years. Because of their multielemental compositions and unique high-entropy mixing states (i.e., solid-solution) that can lead to tunable activity and enhanced stability, these nanoparticles have received notable attention for catalyst design and exploration. However, this strong potential is also accompanied by grand challenges originating from their vast compositional space and complex atomic structure, which hinder comprehensive exploration and fundamental understanding. Through a multidisciplinary view of synthesis, characterization, catalytic applications, high-throughput screening, and data-driven materials discovery, this review is dedicated to discussing the important progress of high-entropy nanoparticles and unveiling the critical needs for their future development for catalysis, energy, and sustainability applications.

High-entropy nanoparticles have received a great amount of attention in recent years because of their multielemental composition (typically five or more elements) and homogeneously mixed solid-solution state, providing not only an enormous number of combinations for materials discovery but also a unique microstructure for property optimization (Fig. 1A) (*1*–*3*). Early reports of multielemental (five or more) alloy nanoparticles suggested the potential of these unique materials (*4*–*6*) but did not provide detailed structural understanding or reveal a general synthesis route for different compositions. In 2016, Mirkin and colleagues made a substantial advance in synthesizing various compositions of multielemental nanoparticles using confined nanoreactors (*7*). However, these materials featured heterogeneous structures with phase separation due to elemental immiscibility. Recent advances in ultrafast synthetic methodologies, such as nonequilibrium thermal-shock-based approaches, have since enabled a variety of high-entropy nanoparticles without phase separation, even among immiscible elemental combinations (Fig. 1B) (*8*). In a typical thermal shock process (e.g., 2000 K in 55 ms), the rapid heating of precursors to a high temperature induces multielemental mixing and alloying to achieve a solid-solution state, whereas the short heating duration and subsequent rapid quenching help to retain and freeze the uniform structure and small particle size (*8*). Since then, various high-entropy nanomaterials, including alloys (e.g., PtPdFeCoNiAuCuSn) (*9*–*14*), metallic glasses (e.g., amorphous CoCrMnNiV) (*15*, *16*), intermetallics [e.g., $L1_0$ type $(Pt_{0.8}Pd_{0.1}Au_{0.1})(Fe_{0.6}Co_{0.1}Ni_{i0.1}Cu_{0.1}Sn_{0.1})$] (*17*, *18*), oxides, fluorides, sulfides, carbides, MXenes (*19*–*25*), and van der Waals materials (e.g., dichalcogenides, halides, and phosphorus trisulfide) (*26*), have all been successfully demonstrated using thermal shock and other nonequilibrium approaches.

Despite having been developed only recently, high-entropy nanoparticles have already shown great promise for a range of emerging energy-related processes and applications, particularly in catalysis (Fig. 1C) (*27*–*36*). The compositional flexibility of high-entropy nanoparticles enables fine-tuning of the catalytic activity, whereas the high-entropy solid-solution mixing potentially offers structural stability 1that is critical for operation under harsh conditions. For example, non-noble $(Co_xMo_{0.7-x})Fe_{0.1}Ni_{0.1}Cu_{0.1}$ nanoparticles have been shown to overcome the immiscibility of Co–Mo, allowing for robust tuning of the Co–Mo ratios and associated surface adsorption properties. As a result, $(Co_{0.25}Mo_{0.45})Fe_{0.1}Ni_{0.1}Cu_{0.1}$ nanoparticles have demonstrated a fourfold improvement in ammonia decomposition compared with noble Ru and are stable at 500°C for 50 hours without noticeable degradation (*36*). In another example, $Pt_{18}Ni_{26}Fe_{15}Co_{14}Cu_{27}$ nanoparticles were developed for electrochemical hydrogen evolution and showed a lower onset potential (11 versus 84 mV), a higher activity (10.98 versus 0.83 A/mg_{Pt}), and excellent stability compared with commercial Pt–C catalysts (*11*). These examples epitomize the strong potential of high-entropy nanoparticles as highly efficient and cost-effective catalysts (*9*, *11*, *12*, *20*, *36*–*39*).

Compared with materials having relatively simple compositions (i.e., one to three elements), high-entropy nanoparticles have two distinct features: (i) a vast compositional space that derives from the multielemental combinations and (ii) complex atomic configurations due to the random multielemental mixing. The former provides huge compositional choices for catalyst design and development, and the latter makes these materials fundamentally different from conventional catalysts in that they feature a diverse range of adsorption sites and a near-continuous binding energy distribution pattern (*40*, *41*). These qualities are particularly attractive for complex or tandem reactions that involve numerous intermediate steps and require multifunctionality (*28*, *38*, *40*, *42*–*44*).

However, along with these opportunities, the vast number of possible compositions and complex atomic arrangements create grand challenges in the design, synthesis, characterization, and application of these unique nanomaterials. First, considering the wide span of physicochemical properties (e.g., atomic size and electronic structure) among the different constituent elements, synthesizing high-entropy nanoparticles in a highly controllable manner is difficult. Moreover, characterizing the detailed structure of high-entropy nanoparticles, such as the reactive surfaces and defects, is challenging or still lacking because of the complex atomic configurations and multiple elements of similar electron contrasts. Additionally, we have very limited knowledge of how elemental composition and synthesis methods affect the structure and properties of high-entropy nanoparticles. Although identifying these relationships for such complex materials is a daunting task, understanding them is critical to guiding material design and optimization.

In response to the increasing interest, rapid development, and large challenges of this field, we aim to highlight the important progress and critical unknowns regarding the synthesis, structure, characterization, and applications of high-entropy nanoparticles. We also discuss the potential and implementation of computationally guided and data-driven

[1]Department of Materials Science and Engineering, University of Maryland, College Park, MD 20742, USA. [2]Department of NanoEngineering, Program of Materials Science and Engineering, University of California San Diego, La Jolla, CA 92093, USA. [3]Department of Physics and Astronomy and California NanoSystems Institute, University of California, Los Angeles, Los Angeles, CA 90095, USA. [4]Center for Nanophase Materials Sciences, Oak Ridge National Laboratory, Oak Ridge, TN 37932, USA. [5]Department of Chemical and Biomolecular Engineering, Johns Hopkins University, Baltimore, MD 21218, USA. [6]Department of Civil and Environmental Engineering and Andlinger Center for Energy and the Environment, Princeton University, Princeton, NJ 08544, USA. [7]School of Chemical Engineering, Purdue University, West Lafayette, IN 47907, USA. [8]Department of Mechanical Engineering and Materials Science, University of Pittsburgh, Pittsburgh, PA 15261, USA. [9]Toyota Research Institute, Los Altos, CA 94022, USA. [10]Center for Materials Innovation, University of Maryland, College Park, MD 20742, USA.
*Corresponding author. Email: binghu@umd.edu
†These authors contributed equally to this work.

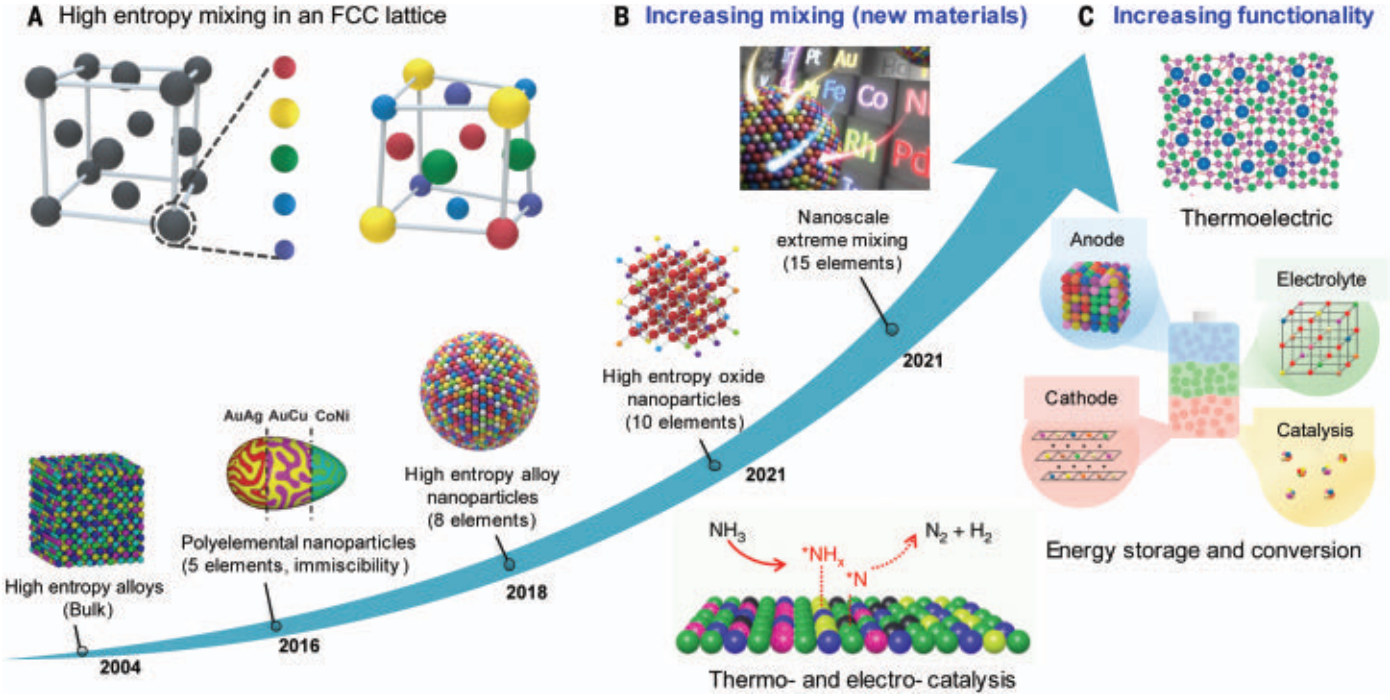


Fig. 1. Development of high-entropy nanoparticles with multielemental composition and enhanced functionality. (**A**) Schematic showing high-entropy mixing in a face-centered cubic lattice. Multiple elements will occupy the same lattice site randomly to form a high-entropy structure such as a high-entropy alloy. (**B**) The study of bulk high-entropy alloys has taken off and gained substantial interests since 2004 (*1*, *3*). In 2016, a multielemental nanoparticle library was synthesized (though with immiscibility, and thus phase segregation), followed by various single-phase, high-entropy nanoparticles with an increasing number and range of elements (*7*, *8*, *14*, *20*). Reprinted from (*14*) with permission from Elsevier. (**C**) These high-entropy nanoparticles have found critical application in thermo- and electro-catalysis, energy storage and conversation, and environmental and thermoelectric technologies (*29*–*31*, *35*, *36*). Reprinted from (*31*) with permission (copyright 2021 American Chemical Society) and from (*35*) with permission. Other portions of the figure are reprinted from (*7*, *8*) with permission, from (*20*) with permission from Springer-Nature, and from (*36*) CC BY 4.0.

accelerated exploration of high-entropy nanoparticles, along with the remaining challenges and future directions for this field. We intend to stimulate continuing and integrated efforts from multiple disciplines to study high-entropy nanoparticles and explore synthesis-structure-property relationships in the multidimensional space. Note that we use the term "high-entropy nanoparticles" to refer to such particles with a complex composition (five or more elements) and solid-solution structure rather than the conventional definition based on the somewhat subjective threshold of 1.5 k_B per atom for metals or per cation for ceramics, where k_B is the Boltzmann constant (*3*, *45*). In this review, although we focus on high-entropy nanoparticles, the basic concepts are expected to be applicable to other nanomaterials as well. We anticipate that with these advances, high-entropy nanoparticles will have a substantial impact in many fields and particularly catalysis, where this new material can potentially replace the longstanding noble metal counterparts.

High-entropy nanoparticle synthesis

From a thermodynamic point of view, the formation of high-entropy nanoparticles is a result of competition between enthalpy and entropy ($\Delta G = \Delta H - T \bullet \Delta S$). The configurational entropy of high-entropy nanoparticles increases with a greater number of elements and acts as a driving force for single-phase mixing (Fig. 2A). The enthalpy of the multielemental interactions (ΔH_{ij}) varies largely depending on the nature of the constituent elements, which directly affects the resulting phase under near-equilibrium conditions (Fig. 2B). For example, elemental combinations that have highly positive values of ΔH_{ij} (i.e., repelling force) cause immiscibility and phase segregation, whereas highly negative values of ΔH_{ij} (i.e., attractive force) promote structural ordering, such as intermetallic formation. If all ΔH_{ij} pairs in the multielement composition are near-zero values, indicating little attraction or repelling between these elements, the entropic term then dominates and promotes homogeneous random elemental mixing and high-entropy formation (Fig. 2B). However, because of the large physicochemical differences among different elements (i.e., the wide range of ΔH_{ij} values), natural single-phase mixing is often challenging and rare (*46*, *47*), with phase-segregated structures being more typical when using near-equilibrium approaches (e.g., wet chemistry) to synthesize multielemental nanoparticles (*7*, *48*, *49*).

The initial breakthrough in the general synthesis of high-entropy alloy (HEA) nanoparticles with a wide compositional range (including many immiscible combinations) and large elemental numbers (up to eight) was realized by a high-temperature "thermal shock" process invented by the Hu group at the University of Maryland (*8*, *50*–*52*) (Fig. 2C). The cooling rate of this synthesis approach is an important parameter because it affects the degree of nonequilibrium and structural ordering that can be achieved by the constituent elements, as described in the well-known temperature-time-transformation diagrams used in physical metallurgy and polymer curing (Fig. 2D) (*8*, *53*, *54*). The generated structures can include metallic glass nanoparticles (random mixing in a disordered lattice), regular HEA nanoparticles (random mixing in a crystalline lattice), intermetallic nanoparticles (chemical ordering between sublattices but random mixing within each sublattice), and heterogeneous nanoparticles (phase separation) (*8*, *17*). Moreover, the short duration and rapid quenching of thermal shock synthesis also assist the formation of small and uniform particles (*8*, *55*), which can be further modulated through defect engineering and appropriate substrates (*56*–*58*). Similar to this "shock"-based concept, a variety

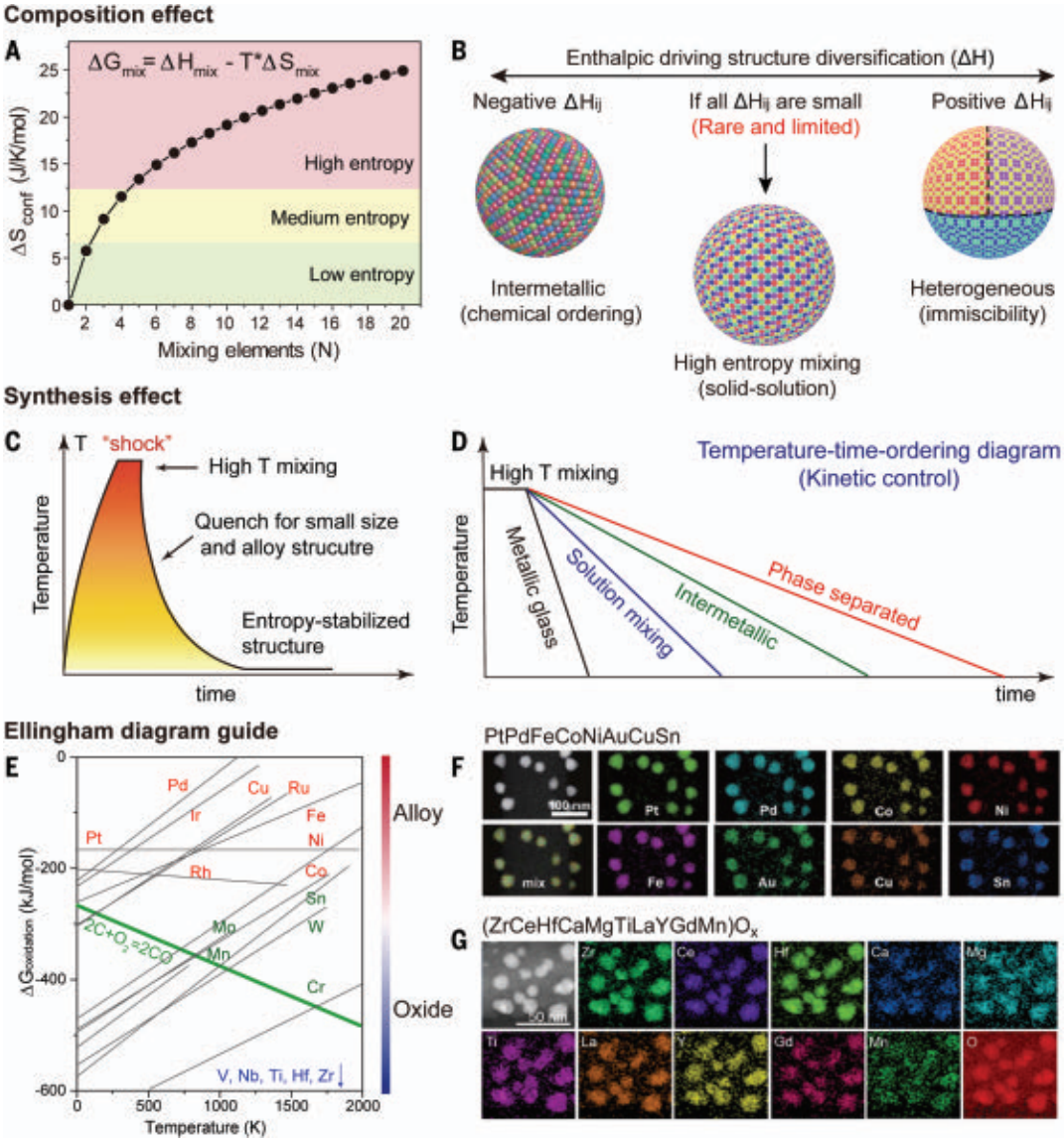


Fig. 2. High-entropy nanoparticle synthesis and structure. Thermodynamic analysis of high-entropy mixing considers both entropy (**A**) and enthalpy (**B**), which are mainly determined by the composition of high-entropy nanoparticles (*8*). (**C**) Thermal shock synthesis of high-entropy nanoparticles features a high-temperature pulse for elemental mixing and then rapid temperature quenching to maintain the high-entropy structure. (**D**) Temperature-time-transformation diagram describing how the cooling rates of high-temperature, kinetically controlled syntheses can be adjusted to form various nanoparticles featuring different degrees of structural and chemical ordering. (**E**) The Ellingham diagram [reprinted from (*14*) with permission from Elsevier] provides a guide for composing either alloy (e.g., PtPdFeCoNiAuCuSn) (*8*) (**F**) or oxide high-entropy nanoparticles (e.g., ZeCeHfCaMgTiLaYGdMnO_x) (*20*) (**G**) according to the oxidation potentials of each element. Reprinted from (*20*) with permission from Springer Nature.

of other methods have also been developed that have enabled a wide range of high-entropy nanoparticles, including vapor phase spark discharge (*13*), rapid radiative heating or annealing (*12*, *59*, *60*), acute chemical reduction (*33*, *43*), low-temperature hydrogen spillover (*61*), sputtering (*9*, *62*–*64*), transient electrosynthesis (*15*), and plasma, laser, and microwave heating (*65*–*67*), all featuring a strong kinetics-driven process. These rapid, shock-type syntheses are also fast enough to enable the efficient manufacturing of nanocatalysts (*10*, *12*, *13*, *68*, *69*).

The Ellingham diagram can be used to guide the thermochemical synthesis of high-entropy nanoparticles by illustrating the oxidation potential of the constituent elements as a function of temperature (Fig. 2E). Despite being initially developed for bulk metallurgy, we found that the Ellingham diagram is also applicable for nanoscale, shock-type reactions (*14*, *20*). Generally, elements closer to the top of the Ellingham diagram, such as noble metals and Fe, Co, and Cu, have smaller oxidation potentials (i.e., are more easily reduced) and can form alloy nanoparticles through high-temperature syntheses, such as octonary HEA nanoparticles of PtPdFeCoNiAuCuSn (Fig. 2F) (*8*). By contrast, elements near the bottom of the diagram, such as Zr, Ti, Hf, and Nb, have larger oxidation potentials and can form high-entropy oxide nanoparticles, such as (ZrCeHfCaMgTiLaYGdMn)O_x (Fig. 2G) (*20*). For the elements in the middle, such as Mo, W, and Mn (as shown in green in Fig. 2E) with moderate oxidation potentials, different synthesis strategies have been explored that can toggle the elements between their metallic and oxide states, thus expanding possible

high-entropy alloy or oxide elemental spaces (*14*, *20*, *70*). In addition to high-entropy oxides (*8*), other high-entropy compounds (e.g., sulfides and carbides) have also been synthesized with a wide range of sizes, shapes, and phases (*8*, *12*, *14*–*18*, *26*–*30*, *71*).

Advanced characterization

High-entropy nanoparticles should display a single-phase structure, demonstrating uniform and random mixing of the constituent elements. However, the characterization of this random mixing of multielements and their synergy is very challenging. Conventional techniques, such as powder x-ray diffraction (XRD, λ = 1.5418 Å), scanning and transmission electron microscopy (SEM and TEM, respectively), and x-ray photoelectron spectroscopy (XPS), can help to determine the basic phase structure, morphology, elemental distribution, and valence state, but may lack the required resolution to decouple the multielemental mixing. Synchrotron x-ray–based techniques, which use a much shorter wavelength (e.g., λ = 0.2113 Å), can provide a high resolution to better understand the atomic arrangement, bonding and coordination, and electronic properties of high-entropy nanoparticles (Fig. 3A). For example, synchrotron XRD can detect the overall phase structure

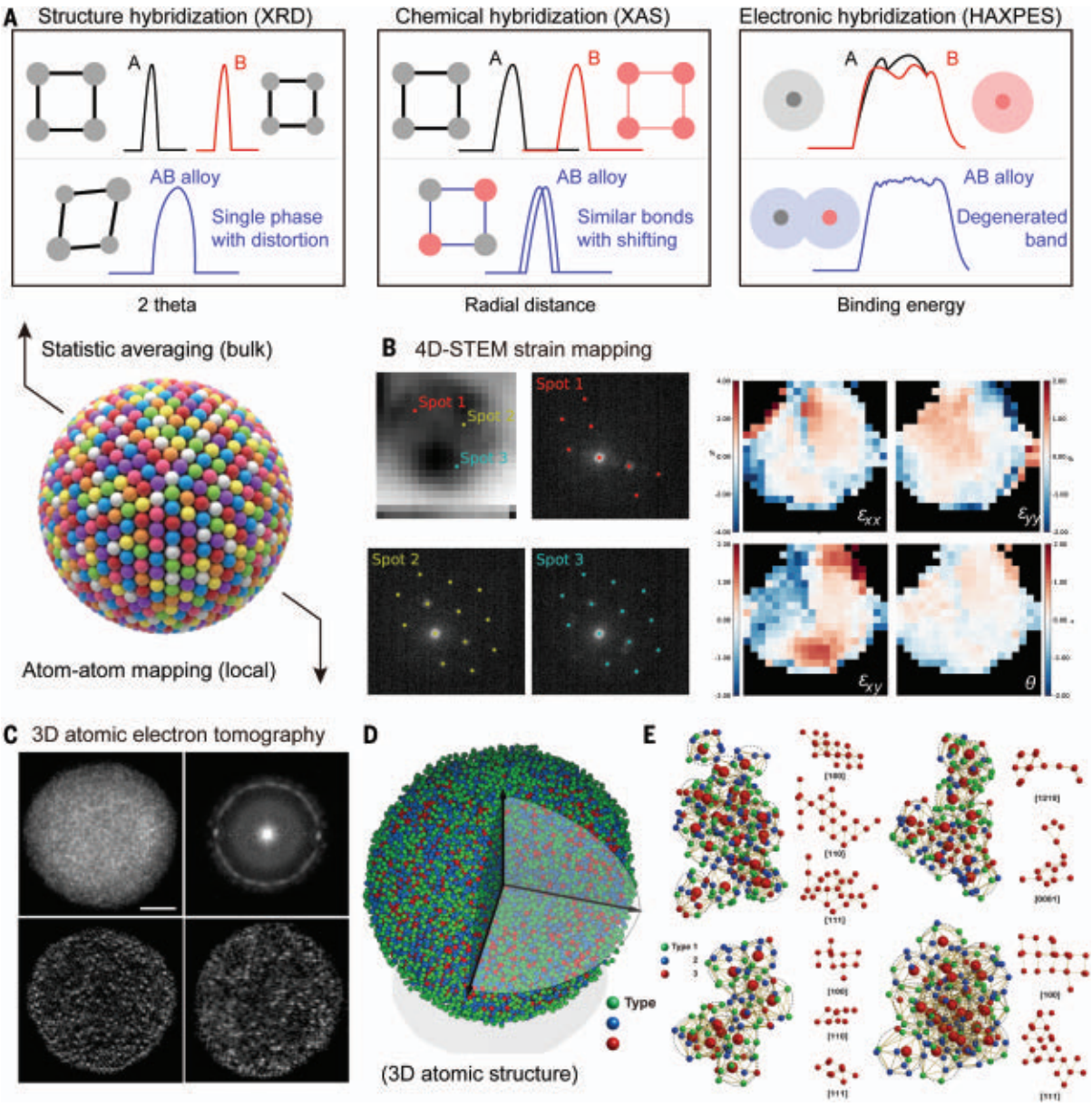


Fig. 3. Advanced characterization of high-entropy nanoparticles. (**A**) Schematic of the macroscopic and bulk characterization of the structural, chemical, and electronic hybridization in high-entropy nanoparticles through x-ray–based techniques, including XRD, XAS, and HAXPES, particularly using synchrotron x-ray sources that provide higher resolution. (**B**) 4D-STEM and strain mapping of a high-entropy nanoparticle, in which local diffraction (e.g., spots 1 to 3) is collected and compared with the average structure to derive the local lattice strain distribution including tensile (red) and compressive (blue). Reprinted from (*14*) with permission from Elsevier. (**C**) Determining the 3D atomic structure of a high-entropy metallic glass nanoparticle by AET. Shown are a representative experimental image (top left), the average 2D power spectrum (top right), and two 2.4-Å-thick slices of the 3D reconstruction in the *x*–*y* and *y*–*z* planes (bottom). Scale bar, 2 nm. (**D**) Experimental 3D atomic model of the high-entropy metallic glass nanoparticle. (**E**) Identification of four types of crystal-like medium-range order that coexist in the metallic glass nanoparticle based on AET results (*16*). Reprinted from (*8*) with permission and from (*36*) (CC BY 4.0).

as well as possible immiscible phases and impurities in high-entropy nanoparticles with greater accuracy (*49*, *72*), confirming whether high-entropy mixing is achieved. x-ray absorption spectroscopy (XAS) is an element-specific technique that can be used to study the atomic and/or local coordination environment of each element, which is critical to understanding the multielemental mixing and possible short-range or local ordering in high-entropy nanoparticles (*14*, *73*, *74*). Finally, hard x-ray photoelectron spectroscopy (HAXPES) can reveal the electronic structure (e.g., valence band and d-band center) in high-entropy nanoparticles, which is closely related to the adsorption and binding energy of key reaction intermediates, helping to rationalize the corresponding catalytic activity (*75*).

Although x-ray techniques can provide statistical analysis, electron microscopy–based techniques are critical to directly visualizing the particle size and distribution, phase, structure, composition, and chemical environments. For example, in situ TEM has been used to study nanoparticles synthesized by the high-temperature shock method, revealing the formation process as well as their dispersion and stability on defective carbon substrates (*76*). Another advanced approach that could meet higher-throughput and higher-resolution needs is four-dimensional scanning transmission electron microscopy (4D-STEM) (Fig. 3B) (*14*, *77*). 4D-STEM uses a small probe (~1 nm) to scan a large geometric region of up to ~1 × 1 μm^2 in area, thus enabling fast and high-resolution characterization of high-entropy nanoparticles for local lattice distortion, structural heterogeneity, and short-range ordering (*78*). As shown in Fig. 3B, local diffraction patterns can be obtained on high-entropy nanoparticles, and corresponding strain maps within the nanoparticles can be generated by comparing the differences between the local and average phase structures (*14*, *77*), indicating potential lattice distortion and strain.

For more advanced characterization of the 3D atomic structure, atomic electron tomography (AET) has proven to be the method of choice (Fig. 3C) (*79*–*81*). Very recently, AET was sufficiently advanced to resolve the 3D atomic structure of a high-entropy metallic glass nanoparticle containing eight elements: Co, Ni, Ru, Rh, Pd, Ag, Ir, and Pt (Fig. 3C) (*16*). Because the image contrast of AET depends on the atomic number, AET is currently only sensitive enough to classify the eight elements into three types: Co and Ni as type 1 (green); Ru, Rh, Pd, and Ag as type 2 (blue); and Ir and Pt as type 3 (red) (Fig. 3, D and E). Figure 3D shows the 3D atomic structure of the high-entropy metallic glass nanoparticle, in which the type 1, 2, and 3 atoms are uniformly distributed. The 3D atomic structure revealed four different crystal-like medium-range orderings, including face-centered cubic, hexagonal close-packed, body-centered cubic, and simple cubic structures that coexist in the high-entropy nanoparticle (Fig. 3E). These results provide direct experimental evidence that support the general framework of the efficient cluster packing model of metallic glasses (*82*) and demonstrate how AET techniques will enable researchers to study the 3D structure of high-entropy nanoparticles at the single-atom level.

Multifunctional catalytic activity

Previously, high-entropy materials, particularly high-entropy alloys, were mostly used for structural engineering applications (*3*). Wang *et al.* for the first time demonstrated that high-entropy alloy nanoparticles can serve as highly efficient catalysts in thermocatalysis (*8*, *36*). In catalysis, the binding of reactants or intermediates to the catalyst surface should be neither too strong nor too weak (the Sabatier principle) to maximize the performance, thus showing a "volcano plot" in the dependence of activity on binding energy (*83*, *84*). As schematically shown in Fig. 4A, the binding energy distribution patterns of individual elements (e.g., Co, Mo, Fe, Ni, and Cu) often exhibit sharp peaks because of their relatively fixed structure and adsorption sites. However, when multiple elements are mixed into high-entropy alloys (e.g., CoMoFeNiCu), their adsorption energy could transform into a broadened, multipeak, nearly continuous spectrum through electronic hybridization. Recently, Löffler *et al.* reported "current-wave" patterns in electrocatalysis on high-entropy catalysts, where multiple inflection points and current plateaus were observed, a strong indication of multiple active site centers in the high-entropy nanoparticles (*40*, *42*).

Because of the unique binding energy distribution, high-entropy nanoparticles can be readily tuned to obtain the desired surface properties for optimal catalytic performance (*28*, *40*, *62*). For example, in the NH_3 decomposition reaction ($2NH_3 \rightarrow N_2 + 3H_2$), it was theoretically proposed that non-noble Co–Mo alloy could outperform Ru because of the optimized *N adsorption based on the theoretical analysis (volcano plot in Fig. 4B) (*85*); however, such a design is hindered by the immiscibility between Co and Mo. Recently, alloyed CoMo-based catalysts were demonstrated in $(Co_xMo_{1-x})_{70}(FeNiCu)_{30}$ HEA nanoparticles synthesized using the thermal shock method (*36*). The Co:Mo elemental ratio can be tuned to optimize the nitrogen adsorption energy (ΔE_N) under the given reaction conditions, achieving superior performance compared with Ru, the most active monometallic catalyst (Fig. 4C). Similar high performances of high-entropy nanocatalysts have also been observed in many other systems (*9*, *15*, *20*, *43*, *62*, *73*, *75*, *86*, *87*), demonstrating the importance of multielemental design and compositional tunability. It should be noted that the diverse and heterogeneous active sites can lead to statistical variations in local activity (<50 nm) but overall repeatable performances (*88*).

Theoretically, the volcano plot can be interpreted as the result of the linear scaling relation (LSR) in first-principle calculation studies (*85*, *89*). The LSR says that in a complex or multistep reaction, the adsorption energies of reaction intermediates (e.g., O* or OH*) are linearly linked or scale linearly (*83*); in other words, strong adsorption of reactants will likely lead to the strong adsorption of products (i.e., difficult to desorb), thus slowing down the reaction substantially (*90*). Many strategies have been proposed to circumvent the LSR in nanoparticle catalyst design, including the introduction of co-adsorbates and tethers, promoters, ligands, and new alloys with complex synergy between the constituent elements (*83*, *85*). High-entropy nanoparticles offer complex atomic configurations, diverse adsorption sites, and tunable binding energies that could lead to a range of new opportunities compared with simple catalysts (*83*). For example, Wu *et al.* reported noble IrPdPtRhRu HEA nanoparticles for the hydrogen evolution reaction ($2H_2O \rightarrow O_2 + 2H_2$) and found that the material displayed superior performance compared with individual metals (Ir, Pd, Pt, Rh, and Ru) (Fig. 4D) (*75*). More importantly, the turnover frequency of the IrPdPtRhRu was far beyond what was expected by traditional LSR theories (blue region in Fig. 4E), suggesting the HEA's ability to circumvent the LSR predictions.

In addition, the broadband adsorption energy landscape of high-entropy nanoparticles is particularly promising for catalysis in tandem and complex reactions, which normally require different active sites and adsorption for multiple reaction intermediates to achieve overall high activity and/or selectivity (*27*, *71*). For example, in the ethanol oxidation reaction, which involves a complex 12-electron transfer and a range of intermediates, high-entropy PtPdRuRhOsIr (PGM-HEA) nanoparticles not only demonstrated a much higher activity than monometallic catalysts and their physical mixture but also enabled a much higher 12-electron selectivity to complete oxidation to CO_2 (Fig. 4F) (*43*, *62*, *71*). In another example, $Ru_{22}Fe_{20}Co_{18}Ni_{21}Cu_{19}$ HEA nanoparticles demonstrated high activity and selectivity in the nitrogen reduction reaction (NRR: $N_2 + 3H_2 \rightarrow 2NH_3$) (*38*). Theoretical analysis found that Fe in the HEA is suitable for N_2 adsorption and dissociation, whereas the nearby Co–Cu and Ru–Ni combinations favor H_2 adsorption and dissociation, illustrating the importance of multifunctional active sites for overall efficient NH_3 synthesis. Similarly, high-performance high-entropy nanoparticles have been reported for other complex

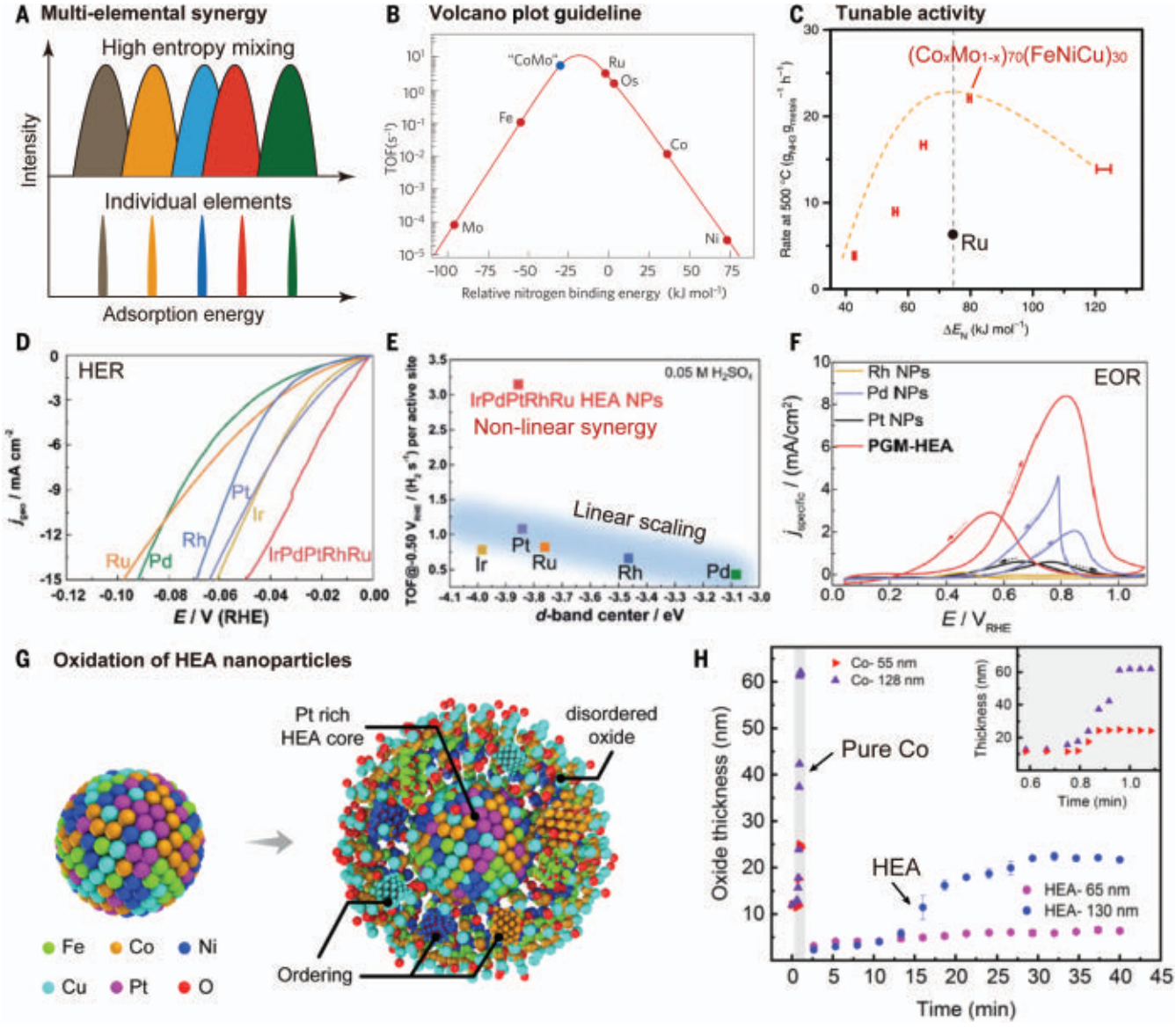


Fig. 4. High-entropy nanoparticles in catalytic reactions. (**A**) Multielemental synergy in high-entropy nanoparticles leads to multiactive sites and a broadband binding energy distribution pattern (*40*, *42*). (**B**) The composition volcano plot is a facile guide for designing high-performance catalysts, in which alloying can enable tuning of the adsorption energy toward peak performance, such as CoMo alloys. Reprinted from (*100*) with permission from Springer Nature. (**C**) Optimized CoMoFeNiCu HEA nanoparticles showed a four times higher conversion rate at 500°C compared with Ru, which was achieved by adjusting the composition ratio between Co and Mo to adjust the ΔE_N. Reprinted from (*36*) (CC BY 4.0). (**D** and **E**) IrPdPtRhRu HEA nanoparticles display superior hydrogen evolution reaction performance (D) and a much higher turnover frequency than that of the individual metals after a linear scaling relation (E), indicating strong nonlinear synergy in HEA catalysts. Reprinted from (*75*). (**F**) PGM-HEA (IrPdPtRhRuOs) nanoparticles show excellent performance for the complex and multistep ethanol oxidation reaction (EOR) compared with individual metals (*43*). (**G**) In situ oxidation of HEA nanoparticles (PtFeCoNiCu) leads to an HEA-core/oxide-shell structure. (**H**) HEA nanoparticles show much slower logarithmic oxidation kinetics compared with pure Co, which catastrophically oxidized in 1 min. Reprinted from (*97*) with permission (copyright 2021 American Chemical Society).

and multistep reactions, such as the CO_2 reduction reaction (*39*, *71*, *91*) and the oxidation of various chemicals such as methanol (*11*, *13*, *43*).

Stability

High-entropy nanoparticles can potentially provide enhanced stability for catalytic applications, similar to their bulk scale counterparts that feature improved structural stability (*3*, *45*, *53*, *92*). Thermodynamically, the high-entropy nature benefits the formation and stabilization of high-entropy nanoparticles ($\Delta G = \Delta H - T\Delta S$), especially at high temperatures where the $T\Delta S$ term is more pronounced (*20*, *73*, *93*). In situ TEM analysis has revealed the stability of high-entropy alloy and oxide nanoparticles, where the size distribution, particle dispersion, and solid-solution phase remain unchanged even when subjected to temperatures up to 1073 K (*13*, *20*, *73*). Kinetically, the high-entropy mixing may also improve the structural stability because of the size mismatch of the different elements and resultant lattice distortion, which can cause large diffusion barriers that help to prevent phase segregation, particularly at low temperature (*2*, *20*, *53*, *70*). As an example, the diffusion coefficient of Ru atoms in RuRhCoNiIr HEA nanoparticles was simulated to be two orders of magnitude lower compared with the diffusion

of Ru in bimetallic Ru–Ni, suggesting better kinetic stability in HEA nanoparticles (*73*). Another important factor affecting catalyst stability is the interfacial bonding between catalysts and supports to avoid particle aggregation. The high-temperature shock syntheses can enable better interfacial stability between the high-entropy nanoparticles and substrates (*76*, *87*, *94*). Experimentally, the stability of high-entropy catalysts has been illustrated by their steady performance in both high-temperature and electrochemical catalytic reactions (*11*–*13*, *15*, *20*, *36*–*38*, *43*, *44*, *73*, *75*). However, the entropy stabilization role could become limited and surface reconstruction can easily occur in harsh conditions (*95*–*97*). For example, Shahbazian-Yassar *et al.* performed in situ oxidation of $Fe_{0.28}Co_{0.21}Ni_{0.20}Cu_{0.08}Pt_{0.23}$ HEA nanoparticles and observed surface oxidation of the non-noble elements while the core of the HEA nanoparticles remained stable with a Pt-rich composition (Fig. 4G). Qualitative analysis revealed that the HEA nanoparticles exhibit logarithmic oxidation kinetics, resulting in a stable HEA-core and oxide-shell structure after 40 minutes of exposure in an oxidative environment at 400°C (Fig. 4H). By contrast, pure Co nanoparticles underwent catastrophic oxidation kinetics and oxidized within ~1 min. The surface reconstruction or transformation of high-entropy catalysts is often more evident in electrochemical reactions, in which the entropic stabilization effect is less profound compared with chemical leaching and electrochemical redox. Nevertheless, many studies have reported stable performance of high-entropy nanoparticles in diverse electrochemical conditions, especially compared with their fewer-element counterparts (*12*, *29*, *75*, *86*, *87*).

High-throughput screening

Despite the superior catalytic performances observed in several cases, it remains unknown how to generally develop high-entropy nanoparticles for targeted catalytic reaction schemes. In addition, identifying catalytic active sites in high-entropy nanoparticles is challenging because of the complex microstructure and binding energy distribution pattern (*40*, *42*, *98*). These issues may be resolved by taking advantage of emerging high-throughput (*64*, *99*, *100*) and data-driven material discovery approaches (*27*, *71*, *101*–*103*).

Computationally, first-principles-based methods have been developed to predict the composition-structure-property relationships of high-entropy nanoparticles (*84*, *100*, *104*). Additionally, high-throughput computation has been demonstrated for phase prediction of multielemental compositions by following empirical rules derived from high-entropy materials (*46*, *73*) or using calculation of phase diagram (CALPHAD) methods with largely reduced parameter spaces (*105*), both of which are capable of screening millions of elemental compositions (Fig. 5A). However, these calculations are mostly based on thermodynamic equilibrium considerations of bulk materials, which may not be readily transferable to high-entropy nanoparticles because of their small size and synthesis under nonequilibrium conditions.

For the prediction of functional properties (e.g., catalysis) of high-entropy nanoparticles, there are additional challenges, such as building precise atomic packing models and determining binding sites (*11*). Recently, Rossmeisl *et al.* developed a high-throughput computation method combined with supervised learning to explore the random atomic configurations in high-entropy nanoparticles and predict their adsorption energies in catalysis (Fig. 5B) (*27*, *71*, *101*). The authors also simulated the near-continuous binding energy distribution pattern (Fig. 5B) for high-entropy catalysts. On the basis of these calculations, high-performance multielemental catalysts for oxygen reduction and CO_2 reduction were experimentally realized (*27*, *71*, *101*, *106*). Additional machine learning (ML)–based methods are being developed to efficiently explore the configurations of adsorbates on multielemental surfaces, including the effects of variable adsorbate coverage, multiple adsorption species, and surface reconstruction on the catalytic properties (*107*).

Experimentally, researchers have demonstrated the combinatorial synthesis and high-throughput screening of multielemental catalysts (*64*, *108*–*110*). For example, Ludwig *et al.* achieved the combinatorial synthesis of hundreds of high-entropy compositions (~342 per batch) on thin-film substrates using co-sputtering of multiple metal sources, along with high-throughput characterization, including energy-dispersive spectroscopy (composition), XRD (structure), and scanning droplet cell (electrochemistry), to rapidly screen these 2D thin-film samples for rapid catalyst discovery (*64*, *111*–*113*). Direct high-throughput synthesis and screening of high-entropy nanoparticles have also been achieved (*9*, *72*). By combinatorial co-sputtering into ionic liquid (~ 40 μl per cavity, in total 64 cavities), Ludwig *et al.* demonstrated synthesis of CrMnFeCoNi-based HEA nanoparticles immobilized on the microelectrodes with various compositions, which led to the discovery of $Cr_9Mn_{60}Fe_9Co_{11}Ni_{11}$ with exceptional activity for oxygen reduction reaction (*9*). Yao *et al.* reported the high-throughput synthesis of ultrafine and homogeneous HEA nanoparticles with different elemental combinations from binary up to octonary PtPdRhRuIrNiCoFe (Fig. 5C) (*72*). In the process, different metal-precursor solutions were ink printed, followed by a high-temperature radiation shock synthesis to obtain uniform microstructures despite different compositions. Scanning droplet cell screening then enabled the discovery of high-performance PtPdFeCoNi HEA catalysts for oxygen reduction reaction, the catalytic performances of which were verified by conventional rotating disk electrode measurement (*72*). The combinatorial synthesis and high-throughput screening pipeline therefore presents a new paradigm for accelerated exploration of high-entropy nanoparticles.

ML acceleration and active exploration

ML is an excellent tool with which to accelerate materials discovery by enabling extensive prediction of unmeasured compositions (the generalization process in ML), guided exploration to quickly find the performance optima (active learning in ML), and quantitatively understanding of composition and process-structure-property relationships (feature analysis in ML) (*27*, *63*, *71*, *101*–*103*, *114*, *115*). As an example, ML prediction has been used to guide the design of ternary medium-entropy PtFeCu catalysts, illustrating the closed-loop process by (i) model building and simulation data generation, (ii) ML and fitting of the simulation data, (iii) extensive exploration and screening by ML in a larger compositional space, and (iv) experimental verification and feedback to previous simulations and ML models (Fig. 5D) (*116*). A similar process has also been demonstrated for multielemental catalysts in the Ag-Ir-Pd-Pt-Ru space by combining computational prediction with ML and using thin-film–based high-throughput synthesis and screening for data feedback and model refining, thus forming a closed-loop optimization protocol to improve the prediction power toward high-performance catalysts (*63*).

Despite these advances, current efforts can at most cover <1% of available compositions in high-entropy nanoparticles (*113*). Therefore, guided optimization and careful sampling are critical for identifying important data points to save exploration efforts. This can be realized by using active learning methods (e.g., Bayesian optimization and reinforcement learning) (*103*, *117*–*119*). For example, Rossmeisl *et al.* used Bayesian optimization with Gaussian process surrogate function models to discover multielemental catalysts based on computational data (*120*). With ~150 iterations based on Bayesian optimization, many important local optima of targeted properties were discovered, illustrating the great promise of active learning in exploring the vast multidimensional space. Such approaches can also be combined with graph network descriptors of the HEA surfaces and neural networks to accelerate the development of surrogate computational models of the surface and adsorption properties (*107*). Active learning can also enable multiobjective optimization, which has not yet been realized in the development of high-entropy nanoparticles but is highly desirable toward the

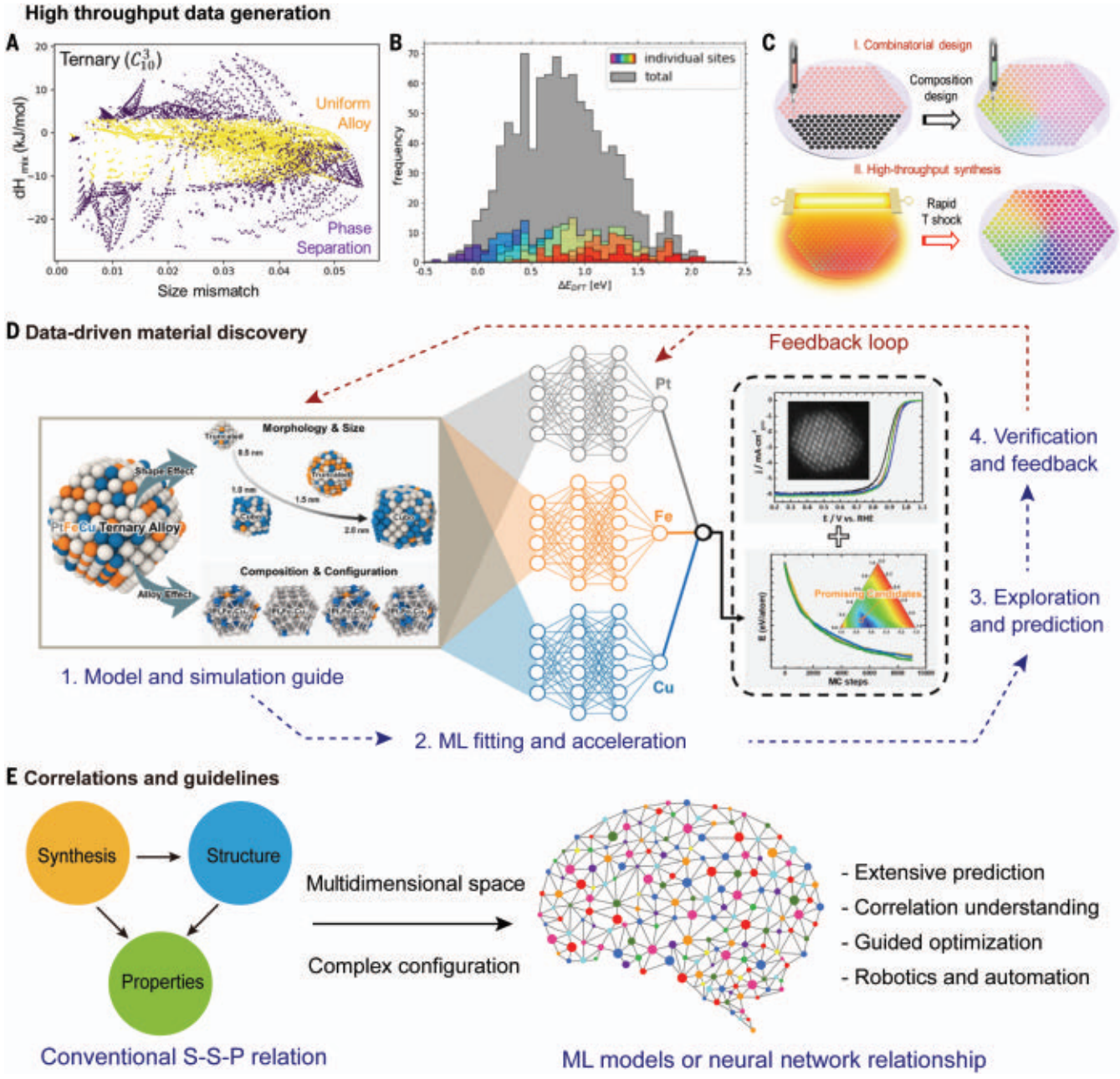


Fig. 5. Data-driven high-entropy nanoparticle discovery. (**A** and **B**) High-throughput computation for structural prediction based on size mismatch and enthalpy (*73*) (A) and adsorption sites and binding energy distribution patterns to predict catalytic properties (B). Reprinted from (*27*) with permission from Elsevier. (**C**) Example of the combinatorial and high-throughput synthesis of high-entropy nanoparticles (*72*). (**D**) Data-driven methodology for the discovery of PtFeCu catalysts consists of (i) modeling and simulation, (ii) ML fitting and acceleration, (iii) composition exploration and prediction, and (iv) experimental verification and feedback. Reprinted from (*116*) with permission from Elsevier. (**E**) Synthesis-structure-property relationships in conventional materials research may be replaced by data-driven approaches featuring ML-trained models for prediction, understanding, and optimization, which could even enable automated discovery.

goal of achieving superior catalysts with simultaneously high activity, selectivity, and stability.

Understanding synthesis-structure-property relationships will always be challenging for material systems as complex as high-entropy nanoparticles. Some initial efforts have used theoretical models and neural networks to decouple the multielemental synergy into ligand effects (i.e., different elements) and coordination effects (i.e., different structures) to correlate the structural features with their catalytic performances (*101*). Therefore, instead of a traditional synthesis-structure-property relationship built upon a clear picture of the catalytic mechanism, data-trained mathematical models may gradually learn and facilitate property prediction, guided optimization, and fundamental understanding of high-entropy nanoparticle catalysts (Fig. 5E). Such trained models (in the form of a Gaussian process model or a neural network) may become a new norm for the study of complex materials such as high-entropy nanoparticles.

Conclusion and outlook

Great progress has been made on high-entropy nanoparticles, but to further advance their

development, continued efforts are needed in many areas, such as synthesis methodologies, advanced characterization, fundamental understanding, and application- and data-driven discoveries, as described below.

Tunable synthesis is currently the most explored aspect of high-entropy nanoparticles and now requires precision. Considering the immiscibility caused by the elemental differences and compositional complexity in high-entropy nanoparticles, syntheses must continue to rely on nonequilibrium approaches in terms of temperature, force, pressure, energy field, etc., to achieve uniform mixing and small particle size. Furthermore, we still need to learn how to balance nonequilibrium syntheses with delicate structural or morphology control in terms of size, phase, shape, facets, and surface decoration, which will require considerable effort and knowledge gained from existing wet chemistry.

An important aspect of high-entropy nanoparticle research that is currently lacking is a fundamental understanding of the surfaces, defects, and elemental distribution in high-entropy nanoparticles, which will have a profound impact on the catalytic properties. We have not yet established basic knowledge of surface or interface elemental segregation, reconstruction, and electronic structure, especially their dynamic evolution under catalytic operation conditions. Integrating state-of-the-art in situ electron microscopies, such as in situ liquid and environmental microscopy, into advanced atomic-resolution chemical analysis and atomic structural imaging will provide valuable insight into the fundamental understanding of active sites and reaction pathways in high-entropy nanomaterials for catalytic applications. Also, we envision the combination of atomic-resolution in situ environmental microscopy with ML-assisted data acquisition and analysis to allow us to capture the critical dynamic changes of high-entropy nanomaterials during catalytic reactions. Information such as the evolution of surface atomic structure, lattice strain, chemical diffusion, and electronic structures of high-entropy nanoparticles will be attained, providing reliable inputs for theoretical calculations and insights into understanding reaction pathways.

High-entropy nanoparticles have great promise for high-performance catalysis and are particularly advantageous for multistep and tandem reactions that require a combination of different active sites. However, it remains an open question how to properly design high-entropy nanoparticles to best fit those reaction schemes. In addition, it is unclear how to identify the active sites and understand the performance origin. Although catalyst discovery based on traditional routes is possible, high-entropy nanoparticle research would greatly benefit from the advancement of high-throughput methodologies and data mining. Currently, combinatorial syntheses and high-throughput screening are mostly limited to thin-film samples and simple electrochemical reactions. Additionally, high-throughput computation often comes at the price of precision for simplicity or computational efficiency, leading to some disparity between screening results and actual trends of performance. Therefore, these data-driven methodologies will likely require the most significant effort in the next stages of research.

Many published research results demonstrate different compositions with interesting properties, but more systematic and standardized reporting is necessary to take full advantage of these "expensive" data (*102*). Therefore, establishing reporting standards for a sharable data repository should be developed so that the knowledge can be better collected and analyzed. Some such efforts are already taking place, such as the establishment of the Materials Data Bank for archiving the 3D atomic coordinates and chemical species of a wide range of materials including multielement and high-entropy nanoparticles determined by AET (*121*). The experimentally determined 3D atomic models of high-entropy nanoparticles can be coupled with computational and ML methods to understand their structure-property relationships at the fundamental level. We expect that with the expanding knowledge of the synthesis-structure-property relationships of high-entropy nanoparticles, an integrated material discovery workflow combining ML-guided optimization and screening will soon become possible to expedite progress in this promising field, particularly multi-objective optimization toward simultaneously high activity, selectivity, stability, and low cost.

REFERENCES AND NOTES

1. J. W. Yeh *et al*., Nanostructured high-entropy alloys with multiple principal elements: Novel alloy design concepts and outcomes. *Adv. Eng. Mater.* **6**, 299–303 (2004). doi: 10.1002/adem.200300567
2. C. M. Rost *et al*., Entropy-stabilized oxides. *Nat. Commun.* **6**, 8485 (2015). doi: 10.1038/ncomms9485; pmid: 26415623
3. E. P. George, D. Raabe, R. O. Ritchie, High-entropy alloys. *Nat. Rev. Mater.* **4**, 515–534 (2019). doi: 10.1038/s41578-019-0121-4
4. B. Niu *et al*., Sol-gel autocombustion synthesis of nanocrystalline high-entropy alloys. *Sci. Rep.* **7**, 3421 (2017). doi: 10.1038/s41598-017-03644-6; pmid: 28611380
5. M. P. Singh, C. Srivastava, Synthesis and electron microscopy of high entropy alloy nanoparticles. *Mater. Lett.* **160**, 419–422 (2015). doi: 10.1016/j.matlet.2015.08.032
6. C.-F. Tsai, P.-W. Wu, P. Lin, C.-G. Chao, K.-Y. Yeh, Sputter deposition of multi-element nanoparticles as electrocatalysts for methanol oxidation. *Jpn. J. Appl. Phys.* **47**, 5755–5761 (2008). doi: 10.1143/JJAP.47.5755
7. P.-C. Chen *et al*., Polyelemental nanoparticle libraries. *Science* **352**, 1565–1569 (2016). doi: 10.1126/science.aaf8402; pmid: 27339985
8. Y. Yao *et al*., Carbothermal shock synthesis of high-entropy-alloy nanoparticles. *Science* **359**, 1489–1494 (2018). doi: 10.1126/science.aan5412; pmid: 29599236
9. T. Löffler *et al*., Discovery of a multinary noble metal-free oxygen reduction catalyst. *Adv. Energy Mater.* **8**, 1802269 (2018). doi: 10.1002/aenm.201802269
10. X. Wang *et al*., Continuous synthesis of hollow high-entropy nanoparticles for energy and catalysis applications. *Adv. Mater.* **32**, e2002853 (2020). doi: 10.1002/adma.202002853; pmid: 33020998
11. H. Li *et al*., Fast site-to-site electron transfer of high-entropy alloy nanocatalyst driving redox electrocatalysis. *Nat. Commun.* **11**, 5437 (2020). doi: 10.1038/s41467-020-19277-9; pmid: 33116124
12. S. Gao *et al*., Synthesis of high-entropy alloy nanoparticles on supports by the fast moving bed pyrolysis. *Nat. Commun.* **11**, 2016 (2020). doi: 10.1038/s41467-020-15934-1; pmid: 32332743
13. J. Feng *et al*., Unconventional alloys confined in nanoparticles: Building blocks for new matter. *Matter* **3**, 1646–1663 (2020). doi: 10.1016/j.matt.2020.07.027
14. Y. Yao *et al*., Extreme mixing in nanoscale transition metal alloys. *Matter* **4**, 2340–2353 (2021). doi: 10.1016/j.matt.2021.04.014
15. M. W. Glasscott *et al*., Electrosynthesis of high-entropy metallic glass nanoparticles for designer, multi-functional electrocatalysis. *Nat. Commun.* **10**, 2650 (2019). doi: 10.1038/s41467-019-10303-z; pmid: 31201304
16. Y. Yang *et al*., Determining the three-dimensional atomic structure of an amorphous solid. *Nature* **592**, 60–64 (2021). doi: 10.1038/s41586-021-03354-0; pmid: 33790443
17. M. Cui *et al*., Multi-principal elemental intermetallic nanoparticles synthesized via a disorder-to-order transition. *Sci. Adv.* **8**, eabm4322 (2022). doi: 10.1126/sciadv.abm4322; pmid: 35089780
18. C.-L. Yang *et al*., Sulfur-anchoring synthesis of platinum intermetallic nanoparticle catalysts for fuel cells. *Science* **374**, 459–464 (2021). doi: 10.1126/science.abj9980; pmid: 34672731
19. T. Wang, H. Chen, Z. Yang, J. Liang, S. Dai, High-entropy perovskite fluorides: A new platform for oxygen evolution catalysis. *J. Am. Chem. Soc.* **142**, 4550–4554 (2020). doi: 10.1021/jacs.9b12377; pmid: 32105461
20. T. Li *et al*., Denary oxide nanoparticles as highly stable catalysts for methane combustion. *Nat. Catal.* **4**, 62–70 (2021). doi: 10.1038/s41929-020-00554-1
21. C. R. McCormick, R. E. Schaak, Simultaneous multication exchange pathway to high-entropy metal sulfide nanoparticles. *J. Am. Chem. Soc.* **143**, 1017–1023 (2021). doi: 10.1021/jacs.0c11384; pmid: 33405919
22. M. Cui *et al*., High-entropy metal sulfide nanoparticles promise high-performance oxygen evolution reaction. *Adv. Energy Mater.* **2002887**, 1–8 (2020).
23. Z. Du *et al*., High-entropy atomic layers of transition-metal carbides (MXenes). *Adv. Mater.* **33**, e2101473 (2021). doi: 10.1002/adma.202101473; pmid: 34365658
24. S. K. Nemani *et al*., High-Entropy 2D Carbide MXenes: $TiVNbMoC_3$ and $TiVCrMoC_3$. *ACS Nano* **15**, 12815–12825 (2021). doi: 10.1021/acsnano.1c02775; pmid: 34128649
25. Q. Dong *et al*., Rapid synthesis of high-entropy oxide microparticles. *Small* **2104761**, e2104761 (2022). doi: 10.1002/smll.202104761; pmid: 35049145
26. T. Ying *et al*., High-entropy van der Waals materials formed from mixed metal dichalcogenides, halides, and phosphorus trisulfides. *J. Am. Chem. Soc.* **143**, 7042–7049 (2021). doi: 10.1021/jacs.1c01580; pmid: 33926192
27. T. A. A. Batchelor *et al*., High-entropy alloys as a discovery platform for electrocatalysis. *Joule* **3**, 834–845 (2019). doi: 10.1016/j.joule.2018.12.015
28. T. Löffler *et al*., Toward a paradigm shift in electrocatalysis using complex solid solution nanoparticles. *ACS Energy Lett.* **4**, 1206–1214 (2019). doi: 10.1021/acsenergylett.9b00531
29. A. Amiri, R. Shahbazian-Yassar, Recent progress of high-entropy materials for energy storage and conversion. *J. Mater. Chem. A Mater. Energy Sustain.* **9**, 782–823 (2021). doi: 10.1039/D0TA09578H
30. Y. Sun, S. Dai, High-entropy materials for catalysis: A new frontier. *Sci. Adv.* **7**, eabg1600 (2021). doi: 10.1126/sciadv.abg1600; pmid: 33980494
31. Y. Chen *et al*., Opportunities for high-entropy materials in rechargeable batteries. *ACS Mater. Lett.* **3**, 160–170 (2021). doi: 10.1021/acsmaterialslett.0c00484
32. Y. Ma *et al*., High-entropy energy materials: Challenges and new opportunities. *Energy Environ. Sci.* **14**, 2883–2905 (2021). doi: 10.1039/D1EE00505G
33. K. Kusada, D. Wu, H. Kitagawa, New aspects of platinum group metal-based solid-solution alloy nanoparticles: Binary to high-entropy alloys. *Chemistry* **26**, 5105–5130 (2020). doi: 10.1002/chem.201903928; pmid: 31863514
34. W.-T. Koo, J. E. Millstone, P. S. Weiss, I.-D. Kim, The design and science of polyelemental nanoparticles. *ACS Nano* **14**,

6407–6413 (2020). doi: 10.1021/acsnano.0c03993; pmid: 32469489

35. B. Jiang *et al*., High-entropy-stabilized chalcogenides with high thermoelectric performance. *Science* **371**, 830–834 (2021). doi: 10.1126/science.abe1292; pmid: 33602853
36. P. Xie *et al*., Highly efficient decomposition of ammonia using high-entropy alloy catalysts. *Nat. Commun.* **10**, 4011 (2019). doi: 10.1038/s41467-019-11848-9; pmid: 31488814
37. H. J. Qiu *et al*., Nanoporous high-entropy alloys for highly stable and efficient catalysts. *J. Mater. Chem. A Mater. Energy Sustain.* **7**, 6499–6506 (2019). doi: 10.1039/C9TA00505F
38. D. Zhang *et al*., Multi-site electrocatalysts boost pH-universal nitrogen reduction by high-entropy alloys. *Adv. Funct. Mater.* **31**, 2006939 (2021). doi: 10.1002/adfm.202006939
39. C. Yang *et al*., Overcoming immiscibility toward bimetallic catalyst library. *Sci. Adv.* **6**, eaaz6844 (2020). doi: 10.1126/sciadv.aaz6844; pmid: 32494647
40. T. Löffler, A. Ludwig, J. Rossmeisl, W. Schuhmann, What makes high-entropy alloys exceptional electrocatalysts? *Angew. Chem. Int. Ed.* **60**, 26894–26903 (2021). doi: 10.1002/anie.202109212; pmid: 34436810
41. Y. Wang, X. Zheng, D. Wang, Design concept for electrocatalysts. *Nano Res.* **15**, 1730–1752 (2021). doi: 10.1007/s12274-021-3794-0
42. T. Löffler *et al*., Comparing the activity of complex solid solution electrocatalysts using inflection points of voltammetric activity curves as activity descriptors. *ACS Catal.* **11**, 1014–1023 (2021). doi: 10.1021/acscatal.0c03313
43. D. Wu *et al*., Platinum-group-metal high-entropy-alloy nanoparticles. *J. Am. Chem. Soc.* **142**, 13833–13838 (2020). doi: 10.1021/jacs.0c04807; pmid: 32786816
44. M. Wu *et al*., Hierarchical polyelemental nanoparticles as bifunctional catalysts for oxygen evolution and reduction reactions. *Adv. Energy Mater.* **10**, 2001119 (2020). doi: 10.1002/aenm.202001119
45. D. B. Miracle, O. N. Senkov, A critical review of high entropy alloys and related concepts. *Acta Mater.* **122**, 448–511 (2017). doi: 10.1016/j.actamat.2016.08.081
46. M. C. Troparevsky, J. R. Morris, P. R. C. Kent, A. R. Lupini, G. M. Stocks, Criteria for predicting the formation of single-phase high-entropy alloys. *Phys. Rev. X* **5**, 011041 (2015). doi: 10.1103/PhysRevX.5.011041
47. R. Guo *et al*., Enthalpy induced phase partition toward hierarchical, nanostructured high-entropy alloys. *Nano Res.* (2021). doi: 10.1007/s12274-021-3912-z
48. P.-C. Chen *et al*., Interface and heterostructure design in polyelemental nanoparticles. *Science* **363**, 959–964 (2019). doi: 10.1126/science.aav4302; pmid: 30819959
49. S. G. Kwon *et al*., Heterogeneous nucleation and shape transformation of multicomponent metallic nanostructures. *Nat. Mater.* **14**, 215–223 (2015). doi: 10.1038/nmat4115; pmid: 25362354
50. Y. Chen *et al*., Ultra-fast self-assembly and stabilization of reactive nanoparticles in reduced graphene oxide films. *Nat. Commun.* **7**, 12332 (2016). doi: 10.1038/ncomms12332; pmid: 27515900
51. L. Hu, Y. Chen, Y. Yao, "Nanoparticles and systems and methods for synthesizing nanoparticles through thermal shock," US Patent Application 20180369771 (2018).
52. Y. Yao, L. Hu, "Thermal shock synthesis of multielement nanoparticles," US Patent Application 11193191B2 (2021).
53. S. A. Kube, J. Schroers, Metastability in high entropy alloys. *Scr. Mater.* **186**, 392–400 (2020). doi: 10.1016/j.scriptamat.2020.05.049
54. M. T. Aronhime, J. K. Gillham, Time-temperature-transformation (Ttt) cure diagram of thermosetting polymeric systems. *Adv. Polym. Sci.* **78**, 83–113 (1986). doi: 10.1007/BFb0035358
55. F. Chen *et al*., High-temperature atomic mixing toward well-dispersed bimetallic electrocatalysts. *Adv. Energy Mater.* **8**, 1800466 (2018). doi: 10.1002/aenm.201800466
56. Y. Yao *et al*., Ultrafast, controllable synthesis of sub-nano metallic clusters through defect engineering. *ACS Appl. Mater. Interfaces* **11**, 29773–29779 (2019). doi: 10.1021/acsami.9b07198; pmid: 31356053
57. Y. Yao *et al*., High temperature shockwave stabilized single atoms. *Nat. Nanotechnol.* **14**, 851–857 (2019). doi: 10.1038/s41565-019-0518-7; pmid: 31406363
58. Y. Zhou *et al*., Tuning the high-temperature wetting behavior of metals toward ultrafine nanoparticles. *Angew. Chem. Int. Ed.* **57**, 2625–2629 (2018). doi: 10.1002/anie.201712202; pmid: 29346707
59. Y. Chen *et al*., Synthesis of monodisperse high entropy alloy nanocatalysts from core@shell nanoparticles. *Nanoscale Horiz.* **6**, 231–237 (2021). doi: 10.1039/D0NH00656D; pmid: 33480921
60. Y. Yang *et al*., Aerosol synthesis of high entropy alloy nanoparticles. *Langmuir* **36**, 1985–1992 (2020). doi: 10.1021/acs.langmuir.9b03392; pmid: 32045255
61. K. Mori *et al*., Hydrogen spillover-driven synthesis of high-entropy alloy nanoparticles as a robust catalyst for CO_2 hydrogenation. *Nat. Commun.* **12**, 3884 (2021). doi: 10.1038/s41467-021-24228-z; pmid: 34162865
62. T. Löffler *et al*., design of complex solid-solution electrocatalysts by correlating configuration, adsorption energy distribution patterns, and activity curves. *Angew. Chem. Int. Ed.* **59**, 5844–5850 (2020). doi: 10.1002/anie.201914666; pmid: 31867829
63. T. A. A. Batchelor *et al*., Complex-solid-solution electrocatalyst discovery by computational prediction and high-throughput experimentation. *Angew. Chem. Int. Ed.* **60**, 6932–6937 (2021). doi: 10.1002/anie.202014374; pmid: 33372334
64. A. Ludwig, Discovery of new materials using combinatorial synthesis and high-throughput characterization of thin-film materials libraries combined with computational methods. *npj Comput. Mater.* **5**, 70 (2019). doi: 10.1038/s41524-019-0205-0
65. F. Waag *et al*., Kinetically-controlled laser-synthesis of colloidal high-entropy alloy nanoparticles. *RSC Advances* **9**, 18547–18558 (2019). doi: 10.1039/C9RA03254A
66. H. Qiao *et al*., Scalable synthesis of high entropy alloy nanoparticles by microwave heating. *ACS Nano* **15**, 14928–14937 (2021). doi: 10.1021/acsnano.1c05113; pmid: 34423972
67. Y. Yao, Q. Dong, L. Hu, Overcoming immiscibility via a milliseconds-long "shock" synthesis toward alloyed nanoparticles. *Matter* **1**, 1451–1453 (2019). doi: 10.1016/j.matt.2019.11.006
68. X. Wang *et al*., Continuous 2000 K droplet-to-particle synthesis. *Mater. Today* **35**, 106–114 (2020). doi: 10.1016/j.mattod.2019.11.004
69. M. Jiao *et al*., Fly-through synthesis of nanoparticles on textile and paper substrates. *Nanoscale* **11**, 6174–6181 (2019). doi: 10.1039/C8NR10137J; pmid: 30874268
70. H. Xu *et al*., Entropy-stabilized single-atom Pd catalysts via high-entropy fluorite oxide supports. *Nat. Commun.* **11**, 3908 (2020). doi: 10.1038/s41467-020-17738-9; pmid: 32764539
71. J. K. Pedersen, T. A. A. Batchelor, A. Bagger, J. Rossmeisl, High-entropy alloys as catalysts for the CO_2 and CO reduction reactions. *ACS Catal.* **10**, 2169–2176 (2020). doi: 10.1021/acscatal.9b04343
72. Y. Yao *et al*., High-throughput, combinatorial synthesis of multimetallic nanoclusters. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 6316–6322 (2020). doi: 10.1073/pnas.1903721117; pmid: 32156723
73. Y. Yao *et al*., Computationally aided, entropy-driven synthesis of highly efficient and durable multi-elemental alloy catalysts. *Sci. Adv.* **6**, eaaz0510 (2020). doi: 10.1126/sciadv.aaz0510; pmid: 32201728
74. D. Morris *et al*., Composition-dependent structure and properties of 5- and 15-element high-entropy alloy nanoparticles. *Cell Rep. Phys. Sci.* **2**, 100641 (2021). doi: 10.1016/j.xcrp.2021.100641
75. D. Wu *et al*., On the electronic structure and hydrogen evolution reaction activity of platinum group metal-based high-entropy-alloy nanoparticles. *Chem. Sci.* **11**, 12731–12736 (2020). doi: 10.1039/D0SC02351E; pmid: 34094468
76. Z. Huang *et al*., Direct observation of the formation and stabilization of metallic nanoparticles on carbon supports. *Nat. Commun.* **11**, 6373 (2020). doi: 10.1038/s41467-020-20084-5; pmid: 33311508
77. Q. Ding *et al*., Tuning element distribution, structure and properties by composition in high-entropy alloys. *Nature* **574**, 223–227 (2019). doi: 10.1038/s41586-019-1617-1; pmid: 31597974
78. B. H. Savitzky *et al*., py4DSTEM: Open source software for 4D-STEM data analysis. *Microsc. Microanal.* **25** (S2), 124–125 (2019). doi: 10.1017/S1431927619001351
79. J. Miao, P. Ercius, S. J. L. Billinge, Atomic electron tomography: 3D structures without crystals. *Science* **353**, aaf2157 (2016). doi: 10.1126/science.aaf2157; pmid: 27708010
80. Y. Yang *et al*., Deciphering chemical order/disorder and material properties at the single-atom level. *Nature* **542**, 75–79 (2017). doi: 10.1038/nature21042; pmid: 28150758
81. J. Zhou *et al*., Observing crystal nucleation in four dimensions using atomic electron tomography. *Nature* **570**, 500–503 (2019). doi: 10.1038/s41586-019-1317-x; pmid: 31243385
82. D. B. Miracle, A structural model for metallic glasses. *Nat. Mater.* **3**, 697–702 (2004). doi: 10.1038/nmat1219; pmid: 15378050
83. J. Pérez-Ramírez, N. López, Strategies to break linear scaling relationships. *Nat. Catal.* **2**, 971–976 (2019). doi: 10.1038/s41929-019-0376-6
84. J. Greeley *et al*., Alloys of platinum and early transition metals as oxygen reduction electrocatalysts. *Nat. Chem.* **1**, 552–556 (2009). doi: 10.1038/nchem.367; pmid: 21378936
85. A. Vojvodic, J. K. Nørskov, New design paradigm for heterogeneous catalysts. *Natl. Sci. Rev.* **2**, 140–143 (2015). doi: 10.1093/nsr/nwv023
86. W.-B. Jung *et al*., Polyelemental nanoparticles as catalysts for a Li-O_2 battery. *ACS Nano* **15**, 4235–4244 (2021). doi: 10.1021/acsnano.0c06528; pmid: 33691412
87. T. Li *et al*., Carbon-supported high-entropy oxide nanoparticles as stable electrocatalysts for oxygen reduction reactions. *Adv. Funct. Mater.* **31**, 2010561 (2021). doi: 10.1002/adfm.202010561
88. E. B. Tetteh *et al*., Zooming-in – Visualization of active site heterogeneity in high entropy alloy electrocatalysts using scanning electrochemical cell microscopy. *Electrochem. Sci. Adv.* •••, 2100105 (2021). doi: 10.1002/elsa.202100105
89. J. Greeley, Theoretical heterogeneous catalysis: Scaling relationships and computational catalyst design. *Annu. Rev. Chem. Biomol. Eng.* **7**, 605–635 (2016). doi: 10.1146/annurev-chembioeng-080615-034413; pmid: 27088666
90. A. Kulkarni, S. Siahrostami, A. Patel, J. K. Nørskov, Understanding catalytic activity trends in the oxygen reduction reaction. *Chem. Rev.* **118**, 2302–2312 (2018). doi: 10.1021/acs.chemrev.7b00488; pmid: 29405702
91. J. Cavin *et al*., 2D high-entropy transition metal dichalcogenides for carbon dioxide electrocatalysis. *Adv. Mater.* **33**, e2100347 (2021). doi: 10.1002/adma.202100347; pmid: 34173281
92. Z. Lei *et al*., Development of advanced materials via entropy engineering. *Scr. Mater.* **165**, 164–169 (2019). doi: 10.1016/j.scriptamat.2019.02.015
93. S. Nie *et al*., Entropy-driven chemistry reveals highly stable denary $MgAl_2O_4$-type catalysts. *Chem Catal.* **1**, 648–662 (2021). doi: 10.1016/j.checat.2021.04.001
94. T. Li *et al*., Interface engineering between multi-elemental alloy nanoparticles and a carbon support toward stable catalysts. *Adv. Mater.* **34**, e2106436 (2022). doi: 10.1002/adma.202106436; pmid: 34875115
95. Y. J. Li, A. Savan, A. Kostka, H. S. Stein, A. Ludwig, Accelerated atomic-scale exploration of phase evolution in compositionally complex materials. *Mater. Horiz.* **5**, 86–92 (2018). doi: 10.1039/C7MH00486A
96. Y. J. Li, A. Kostka, A. Savan, A. Ludwig, Atomic-scale investigation of fast oxidation kinetics of nanocrystalline CrMnFeCoNi thin films. *J. Alloys Compd.* **766**, 1080–1085 (2018). doi: 10.1016/j.jallcom.2018.07.048
97. B. Song *et al*., In situ oxidation studies of high-entropy alloy nanoparticles. *ACS Nano* **14**, 15131–15143 (2020). doi: 10.1021/acsnano.0c05250; pmid: 33079522
98. P. Majumdar, J. Greeley, Generalized scaling relationships on transition metals: Influence of adsorbate-coadsorbate interactions. *Phys. Rev. Mater.* **2**, 045801 (2018). doi: 10.1103/PhysRevMaterials.2.045801
99. A. Jain, Y. Shin, K. A. Persson, Computational predictions of energy materials using density functional theory. *Nat. Rev. Mater.* **1**, 15004 (2016). doi: 10.1038/natrevmats.2015.4
100. J. K. Nørskov, T. Bligaard, J. Rossmeisl, C. H. Christensen, Towards the computational design of solid catalysts. *Nat. Chem.* **1**, 37–46 (2009). doi: 10.1038/nchem.121; pmid: 21378799
101. Z. Lu, Z. W. Chen, C. V. Singh, Neural network-assisted development of high-entropy alloy catalysts: Decoupling ligand and coordination effects. *Matter* **3**, 1318–1333 (2020). doi: 10.1016/j.matt.2020.07.029
102. M. Aykol, P. Herring, A. Anapolsky, Machine learning for continuous innovation in battery technologies. *Nat. Rev. Mater.* **5**, 725–727 (2020). doi: 10.1038/s41578-020-0216-y
103. P. M. Attia *et al*., Closed-loop optimization of fast-charging protocols for batteries with machine learning. *Nature* **578**, 397–402 (2020). doi: 10.1038/s41586-020-1994-5; pmid: 32076218
104. S. Curtarolo *et al*., AFLOWLIB.ORG: A distributed materials properties repository from high-throughput ab initio calculations. *Comput. Mater. Sci.* **58**, 227–235 (2012). doi: 10.1016/j.commatsci.2012.02.002
105. O. N. Senkov, J. D. Miller, D. B. Miracle, C. Woodward, Accelerated exploration of multi-principal element alloys

with solid solution phases. *Nat. Commun.* **6**, 6529 (2015). doi: 10.1038/ncomms7529; pmid: 25739749
106. S. Nellaiappan *et al*., High-entropy alloys as catalysts for the CO_2 and CO reduction reactions: Experimental realization. *ACS Catal.* **10**, 3658–3663 (2020). doi: 10.1021/acscatal.9b04302
107. S. Deshpande, T. Maxson, J. Greeley, Graph theory approach to determine configurations of multidentate and high coverage adsorbates for heterogeneous catalysis. *npj Comput. Mater.* **6**, 1–6 (2020).
108. E. J. Kluender *et al*., Catalyst discovery through megalibraries of nanomaterials. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 40–45 (2019). doi: 10.1073/pnas.1815358116 30559201
109. X. Liu *et al*., Inkjet printing assisted synthesis of multicomponent mesoporous metal oxides for ultrafast catalyst exploration. *Nano Lett.* **12**, 5733–5739 (2012). doi: 10.1021/nl302992q; pmid: 23051615
110. J. Li *et al*., Combinatorial screening of Pd-based quaternary electrocatalysts for oxygen reduction reaction in alkaline media. *J. Mater. Chem. A Mater. Energy Sustain.* **5**, 67–72 (2017). doi: 10.1039/C6TA08088J
111. L. Banko *et al*., Unravelling composition–activity–stability trends in high entropy alloy electrocatalysts by using a data-guided combinatorial synthesis strategy and computational modeling. *Adv. Energy Mater.* **12**, 2103312 (2022). doi: 10.1002/aenm.202103312
112. O. A. Krysiak *et al*., Searching novel complex solid solution electrocatalysts in unconventional element combinations. *Nano Res.* (2021). doi: 10.1007/s12274-021-3637-z
113. L. Banko, O. A. Krysiak, B. Xiao, T. Löffler, A. Savan, J. K. Pedersen, J. Rossmeisl, W. Schuhmann, A. Ludwig, Combinatorial materials discovery strategy for high entropy alloy electrocatalysts using deposition source permutations. arXiv:2106.08776 [cond-mat.mtrl-sci] (2021).
114. Z. Zhou *et al*., Machine learning guided appraisal and exploration of phase design for high entropy alloys. *npj Comput. Mater.* **5**, 128 (2019). doi: 10.1038/s41524-019-0265-1
115. H. S. Stein, D. Guevarra, P. F. Newhouse, E. Soedarmadji, J. M. Gregoire, Machine learning of optical properties of materials - predicting spectra from images and images from spectra. *Chem. Sci.* **10**, 47–55 (2018). doi: 10.1039/C8SC03077D; pmid: 30746072
116. H. Chun *et al*., First-principle-data-integrated machine-learning approach for high-throughput searching of ternary electrocatalyst toward oxygen reduction reaction. *Chem Catal.* **1**, 855–869 (2021). doi: 10.1016/j.checat.2021.06.001
117. T. Lookman, P. V. Balachandran, D. Xue, R. Yuan, Active learning in materials science with emphasis on adaptive sampling using uncertainties for targeted design. *npj Comput. Mater.* **5**, 21 (2019). doi: 10.1038/s41524-019-0153-8
118. R. Yuan *et al*., Accelerated discovery of large electrostrains in $BaTiO_3$-based piezoelectrics using active learning. *Adv. Mater.* **30**, 1702884 (2018). doi: 10.1002/adma.201702884; pmid: 29315814
119. D. Xue *et al*., Accelerated search for materials with targeted properties by adaptive design. *Nat. Commun.* **7**, 11241 (2016). doi: 10.1038/ncomms11241; pmid: 27079901
120. J. K. Pedersen *et al*., Bayesian optimization of high-entropy alloy compositions for electrocatalytic oxygen reduction. *Angew. Chem. Int. Ed.* **60**, 24144–24152 (2021). doi: 10.1002/anie.202108116; pmid: 34506069
121. J. Miao *et al*., "Materials Data Bank: A database for archiving 3D atomic coordinates and chemical species of materials" (2022); https://www.materialsdatabank.org.

ACKNOWLEDGMENTS

Funding: L.H. acknowledges support from the National Science Foundation (NSF CMMI-1635221) and the US Department of Energy (DOE), Advanced Research Projects Agency – Energy (ARPA-E). J.L. acknowledges support from the NSF (DMR-2026193 for compositionally complex fluorite–based oxides and DMR-2011967 for interfacial science). J.M. acknowledges support from the DOE, Office of Science, Basic Energy Sciences, Division of Materials Sciences and Engineering, under award no. DE-SC0010378 and by STROBE: A National Science Foundation Science and Technology Center under award no. DMR-1548924. C.W. acknowledges support from the DOE, ARPA-E, and the Petroleum Research Fund (PRF) of the American Chemical Society. G.W. acknowledges support from the US National Science Foundation (NSF DMR grant no. 1905572). J.G., I.K., C.W., and L.H. acknowledge the DOE Office of Science, Office of Basic Energy Sciences (BES), Chemical, Biological, and Geosciences Division, Data Science Initiative grant no. DE-SC0020381. Use of computational resources from the National Energy Research Scientific Computing Center is also acknowledged. M.C. acknowledges support from the DOE, Office of Basic Energy Sciences, under early career award no. ERKCZ55 and the Center for Nanophase Materials Sciences, which is a DOE Office of Science User Facility. **Author contributions:** Y.Y., Q.D., A.B., and L.H. wrote the paper with input from J.L., J.M., M.C., C.W., I.G.K., Z.J.R., J.G., G.W., and A.A. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** References to all data are provided in the manuscript.

10.1126/science.abn3103

RESEARCH ARTICLE SUMMARY

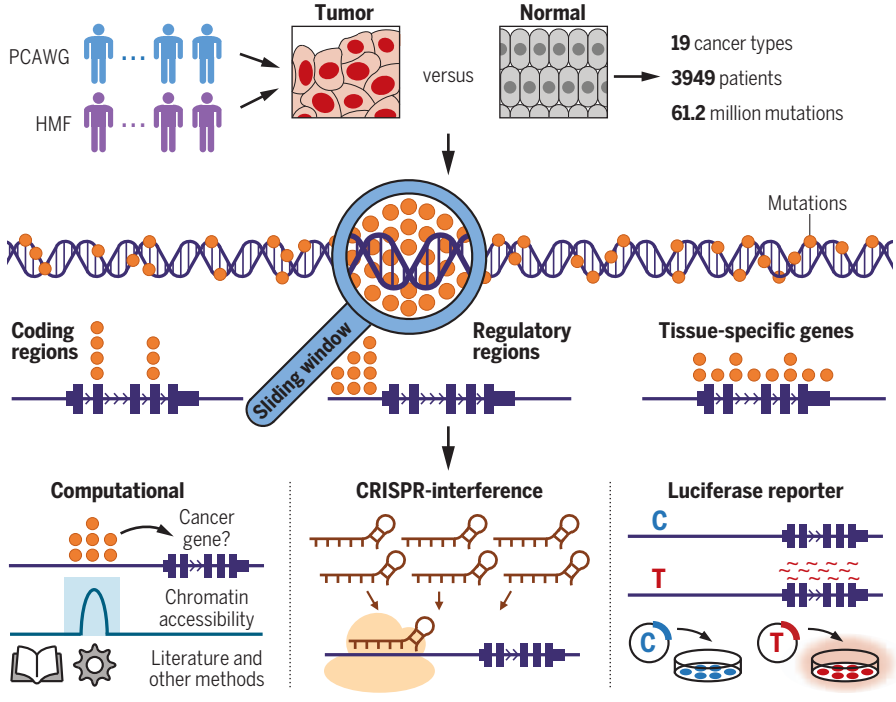
CANCER GENOMICS

Genome-wide analysis of somatic noncoding mutation patterns in cancer

Felix Dietlein*, Alex B. Wang†, Christian Fagre†, Anran Tang†, Nicolle J. M. Besselink, Edwin Cuppen, Chunliang Li, Shamil R. Sunyaev, James T. Neal‡, Eliezer M. Van Allen*‡

INTRODUCTION: A central hallmark of tumor development is that cancer cells acquire somatic mutations in their genomes that are not present in normal tissue. Some mutations are drivers and contribute to the growth of tumor cells, but many others are passengers without apparent effects on tumor biology. Over the past decade, driver mutations have been comprehensively characterized in protein-coding genomic regions by analyzing sequencing data from thousands of tumor-normal pairs. This characterization in protein-coding regions has yielded a wealth of insights into tumor biology, including many genome-inspired drug targets. However, the role of somatic mutations in the other 98% of the cancer genome—the noncoding genome—remains incompletely understood.

RATIONALE: Many statistical approaches detect drivers as recurrent mutation events by comparing the number of mutations with and without effects on protein-coding sequences in each gene. These approaches are therefore inapplicable outside of protein-coding regions, where the roles of somatic mutations remain less well understood. The noncoding genome encompasses a diverse spectrum of elements, including regulatory regions of gene expression that differ in their locations and activities between tumor types. To expand our understanding of mutations beyond protein-coding regions, we designed and implemented a genome-wide, sliding-window approach that detects mutation events irrespective of their locations in regulatory elements or effects on protein-coding sequences.



Genome-wide compendium of somatic mutation patterns in human cancer. We analyzed 61.2 million mutations from 3949 patients of 19 cancer types (top). Using a sliding-window approach, we detected mutation events across the entire cancer genome and classified them by their genomic locations (middle). For systematic follow-up, we used both computational and experimental strategies (bottom). PCAWG, Pan-Cancer Analysis of Whole Genomes; HMF, Hartwig Medical Foundation.

RESULTS: We developed a composite of three methods to detect recurrent mutation events across the whole genomes of 3949 patients with 19 cancer types and 61.2 million somatic mutations. This approach automatically stratified mutation events into different categories on the basis of their position in the genome. In protein-coding regions, we identified an average of 7.5 events per cancer type and recovered well-established driver mutations. In the noncoding genome, 3.7 events per cancer type occurred adjacent to genes exclusively expressed in specific tissue types (*ALB* in liver, *KLK3* in prostate, *SFTPB* in lung, *SLC5A12* in kidney, *TG* in thyroid tissue, and many others). These tissue-specific events were unlikely to be prototypical drivers because they stemmed from a mutagenic process that was exclusively active around these genes, instead reflecting possible imprints of the expression programs of the tumor cells of origin. Moreover, we found 3.8 noncoding events per cancer type in regulatory regions of expression, many involving cancer-relevant genes (*BCL6*, *FGFR2*, *RAD51B*, *SMC6*, *TERT*, *XBP1*, and many others). In contrast to most events in regulatory regions, breast cancer mutations near *XBP1* mainly accumulated in a regulatory region outside of its promoter. We validated their regulatory effects on gene expression by performing CRISPR-interference screening and luciferase reporter assays, illuminating the potential of genome-wide approaches paired with harmonized sequencing cohorts to comprehensively capture mutation patterns in both known and unknown elements of the noncoding genome.

CONCLUSION: Our study establishes a genome-wide compendium of the diverse mutation patterns that shape the genomes of 19 major cancer types, including events near genes with known roles in tumor biology and some exhibiting experimentally validated effects on gene expression. Our results demonstrate that noncoding mutations are associated with a broad spectrum of different biological processes and that their location in the genome is essential for their accurate interpretation. Broadly, our study provides a blueprint for interpreting whole-genome sequencing data and lays the foundation for future experimental endeavors to implicate noncoding mutations in tumor development, ultimately paving the way for therapies tailored to the noncoding cancer genome. ■

*Corresponding author. Email: EliezerM_VanAllen@dfci.harvard.edu (E.M.V.A.); Felix_Dietlein@dfci.harvard.edu (F.D.)
†These authors contributed equally to this work.
‡Co-senior authors.
Cite this article as F. Dietlein *et al.*, *Science* **376**, eabg5601 (2022). DOI: 10.1126/science.abg5601

READ THE FULL ARTICLE AT
https://doi.org/10.1126/science.abg5601

RESEARCH ARTICLE

CANCER GENOMICS

Genome-wide analysis of somatic noncoding mutation patterns in cancer

Felix Dietlein[1,2]*, Alex B. Wang[2]†, Christian Fagre[2]†, Anran Tang[1,2]†, Nicolle J. M. Besselink[3], Edwin Cuppen[3,4], Chunliang Li[5], Shamil R. Sunyaev[6,7], James T. Neal[2]‡, Eliezer M. Van Allen[1,2]*‡

We established a genome-wide compendium of somatic mutation events in 3949 whole cancer genomes representing 19 tumor types. Protein-coding events captured well-established drivers. Noncoding events near tissue-specific genes, such as *ALB* in the liver or *KLK3* in the prostate, characterized localized passenger mutation patterns and may reflect tumor-cell-of-origin imprinting. Noncoding events in regulatory promoter and enhancer regions frequently involved cancer-relevant genes such as *BCL6*, *FGFR2*, *RAD51B*, *SMC6*, *TERT*, and *XBP1* and represent possible drivers. Unlike most noncoding regulatory events, *XBP1* mutations primarily accumulated outside the gene's promoter, and we validated their effect on gene expression using CRISPR-interference screening and luciferase reporter assays. Broadly, our study provides a blueprint for capturing mutation events across the entire genome to guide advances in biological discovery, therapies, and diagnostics.

Tumors carry different types of somatic mutations in their genomes. Most of these mutations are random "passengers" that are propagated through clonal evolution without contributing to tumor development (*1*). However, a few are "drivers" that contribute to the uncontrolled growth and proliferation of cancer cells (*1*) and therefore represent targets for many therapies in precision medicine.

Over the past decade, the characterization of somatic drivers has focused primarily on protein-coding regions (*2*), where such mutations change the amino acid sequences of oncogenes and tumor suppressor genes. Statistical algorithms have been established to detect drivers as recurrent "mutation events" in large sequencing cohorts of tumor patients (*3*–*5*). Applying these algorithms to the sequencing data of thousands of tumor-normal pairs has helped considerably to elucidate which mutations contribute to tumor development in coding regions (*2*), whereas the role of noncoding somatic mutations in the remaining ~98% of the genome remains less well understood (*6*).

In the noncoding genome, the detection and interpretation of mutation events are complex. Many algorithms have been established that detect mutation events based on nonsynonymous and synonymous amino acid changes in coding regions (*3*, *4*), rendering them inapplicable to noncoding regions in whole-genome sequencing (WGS) data. Furthermore, the noncoding genome comprises a diverse spectrum of genomic elements, ranging from active regulatory elements of gene expression to inactive heterochromatic regions (*7*, *8*). Therefore, mutation events in different parts of the noncoding genome mirror separate biological processes, as revealed by recent studies such as the Pan-Cancer Analysis of Whole Genomes (PCAWG) (*9*). Although several mutation events represent possible noncoding drivers, such as those identified in the promoters and enhancers of cancer-relevant genes, others are less likely to be drivers, such as those resulting from mutagenic processes around tissue-specific genes (*9*, *10*).

To address these specific challenges in noncoding regions, we implemented a genome-wide approach that identifies somatic mutation events in point mutations and in short insertions and deletions across the entire cancer genome irrespective of their positions in the genome or their effects on protein-coding sequences. This approach automatically stratifies mutation events based on their genomic locations, thus capturing their different propensities to represent possible drivers or localized passenger mutation patterns. By applying this strategy to a harmonized cohort of 3949 somatic whole cancer genomes and combining it with systematic computational and experimental follow-up, our study establishes a genome-wide compendium of mutation events in 19 major cancer types.

Results

Genome-wide detection of somatic mutation events in whole cancer genomes

For genome-wide detection and classification of somatic mutation events, we proceeded in three steps (Fig. 1, A to C, and fig. S1). First, we tiled the genome with three interval sizes (1, 10, and 100 kb; see illustration in fig. S2) and performed three significance tests in each interval: test 1 to determine whether a genomic region contained more mutations than expected based on its epigenomic signal; test 2 to compare mutation counts between different cancer types in each region; and test 3 to determine whether more mutations clustered together than expected. Second, we integrated *P* values from these three tests and different interval lengths into a continuous genome-wide signal of significance based on Brown's method (*11*), and then adjusted this signal by weighted multiple hypothesis correction based on cancer-specific expression data (*12*). Third, we identified all statistically significant events in this genome-wide signal [false discovery rate (FDR) < 0.1] and automatically classified them based on their genomic locations into protein-coding regions (mutations in exons of oncogenes and tumor suppressor genes), regulatory regions [promoters and enhancers overlapping with signals of H3K4me3 and H3K27ac histone chromatin immunoprecipitation sequencing (ChIP-seq) (*7*)], or mutagenic processes around tissue-specific genes (genes exclusively expressed in a specific cancer type). In this way, we captured their different propensities to be possible drivers or passengers building on insights gained from prior studies (*9*). We excluded events with mutational hotspots in secondary DNA hairpin structures or low genomic mappability; events not meeting any of these criteria were labeled as "other" (Fig. 1, A to C).

Q-Q plots demonstrated that the three significance tests and their combined *P* values were accurately calibrated to their background signals and exhibited no inflation of low *P* values (Fig. 1, D and E, and figs. S3 and S4). Histograms revealed that the background models of the three tests matched the observed distributions of mutation rates and positional clustering in the upper distribution tails (fig. S3). These results further suggested that the three tests did not rely on cancer-type-specific assumptions, and that our genome-wide analysis was applicable across a wide range of different cancer types. The materials and methods and supplementary text include a comprehensive explanation of the rationale behind our statistical framework in the context of prior approaches, additional analyses of the performance and accuracy of the three significance tests (figs. S5 to S11), the necessity of combining different tests (fig. S12) and interval sizes (fig. S2) to capture a broad spectrum

[1]Department of Medical Oncology, Dana-Farber Cancer Institute, Harvard Medical School, Boston, MA 02215, USA. [2]Cancer Program, Broad Institute of Massachusetts Institute of Technology and Harvard, Cambridge, MA 02142, USA. [3]Center for Molecular Medicine and Oncode Institute, University Medical Center Utrecht, 3584 CX Utrecht, Netherlands. [4]Hartwig Medical Foundation, 1098 XH Amsterdam, Netherlands. [5]Department of Tumor Cell Biology, St. Jude Children's Research Hospital, Memphis, TN 38105, USA. [6]Division of Genetics, Brigham and Women's Hospital, Harvard Medical School, Boston, MA 02115, USA. [7]Department of Biomedical Informatics, Harvard Medical School, Boston, MA 02115, USA.
***Corresponding author. Email: EliezerM_VanAllen@dfci.harvard.edu (E.M.V.A.); Felix_Dietlein@dfci.harvard.edu (F.D.)**
†These authors contributed equally to this work.
‡Co-senior authors.

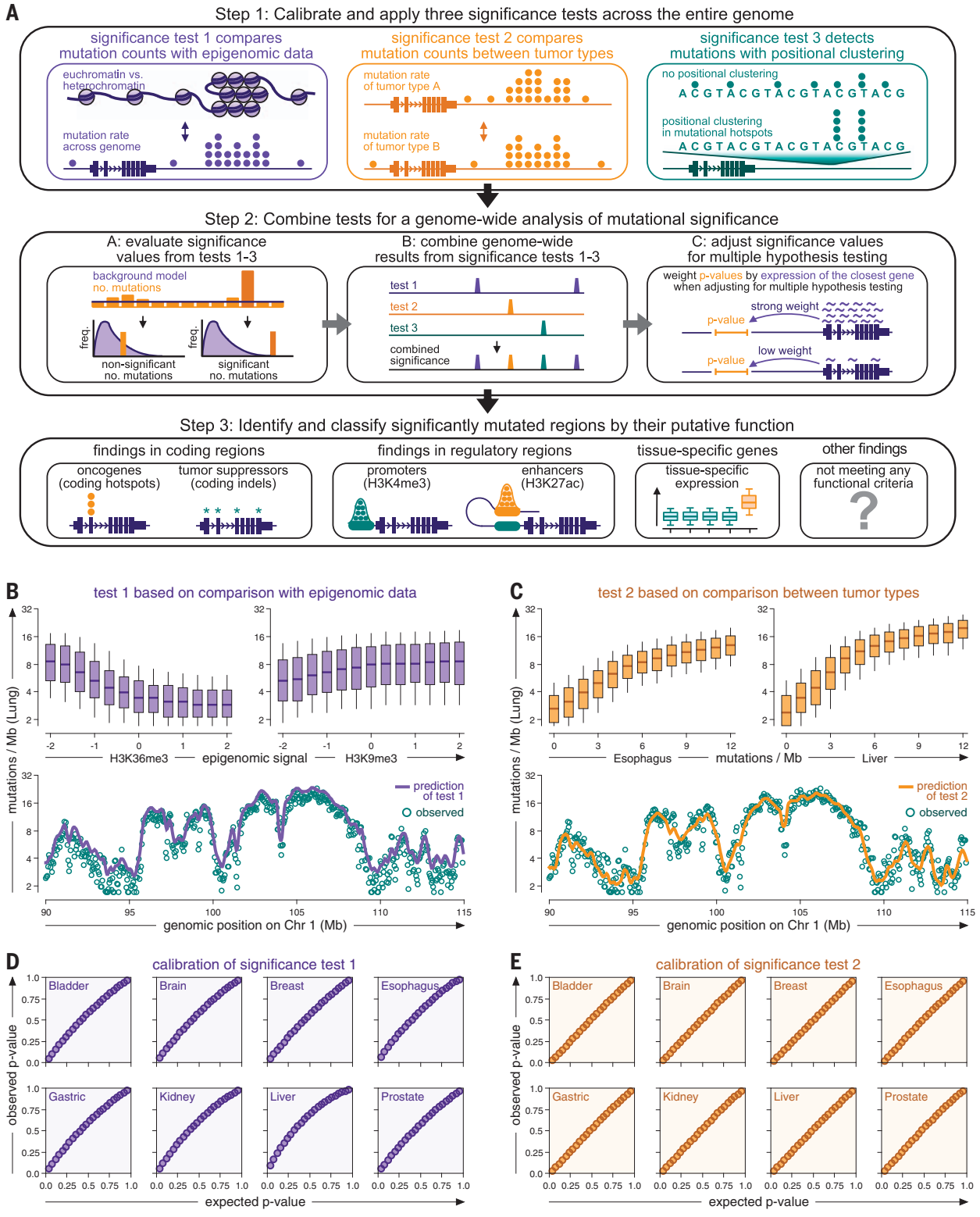


Fig. 1. Genome-wide analysis of somatic mutation events in whole cancer genomes. (**A**) Genome-wide detection of somatic mutation events in whole cancer genome sequencing data. Step 1 combines three complementary test strategies. Step 2 integrates the results of tests 1 to 3 into a joint, genome-wide signal and identifies significant mutation events. Step 3 classifies mutation events according to their genomic location. (**B** and **C**) Top: Boxplots comparing mutation rates of a representative cancer type (lung cancer) against epigenomic signals [(B), the rationale of test 1] and mutation rates of other cancer types [(C), the rationale of test 2]. Boxes indicate 25/75% interquartile ranges, vertical lines extend to 10/90% percentiles, and horizontal lines reflect distribution medians. Bottom: Observed (teal dots) and predicted (continuous line) mutation rates (10-kb intervals) plotted against their position on chromosome 1 (function smoothed by Gaussian kernel). (**D** and **E**) Q-Q plots comparing observed (*y*-axis) and expected (*x*-axis) *P* values for test 1 (D) and test 2 (E).

of mutation events, and a comparison with alternative implementations (figs. S13 to S17).

A genome-wide compendium of mutation events in 19 cancer types

For a harmonized analysis of 6.12×10^7 somatic mutations in 3949 whole genomes from 19 cancer types, we assembled high-confidence samples, regions, mutations, and cancer types from two sequencing consortia, PCAWG (*9*) and the Hartwig Medical Foundation [HMF (*13*)]. A detailed description of our filtering criteria and the cancer types included in this study is provided in the materials and methods and figs. S18 to S21. In 19 cancer types, our genome-wide approach detected 142 events in coding regions (average 7.5 per cancer type; 45 in oncogenes and 97 in tumor suppressors), 73 events in regulatory regions (average 3.8 per cancer type; 49 in promoters and 24 in enhancers), 70 events around tissue-specific genes (average 3.7 per cancer type; 70 genes exclusively expressed in a specific cancer type, such as albumin in the liver), and 87 "other" events (average 4.6 per cancer type; the exact role of these findings was less clear) (Fig. 2, A and B; figs. S22 to S24; and tables S1 to 20). To refer to the genomic location of our findings, we annotated them by their closest genes (table S1). For confirmation, we used the activity-by-contact model (*14*) based on three-dimensional genomic distance, which returned the same genes for 91% of coding, regulatory, and tissue-specific findings (fig. S12, G to I).

Events in protein-coding regions

Findings in protein-coding regions largely captured well-established driver mutations, with 93.0% (132/142) involving canonical cancer genes (Fig. 2C) and 96.5% (137/142) matching the results obtained by two established methods for identifying coding drivers [MutSigCV (*3*) and dNdScv (*4*)] (fig. S25, A and B). This low rate of false positives in coding regions supports the robustness of our approach in the entire genome because it uses the same statistics in both coding and noncoding regions. Furthermore, significance values returned by our genome-wide approach in protein-coding regions correlated with the ratio of nonsynonymous to synonymous mutations (fig. S25C), an established marker of positive selection (*4*). We obtained a similar result in the rest of the genome by predicting the pathogenicity of noncoding mutations based on two bioinformatics scores (*15*, *16*) (fig. S25, D to F).

Events in regulatory regions

Events in regulatory regions were significantly enriched for canonical cancer genes ($P < 0.001$, Fisher's exact test), with 37.0% (27/73) of the findings linked to genes in the Cancer Gene Census (*17*) or the Oncology Knowledge Base (*18*), compared with the 4.1% (the percentage of cancer genes among all genes) that would be expected to occur by chance (Fig. 2C). Because of the link between these regions and gene expression, some findings in this category have been discussed as plausible noncoding drivers in the literature (*6*, *9*, *10*, *19*). This includes mutation events in the *TERT* promoter (telomere regulation), which we identified in bladder, brain, head and neck, kidney, liver, and thyroid cancer, and mutations at *MIR21* (cancer-promoting microRNA gene), which we detected in breast, esophagus, gastric, and lung cancer. Furthermore, consistent with these prior studies (*6*, *9*, *10*, *19*), we found noncoding mutations upstream of *FOXA1* in breast cancer and downstream of *FOXA1* in prostate cancer, in addition to many coding mutations in the same gene.

Our study expanded this category by 46 additional findings in promoters and enhancers of genes potentially relevant to cancer (Figs. 2, A and B, and 3A and figs. S22 and S23). For example, we identified recurrent events in the promoters of leukemia-related genes, including *BACH2*, *BTG2*, *CXCR4*, *BCL6*, *BCL7A*, and *IRF8*. Other mutations accumulated in promoters of the cancer-associated genes *FGFR2* in bladder and lung cancer; *B2M*, *KLF6*, and *SRCAP* (chromatin remodeling complex) in lung cancer; and *MDM4*, *PIK3C2B*, *CDCA4* (cell cycle gene), and *BTG3* (antiproliferation factor) in bladder cancer. We found additional events in the promoters of *MED16* (coactivator of RNA polymerase II transcription) in liver cancer, as well as *STAG1* (cohesion of sister chromatids during the S-phase), *SMC6* (maintenance of telomere length), and *GEN1* (double-strand break repair) in breast cancer.

Other additional findings were in enhancers, including *RAD51B* (canonical cancer gene involved in double-strand break repair) in bladder and breast cancer, *ETS2* (transcription factor related to proliferation, apoptosis, and telomere maintenance) in colorectal cancer, *ST6GAL1* (glycosyltransferase inducing an invasive phenotype) in leukemia, and *XBP1* (established function as an estrogen-induced transcription factor) in breast cancer. Some mutations in this category recurred as hotspots in the same genomic position, including *BTG3*, *FGFR2*, *MED16*, *PIK3C2B*, *SMC6*, *STAG1*, and *TERT* (fig. S26A and table S21), although the occurrence of this mutation pattern was rare in noncoding regulatory regions compared with its high frequency in coding regions.

Events near tissue-specific genes

In contrast to protein-coding and regulatory regions, findings around tissue-specific genes are unlikely to represent candidate driver events themselves because of their reported link to localized mutagenic processes (*9*, *10*) and lack of enrichment for known cancer genes (Fig. 2C). However, according to the MalaCards database (*20*), 42.9% (30/70) of tissue-specific genes linked to mutation events exhibited physiological roles in their associated normal tissues, compared with the 3.9% (the percentage of genes included in the MalaCards database) that would be expected to occur by chance (fig. S26, B and C). Therefore, mutation events in this category were significantly enriched around genes with reported physiological roles independent of cancer signaling ($P < 0.001$, Fisher's exact test), concordant with their unique expression in a specific tissue type. Some of our findings near tissue-specific genes have been observed in previous studies, either as primary results (*10*) or as incidental findings annotated as nondrivers (*9*). These included *LIPF* in gastroesophageal cancer, *ALDOB* in kidney and liver cancer, *SFTPB* and *SFTPC* in lung cancer, *CPB1* and *PNLIP* in pancreatic cancer, *TG* in thyroid cancer, and 12 tissue-specific genes in liver cancer (including *ALB*, *CYP3A5*, *FGA*, and *MIR122*).

Our study expanded this category by 54 additional findings (Figs. 2, A and B, and 3B and figs. S22 and S23), including *TMEFF2* (survival factor for neurons) and *HCN1* (hyperpolarization-activated cation channel in neurons) in brain tumors, as well as *STC2* (glycoprotein induced by estrogen), *TRPS1* (repressor of GATA-regulated genes), *ANKRD30A* (serologically defined breast cancer antigen), and *MGP* (estrogen-regulated matrix protein involved in cellular differentiation) in breast cancer. Other additional events in this category included *KLK3* (prostate-specific antigen, a serum marker for prostate cancer), *PLPP1* (androgen-regulated phosphatase expressed on the cell surface), and *TMPRSS2* (androgen-regulated serine protease) in prostate cancer, and *GCG* (glucagon, a pancreatic hormone) in neuroendocrine tumors. Furthermore, we identified tissue-specific events around *SLC5A12* (lactate reabsorption in proximal tubules), *KCNJ15* (potassium channel in the kidney), *GLYAT* (glycine-acyltransferase), and *PCK1* (gluconeogenesis) in kidney cancer, as well as *MUC6* (mucin; protects epithelium from gastric acid) and *AGR2* (expressed in mucus-secreting tissues and overexpressed in Barrett's esophagus) in gastroesophageal tumors. Moreover, liver cancer exhibited the largest number of additional mutation events in the tissue-specific category, including 18 genes encoding liver-specific proteins (including *C3*, *CRP*, and *TF*) and 17 genes associated with liver metabolism and detoxification (including *AKR1C1*, *BAAT*, *CYP2E1*, *G6PC*, and *HEXB*).

Other events

For some events, the status remained less clear. For example, in agreement with the prior literature, we identified events at the neighboring genes *NEAT1* and *NEAT2* in breast, bladder, esophagus, kidney, and liver cancer. Our genome-wide approach placed them in the regulatory category (fig. S27), whereas PCAWG interpreted

Fig. 2. Mutation events identified in a genome-wide analysis of the PCAWG and HMF consortia. (**A** and **B**) Top: Pie charts showing the number of mutation events per category (purple: coding, orange: regulatory, teal: tissue-specific, gray: other) in aggregate (A) and individual cancer types (B). Bottom: Genomic positions (*y*-axis) plotted against their significance in a genome-wide analysis (*x*-axis) and colored by categories (B). The position (*y*-axis) of findings recurring in more than one cancer type is plotted against the number of cancer types (*x*-axis) (A). *NEAT1* and *MALAT1* are marked by asterisks because their classification was ambiguous. (**C**) Mutation events sorted by their significance in a genome-wide analysis (*x*-axis, orange) and plotted against the number of findings involving known cancer genes (*y*-axis, top). Random overlap between findings and cancer genes serves as a negative control (purple).

them as being the result of a transcription-related mutational process (*9*), and other studies arrived at different conclusions regarding their relevance in tumor signaling (*19*, *21*).

Furthermore, some noncoding events did not fall into the protein-coding, regulatory, or tissue-specific categories. This "other" category exhibited mild enrichment for canonical cancer genes (Fig. 2C) and included *MAD1L1* and *MAD2L1* (mitotic spindle assembly checkpoint) in brain and ovarian tumors; *NF1* (tumor suppressor) in breast tumors; *DCC* (known

cancer gene) in esophageal cancer; *KCNJ15* (potassium channel) in kidney cancer; *TCL1A*, *BCR*, and *NFKBIE* (known cancer genes) in leukemia; as well as *ABHD5* (lipid binding), *LIPG* (lipase), *FN1* (fibronectin), *HNF4A* (hepatocyte nuclear factor), *MAP2K6* (mitogen-activated kinase), and *ERRFI1* (ERBB receptor feedback inhibitor) in liver cancer. In addition, *APC* and *SMAD4* in colorectal cancer harbored noncoding splice site mutations outside of canonical exon-intron boundaries (fig. S22D).

Altogether, our study establishes a genome-wide compendium of somatic mutation events for 19 cancer types, categorized by their genomic locations and different biology, including many findings from recent studies and several additional results (see table S1 for literature references). A complete list of our findings in each cancer type is provided in tables S2 to S20, annotated by their genomic locations, mutation frequencies, status as known cancer genes, and significance values returned by our genome-wide approach.

Systematic follow-up on mutation events identified in our genome-wide analysis

We performed three systematic follow-up analyses to examine the ability of our approach to detect mutation events in the noncoding genome and evaluate the plausibility of our results.

Inspection of the genomic territory around mutation events

Although our genome-wide approach examined the entire genome, 76.6% (285/372) of the mutation events occurred in coding, regulatory, or tissue-specific regions (Fig. 2, A and B, and figs. S22 and S23), which account for 10.2% of the genome. Furthermore, they accumulated in regulatory and transcribed regions based on ChIP-seq data from normal tissue (*7*) (fig. S28A), and this enrichment was even more pronounced in chromatin accessibility data [assay for transposase-accessible chromatin using sequencing (ATAC-seq)] from the same type of tumor tissue, when available (*8*) (fig. S28B). Moreover, mutation events exhibited strong enrichment around the following four markers (figs. S29 and S30 and tables S2 to S20): (i) ATAC-seq peaks that existed in tumor but not in normal tissue (fig. S29, A and B), (ii) ATAC-seq peaks that correlated with the expression of their closest gene (fig. S29, C and D), (iii) methylation markers that correlated negatively with the expression of their associated genes (fig. S29, E and F), and (iv) genome-wide association study (GWAS) peaks from germline data (fig. S29, G and H).

The accumulation of events around these four markers prompted us to investigate whether the performance of our genome-wide analysis could be improved by restricting it to regions around these four markers. However, this restricted version missed a substantial number of findings (fig. S30H), including many events associated with known cancer genes. Furthermore, the applicability of the four markers varied between cancer types, depending on the availability of ATAC-seq data (*8*). Similar results were obtained when restricting our analysis to five databases of established promoter and enhancer regions (*22*–*26*) (fig. S30, C and D), illuminating the potential of a genome-wide approach.

Compatibility with prior findings and methods

Previous studies, including PCAWG, reported 30.1% (43/143) of the noncoding mutation events in the tissue-specific and regulatory categories observed herein (*6*, *9*, *10*, *19*), compared with the 1.47% (the percentage of genes for which noncoding findings had been reported previously) that would be expected by chance ($P < 0.001$, Fisher's exact test). Conversely, our genome-wide analysis identified 39 of the noncoding findings from prior work (39/65 previous findings; 30/39 previous findings with an FDR < 10^{-4}) (tables S22 and S23). Tissue-specific events in this comparison were interpreted differently in prior studies that either reported them as primary results (*10*) or incidental, nondriver findings (*9*). Furthermore, our WGS dataset overlapped with that of previous studies, so that shared findings affirm the general compatibility of our genome-wide approach in regions evaluated by both our study and prior work.

For further comparison, we ran four existing and available methods [DriverPower (*27*), Larva (*28*), MutSpot (*29*), and OncodriveFML (*5*)] on the entire WGS dataset. This revealed that our genome-wide approach identified nearly all the noncoding events detected by these four methods in the genomic territory included in our analysis (figs. S31 and S32). This comparison further highlighted the importance of excluding low-quality mutations and low-coverage regions from our genome-wide analysis for technical considerations (figs. S18 and S32), given that not all parts of the genome are amenable to WGS.

Analysis of the statistical power of our genome-wide approach to detect mutation events

This analysis demonstrated that the power of our approach varied substantially between cancer types, depending on their background mutation rates, the available number of samples, and the size of the genomic territory included in the analysis (fig. S33). Additional technical factors beyond those captured in this model may interfere with the statistical power (*9*). Although combining the HMF and PCAWG consortia increased the statistical power of our study considerably, the amount of whole-genome data was still smaller than the amount of whole-exome data generated over more than a decade and used to characterize mutations in coding regions (*2*). Therefore, there may be noncoding events in addition to those identifiable in the available data (fig. S33), as was concordantly concluded in a power analysis by the PCAWG study (*9*).

Characterization of mutation and expression patterns of tissue-specific genes

We next studied the pattern of mutation events near or within tissue-specific genes in more detail (fig. S34). We first focused on liver cancer, which contained the largest number of events in this category. Consistent with previous studies connecting this category of mutations with localized mutagenic processes (*9*, *10*), noncoding regions around tissue-specific genes were enriched for insertions and deletions ("indels") (Fig. 4A). These indels were longer than those in the rest of the genome (83.2 versus 22.4% of deletions had target lengths >1 bp; 30.1 versus 15.5% for insertions) (Fig. 4, B and C, and fig. S34A). In addition, we observed that indels around tissue-specific genes accumulated in A/T-rich nucleotide contexts and resembled Catalogue of Somatic Mutations in Cancer (COSMIC) indel signatures ID4 and ID8 (*30*), a pattern that rarely occurred in the rest of the genome (fig. S34, B to H). Comparison of mutations around tissue-specific versus highly expressed genes yielded the same differences (fig. S34, I and J), suggesting that mutation events in this category only occurred around genes exhibiting unique expression in a particular tissue type and not around highly expressed genes in general. Concordantly, expression and mutation rates exhibited positive correlation in noncoding regions around tissue-specific genes, the opposite of their relationship in the rest of the genome (fig. S35, A and B). In addition to mutations, other recurrent events accumulated in proximity to tissue-specific genes, including hypermethylation (fig. S35, C and D) and copy number loss (fig. S35, E to H). We obtained similar results in cancer types other than liver (fig. S34K).

However, mutation events did not occur ubiquitously around all tissue-specific genes, with most cancer types harboring >100 tissue-specific genes but five or fewer tissue-specific events (fig. S36A). Furthermore, the number of events in this category differed greatly between cancer types (Fig. 2B and fig. S22, A and B), and the fraction of indels and their lengths varied considerably between individual tissue-specific genes (fig. S36, B and C). These observations suggest that some but not all tissue-specific genes harbor a mutation pattern in their surrounding noncoding territory that deviates from the rest of the genome. These differences manifested as mutation events detected by our genome-wide approach and characterized the specific genomic regions

Fig. 3. Categories of mutation events exhibit different mutation patterns. Positional clustering of mutations (*y*-axis, percentage of maximum) plotted against genomic positions (*x*-axis) around mutation events that fall into regulatory regions [(**A**), orange] or overlap with tissue-specific genes [(**B**), teal]. Genomic boundaries of the closest gene are marked at the bottom of each plot, and white arrowheads mark the direction of its transcription.

and genes where this localized mutation pattern occurred.

Finally, we explored whether characterizing mutation events around tissue-specific genes could offer insights into tumor biology. We hypothesized that these events might be connected to the cell of origin from which a tumor developed, given that these genes exhibited (i) tissue-specific expression (Fig. 4D), (ii) lower expression in tumor cells than in normal cells (Fig. 4, E and F, and fig. S36, D to F), and (iii) physiological roles in their respective tissues (fig. S26, B and C). Consistent with this hypothesis, many tissue-specific genes were heterogeneously expressed in single-cell data from normal tissues (fig. S37, A and B), particularly those harboring mutation events (fig. S37, C and D). For instance, in single-cell expression data for liver (*31*), most tissue-specific genes with mutation events were differentially expressed (87.5%; 35/40) between cells from different histological zones (Fig. 4, G to I, and fig. S38, A to D) compared with 15.5% for arbitrary genes expressed in the liver ($P < 0.001$, Fisher's exact test). Similarly, in single-cell expression data for kidney (*32*), all tissue-specific genes with mutation events were expressed in a specific cell type (proximal tubule cells, 100%; 5/5) (fig. S38, E and F) compared with 26.4% for arbitrary, heterogeneously expressed genes ($P = 0.001$, Fisher's exact test). Likewise, papillary and clear-cell kidney tumors, which originate from proximal tubule cells, carried mutations around tissue-specific genes more frequently than chromophobe kidney tumors that originate from collecting-duct epithelial cells (*33*) (60.9 versus 14.0%; $P < 0.001$, Fisher's exact test) (fig. S38G).

Our analyses thus established a general, reciprocal link among a localized mutation pattern in tumor genomes, tissue-specific expression in bulk expression data, and heterogeneous expression in single-cell data of the related normal tissue. Therefore, the localized mutation pattern around tissue-specific genes may reflect a potential imprint of the characteristic expression program of the cell type from which a tumor originated (Fig. 4, G to I, and figs. S37 and S38), which could be of use in diagnostics.

Evaluation of mutation events in promoter and enhancer regions

We next used the following analyses to further assess the noncoding mutation events in regulatory promoter and enhancer regions.

Transcription factor binding sites

We used a permutation test to identify recurrent mutations that changed transcription factor binding motifs in the JASPAR database (*34*) (see the materials and methods). This test revealed that mutations changed binding motifs in 15.1% (11/73) of our findings in the

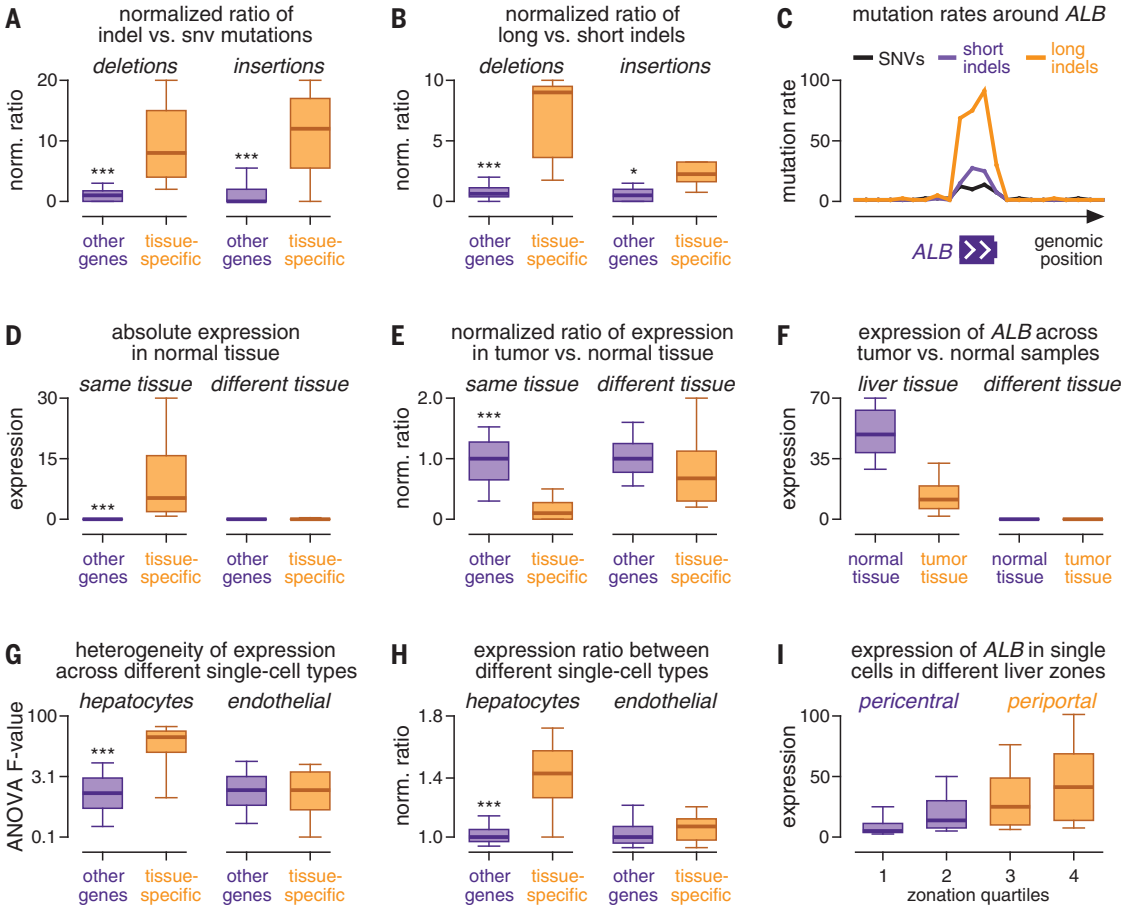


Fig. 4. Characterization of the expression and mutation patterns of tissue-specific genes. (**A** and **B**) Box plots comparing the ratio of the number of indels to single-nucleotide variants (SNVs) (A) and the ratio of the number of long to short indels (B) between tissue-specific genes (orange) and other genes (purple). (**C**) Mutation rates of SNVs (black), short indels (purple), and long indels (orange) (*y*-axis, percentage of maximum) plotted against their genomic position around *ALB* (*x*-axis). (**D** and **E**) Box plots comparing the expression (D) and expression ratio in tumor versus normal tissue (E) of tissue-specific genes (orange) and other genes (purple). (**F**) Box plots comparing *ALB* expression (*y*-axis) between samples from tumor tissue (orange) and normal tissue (purple). (**G** and **H**) Box plots comparing heterogeneous expression of tissue-specific genes (orange) and other genes (purple) in single-cell data of hepatocytes (left) and endothelial cells (right) based on an analysis of variance (ANOVA) (G) and the expression ratio between cell types (H). (**I**) Box plots comparing *ALB* expression in cells from different histological zones of the liver (*x*-axis). Boxes in (A) to (I) indicate the 25/75% interquartile range, vertical lines extend to 10/90% percentiles, and horizontal lines reflect distribution medians. Significant differences (Mann-Whitney *U* test) are marked with asterisks: $*P < 0.05$, $**P < 0.01$, $***P < 0.001$.

regulatory category (fig. S39A), mainly in two binding motifs (81.8%; 9/11): Mutations in the *ELK4* motif produced two binding sites in the *TERT* promoter in many cancer types (*35*) (fig. S39A), whereas mutations in the *EGR1* motif (*36*) removed transcription factor binding sites from the promoters of antiproliferative genes such as *BTG3* or *STAG1* (fig. S39, A and B). We found an additional hotspot in the *FOXA1* promoter that produced a binding site for *E2F1* (*19*) (fig. S39A). In addition to these single-gene analyses, we analyzed mutations across regulatory regions in aggregate and detected additional changes to transcription factor binding sites in regulatory regions (fig. S40).

Differential expression

Differential expression analysis required matched mutation and expression data from the same tumor samples, and the limited availability of such data restricted our search to 12 cancer types (fig. S41, A to C, and materials and methods). In addition, we identified potential confounders of differential expression (fig. S41, D to G), including copy number, methylation, and the positive correlation between expression and mutation rates around tissue-specific genes, which was opposite to their negative correlation in the rest of the genome (fig. S35, A and B). Keeping these intrinsic limitations in mind, the genes linked to 49 mutation events (23 coding, seven regulatory, 19 tissue specific) were associated with differential expression between mutated and nonmutated samples after multiple hypothesis correction (fig. S42). For seven of 12 cancer types, the number of differentially expressed genes was higher than would be expected by chance (fig. S43, A to D). In addition to evaluating differential expression for each mutation event separately, we performed two aggregate analyses and detected additional potential associations between noncoding mutations and differential expression (fig. S43, E and F).

Physical interactions

Noncoding mutation events involved many genes that exhibited direct physical interactions with established driver genes identified from analyses of coding regions, suggesting that they targeted the same pathway (*37*) (fig. S44A and materials and methods).

Differences in survival

We tested whether findings in the regulatory category were associated with differences in the survival of mutated and nonmutated cancer patients. Using a log-rank test, we detected significant differences for *TERT* in brain ($P = 3 \times 10^{-5}$) and thyroid ($P = 5 \times 10^{-2}$) cancer, *B2M*

and *FGFR2* in lung cancer ($P = 9 \times 10^{-4}$ and 1×10^{-2}, respectively), *ARRDC3* in kidney cancer ($P = 4 \times 10^{-2}$), *PIK3C2B* in bladder cancer ($P = 8 \times 10^{-3}$), *BCL6* in leukemia ($P = 1 \times 10^{-2}$), and *XBP1* in breast cancer ($P = 8 \times 10^{-4}$) (fig. S44B). These analyses provide additional support for the plausibility of some of the mutation events in this category, in addition to their location in regulatory regions and enrichment for canonical cancer genes (Fig. 2C).

Experimental evaluation of regulatory regions and noncoding mutations around XBP1

Although many events in the regulatory category fell into the promoter regions of known cancer genes (Figs. 2 and 3A), some events occurred outside of canonical regulatory regions. For example, *XBP1* mutations, which were present in ~6% of the breast cancer patients in our WGS cohort, did not primarily target the *XPP1* promoter but rather clustered in a narrow, noncoding region downstream of *XBP1* (Fig. 3A and fig. S45A), a pattern unlikely to occur by random chance (fig. S45B).

Previous studies have connected *XBP1* to breast cancer (*38*, *39*) and estrogen receptor signaling (*40*, *41*). Concordantly, Gene Set Enrichment Analysis showed estrogen receptor–dependent signaling to be the most differentially expressed pathway (FDR = 7×10^{-4}) between breast cancer samples with high versus low *XBP1* expression (Fig. 5 and fig. S46). Furthermore, *XBP1* was only expressed in prediction analysis of microarray 50 [PAM50 (*42*)] expression types related to hormone receptor signaling (luminal A/B, *HER2*-enriched types) but not in other breast tumors (basal-like type) (fig. S47). In addition, the average ATAC-seq signal around *XBP1* was 1.83-fold higher in receptor-positive versus receptor-negative breast tumors ($P < 0.001$, basal-like versus non–basal-like PAM50 subtype, Mann-Whitney *U* test) (fig. S46, D and E), suggesting that regulatory regions around *XBP1* exhibited primary activity in the hormone receptor–related subtype. We confirmed somatic mutations around *XBP1* using Sanger sequencing in breast tumors from our WGS cohort (fig. S48).

We used two experimental assays to further assess mutations near *XBP1* and to provide proof-of-principle support for the possible

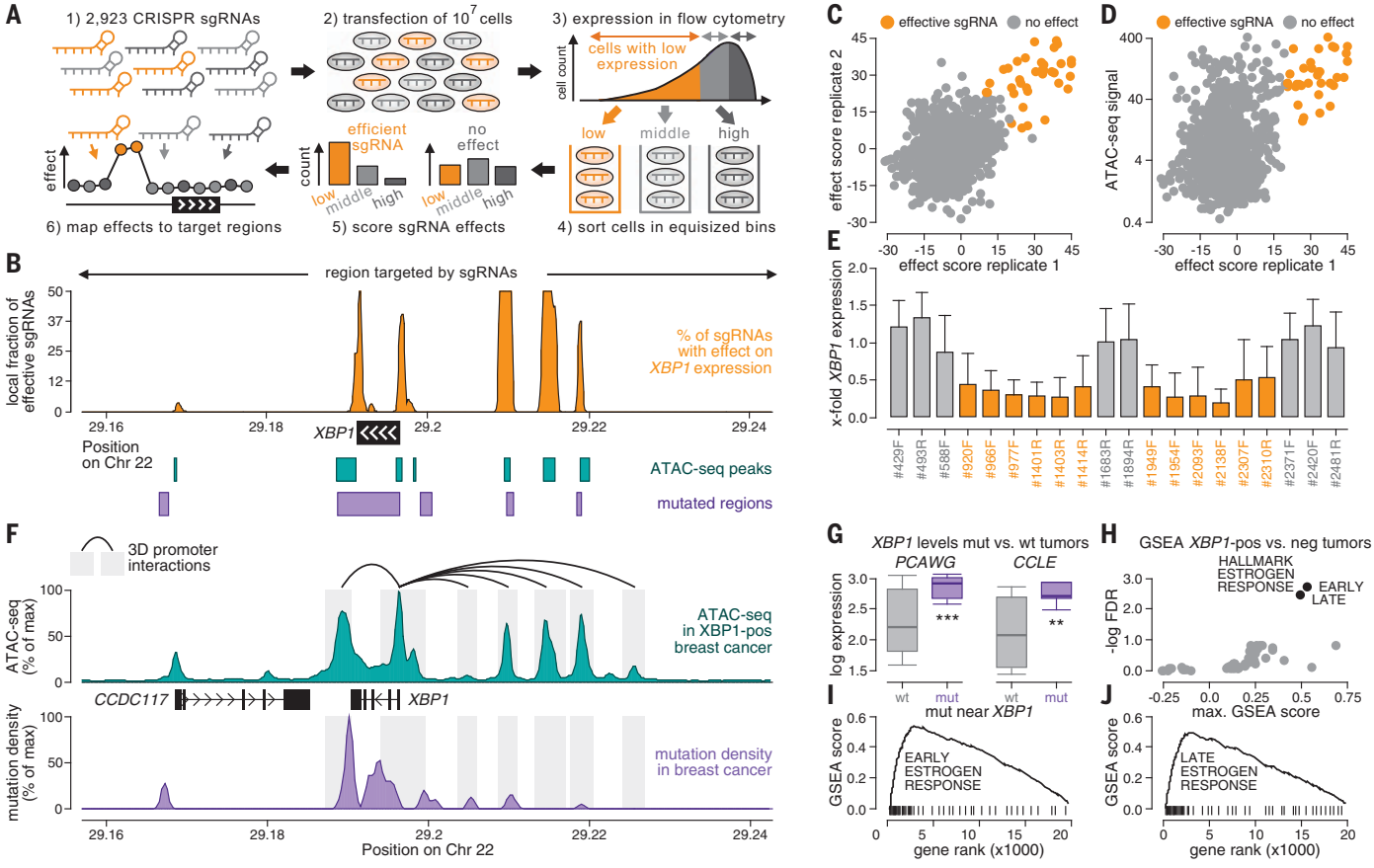


Fig. 5. Noncoding somatic mutations occur in regulatory regions around *XBP1*. (**A**) CRISPRi screening of regions around *XBP1* using a library of 2923 sgRNAs in breast cancer cells (CAMA1). Regulatory regions were localized based on sgRNAs, for which KRAB-mediated silencing of their target region led to decreased *XBP1* expression in flow cytometry (orange). (**B**) Fractions of effective sgRNAs (*y*-axis) plotted against their position around *XBP1* (*x*-axis). Positions of ATAC-seq peaks (teal, bottom), noncoding mutations (purple, bottom), and target regions of the sgRNAs (top) are annotated. (**C** and **D**) Efficacies of sgRNAs (sliding window of 10 adjacent sgRNAs) compared between experimental replicates [*x*-axis versus *y*-axis (C)] and the ATAC-seq signal of their target regions in breast cancer [*y*-axis (D)]. (**E**) Bar graphs displaying the *XBP1* expression ratio before and after CRISPRi in regulatory regions (orange) and nonregulatory regions (gray) for individual sgRNAs. Error bars reflect the SD across cells. (**F**) Mutation densities (purple), ATAC-seq signals (teal), and three-dimensional interactions in the breast cancer genome of MCF7 (ChIA-PET, black) plotted against their genomic position around *XBP1* (*x*-axis). (**G**) *XBP1* expression compared between breast tumors with [purple, mutated (mut)] and without [gray, wild-type (wt)] mutations around *XBP1* in PCAWG (left) and CCLE (right). Boxes indicate the 25/75% interquartile range, vertical lines extend to 10/90% percentiles, and horizontal lines reflect distribution medians of *XBP1* expression. Significant differences (Mann-Whitney *U* test) are annotated with asterisks: $*P < 0.05$, $**P < 0.01$, $***P < 0.001$. (**H**) Gene Set Enrichment Analysis analyzing expression differences in tumors with high versus low *XBP1* expression by computing an enrichment score (*x*-axis) and a significance value (*y*-axis) for each hallmark signature. (**I** and **J**) Gene ranks (*x*-axis) are plotted against enrichment scores (*y*-axis) for early (I) and late (J) estrogen response signatures (black).

biological relevance of mutation events outside of canonical regulatory regions (Fig. 5 and figs. S49 to S55).

As a first experiment, we performed a CRISPR interference (CRISPRi) screen to localize positive regulatory regions around *XBP1* (Fig. 5A). We tiled the genomic region around *XBP1* with a library of 2923 single-guide RNAs (sgRNAs), including the territory outside of canonical promoters and enhancers, and repressed the target regions of these sgRNAs through Krüppel associated box (KRAB)–mediated silencing in breast cancer cells (CAMA1). We then used flow cytometry [CRISPRi-Flow fluorescence in situ hybridization (CRISPRi-FlowFISH)] to quantify to what extent repression of a candidate regulatory region down-regulated *XBP1* expression (*14*) (Fig. 5A and fig. S49). This screen identified five positive regulatory regions (four upstream and one downstream of *XBP1*) in which KRAB-mediated repression down-regulated *XBP1* expression (Fig. 5B). These regulatory regions were consistent between experimental replicates (Fig. 5, C to E), and CRISPRi-FlowFISH screening results correlated with an independent experimental assay (quantitative polymerase chain reaction, $R = 0.59$; 29 sgRNAs tested in both assays) (fig. S50). In particular, many breast cancer mutations accumulated in the regulatory region that this experiment identified downstream of *XBP1*.

Companion analysis of ATAC-seq data from 74 breast tumors (*8*) confirmed the five regulatory regions from our screening experiment at a higher resolution, where they colocalized with five distinct ATAC-seq peaks around *XBP1* (Fig. 5F). These peaks were exclusive to breast tumors with high *XBP1* expression (Fig. 5F and fig. S46E), and their ATAC-seq signals correlated with *XBP1* expression (fig. S51, A to C), with the highest correlation being observed in the ATAC-seq peak downstream of *XBP1* ($R = 0.80$). In addition, regulatory regions physically interacted with the *XBP1* promoter in the three-dimensional structure of the MCF7 breast cancer genome (*43*) (Fig. 5F), and breast cancer–specific transcription factors bound to upstream regulatory regions of *XBP1* in breast cancer ChIP-seq data (fig. S51, D and E). Thus, our first experimental strategy demonstrated that important noncoding mutation events can occur outside of canonical regulatory regions, illuminating the potential of a genome-wide approach to capture somatic mutation events in both known and unknown elements of the noncoding genome.

As a second experiment, we used a luciferase reporter assay to examine the effect of mutations observed in breast cancer genomes near *XBP1* on transcriptional activity directly (figs. S52 and S53A). For this purpose, we cloned the mutated and nonmutated 193-bp sequences around 10 mutations near *XBP1* that were observed in our WGS cohort into the regulatory region of a luciferase reporter plasmid. We measured their luciferase signal in breast cancer cells (CAMA1) as a marker of their effect on transcriptional activity. For five of 10 mutations, we obtained significantly higher luciferase activity ($P < 0.05$; Mann-Whitney *U* test) for mutated sequences compared with their corresponding nonmutated sequences (fig. S52, A and B). For three mutations, we measured a >1.5-fold higher luciferase signal, which was similar to that reported for established noncoding mutations, including those in the *TERT* and *FOXA1* promoters (~2-fold) (*19*, *35*). Furthermore, despite variation between independent experiments, results correlated robustly between replicates (fig. S52C).

Differential expression analysis concordantly revealed that breast tumors with mutations around *XBP1* were associated with elevated *XBP1* expression relative to that observed in nonmutated tumors, both in tumor patients [PCAWG (*9*)] and in the Cancer Cell Line Encyclopedia [CCLE (*44*)] (Fig. 5, G to J, and fig. S42). Likewise, analysis of matched RNA sequencing (RNA-seq) and ATAC-seq data from two samples (three *XBP1* mutations) in our WGS cohort revealed that *XBP1* mutations correlated with increased fractions of mutated reads in RNA-seq and ATAC-seq data compared with their corresponding WGS data (two of three mutations examined) (fig. S53, B and C). In addition, mutations near *XBP1* exhibited differential pathogenicity compared with mutations in the rest of the genome based on two bioinformatics scores (*15*, *16*) (fig. S53, D and E). Thus, the second experimental strategy confirmed that specific mutations observed in breast cancer patients near *XBP1* were associated with increased expression and activity of their downstream regulatory region.

The supplementary materials contain additional analyses related to the phenotypes associated with *XBP1* mutations, including tumor cell proliferation (fig. S54), drug efficacy (fig. S55, A and B), the activity of related pathways (fig. S55, C and D), and patient survival (fig. S55E).

Discussion

Our study establishes a genome-wide compendium of somatic mutation events in 19 major cancer types and advances the field related to four major challenges.

First, noncoding regions comprise a heterogeneous spectrum of genomic elements, and mutation events in different parts of the noncoding genome relate to diverse aspects of tumor biology. To capture these biological differences, our approach automatically stratified mutation events based on their genomic location: Events in protein-coding regions corresponded to established coding drivers that alter protein structures of cancer-related genes. Some mutations in regulatory regions have been discussed as plausible noncoding drivers that could change protein levels of cancer-related genes with low expression in normal tissue to recruit them for oncogenesis (*6*, *9*, *10*, *19*). Events near tissue-specific genes characterized localized passenger mutation patterns linked to characteristic expression programs and physiological processes in the tumor cell of origin and are unlikely to represent prototypical oncogenic drivers. Some noncoding events could not be associated with any of these categories, so their status remains less clear. In addition, although our classification was guided by the insights from prior studies (*9*, *10*), the exact terminology and criteria differed between studies: Our category of tissue-specific genes (based on their expression pattern) was largely equivalent to PCAWG's annotation of "transcriptional processes" (based on a review of their fraction of long indels), our category of regulatory regions was mostly labeled as "candidate drivers" by PCAWG, and our upfront filter of low-quality mutations and regions was consistent with the "technical artifacts" filter used by PCAWG. Despite broad overall consistency, these classifications diverged for individual results observed in both our study and prior work. Therefore, careful follow-up is required to determine the biology of individual mutation events in detail beyond their genomic location and capture the multifaceted functional effects of somatic mutations in noncoding regions.

Our second challenge was that the current understanding of regulatory regions and other functional elements in the noncoding cancer genome is likely incomplete given that their activity and location can vary between cell types, between tumor and normal tissue, and even between patients with the same tumor type (*8*, *45*). Therefore, databases of regulatory regions (*22*–*26*) and ChIP-seq signals from normal tissue (*7*) may not capture the full diversity and versatility of functional elements in noncoding cancer genomes, and differences in the epigenomic structure of tumor and normal cells may be critical for characterizing mutation events in tumor-specific regulatory regions. Several analyses in our study, including experimental evaluation of *XBP1* mutations, highlighted that important noncoding mutation events can occur outside of canonical regulatory elements. Although tumor-specific ATAC-seq and methylation data improved the enrichment for putative functional events, many mutation events linked to cancer genes still fell outside of these regions. To address this challenge, our genome-wide analysis locates mutation events across the entire genome instead of restricting its search to canonical functional regions. In contrast to previous annotation-unbiased approaches (*9*), our approach

tiles the genome with multiple interval sizes. This proved critical for its use and performance in the noncoding genome, which harbors no predefined genomic boundaries and is ~50-fold larger than exons in coding regions. Our results may inform future experimental and clinical characterizations of tumor-specific regulatory elements, prioritize regions for hybrid-capture sequencing, and enable profiling of these mutation events at a higher read coverage.

The third challenge was that detecting somatic mutation events is technically more challenging in noncoding than in coding regions. To detect mutation events based on mutational excess, many established statistical concepts use synonymous mutations as a control of the regional background mutation rate in coding regions (*3*, *4*). These concepts are inapplicable to the noncoding genome because synonymous mutations are available in coding regions only. Therefore, methods for identifying mutation events in the noncoding genome are required to use epigenomic features to calibrate their statistical models and detect mutational excess, which is a statistically more complex problem. Furthermore, the search for activating mutations in coding regions has been guided by hotspots of mutations that recur in the same position, and these are less frequently observed in noncoding regions (*9*), possibly because noncoding mutations might converge on similar biological effects in independent genomic positions. The statistical power to detect noncoding mutation events is further limited by the large number of hypotheses resulting from the size of the noncoding genome and its lack of predefined genomic regions. In addition, although thousands of whole cancer genomes have been sequenced, the amount of WGS data that captures noncoding somatic mutations is still smaller than that available for mutations in protein-coding regions. To account for these technical difficulties, we harmonized data from two WGS consortia (*9*, *13*) and implemented a statistical approach allowing us to detect mutation events irrespective of their effects on protein-coding sequences or location within predefined genomic regions. Our approach incorporates established principles from other fields and methods (*4*, *9*–*12*, *46*, *47*) but differs in critical aspects from many existing methods. For example, instead of negative binomial regression, our genome-wide analysis is based on a segmented statistical model, which gives it greater flexibility to account for overdispersion of mutation counts and complex relationships between epigenomic and mutation data. Furthermore, instead of using synonymous mutations in coding regions for comparison, our analysis compares mutation counts of the tumor type being studied with epigenomics data and sequencing data from unrelated tumor types. Prospective histone modification ChIP-seq data from large cohorts of tumor samples could be integrated into our approach and might improve its calibration to tumor-specific background mutation rates.

The final challenge was that there is currently no consensus on which events in the noncoding genome represent genuine drivers (*6*). In coding regions, many statistical tools detect mutation events based on established markers of positive selection (such as the ratio of nonsynonymous to synonymous mutations or equivalent measures), and their findings thus uniformly harbor signs of positive selection by design (*3*, *4*). In noncoding regions, positive selection markers have not been established, and mutation events are identified based on their deviations from a careful statistical background model, including events resulting from positive selection or localized mutagenic processes. Therefore, the performance of statistical models in noncoding regions cannot be evaluated by classifying findings into true versus false positives, which is a common procedure used in coding regions (*2*, *4*). Furthermore, experimental validation of the "driverness" of mutation events identified by statistical methods remains a general limitation of the field, particularly in noncoding regions, because experimental assays to capture the oncogenic effects of noncoding mutations beyond expression changes are limited. To address these challenges, our study included multiple pan-cancer follow-up strategies, including literature support of the genes linked to noncoding mutation events, comparison with other methods, and analysis of statistical power. Furthermore, we benchmarked mutation events against orthogonal ChIP-seq, ATAC-seq, RNA-seq, drug response, transcription factor binding, protein interaction, and patient survival data. We also established four markers to identify events in candidate regulatory regions outside of traditional ChIP-seq signals and databases. In addition to these computational strategies, our study combined two experimental assays to further assess *XBP1* by characterizing regulatory regions of gene expression (CRISPRi screen) and assessing the effects of noncoding mutations in these regions on expression (luciferase reporter assay). These assays gauge orthogonal effects because point mutations in luciferase reporter experiments change only a few nucleotides, whereas sgRNAs in CRISPRi experiments can affect up to several kilobases around their target regions through KRAB-mediated silencing (*48*) and thus do not mimic the effect of point mutations. In particular, this combined strategy enables experimental follow-up irrespective of the location of mutations in canonical regulatory regions, and could therefore guide future experimental endeavors.

Moving forward, our findings could be further evaluated in prospective multiomics datasets derived from the same patients as mutation sequencing data. These data would allow a deeper characterization of our findings in the context of differential expression (matched expression data), tumor-specific, long-distance promoter-enhancer interactions (matched chromosome conformation capture data), and changes in transcription factor binding (matched transcription factor ChIP-seq data). Furthermore, some of our noncoding findings may be of direct clinical interest because they converge on genes that have been previously explored as direct or indirect targets of cancer therapies, such as *TERT* and imetelstat, *FOXA1* and fulvestrant, *FGFR2* and infigratinib, *BCR* and ibrutinib, or *RAD51B*, *GEN1*, or *STAG1* and olaparib. Additionally, our study revealed that *XBP1* mutations potentially created additional therapeutic avenues. However, many other noncoding findings were linked to genes that have not been nominated as drug targets. These could provide critical starting points for the development of personalized therapies based on noncoding cancer genomes, particularly for patients with resistance to primary treatment or no druggable options in protein-coding regions.

Broadly, given the growing use of somatic WGS in the clinical setting and in biobank-scale datasets, our study establishes a critical step toward expanding our understanding of somatic mutations from protein-coding regions to the remaining ~98% of the genome. It also provides a blueprint for prioritizing noncoding mutations for translational investigation and therapeutic development.

Materials and methods summary

We combined three complementary significance tests for the genome-wide detection of somatic mutation events, which are local accumulations or clusters of somatic mutations that deviate from the pattern observed in the rest of the genome. These three tests integrated and extended principles established in other fields or methods (*4*, *9*–*12*, *46*, *47*), as outlined below.

Significance test 1 models the mutational background based on epigenomic signals, taking into account differences in mutation rates between euchromatic and heterochromatic regions (*47*) (see section 1.2 of the materials and methods). Using this background model, test 1 identifies genomic regions with larger numbers of mutations than would be expected by chance. A similar principle to that of test 1 had been applied in some previous studies that accounted for epigenomic signals by using negative binomial regression to detect mutational significance in coding (*4*) or noncoding (*10*) regions. Significance test 1 generalizes these approaches by using a four-component mixture model [H3K4me1, H3K9me3, H3K27me3, and H3K36me3 histone ChIP-seq

data (*7*)] that allows for nonexponential relationships between mutation rates and epigenomic signals.

Significance test 2 compares the number of mutations per genomic interval between unrelated cancer types and identifies genomic regions with an unusually large number of mutations in a particular cancer type (see section 1.2 of the materials and methods). In this way, test 2 detects accumulations of mutations that are specific to a certain cancer type and could reflect a specific biology in that type of tumor tissue. To take into consideration nonlinear dependencies of mutation counts between cancer types, test 2 uses a segmented statistical model to arrange genomic regions into bins and estimate the background mutation rate within each bin separately. Furthermore, it accounts for differences in mutation rates between tumor types using regional distribution variance. Although test 1 used epigenomic data from normal tissue, test 2 serves as a proxy for tumor-specific epigenomic data given that the epigenomic structure differs between tumor and normal tissue. The importance of these differences has been highlighted in the context of somatic mutations by previous studies (*8*, *45*).

Significance test 3 detects positional clustering of mutations around biologically relevant positions in the cancer genome (see section 1.3 of the materials and methods). In addition to the biological function of genomic positions, other factors, including nucleotide contexts, coverage fluctuation, read mappability, and kataegis events, affect positional clustering. Concepts similar to those of test 3 have been used in other methods for analyzing coding and noncoding regions (*9*, *29*). Therefore, test 3 examines whether mutations occur in different positions than expected by chance, but it does not analyze whether the total number of mutations deviates from the expectation and thus does not require calibration against regional fluctuations of the background mutation rates.

To combine signals from tests 1 through 3, we tiled the genome into 1-, 10-, and 100-kb intervals with 25% overlap and performed the three tests in each of these intervals (all mutations and indels only). This strategy of an unbiased, genome-wide analysis builds on established principles from noncancer germline studies (*46*) and an annotation-unbiased strategy in PCAWG that analyzes 2-kb intervals (*9*). For each 10- and 100-kb interval, we obtained multiple P values from the interval and its subintervals (linked P values of its consecutive, nonoverlapping 1- and 10-kb subintervals; see sections 1.2 and 1.4 of the materials and methods). We then combined them using Brown's method (*11*), which was also used in previous cancer genomics studies, including PCAWG (*9*), and then adjusted them using weighted multiple hypothesis correction (*12*). To derive a genome-wide signal of significance, we selected maximally significant, nonoverlapping intervals, as described previously (*10*), and favored 10- over 100-kb intervals because they allowed us to optimize the resolution of our signal (see section 1.4 of the materials and methods). In this genome-wide signal, we identified mutation events as significant regions with an FDR < 0.1 (peak value < 0.05).

To classify mutation events, we annotated them based on their closest gene and their putative function (see section 1.5 of the materials and methods): coding regions [regions with the most mutations in exons or splice sites in exon-intron boundaries and findings detected by MutSigCV (*3*) or dNdScv (*4*)]; regulatory regions [regions with the most mutations in H3K4me3 or H3K27ac ChIP-seq peaks from Roadmap (*7*)]; tissue-specific genes (mutations around genes that are expressed in a particular tumor type); and "other" findings (mutations with unclear functions that fit no other criteria). We excluded regions with low-alignability mutations or hotspots in DNA loops (see section 1.5 of the materials and methods).

A more detailed description of the significance tests and statistical framework can be found in the materials and methods.

REFERENCES AND NOTES

1. M. R. Stratton, P. J. Campbell, P. A. Futreal, The cancer genome. *Nature* **458**, 719–724 (2009). doi: 10.1038/nature07943; pmid: 19360079
2. M. H. Bailey *et al*., Comprehensive characterization of cancer driver genes and mutations. *Cell* **174**, 1034–1035 (2018). doi: 10.1016/j.cell.2018.07.034; pmid: 30096302
3. M. S. Lawrence *et al*., Mutational heterogeneity in cancer and the search for new cancer-associated genes. *Nature* **499**, 214–218 (2013). doi: 10.1038/nature12213; pmid: 23770567
4. I.Martincorena *et al*., Universal patterns of selection in cancer and somatic tissues. *Cell* **171**, 1029–1041.e21 (2017). doi: 10.1016/j.cell.2017.09.042; pmid: 29056346
5. L. Mularoni, R. Sabarinathan, J. Deu-Pons, A. Gonzalez-Perez, N. López-Bigas, OncodriveFML: A general framework to identify coding and non-coding regions with cancer driver mutations. *Genome Biol.* **17**, 128 (2016). doi: 10.1186/s13059-016-0994-0; pmid: 27311963
6. K. Elliott, E. Larsson, Non-coding driver mutations in human cancer. *Nat. Rev. Cancer* **21**, 500–509 (2021). doi: 10.1038/s41568-021-00371-z; pmid: 34230647
7. B. E. Bernstein *et al*., The NIH Roadmap Epigenomics Mapping Consortium. *Nat. Biotechnol.* **28**, 1045–1048 (2010). doi: 10.1038/nbt1010-1045; pmid: 20944595
8. M. R. Corces *et al*., The chromatin accessibility landscape of primary human cancers. *Science* **362**, eaav1898 (2018). doi: 10.1126/science.aav1898; pmid: 30361341
9. E. Rheinbay *et al*., Analyses of non-coding somatic drivers in 2,658 cancer whole genomes. *Nature* **578**, 102–111 (2020). doi: 10.1038/s41586-020-1965-x; pmid: 32025015
10. M. Imielinski, G. Guo, M. Meyerson, Insertions and deletions target lineage-defining genes in human cancers. *Cell* **168**, 460–472.e14 (2017). doi: 10.1016/j.cell.2016.12.025; pmid: 28089356
11. M. B. Brown, 400: A method for combining non-independent, one-sided tests of significance. *Biometrics* **31**, 987–992 (1975). doi: 10.2307/2529826
12. N. Ignatiadis, B. Klaus, J. B. Zaugg, W. Huber, Data-driven hypothesis weighting increases detection power in genome-scale multiple testing. *Nat. Methods* **13**, 577–580 (2016). doi: 10.1038/nmeth.3885; pmid: 27240256
13. P. Priestley *et al*., Pan-cancer whole-genome analyses of metastatic solid tumours. *Nature* **575**, 210–216 (2019). doi: 10.1038/s41586-019-1689-y; pmid: 31645765
14. C. P. Fulco *et al*., Activity-by-contact model of enhancer-promoter regulation from thousands of CRISPR perturbations. *Nat. Genet.* **51**, 1664–1669 (2019). doi: 10.1038/s41588-019-0538-0; pmid: 31784727
15. M. Kircher *et al*., A general framework for estimating the relative pathogenicity of human genetic variants. *Nat. Genet.* **46**, 310–315 (2014). doi: 10.1038/ng.2892; pmid: 24487276
16. H. A. Shihab *et al*., An integrative approach to predicting the functional effects of non-coding and coding sequence variation. *Bioinformatics* **31**, 1536–1543 (2015). doi: 10.1093/bioinformatics/btv009; pmid: 25583119
17. P. A. Futreal *et al*., A census of human cancer genes. *Nat. Rev. Cancer* **4**, 177–183 (2004). doi: 10.1038/nrc1299; pmid: 14993899
18. D. Chakravarty *et al*., OncoKB: A precision oncology knowledge base. *JCO Precis. Oncol.* **2017**, PO.17.00011 (2017). pmid: 28890946
19. E. Rheinbay *et al*., Recurrent and functional regulatory mutations in breast cancer. *Nature* **547**, 55–60 (2017). doi: 10.1038/nature22992; pmid: 28658208
20. N. Rappaport *et al*., MalaCards: An integrated compendium for diseases and their annotation. *Database* **2013**, bat018 (2013). doi: 10.1093/database/bat018; pmid: 23584832
21. A. Fujimoto *et al*., Whole-genome mutational landscape and characterization of noncoding and structural mutations in liver cancer. *Nat. Genet.* **48**, 500–509 (2016). doi: 10.1038/ng.3547; pmid: 27064257
22. J. E. Moore *et al*., Expanded encyclopaedias of DNA elements in the human and mouse genomes. *Nature* **583**, 699–710 (2020). doi: 10.1038/s41586-020-2493-4; pmid: 32728249
23. S. Fishilevich *et al*., GeneHancer: Genome-wide integration of enhancers and target genes in GeneCards. *Database* **2017**, bax028 (2017). doi: 10.1093/database/bax028; pmid: 28605766
24. R. Andersson *et al*., An atlas of active enhancers across human cell types and tissues. *Nature* **507**, 455–461 (2014). doi: 10.1038/nature12787; pmid: 24670763
25. D. R. Zerbino, S. P. Wilder, N. Johnson, T. Juettemann, P. R. Flicek, The ensembl regulatory build. *Genome Biol.* **16**, 56 (2015). doi: 10.1186/s13059-015-0621-5; pmid: 25887522
26. A. Visel, S. Minovitsky, I. Dubchak, L. A. Pennacchio, VISTA Enhancer Browser—A database of tissue-specific human enhancers. *Nucleic Acids Res.* **35** (Database), D88–D92 (2007). doi: 10.1093/nar/gkl822; pmid: 17130149
27. S. Shuai, S. Gallinger, L. Stein; PCAWG Drivers and Functional Interpretation Working Group; PCAWG Consortium, Combined burden and functional impact tests for cancer driver discovery using DriverPower. *Nat. Commun.* **11**, 734 (2020). doi: 10.1038/s41467-019-13929-1; pmid: 32024818
28. L. Lochovsky, J. Zhang, Y. Fu, E. Khurana, M. Gerstein, LARVA: An integrative framework for large-scale analysis of recurrent variants in noncoding annotations. *Nucleic Acids Res.* **43**, 8123–8134 (2015). doi: 10.1093/nar/gkv803; pmid: 26304545
29. Y. A. Guo, M. M. Chang, A. J. Skanderup, MutSpot: Detection of non-coding mutation hotspots in cancer genomes. *NPJ Genom. Med.* **5**, 26 (2020). doi: 10.1038/s41525-020-0133-4; pmid: 32550006
30. L. B. Alexandrov *et al*., The repertoire of mutational signatures in human cancer. *Nature* **578**, 94–101 (2020). doi: 10.1038/s41586-020-1943-3; pmid: 32025018
31. N. Aizarani *et al*., A human liver cell atlas reveals heterogeneity and epithelial progenitors. *Nature* **572**, 199–204 (2019). doi: 10.1038/s41586-019-1373-2; pmid: 31292543
32. J. Liao *et al*., Single-cell RNA sequencing of human kidney. *Sci. Data* **7**, 4 (2020). doi: 10.1038/s41597-019-0351-8; pmid: 31896769
33. N. Kawano *et al*., Composite distal nephron-derived renal cell carcinoma with chromophobe and collecting duct carcinomatous elements. *Pathol. Int.* **55**, 360–365 (2005). doi: 10.1111/j.1440-1827.2005.01837.x; pmid: 15943794
34. A. Sandelin, W. Alkema, P. Engström, W. W. Wasserman, B.Lenhard, JASPAR: An open-access database for eukaryotic transcription factor binding profiles. *Nucleic Acids Res.* **32**, D91–D94 (2004). doi: 10.1093/nar/gkh012; pmid: 14681366
35. F. W. Huang *et al*., Highly recurrent TERT promoter mutations in human melanoma. *Science* **339**, 957–959 (2013). doi: 10.1126/science.1229259; pmid: 23348506
36. A. Krones-Herzig *et al*., Early growth response 1 acts as a tumor suppressor in vivo and in vitro via regulation of p53. *Cancer Res.* **65**, 5133–5143 (2005). doi: 10.1158/0008-5472.CAN-04-3742; pmid: 15958557

37. D. Szklarczyk *et al*., The STRING database in 2017: Quality-controlled protein-protein association networks, made broadly accessible. *Nucleic Acids Res.* **45** (D1), D362–D368 (2017). doi: 10.1093/nar/gkw937; pmid: 27924014
38. A. Gupta *et al*., NCOA3 coactivator is a transcriptional target of XBP1 and regulates PERK-eIF2α-ATF4 signalling in breast cancer. *Oncogene* **35**, 5860–5871 (2016). doi: 10.1038/onc.2016.121; pmid: 27109102
39. S. Chen *et al*., The emerging role of XBP1 in cancer. *Biomed. Pharmacother.* **127**, 110069 (2020). doi: 10.1016/j.biopha.2020.110069; pmid: 32294597
40. S. Sengupta, C. G. Sharma, V. C. Jordan, Estrogen regulation of X-box binding protein-1 and its role in estrogen induced growth of breast and endometrial cancer cells. *Horm. Mol. Biol. Clin. Investig.* **2**, 235–243 (2010). doi: 10.1515/hmbci.2010.025; pmid: 21297881
41. C. Wang *et al*., Estrogen receptor antagonist fulvestrant inhibits proliferation and promotes apoptosis of prolactinoma cells by regulating the IRE1/XBP1 signaling pathway. *Mol. Med. Rep.* **18**, 4037–4041 (2018). doi: 10.3892/mmr.2018.9379; pmid: 30106152
42. J. S. Parker *et al*., Supervised risk predictor of breast cancer based on intrinsic subtypes. *J. Clin. Oncol.* **27**, 1160–1167 (2009). doi: 10.1200/JCO.2008.18.1370; pmid: 19204204
43. G. Li *et al*., Extensive promoter-centered chromatin interactions provide a topological basis for transcription regulation. *Cell* **148**, 84–98 (2012). doi: 10.1016/j.cell.2011.12.014; pmid: 22265404
44. J. Barretina *et al*., The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. *Nature* **483**, 603–607 (2012). doi: 10.1038/nature11003; pmid: 22460905
45. I. Sur, J. Taipale, The role of enhancers in cancer. *Nat. Rev. Cancer* **16**, 483–493 (2016). doi: 10.1038/nrc.2016.62; pmid: 27364481
46. V. Tam *et al*., Benefits and limitations of genome-wide association studies. *Nat. Rev. Genet.* **20**, 467–484 (2019). doi: 10.1038/s41576-019-0127-1; pmid: 31068683
47. B. Schuster-Böckler, B. Lehner, Chromatin organization is a major influence on regional mutation rates in human cancer cells. *Nature* **488**, 504–507 (2012). doi: 10.1038/nature11273; pmid: 22820252
48. A. C. Groner *et al*., KRAB-zinc finger proteins and KAP1 can mediate long-range transcriptional repression through heterochromatin spreading. *PLOS Genet.* **6**, e1000869 (2010). doi: 10.1371/journal.pgen.1000869; pmid: 20221260
49. F. Dietlein, A. B. Wang, C. Fagre, A. Tang, N. Besselink, E. Cuppen, C. Li, S. R. Sunyaev, J. T. Neal, E. M. Van Allen, Source code for: Genome-wide analysis of somatic noncoding mutation patterns in cancer. Zenodo (2022); doi: 10.5281/zenodo.5913867

ACKNOWLEDGMENTS

We thank M. Meyerson, M. Brown, G. Getz, E. Rheinbay, P. Priestley, S. Chen, X. Zhou, H. Cao, M. Lupien, J. Kreisberg, and J. Ma for valuable feedback and suggestions; B. Reardon, S. Camp, B. Jiang, C. Ricker, and K. Mandl for proofreading our manuscript and improving its readability; C. Otis, P. Rogers, and the Flow Cytometry Facility of the Broad Institute for technical assistance; and J. Hyle and Y. Zhang from the St. Jude Children's Research Hospital for experimental support. This publication and the underlying study have been made possible partly through data requested (DR-050) and made available by HMF and the Center of Personalized Cancer Treatment (CPCT). Furthermore, the results presented in this study are in part based on data generated by the The Cancer Genome Atlas (TCGA) Research Network (https://www.cancer.gov/tcga) and the PCAWG and International Cancer Genome Consortium (ICGC) networks through the Data Access Compliance Office (DACO-1078465). We further acknowledge the contributions of the many clinical networks across ICGC and TCGA in providing samples and data to the PCAWG Consortium and the contributions of the Technical Working Group and the Germline Working Group of the PCAWG Consortium for the collation, realignment, and harmonized variant-calling of the cancer genomes used in this study. Finally, we thank the patients and their families for their participation in the individual CPCT, HMF, ICGC, and TCGA projects. **Funding:** F.D. was supported by the National Institutes of Health (grant no. K99 CA262152), the Claudia Adams Barr Program for Innovative Cancer Research, the EMBO Long-Term Fellowship Program (grant no. ALTF 502-2016), and the AWS Cloud Credits for Research Program. J.T.N. was supported by a Merkin Institute Fellowship. E.M.V.A. was supported by the National Institutes of Health (grant nos. R01 CA227388 and R21 CA242861). F.D., J.T.N., and E.M.V.A. were supported by ASPIRE Awards from The Mark Foundation for Cancer Research. **Author contributions:** F.D., A.B.W., C.F., A.T., S.R.S., J.T.N., and E.M.V.A. wrote the manuscript and prepared the figures with the help of all authors. F.D., A.T., S.R.S., and E.M.V.A. designed and performed computational analyses for identifying mutation events in whole-genome sequencing data. F.D., A.T., S.R.S., and E.M.V.A. designed and performed computational analyses for classifying and interpreting noncoding mutation events. A.B.W., C.F., N.B., E.C., C.L., and J.T.N. designed, performed, and interpreted experiments to evaluate noncoding mutations around XBP1. F.D., A.B.W., C.F., A.T., N.B., E.C., C.L., S.R.S., J.T.N., and E.M.V.A. reviewed the manuscript, figures, and results. **Competing interests:** E.M.V.A. is a consultant for Tango Therapeutics, Genome Medical, Invitae, Foresite Capital, Enara Bio, Monte Rosa Therapeutics, Manifold Bio, Janssen, Dynamo, and Illumina; received research support from Novartis and Bristol-Myers Squibb and travel support from Roche and Genentech; and is an equity holder of Syapse, Tango Therapeutics, Syapse, Enara Bio, Monte Rosa Therapeutics, and Genome Medical. J.T.N. is a consultant for AbbVie Inc. The other authors declare no competing interests. **Data and materials availability:** Code used to perform our analyses is available on Zenodo (*49*). Access to controlled data from the PCAWG project in the ICGC was obtained through DACO-1078465. Clinical annotations and somatic whole-genome sequencing data from the CPCT and the HMF were obtained through a data access request (DR-050). Access requests for these data can be submitted under https://www.hartwigmedicalfoundation.nl/en/applying-for-data (HMF, CPCT) and https://daco.icgc.org (PCAWG, ICGC). All other data used in this study are publicly available without restrictions.

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abg5601
Materials and Methods
Supplementary Text
Figs. S1 to S55
Tables S1 to S23
References (*50*–*221*)
MDAR Reproducibility Checklist

13 January 2021; accepted 7 March 2022
10.1126/science.abg5601

RESEARCH ARTICLE SUMMARY

IMMUNOGENOMICS

Single-cell RNA-seq reveals cell type–specific molecular and genetic associations to lupus

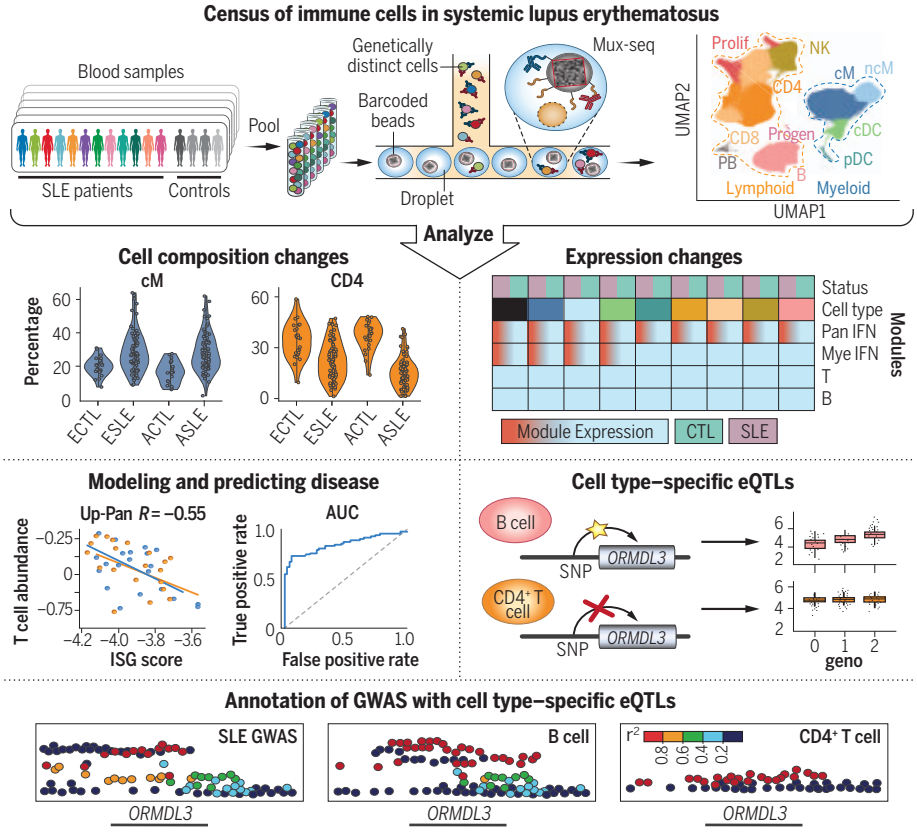
Richard K. Perez†, M. Grace Gordon†, Meena Subramaniam†, Min Cheol Kim, George C. Hartoularos, Sasha Targ, Yang Sun, Anton Ogorodnikov, Raymund Bueno, Andrew Lu, Mike Thompson, Nadav Rappoport, Andrew Dahl, Cristina M. Lanata, Mehrdad Matloubian, Lenka Maliskova, Serena S. Kwek, Tony Li, Michal Slyper, Julia Waldman, Danielle Dionne, Orit Rozenblatt-Rosen, Lawrence Fong, Maria Dall'Era, Brunilda Balliu, Aviv Regev, Jinoos Yazdany, Lindsey A. Criswell, Noah Zaitlen*, Chun Jimmie Ye*

INTRODUCTION: Systemic lupus erythematosus (SLE) is a heterogeneous autoimmune disease with elevated prevalence in women and individuals of Asian, African, and Hispanic ancestry. Bulk transcriptomic profiling has implicated increased type 1 interferon signaling, dysregulated lymphocyte activation, and failure of apoptotic clearance as hallmarks of disease. Many genes participating in these processes are proximal to the ~100 loci associated with SLE. Despite this progress, a comprehensive census of circulating immune cells in SLE remains incomplete, and annotating the cell types and contexts that mediate genetic associations remains challenging.

RATIONALE: Historically, flow cytometry and bulk transcriptomic analyses were used to profile the composition and gene expression of circulating immune cells in SLE. However, flow cytometry is biased by its use of a limited set of known markers, whereas bulk transcriptomic profiling does not have sufficient power to detect cell type–specific expression differences. Single-cell RNA sequencing (scRNA-seq) of peripheral blood mononuclear cells (PBMCs) holds potential as a comprehensive and unbiased approach to simultaneously profile the composition and transcriptional states of circulating immune cells. However, application of scRNA-seq to population cohorts has been limited by sample throughput, cost, and susceptibility to technical variability. To overcome these limitations, we previously developed multiplexed scRNA-seq (mux-seq) to enable systematic and cost-effective scRNA-seq of population cohorts.

RESULTS: We used mux-seq to profile more than 1.2 million PBMCs from 162 SLE cases and 99 healthy controls of either Asian or European ancestry. SLE cases exhibited differences in both the composition and state of PBMCs. Analysis of lymphocyte composition revealed a reduction in naïve $CD4^+$ T cells and an increase in repertoire-restricted *GZMH*$^+$ $CD8^+$ T cells. Analysis of transcriptomic profiles across eight cell types revealed that classical monocytes expressed the highest levels of both pan–cell type and myeloid-specific type 1 interferon–stimulated genes (ISGs). The expression of ISGs in monocytes was inversely correlated with naïve $CD4^+$ T cell abundance. Cell type–specific expression features accurately predicted case-control status and stratified patients into molecular subtypes. By integrating genotyping data and using a novel matrix decomposition method, we mapped shared and cell type–specific cis-expression quantitative trait loci (cis-eQTLs) across eight cell types. Cell type–specific cis-eQTLs were enriched for regions of open chromatin specific to the same or related cell types. Joint analysis of cis-eQTLs and genome-wide association study results enabled identification of cell types relevant to immune-mediated diseases, fine-mapping of disease-associated loci, and discovery of novel SLE associations. Interaction analysis identified variants whose effects on gene expression are further modified by interferon activation across patients.

CONCLUSION: SLE remains challenging to diagnose and treat. The heterogeneity of disease manifestations and treatment response highlight the need for improved molecular characterization. In a large multiethnic cohort, we demonstrate mux-seq as a systematic approach to characterize cellular composition, identify cell type–specific transcriptomic signatures, and annotate genetic variants associated with SLE. ■



Detection of cellular and genetic correlates of SLE. Genetic multiplexing enabled single-cell profiling of hundreds of individuals with and without SLE. These profiles revealed that SLE patients exhibit changes in cell composition and cell type–specific gene expression, which were used to model disease status and severity. Additionally, cell type–specific cis-eQTL maps were produced and used to annotate and contextualize genetic loci associated with SLE.

The list of author affiliations is available in the full article online.
*Corresponding author. Email: jimmie.ye@ucsf.edu (C.J.Y.); nzaitlen@g.ucla.edu (N.Z.)
†These authors contributed equally to this work.
Cite this article as R. K. Perez *et al.*, *Science* **376**, eabf1970 (2022). DOI: 10.1126/science.abf1970

READ THE FULL ARTICLE AT
https://doi.org/10.1126/science.abf1970

RESEARCH ARTICLE

IMMUNOGENOMICS

Single-cell RNA-seq reveals cell type–specific molecular and genetic associations to lupus

Richard K. Perez[1]†, M. Grace Gordon[2,3,4,5]†, Meena Subramaniam[2,4]†‡, Min Cheol Kim[1,3,4,6,7], George C. Hartoularos[2,3,4], Sasha Targ[1,2,6], Yang Sun[3,4], Anton Ogorodnikov[3,4], Raymund Bueno[3,4], Andrew Lu[8], Mike Thompson[9], Nadav Rappoport[10], Andrew Dahl[11], Cristina M. Lanata[3,12]§, Mehrdad Matloubian[3,12], Lenka Maliskova[4,13], Serena S. Kwek[14], Tony Li[14], Michal Slyper[15]¶, Julia Waldman[15], Danielle Dionne[15], Orit Rozenblatt-Rosen[15]¶, Lawrence Fong[14], Maria Dall'Era[1], Brunilda Balliu[16], Aviv Regev[15,17,18]¶, Jinoos Yazdany[3], Lindsey A. Criswell[3,4,12]#, Noah Zaitlen[19]*, Chun Jimmie Ye[3,4,12,13,20,21,22]*

Systemic lupus erythematosus (SLE) is a heterogeneous autoimmune disease. Knowledge of circulating immune cell types and states associated with SLE remains incomplete. We profiled more than 1.2 million peripheral blood mononuclear cells (162 cases, 99 controls) with multiplexed single-cell RNA sequencing (mux-seq). Cases exhibited elevated expression of type 1 interferon–stimulated genes (ISGs) in monocytes, reduction of naïve CD4$^+$ T cells that correlated with monocyte ISG expression, and expansion of repertoire-restricted cytotoxic *GZMH*$^+$ CD8$^+$ T cells. Cell type–specific expression features predicted case-control status and stratified patients into two molecular subtypes. We integrated dense genotyping data to map cell type–specific cis–expression quantitative trait loci and to link SLE-associated variants to cell type–specific expression. These results demonstrate mux-seq as a systematic approach to characterize cellular composition, identify transcriptional signatures, and annotate genetic variants associated with SLE.

Systemic lupus erythematosus (SLE) is a heterogeneous autoimmune disease affecting multiple organ systems, with elevated prevalence in women (*1*) and individuals of Asian, African, and Hispanic ancestries (*2*). Bulk transcriptomic profiling has implicated increased type 1 interferon signaling, dysregulated lymphocyte activation, and failure of apoptotic clearance as hallmarks of disease (*3*). Many genes participating in these immunological processes are proximal to the ~100 known genetic variants associated with SLE (*4*). Despite this progress, a comprehensive census of circulating immune cells in SLE remains incomplete, and annotating the cell types and cell contexts mediating genetic associations remains challenging.

Historically, different approaches have been used to characterize the role of circulating immune cells in SLE. Flow cytometry analyses, which quantify composition on the basis of known cell surface markers, reported B and T cell lymphopenia (*5*). Bulk transcriptomic analyses of peripheral blood mononuclear cells (PBMCs) universally found elevated expression of interferon-stimulated genes (ISGs) and molecularly stratified patients according to expression features (*3*, *6*). However, flow cytometry is biased by its use of a limited set of markers, whereas bulk transcriptomic profiling does not have sufficient power to detect cell type–specific expression differences. Bulk transcriptomic analysis of sorted cell populations can identify cell type–specific expression signatures in SLE (*7*). However, it does not capture cell type frequencies, obscures heterogeneity within sorted populations, and is challenging to scale to well-powered cohorts for detecting subtle disease-associated differences in gene expression.

Single-cell RNA sequencing (scRNA-seq) of PBMCs holds potential as a comprehensive and unbiased approach to simultaneously profile the composition and cell type–specific transcriptional states of circulating immune cells. When integrated with dense genotyping data, there are further opportunities to fine-map disease-associated variants and identify the cell types and states where they exert their effects. Despite its potential, application of scRNA-seq to population cohorts has been limited by low sample throughput, high cost, and susceptibility to technical variability. To overcome these limitations, we previously developed multiplexed scRNA-seq (mux-seq) to enable systematic and cost-effective scRNA-seq of population cohorts (*8*).

A census of circulating immune cells in SLE

We used mux-seq (*8*) to profile more than 1.2 million PBMCs from 264 unique samples obtained from the California Lupus Epidemiology Study (CLUES) (*9*) and the ImmVar Consortium (*10*–*12*). The 264 samples corresponded to 162 SLE cases, including 19 disease flare cases and 10 matched samples post–flare treatment, along with 99 healthy controls (fig. S1A). Most samples were from women of either European or Asian ancestry. The 264 samples and 91 replicates were profiled in 23 pools across four batches (fig. S1B). Surface protein expression for cells from processing batches 3 (155,034 cells) and 4 (375,261 cells) were also profiled using 16 and 99 DNA-conjugated antibodies, respectively. After quality control and doublet removal using freemuxlet (*8*) (mean doublet rate 22.12%; fig. S1C), 1,444,450 cells remained. Additional removal of doublets using Scrublet (*13*) (67,969 droplets), contaminating platelets, and red blood cells (112,805 cells) yielded a total of 1,263,676 cells remaining in the final dataset (fig. S1C). Genotype-based sample demultiplexing resulted in an average of 3560 singlets (standard deviation, 1103) assigned to each sample (fig. S1D).

Compositional analysis reveals CD4$^+$ T cell lymphopenia in SLE

Louvain clustering (*14*) of normalized and batch-corrected single-cell transcriptomic profiles identified 23 clusters, which were assigned to 11 cell types: CD14$^+$ classical and CD16$^+$ nonclassical monocytes (cM and ncM);

[1]School of Medicine, University of California, San Francisco, CA, USA. [2]Biological and Medical Informatics Graduate Program, University of California, San Francisco, CA, USA. [3]Division of Rheumatology, Department of Medicine, University of California, San Francisco, CA, USA. [4]Institute for Human Genetics, University of California, San Francisco, CA, USA. [5]Department of Bioengineering and Therapeutic Sciences, University of California, San Francisco, CA, USA. [6]Medical Scientist Training Program, University of California, San Francisco, CA, USA. [7]UC Berkeley–UCSF Graduate Program in Bioengineering, San Francisco, CA, USA. [8]UCLA-Caltech Medical Scientist Training Program, Los Angeles, CA, USA. [9]Department of Computer Science, University of California, Los Angeles, CA, USA. [10]Department of Software and Information Systems Engineering, Ben-Gurion University of the Negev, Be'er Sheva, Israel. [11]Section of Genetic Medicine, Department of Medicine, University of Chicago, Chicago, IL 60637, USA. [12]Rosalind Russell/Ephraim P. Engleman Rheumatology Research Center, University of California, San Francisco, CA, USA. [13]Department of Epidemiology and Biostatistics, University of California, San Francisco, CA, USA. [14]Division of Hematology/Oncology, Department of Medicine, University of California, San Francisco, CA, USA. [15]Klarman Cell Observatory, Broad Institute, Cambridge, MA, USA. [16]Department of Computational Medicine, David Geffen School of Medicine, University of California, Los Angeles, CA, USA. [17]Koch Institute for Integrative Cancer Research, Massachusetts Institute of Technology, Cambridge, MA 02139, USA. [18]Howard Hughes Medical Institute, Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139, USA. [19]Center for Neurobehavioral Genetics, Semel Institute for Neuroscience and Human Behavior, University of California, Los Angeles, CA, USA. [20]Parker Institute for Cancer Immunotherapy, San Francisco, CA, USA. [21]Chan Zuckerberg Biohub, San Francisco, CA 94158, USA. [22]Bakar Computational Health Sciences Institute, University of California, San Francisco, CA, USA.
*Corresponding author. Email: jimmie.ye@ucsf.edu (C.J.Y.); nzaitlen@g.ucla.edu (N.Z.)
†These authors contributed equally to this work.
‡Present address: Immunai Inc., New York, NY, USA.
§Present address: National Human Genome Research Institute, Bethesda, MD, USA.
¶Present address: Genentech, San Francisco, CA, USA.
#Present address: Genomics of Autoimmune Rheumatic Disease Section, National Human Genome Research Institute, Bethesda, MD, USA.

conventional and plasmacytoid dendritic cells (cDC and pDC); $CD4^+$ and $CD8^+$ T cells (CD4 and CD8); natural killer cells (NK); B cells (B); plasmablasts (PB); proliferating T and NK cells (Prolif); and progenitor cells (Progen) (fig. S2A). Regions of the uniform manifold approximation and projection (UMAP) (*15*) were occupied by cells of different cell types (Fig. 1A), and to a lesser extent, different case-control status and ancestry (Fig. 1B and fig. S2B). Different pools and processing batches had no observable effects on the distribution of cells (fig. S2, C and D).

We first assessed changes in cellular composition in SLE by comparing the frequencies of 11 cell types between cases and controls of Asian and European ancestry separately. Cell type percentage estimates from mux-seq were reproducible between biological replicates (median Pearson R_{cases} = 0.79 and $R_{controls}$ = 0.85) (fig. S2E) and correlated with estimates obtained from surface protein profiling for batch 4 (median Spearman R = 0.88). Relative to controls, cases were most notably marked by a decrease in CD4 percentage [weighted least squares (WLS); Asian, −20.4%; European, −10.0%; Fisher's method $P_{meta:Fisher} < 5.58 \times 10^{-16}$] and an increase in cM (Asian, +11.9%; European, +8.8%; $P_{meta:Fisher} < 9.75 \times 10^{-7}$) and Prolif percentages (Asian, +0.55%; European, +0.38%; $P_{meta:Fisher} < 1.93 \times 10^{-3}$; Fig. 1C and table S1). Although most changes were correlated between ancestries (Pearson R = 0.97), Asian cases were marked by a greater reduction in CD4 percentage [$\log_2$(fold change) = −0.36, $P_{WLS} < 5.60 \times 10^{-5}$; Fig. 1D]. Cases not receiving therapy (N = 21) exhibited changes in composition similar to cases receiving therapy (Pearson R_{Asian} = 0.89 and $R_{European}$ = 0.92; fig. S2H). Relative to cases not receiving oral steroids (OS; N = 78), cases treated with OS (N = 82) exhibited an increase in CD8 percentage (Asian, +5.2%; European, +3.9%; $P_{meta:Fisher} < 4.23 \times 10^{-3}$) and a decrease in ncM percentage (Asian, −1.3%; European, −1.0%; $P_{meta:Fisher} < 3.54 \times 10^{-3}$; fig. S2F). Cases treated with azathioprine (AZ, N = 15) had a decrease in NK percentage (Asian, −4.3%; European, −7.7%; $P_{meta:Fisher} < 6.68 \times 10^{-5}$) and an increase in PB percentage (Asian, +0.2%; European, +0.3%; $P_{meta:Fisher} < 1.36 \times 10^{-3}$; fig. S2F) relative to cases not receiving AZ. Cases treated with mycophenolate mofetil (N = 54), hydroxychloroquine (N = 113), methotrexate (N = 13), or a calcineurin inhibitor (N = 10) did not exhibit significant differences in composition compared with cases not receiving each of these therapies. These results suggest that the decrease in $CD4^+$ T cell percentages and increase in classical monocyte percentages in patients with SLE are not due to therapy.

We next assessed whether changes in CD4 and cM percentages were due to changes in the absolute abundance of either population. We analyzed lymphocyte and monocyte abundances reported in the UCSF electronic health record (EHR) complete blood count. Reported abundances in the EHR were highly correlated with the estimated abundances from mux-seq (Pearson R_{lympho} = 0.97 and R_{mono} = 0.87; fig. S2G). Comparing an additional 100 cases with 154 controls matched for self-reported ancestry, age, and sex, cases exhibited a significant reduction in lymphocyte abundance [ordinary least squares (OLS); Asian, -7.4×10^8 cells/liter, $P_{OLS} < 3.46 \times 10^{-9}$; European, -5×10^8 cells/liter, $P_{OLS} < 1.07 \times 10^{-6}$; Fig. 1E] but no difference in monocyte abundance (Asian, P_{OLS} = 0.61; European, P_{OLS} = 0.98). To assess whether a causal relationship exists between lymphocyte decrease and SLE, we performed generalized summary data–based Mendelian randomizations using summary statistics for genetic associations to immune cell composition (*16*, *17*). The mediation effect of variants associated with lymphocyte abundance ($\beta_{lympho \to SLE}$ = −0.39, $P_{lympho \to SLE} < 0.008$), but not monocyte abundance ($\beta_{mono \to SLE}$ = 0.009, $P_{mono \to SLE} < 0.92$), was negative on SLE risk. A reverse causation analysis did not show mediation of SLE risk on lymphopenia ($P_{SLE \to lympho} < 0.24$, $P_{SLE \to mono} < 0.20$; Fig. 1F), although an alternative explanation of horizontal pleiotropy cannot be excluded.

Decrease of circulating naïve $CD4^+$ T cells in SLE

Previous studies revealed impaired activation of T and B memory cells and elevated expression of ISGs in lymphocytes from patients with SLE (*18*). To characterize changes in frequencies and transcriptomic profiles of lymphoid populations in SLE, we reclustered lymphoid cells and assigned the resulting 26 clusters to 14 subpopulations (Fig. 2A). Within non-T cells, we identified two NK and four B cell subpopulations. The NK compartment consisted of NK_{Bright} cells expressing high levels of *GNLY* and moderate levels of *NKG7* and NK_{Dim} cells expressing high levels of *NKG7* and *CD16* (*FCGR3A*) (Fig. 2B). The B cell compartment consisted of naïve cells expressing *TCL1A* ($B_{Naïve}$), memory cells expressing *BANK1* (B_{Mem}), plasma cells expressing *MZB1* (B_{Plasma}), and an atypical memory subpopulation expressing *FCRL5*, *CD11c*, and *TBX21* and lacking expression of *CD21* ($B_{Atypical}$; Fig. 2B). Atypical B cells may also contain age-associated B cells that share some (*CD11c*$^+$, *TBX21*$^+$, *CD21*$^-$) but not all of the expression markers [*FCRL5* (*19*)]. As a percentage of lymphocytes, neither NK nor B cell subpopulations significantly differed by case-control status.

In the $CD4^+$ T cell compartment, we identified canonical subpopulations of naïve cells expressing *CCR7* ($CD4_{Naïve}$), effector memory cells lacking *CCR7* expression while expressing OX40 receptor (*TNFRSF4*) and *IL7R* ($CD4_{EM}$), and regulatory cells expressing the canonical transcription factor *FOXP3* and its direct target *RTKN2* (*20*) ($CD4_{Reg}$; Fig. 2, A and B). Relative to controls, the most pronounced difference in cases was a reduction of $CD4_{Naïve}$ percentage (WLS; Asian, −21.7%; European, −11.8%; Fisher's method $P_{meta:Fisher} < 8.63 \times 10^{-21}$; Fig. 2C and table S2), with Asian cases exhibiting significantly more reduction than European cases ($P_{WLS} < 5.20 \times 10^{-5}$). No significant association between $CD4_{Naïve}$ percentage and age (Spearman P = 0.76; fig. S3A) or treatment (fig. S3B) was detected. Cases not on therapy (N = 21) exhibited a similar decrease in $CD4_{Naïve}$ percentage relative to controls (Asian, −25.6%; European, −9.7%; $P_{meta:Fisher} < 2.66 \times 10^{-7}$; fig. S3E).

Clonal expansion of cytotoxic *GZMH*$^+$ T cells in SLE

Within the $CD8^+$ T cell compartment, we identified naïve cells expressing *CCR7* ($CD8_{Naïve}$) and three effector memory subpopulations, including mucosal-associated invariant T cells expressing *KLRB1* and *GZMK* ($CD8_{MAIT}$) and two clusters lacking expression of *KLRB1* and expressing the chemokine *CCL5*, effector molecule *PRF1*, and exhaustion marker *LAG3* (Fig. 2, A and B). The two non-MAIT clusters could be distinguished by the expression of granzymes ($CD8_{GZMH}$, *GZMH* and *GZMB*; $CD8_{GZMK}$, *GZMK*) and mirrored the NK subpopulations (NK_{Dim}, *GZMH* and *GZMB*; NK_{Bright}, *GZMK*) (Fig. 2B and fig. S3C). Within the $CD8_{GZMH}$ population, 6% were $CD4^+CD8^-$ cells according to CD4 surface expression in the subset of samples that were also profiled using DNA-conjugated antibodies. Relative to controls, the $CD8_{GZMH}$ percentage was significantly increased in cases (Asian, +8.6%; European, +6.0%; $P_{meta:Fisher} < 3.43 \times 10^{-4}$; Fig. 2C and table S2) and was observed at similar percentages in flaring and untreated cases (fig. S3, C to E). Additionally, we observed a reduction in $CD8_{MAIT}$ percentage in cases (Asian, −1.1%; European, −0.7%; $P_{meta:Fisher} < 6.93 \times 10^{-6}$; Fig. 2C and table S2).

In addition to increased frequency within lymphocytes, $CD8_{GZMH}$ cells were a transcriptionally heterogeneous population with elevated expression of cytotoxic, exhaustion, and ISG signatures in SLE cases relative to controls (Fig. 2D). The expression of these signatures was not associated with treatment (fig. S3F). Additionally, only the ISG signature was inversely correlated with age (Pearson R = −0.39, $P < 6.57 \times 10^{-7}$). Across cells, the correlation between cytotoxic and ISG signature genes (mean $R_{Pearson}$ = 0.16) and between cytotoxic and exhaustion signature genes (mean $R_{Pearson}$ = 0.10) were generally low (Fig. 2E). Thus, in cases, these pathways are unlikely to be jointly activated in the same cells. This was in stark contrast

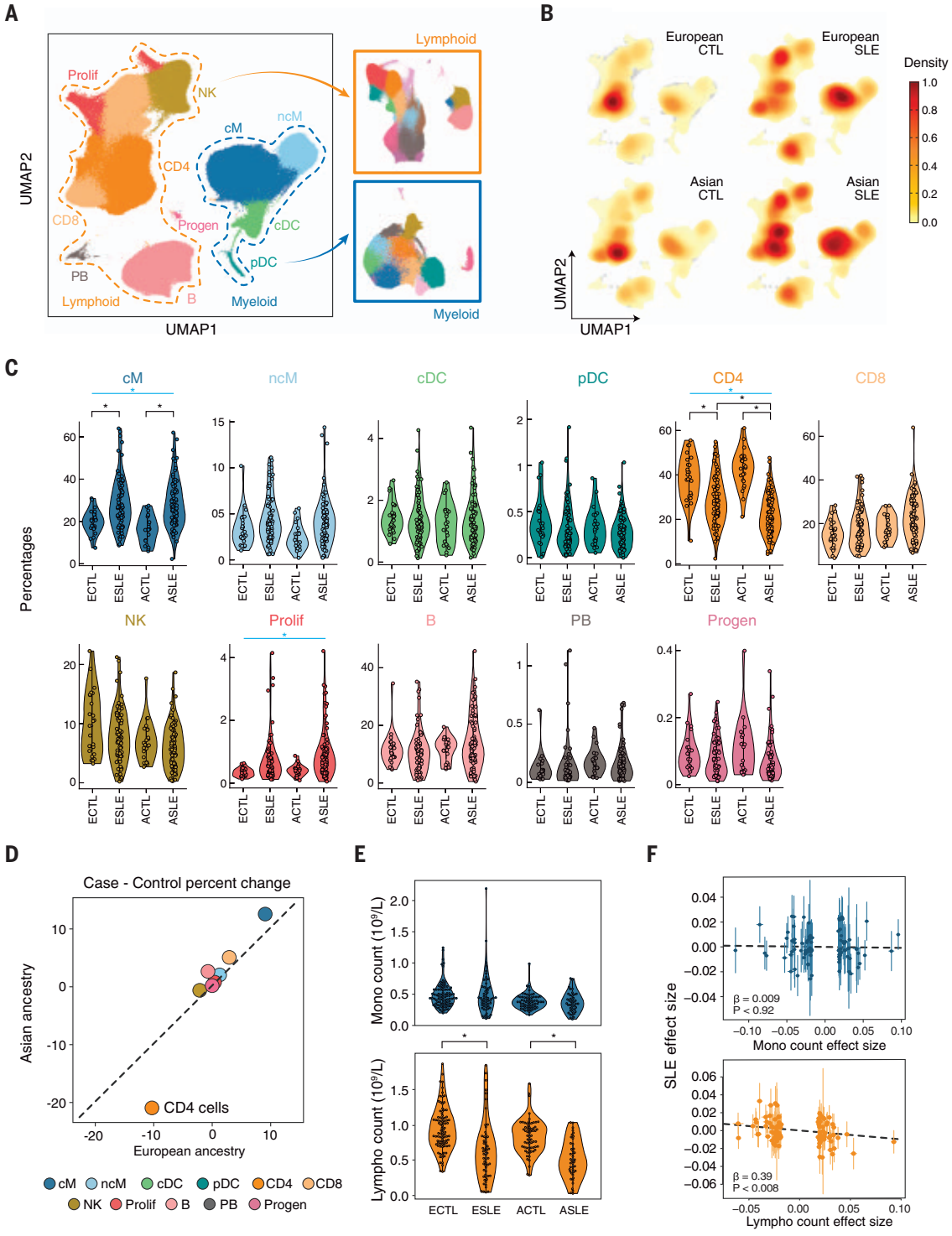


Fig. 1. Changes in composition of circulating immune cells in SLE. (**A**) UMAP and assignment of 1.2 million cells to 11 cell types: classical and nonclassical monocytes (cM and ncM); conventional and plasmacytoid dendritic cells (cDC and pDC); $CD4^+$ and $CD8^+$ T cells (CD4 and CD8); natural killer cells (NK); B cells (B); plasmablasts (PB); proliferating lymphocytes (Prolif); $CD34^+$ progenitors (Progen). Subclustering of lymphoid (orange box) and myeloid (blue box) cell populations. (**B**) Cell density plots of cases and controls separated by ancestry. (**C**) Percentage (*y* axis) versus case-control status (*x* axis) for each cell type separated by ancestry. Cell types with significant percentage changes between cases and controls are highlighted. $^*P_{adjusted} < 0.05$ [weighted least squares (WLS)]; blue bar indicates significant meta-analysis by Fisher's method. (**D**) Correlation in percentage change versus controls between European (*x* axis) and Asian (*y* axis) cases. Colors are the same as in (C). (**E**) Monocyte (top) and lymphocyte (bottom) abundances (*y* axes) versus case-control status (*x* axis) from the UCSF EHR. Significant differences between cases and controls are highlighted. $^*P_{adjusted} < 0.05$ (OLS). (**F**) Scatterplot of effect sizes on SLE status (*y* axis) versus effect sizes on monocyte (top) or lymphocyte (bottom) abundance (*x* axes) for genetic variants associated with both traits reported (*4*, *17*). ECTL, European control; ESLE, European case; ACTL, Asian control; ASLE, Asian case.

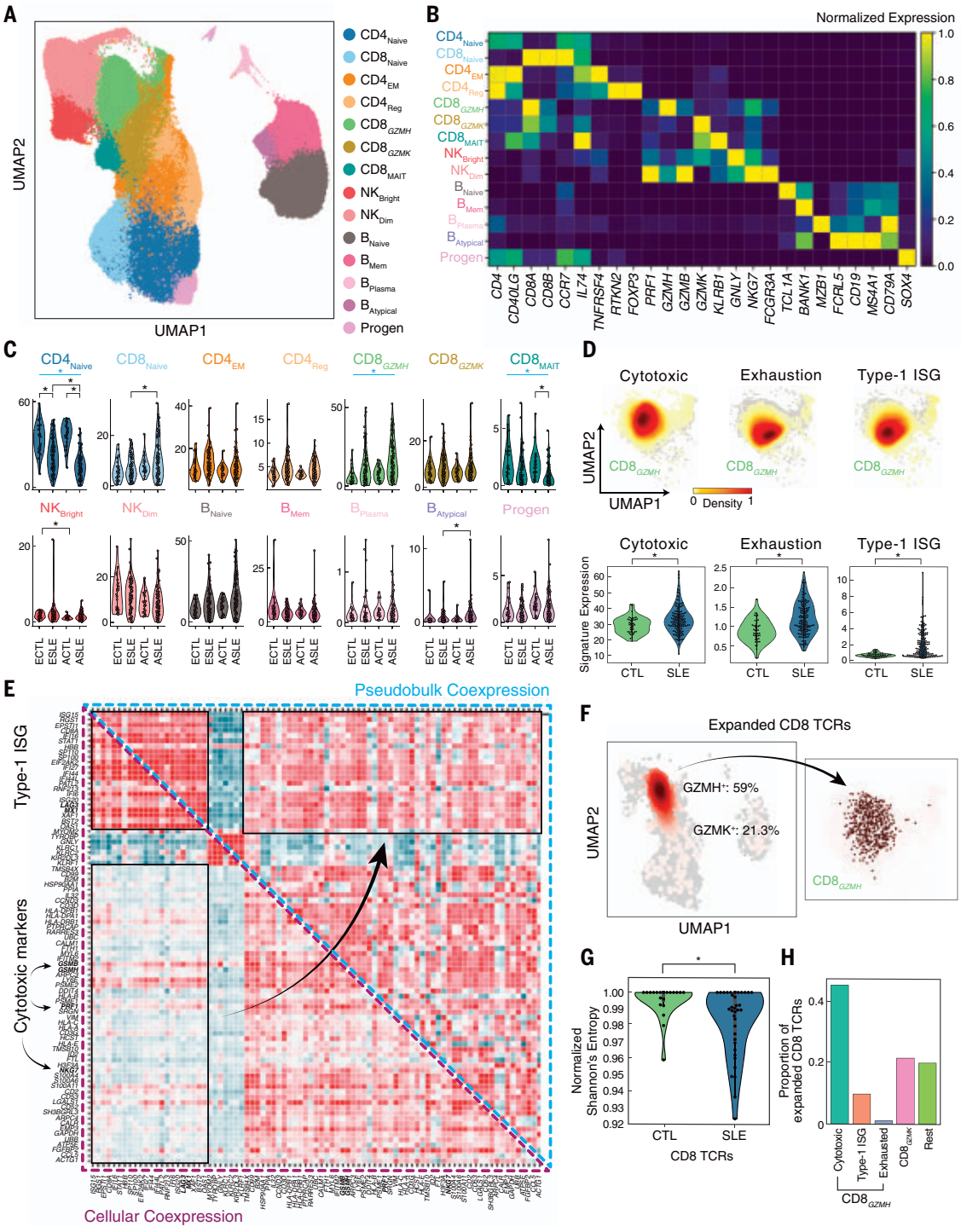


Fig. 2. Reduction of naïve CD4$^+$ and expansion of cytotoxic CD8$^+$ T cells in SLE. (**A**) UMAP of lymphoid cells reclustered into 14 subpopulations: naïve, effector memory, and regulatory CD4$^+$ T cells ($CD4_{Naïve}$, $CD4_{EM}$, $CD4_{Reg}$); naïve, *GZMH*$^+$ cytotoxic, *GZMK*$^+$ cytotoxic, and mucosal-associated invariant CD8$^+$ T cells ($CD8_{Naïve}$, $CD8_{GZMH}$, $CD8_{GZMK}$, $CD8_{MAIT}$); CD56^{bright} and CD56^{dim} natural killer cells (NK_{Bright}, NK_{Dim}); naïve, memory, plasma, and atypical B cells ($B_{Naïve}$, B_{Mem}, B_{Plasma}, $B_{Atypical}$); and CD34$^+$ progenitors (Progen). (**B**) Expression of marker genes (columns) used to annotate each subpopulation (rows) colored by normalized expression level. (**C**) Percentage (*y* axis) versus case-control status (*x* axis) for each lymphoid subpopulation separated by ancestry. Subpopulations with significant percentage changes between cases and controls are highlighted. *$P_{adjusted}$ < 0.05 (WLS); blue bar indicates significant meta-analysis by Fisher's method. (**D**) Density plot showing average expression of cytotoxic, exhaustion, and type 1 interferon–stimulated gene (ISG) signatures in $CD8_{GZMH}$ cells (top) and across individuals (bottom) separated by case-control status and ancestry. *P < 0.05 (WLS). (**E**) Coexpression of top 300 differentially expressed genes between cases and controls in $CD8_{GZMH}$ cells computed across single cells (lower triangular matrix) or across donor-specific pseudobulk expression profiles (upper triangular matrix). (**F**) All (light pink) and expanded (red) TCR sequences detected shown on UMAP of all cells (left) and *GZMH*$^+$ cells (right). (**G**) Normalized Shannon's entropies of CD8$^+$ TCR repertoire diversity (*y* axis) in cases and controls (*x* axis). *P < 0.05 (WLS). (**H**) Percentage of expanded CD8$^+$ TCRs identified as *GZMH*$^+$ cells expressing cytotoxic, ISG, and exhaustion signatures, *GZMK*$^+$ cells (GZMK), and all other cells (Rest). ECTL, European control; ESLE, European case; ACTL, Asian control; ASLE, Asian case.

to the high correlation between signature genes calculated across $CD8_{GZMH}$ pseudobulk expression profiles from different individuals, highlighting the limitation of bulk analysis in uncovering additional heterogeneity within a seemingly homogeneous population (Fig. 2E).

To further investigate the clonality of the $CD8_{GZMH}$ and $CD8_{GZMK}$ populations, we amplified and sequenced the CDR3 region of the T cell receptor (TCR), recovering paired *TCRA* and *TCRB* sequences from 10.2% of CD4 and 8.7% of CD8 cells with no differences in the number of unique TCRs detected between cases (N = 83) and controls (N = 20) ($P_{Wilcoxon}$ = 0.72). Of the expanded CD8 clones, 59% were from $CD8_{GZMH}$ cells and 21% from $CD8_{GZMK}$ cells (Fig. 2F). Relative to controls, cases exhibited a restricted repertoire in CD8 cells ($P_{Wilcoxon}$ < 0.01; Fig. 2G) but not CD4 cells ($P_{Wilcoxon}$ = 0.91; fig. S3, G and H). Within the $CD8_{GZMH}$ subpopulation, cells expressing the cytotoxic signature were expanded at a ~4:1 ratio to cells expressing the ISG signature (44.8% versus 9.7%, Fig. 2H). As a positive control, clones expressing the invariant *TRAV1-2* and *TRAJ33* chains were enriched within the $CD8_{MAIT}$ cluster relative to other cell types (Tukey's HSD P < 0.001; fig. S3I).

Expression changes across 11 peripheral immune cell types in SLE

Bulk transcriptomic analyses of PBMCs have consistently reported the association between SLE and elevated expression of ISGs, which is normally observed during acute viral infections (*21*). Longitudinal bulk analysis of 158 pediatric cases confirmed elevated expression of ISGs in patients with more severe acute presentations and increased renal and neurological involvement (*3*). However, bulk analysis has limited power to pinpoint the cell types producing the ISG signature or to identify additional cell type–specific signatures. Recent analysis of 33 pediatric cases demonstrated the potential of scRNA-seq to assign cell type specificity to previously identified ISGs from bulk analysis (*6*).

Transcriptional differences were characterized for each of 11 circulating immune cell types between SLE cases and controls. We found that 302 genes were differentially expressed (DE) in at least one cell type between cases and controls of either Asian or European ancestry, not confounded by medication [|log(fold change)| > 0.5; $P_{adjusted}$ < 0.05; table S3 and fig. S4, A and G]. Hierarchical clustering of pseudobulk expression profiles of these DE genes across cell types resulted in six modules (Fig. 3A). Relative to controls, cases up-regulated the expression of a module of ISGs across all cell types (Pan_{up}) and a myeloid-specific module (Mye_{up}) containing *IFITM1/3, IFITM3*, *APOBEC3A*, *RNASE2*, and *IFIT2*. Both modules were enriched for type 1 interferon signaling and innate immune pathways (Fig. 3B). Additionally, we identified a down-regulated module across all cell types enriched for the interaction between lymphoid and non-lymphoid cells (Pan_{down}), a myeloid-specific down-regulated module (Mye_{down}) enriched for hedgehog signaling, a T cell–specific up-regulated module (T_{up}) enriched for leukocyte activation, and a B cell–specific up-regulated module (B_{up}) enriched for AP-1 transcriptional response and Toll-like receptor signaling (Fig. 3B).

Our results were validated by single-cell transcriptomic analyses of PBMCs activated in vitro by recombinant interferon-β (rIFNB1) (*8*) and from pediatric patients with SLE (*6*). For each cell type, particularly myeloid populations, expression fold changes between cases and controls were highly correlated with fold changes between rIFNB1-stimulated and unstimulated cells (fig. S4B). Of the 100 ISGs previously identified from bulk analysis and analyzed in pediatric SLE (*6*), 64 were DE in at least one cell type and mainly resided in the Pan_{up} (46/79) and Mye_{up} (8/64) modules. Interestingly, 11 genes were DE only across PBMC pseudobulks, illustrating a likely confounding effect of bulk analysis due to differences in cellular composition between cases and controls (table S4). The large sample size of our cohort resulted in the identification of 238 previously undescribed DE genes in adult SLE, 56 of which were myeloid-specific.

Pronounced type 1 interferon response in classical monocytes

Myeloid cells exhibited the most DE genes between cases and controls, consisting of known and novel genes associated with SLE. To further investigate their heterogeneity, we reclustered myeloid cells into six clusters differentiating the monocyte lineage (cM, $CD14^+$ classical; ncM, $FCGR3A^+$ nonclassical; ncM_{comp}, $C1QA^+/FCGR3A^+$ complement-expressing nonclassical) and the dendritic cell lineage (cDC1, $CLEC10A^+$ conventional type 1; cDC2, $CLEC9A^+$ conventional type 2; pDC, $IRF7^+$ plasmacytoid; Fig. 3, C and D, and fig. S4, C and D). Although pDCs can derive from either myeloid or lymphoid progenitors, their expression profiles were more similar to, and thus jointly analyzed with, other myeloid populations (*22*). We also detected AXL^+ dendritic cells within both cDC1s and pDCs, consistent with their suggested distribution as a transitioning population between cDCs and pDCs (*23*) (fig. S4E). As a percentage of myeloid cells relative to controls, cases exhibited reduced percentages of pDCs (WLS; Asian, −0.6%; European, −1.8%; Fisher's method $P_{meta:Fisher} < 2.33 \times 10^{-24}$), cDC1s (Asian, −2.0%; European, −1.9%; $P_{meta:Fisher} < 2.65 \times 10^{-14}$), and cDC2s (Asian, −0.2%; European, −0.1%; $P_{meta:Fisher} < 2.51 \times 10^{-7}$) and increased percentages of cMs (Asian, +3.6%; European, +3.7%; $P_{meta:Fisher} < 1.78 \times 10^{-5}$) and ncM_{comp}s (Asian: +0.5%; European, +0.2%; $P_{meta:Fisher} < 1.67 \times 10^{-3}$; Fig. 3E and table S5).

Next, we used RNA velocity to assess the transcriptional heterogeneity of each myeloid cell type along a trajectory of inferred activation (*24*, *25*). In cMs, ncMs, and ncM_{comp}s, velocity analysis of DE genes revealed that inferred activation largely reflected the degree of average ISG expression (Mye_{up}; Fig. 3F) with regions of high activation enriched for cells from SLE cases (Fig. 3G). These results were similar in cDC populations (fig. S4F). Ordering cMs along inferred activation showed higher activation from cases with higher SLE Disease Activity Index (SLEDAI) (*26*) defined using clinical features (t test; Asian, $P < 5 \times 10^{-4}$; European, $P < 3.2 \times 10^{-7}$; Fig. 3H). The average inferred activation was better correlated with SLEDAI in European ($R_{Pearson}$ = 0.66) than in Asian cases ($R_{Pearson}$ = 0.52; Fig. 3I). A wide range of average inferred activations were observed in patients of either ancestry with lower disease activity (SLEDAI between 0 and 4), which suggests that clinical measures underlying SLEDAI do not fully capture the molecular heterogeneity of SLE.

Expression modules enable clinical prediction and patient stratification

Previous work in mouse models has shown that type 1 interferons up-regulate the expression of *CD69*, thereby inhibiting lymphocyte egress from lymphoid tissue (*27*). We hypothesized that the pleiotropic effects of type 1 interferons in patients with SLE may underlie the monocyte-dominant expression of ISGs and inhibit $CD4^+$ T cells from exiting lymphoid tissue, resulting in the observed decrease of circulating naïve $CD4^+$ T cells. Consistent with this hypothesis, both the Pan_{up} and Mye_{up} gene module scores were highly correlated with $CD4_{Naïve}$ abundance (Asian, Pearson R_{Panup} = −0.52; European, R_{Panup} = −0.57; $P_{meta:Fisher} < 1.04 \times 10^{-3}$; Asian, R_{Myeup} = −0.35; European, R_{Myeup} = −0.48; $P_{meta:Fisher}$ < 0.02; Fig. 4A and fig. S5A).

One of the diagnostic difficulties of SLE is the extensive heterogeneity in disease manifestations. Consistent with this heterogeneity, individual clinical features weakly correlated with module scores (Fig. 4B). We therefore used the expression of individual module genes over pseudobulks of the relevant cell types as features for clinical prediction and molecular stratification of SLE. Although the 302 expression features had good out-of-sample predictive power for case-control status [area under the curve (AUC) = 0.84; Fig. 4C], they had only modest predictive power for individual clinical features, reflective of the modest correlation between clinical features and module scores (Fig. 4D and fig. S5B). To molecularly stratify cases, we performed principal components

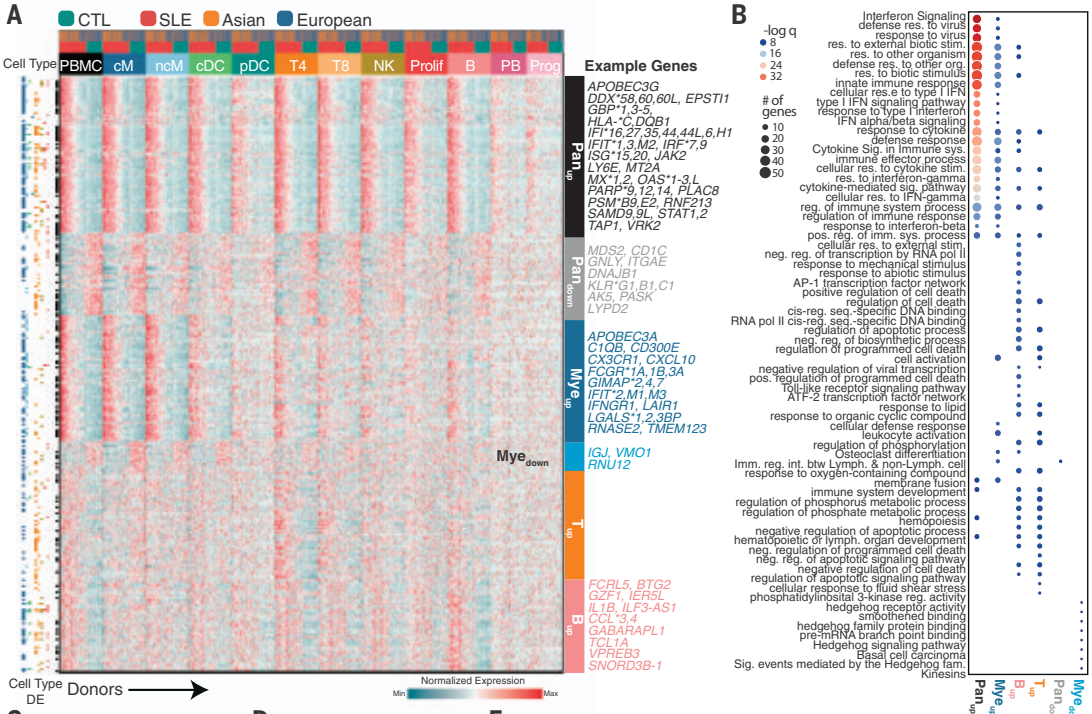


Fig. 3. Type 1 interferon–stimulated gene expression of myeloid cells in SLE. (**A**) Heatmap of pseudobulk gene expression profiles of 302 differentially expressed genes detected in at least one of 11 cell types. For each gene, colored row bars indicate cell types in which it was differentially expressed. Colored columns indicate cell type, case-control status, and ancestry. Labeled modules were identified using hierarchical clustering. (**B**) Top GSEA (gene set enrichment analysis) pathway enrichment results for each module. Each dot color represents the –log(*q*) value; dot size represents the number of genes overlapping with the gene ontology. (**C**) Identification of six myeloid cell types including classical, nonclassical, and complement-expressing nonclassical monocytes (cM, ncM, ncM_{comp}), conventional type 1, conventional type 2, and plasmacytoid dendritic cells (cDC1, cDC2, pDC). (**D**) Marker genes used for annotating each cell type. (**E**) Percentages of myeloid cells (*y* axis) versus case-control status and ancestry (*x* axis) for each myeloid subpopulation. Myeloid subpopulations with significant percentage changes between cases and controls are highlighted. *P < 0.01, ***P < 0.0001 (WLS); blue bar indicates significant meta-analysis by Fisher's method. (**F** and **G**) RNA velocity stream plots for cM (right UMAP) and ncM and ncM_{comp} (left UMAP) subpopulations colored by the average expression of Mye_{up} genes enriched for type 1 ISGs (F) and the relative density of cells from SLE cases versus healthy controls (G). (**H**) Distribution of the degree of inferred activation for individuals across disease activities (HC, healthy controls; inactive, SLEDAI between 0 and 4; active, SLEDAI of 5 or more). (**I**) Average inferred activation across cells per sample (*y* axis) versus disease activity (*x* axis) for Asian (left) and European (right) samples separately. ECTL, European control; ESLE, European case; ACTL, Asian control; ASLE, Asian case.

analysis (PCA) of expression features followed by K-means clustering to identify two clusters that broadly separated donors by case-control status (Fig. 4E), severity of SLEDAI score (Fig. 4F), and along principal component 1 (PC1). Cases in the High cluster had significantly higher inferred activation of monocytes relative to cases in the Low cluster ($P_{Wilcoxon} < 6.20 \times 10^{-9}$; fig. S5C). PC1 correlated most with genes in the Pan_{up}, Mye_{up}, and B_{up} modules, including the myeloid-specific expression of *IFITM3*, a gene previously described to stratify pediatric SLE cases (*3*) (Fig. 4E). To assess the correspondence between molecular clusters and clinical features, we projected 94 held-out cases each to a molecular cluster on the basis of expression features (Fig. 4G). Cases assigned to the High cluster were enriched for disease flare (15/19 flare cases, fig. S5D) and portended a factor of 5 increase in the odds of having anti-Smith antibodies ($P_{adjusted:Fisher} < 0.05$; Fig. 4H).

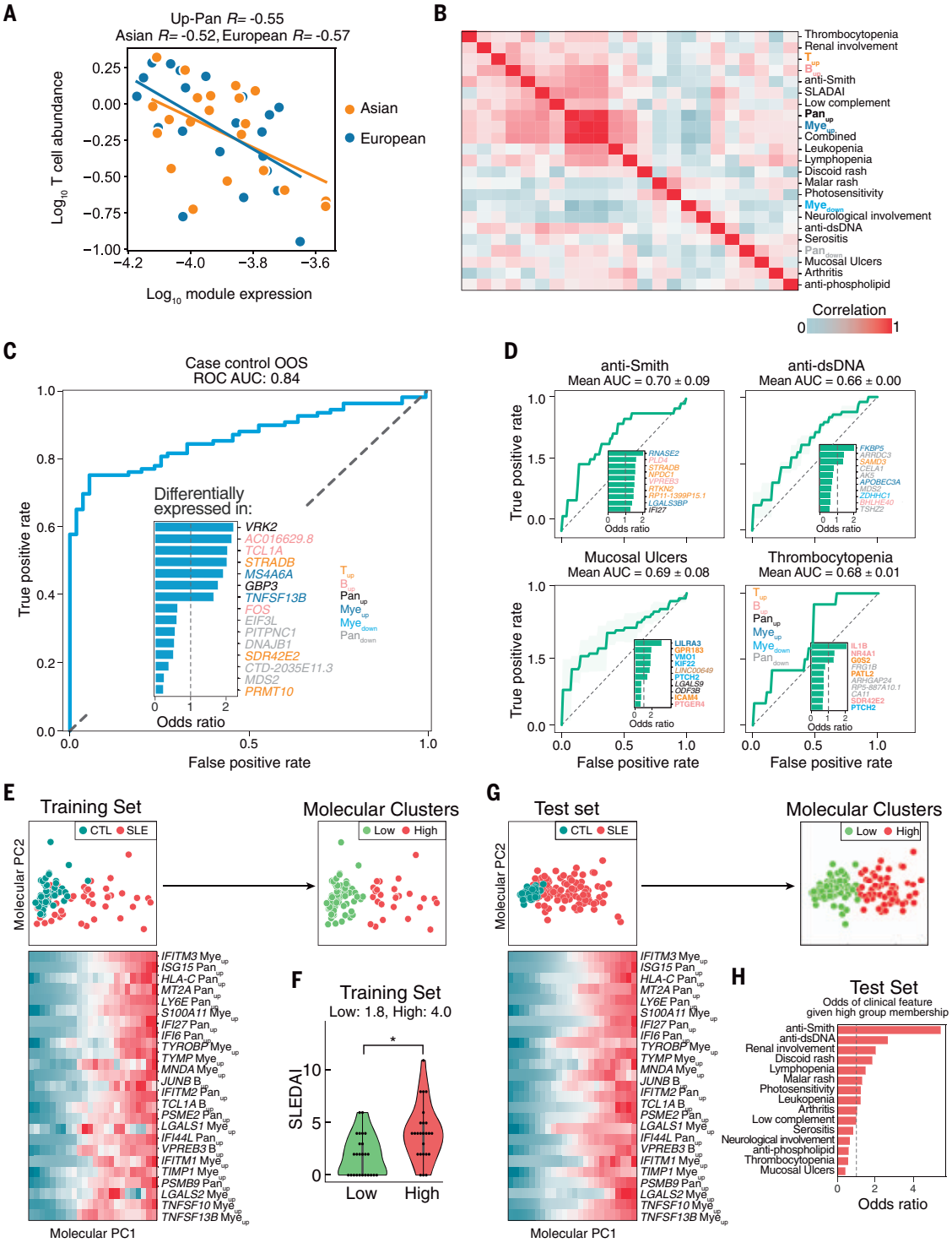


Fig. 4. Prediction of disease status and molecular stratification of SLE. (**A**) Correlation between $\log_{10}$(expression of Pan_{up}) (*x* axis) and $\log_{10}$(abundance of $CD4_{Naïve}$ cells) in processing batch 4 cases only. (**B**) Correlation matrix between average expression of each of six gene modules and clinical features. (**C** and **D**) Receiver operating curve for out-of-sample (OOS) prediction of case-control status (C) and individual clinical variables (D) using a logistic regression model trained on 302 expression features. Inset depicts the most important molecular features inferred by the model, colored by the module to which each feature belongs. (**E**) Principal components analysis of training set based on 302 expression features. Green, control; red, case. Heatmap shows the top 25 most correlated expression features to molecular principal component PC1. Expression was binned and averaged across 24 equal steps across molecular PC1. K-means clustering of samples based on principal components yielded two molecular subphenotypes (Low, High). (**F**) Distribution of SLEDAI scores (*y* axis) for each molecular subphenotype (*x* axis) in the training cases. *$P < 0.05$ (Wilcoxon rank-sum test). (**G**) Projection of OOS test set onto molecular PC1 and PC2 and colored by case-control status (left) and molecular cluster membership (right). Heatmap shows the top 25 most correlated expression features to molecular PC1 in the test set. (**H**) Odds ratio of having a clinical feature given membership in the High molecular cluster versus the Low molecular cluster.

These results show that cell type–specific expression profiles obtained using mux-seq can be used to link cell-intrinsic states with changes in composition, predict case-control status, and molecularly stratify patients with SLE.

Identification of cell type–specific cis-eQTLs across eight immune cell types

We next integrated mux-seq data with genotyping data to map cell type– and cell context–specific cis–expression quantitative trait loci (cis-eQTLs) that may mediate SLE disease associations. Across the eight most abundant cell types, linear regression followed by meta-analysis (*28*, *29*) of three cohorts (92 CLUES Europeans, 98 CLUES Asians, 46 ImmVar Europeans) identified 3331 genes with at least one cis-eQTL in a cell type [false discovery rate (FDR) < 0.05], which we termed cell type–by–cell type cis-eQTLs (CBC-eQTLs) (table S6). Analysis of the genetic architecture of gene expression (*30*) resulted in estimates of average cis heritability ranging from 0.03 to 0.09 per cell type and average cis genetic correlations (rG) ranging from 0.25 to 0.75 for pairs of cell types. Because cells were simultaneously processed, we also estimated shared residual effects (rE) between cell types (e.g., shared environmental and trans-genetic effects) ranging from 0.03 to 0.12. Clustering of rG and rE reflected known lineages between circulating immune cell types (Fig. 5A).

The rG and rE estimates suggest that pleiotropic genetic and shared residual effects are common across immune cell types, which may confound the ability to detect cell type–specific signals among CBC-eQTLs. To account for pleiotropy, we decomposed per–cell type expression profiles into a shared component across all cell types and eight cell type–specific components, then mapped cis-eQTLs associated with each component (*31*). We identified 535 genes with at least one cell type–specific cis-eQTL (cs-eQTL) (FDR < 0.05) and 1207 shared cis-eQTLs (sh-eQTLs) (Fig. 5B and table S7). The effect sizes of CBC-, sh-, and cs-eQTLs were correlated between individuals of European and Asian ancestries (fig. S6, A and B), which separated by genotype principal components (fig. S6C). Relative to CBC-eQTLs, cs-eQTLs for each cell type were significantly and specifically enriched for regions of chromatin accessibility in the same or closely related cell types (*32*), which suggests that decomposition analysis is more likely to identify cis-eQTLs overlapping cell type–specific cis-regulatory elements (Fig. 5C).

Identification and annotation of cell type–specific SLE-associated loci

We next integrated GWAS summary statistics from nine immune-mediated and seven non–immune-mediated traits/diseases to identify cell types where cs-eQTLs harbored the most GWAS associations. Linkage disequilibrium (LD) score regression (*33*) revealed enrichment of disease heritability for relevant cell types across autoimmune diseases (Fig. 5D). The highest enrichment for SLE variants was in cMs and B cells, consistent with our finding that cMs are the highest expressers of type 1 ISGs and with previous work demonstrating that activated B cells produce autoantibodies and secrete cytokines related to disease pathogenesis (*34*, *35*) (Fig. 5D).

We next performed Bayesian genetic colocalization analyses using sh- and cs-eQTLs to fine-map 43 loci associated with SLE (*4*, *36*). Among the five loci colocalized with sh-eQTLs [posterior probability (PP) > 0.6] was the *UBE2L3* locus. Previously identified *UBE2L3* cis-eQTLs in lymphoblastoid cell lines, B cells, and monocytes were replicated by colocalization analysis using CBC-eQTLs (B, cM, ncM; PP > 50%). However, analysis using sh- and cs-eQTLs predicted colocalization of the SLE association and an *UBEL2L3* sh-eQTL (PP = 88.5%), which suggests that this association is shared across cell types (fig. S6D).

Among the seven SLE-associated loci colocalizing with cs-eQTLs was 17q21, a locus associated with asthma (*37*), Crohn's disease (*38*), and type 1 diabetes (*39*). This locus has been difficult to dissect as it encompasses three genes, *IKZF3*, *GSDMB*, and *ORMDL3*, implicated in lymphocyte development (*40*), pyroptosis (*41*), and inflammation (*42*). *ORMDL3*, a regulator of sphingolipid biosynthesis, is linked to the autophagy pathway associated with multiple autoimmune diseases (*43*) and is implicated in the development and differentiation of lymphocytes in SLE pathogenesis (*44*). *ORMDL3* was ubiquitously expressed across cell types with the highest expression in lymphoid populations (Fig. 5, E and F). Colocalization was predicted between SLE associations and both *ORMDL3* sh-eQTLs (PP > 88%) and cs-eQTLs in Bs, CD8s, and pDCs (PP > 96.1%, 92.0%, and 92.1%, respectively) (Fig. 5G). Although *GSDMB* and *IKZF3* were also expressed in most cell types (Fig. 5F), neither gene had a cs-eQTL and the highest posterior probability of colocalization was observed between SLE associations and *GSDMB* sh-eQTLs at 63.8%. Further, conditional analysis (*45*) confirmed that the SLE associations observed near *IKZF3* (Fig. 5G) were independent of the *GSDMB* and *ORMDL3* associations, and that the conditioned SLE associations still colocalized with the *ORMDL3* cs- and sh-eQTLs. The minor allele (T) of rs7216389, a tagging variant in the locus associated with asthma and SLE ($P < 6.09 \times 10^{-7}$) (*4*), conferred an increase of *GSDMB* and *ORMDL3* expression across all cell types, but an additional increase of *ORMDL3* expression in CD8s and Bs that suggested cell type–specific genetic effects in these cell types was not observed for *GSDMB* (Fig. 5G). These results are consistent with previous observations in CD8s and Bs where SNPs in high LD with rs7216389 impacted regulatory elements affecting *ORMDL3* expression (*46*).

We further used expression decomposition to perform a modified transcriptome-wide association study (TWAS) using CONTENT (*47*). Across SLE, Crohn's disease, and rheumatoid arthritis, joint modeling of shared and cell type–specific gene expression identified 93 genes associated with SLE (73 novel), more than twice the number identified by CBC approaches (Fig. 5H). Results were significantly enriched for known SLE associations where 51% of candidate genes, defined as the most proximal gene to each SLE association (*6*), were replicated in the TWAS with $P < 0.05$ ($P_{Enrichment} < 1.2 \times 10^{-24}$). Both the joint and CBC analyses enabled by mux-seq significantly outperformed a standard TWAS using pseudobulk PBMC transcriptomic profiles. These analyses highlight the advantage of leveraging cell type–specific cis-eQTLs to annotate GWAS associations, detangle GWAS signals in gene-dense loci, and power TWAS analysis to identify novel associations.

Modification of genetic effects on gene expression by interferon activation

We next assessed whether variable type 1 interferon activation observed in patients with SLE could modify genetic effects on gene expression in vivo, consistent with our previous in vitro work (*11*, *48*). In SLE cases, we identified 35 genes with a cis-eQTL interacting with the Pan_{up} ISG signature, a proxy for type 1 interferon activation, which we call IFN-eQTL (FDR < 0.1). IFN-eQTL effect size estimates correlated between samples of Asian and European ancestries (fig. S7). Previous interferon response cis-eQTLs (reQTLs) identified in monocyte-derived dendritic cells in vitro (*48*) were significant in cMs but not in other cell types (Fig. 6A).

Among the IFN-eQTLs, we replicated rs11080327 (A>G) as an IFN-eQTL for *SLFN5* in myeloid (cM, $P < 2.5 \times 10^{-10}$; ncM, $P < 0.001$) and B cells ($P < 5.8 \times 10^{-6}$) but not in NK or T cells (Fig. 6B). These results are consistent with the identification of rs11080327 as a cis-eQTL in lymphoblastoid cell lines (*49*) and as a cis-reQTL in monocyte-derived dendritic cells stimulated with rIFNB1 (*11*). We then performed multiplexed single-cell ATAC-seq of PBMCs from five healthy donors either unstimulated or stimulated with rIFNB1. In most cell types, we observed less accessibility in genomic regions near rs11080327 at baseline and a genotype-dependent increase of accessibility after stimulation (Fig. 6C). This was most pronounced in cMs, where the strongest IFN-eQTL was observed. These results are consistent with luciferase reporter assays

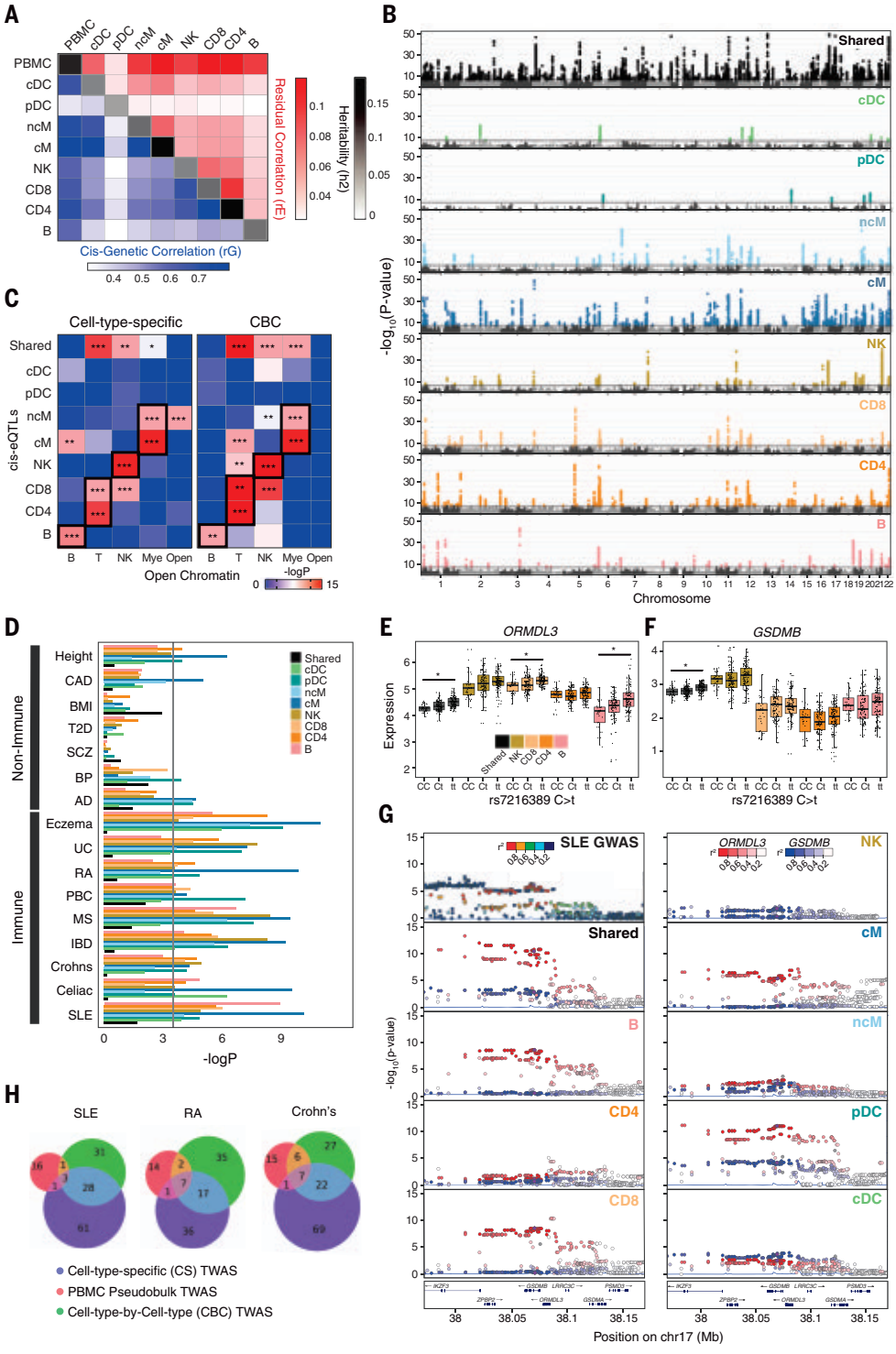


Fig. 5. Cell type–specific genetic determinants of gene expression. (**A**) Cis-genetic correlation (rG; lower triangular plot), shared residual correlation (rE; upper triangular plot), and heritability (h2; diagonal) of eight cell types and PBMCs. Cis is defined 100 kb within the transcription start site. (**B**) Manhattan plots of shared eQTLs (sh-eQTLs; black) and cell type–specific cis-eQTLs (cs-eQTLs; colored) determined by mapping cis-eQTLs associated with shared and cell type–specific expression components from decomposition analysis. Associations are reported as $-\log_{10}(P \text{ value})$ (y axis) ordered by chromosomes (x axis). (**C**) Enrichment of cs-eQTLs (left) and cell type–by–cell type eQTLs (CBC-eQTLs; right) for disjoint sets of cell type–specific regions of open chromatin. $^{*}P < 0.01$, $^{**}P < 0.001$, $^{***}P < 0.0001$ (Mann-Whitney test). (**D**) Enrichment of shared or cs-eQTLs among GWAS associations for seven non–immune-mediated (CAD, coronary artery disease; BMI, body mass index; T2D, type 2 diabetes; SCZ, schizophrenia; BP, bipolar disease; AD, Alzheimer's disease) and nine immune-mediated diseases or traits (UC, ulcerative colitis; RA, rheumatoid arthritis; PBC, primary biliary cirrhosis; MS, multiple sclerosis; IBD, inflammatory bowel disease; SLE, systemic lupus erythematosus). The Bonferroni corrected significance threshold is shown as a black line. (**E** and **F**) Boxplots of decomposed shared and cell type–specific expression of *ORMDL3* (E) and *GSDMB* (F) in all individuals grouped by genotype for rs7216389. *COLOC posterior probability > 0.7. (**G**) LocusZoom plots of SLE GWAS, sh-eQTLs, and cs-eQTLs associated with *ORMDL3* (red) and *GSDMB* (blue) expression. (**H**) Number of associations identified by a modified transcriptome-wide association analysis (TWAS) using decomposed shared and cell type–specific expression matrices (blue), CBC expression matrices (green), or pseudobulk PBMCs (red).

that reveal the region overlapping rs11080327 to be harboring a cis-regulatory element that is activated in response to type 1 interferon (*11*). Overall, our findings illustrate that variability in cell activation in vivo could modify genetic effects on gene expression, which in turn suggests that genetic differences may not only predispose individuals to SLE but may also affect an individual's response to a disease state.

Discussion

SLE remains a challenging autoimmune disease to diagnose and treat. The paucity of targeted therapies, in conjunction with the heterogeneity of disease manifestations and treatment response, highlight the need for improved molecular characterization. In a large ancestrally diverse cohort, we demonstrated the use of mux-seq as a systematic approach to characterize changes in cell type composition and cell type–specific gene expression in adult SLE. We further showed how integration of population genetics with single-cell RNA sequencing could be used to annotate genetic variants with cell type–specific effects on gene expression associated with SLE and other autoimmune diseases.

Using mux-seq, we linked compositional changes to variation in immune cell transcriptional

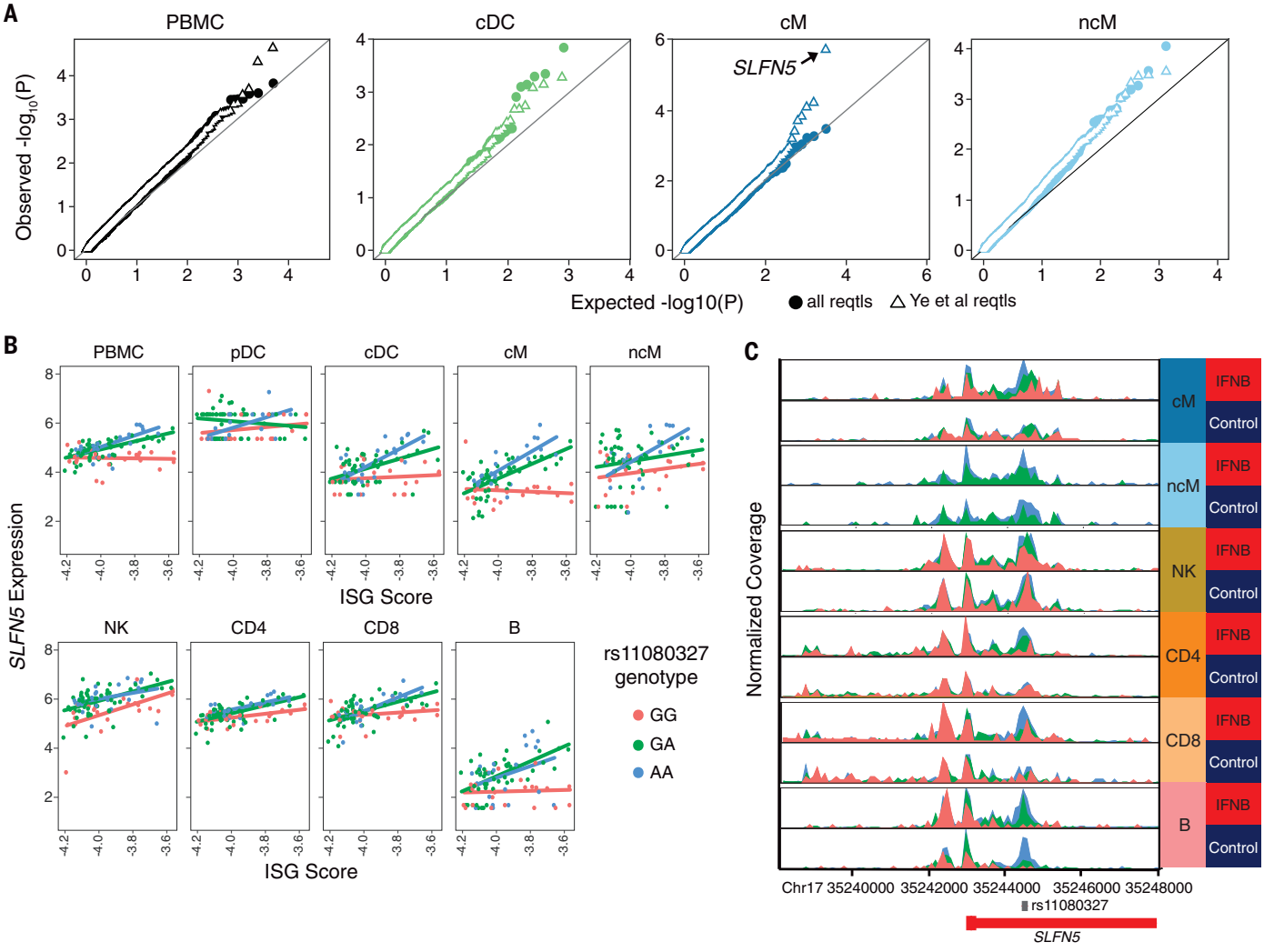


Fig. 6. Interferon modifies cell type–specific genetic effects on gene expression. (**A**) Quantile-quantile plot of expected $-\log_{10}(P$ value) (x axis) versus observed $-\log_{10}(P$ value) (y axis) of cis-IFN-QTLs (solid circles). Previously identified (*48*) response-QTLs (reQTLs) from monocyte-derived dendritic cells are highlighted (open triangles). (**B**) Normalized expression of *SLFN5* expression (y axis) versus ISG score (x axis) separated by rs11080327 genotype (color). Line indicates best linear regression fit for each genotype. (**C**) Gene locus plot of *SLFN5* scATAC-seq peaks for six peripheral immune cell types in unstimulated and rIFNB1-stimulated conditions, separated by genotype. Location of rs11080327 is indicated.

states in SLE. Compositionally, the decrease of naïve $CD4^+$ T cells in cases, particularly those of Asian ancestry, appears to explain the known lymphopenia observed in patients with SLE and importantly was not associated with immunosuppressant treatment, consistent with reports suggesting that mycophenolate mofetil, hydroxychloroquine, and steroids have either no or transient effects on the composition of white blood cells (*50*). Transcriptionally, cMs and ncMs produced the most prominent type 1 ISG signature, including genes specific to myeloid cells, consistent with observations in pediatric SLE (*6*). This finding justifies further investigation into the heterogeneity of type 1 interferon response across leukocyte subsets, particularly in SLE patients being treated with antagonists against the type 1 interferon receptors that have shown mixed results in clinical trials (*51*). Although both cDCs and pDCs also express ISGs, their scarcity in circulation limited their contribution to the overall ISG signature. We did not detect *IFNB1* or *IFNA* transcripts in pDCs or other myeloid cell types; thus, the source of type 1 interferons in SLE remains elusive and is likely not among circulating immune cells (*52*). The inverse correlation between naïve $CD4^+$ T cell abundance and monocyte ISG expression suggests the following model of the pleiotropic effects of type 1 interferons in vivo: ISGs are produced through the interferon signaling cascade and T cells are sequestered at sites of inflammation through the regulation of *CD69* and *S1PR1* (*27*). Whereas age was inversely correlated with the ISG signature, consistent with previous reports, naïve CD4 T cell abundance was not correlated with age and remains inversely correlated to the ISG signature after adjusting for age (*53*). Thus, age is likely not a primary factor for causing SLE, consistent with healthy female first-degree relatives showing a similar inverse correlation between age and serum IFN-α (*7*). Matched profiling of cells from disease-damaged tissue and blood in cases could further shed light on the source of type 1 interferons and confirm the role of lymphocyte trafficking in SLE.

A striking observation from our data is the expansion of $GZMH^+$ but not $GZMK^+$ cytotoxic $CD8^+$ T cells in SLE, in some cases consisting of ~50% of all lymphocytes. Two cytotoxic $CD8^+$ T cell populations were also observed in pediatric SLE (*6*), but the frequency of $GZMH^+$ $CD8^+$ T cells was not reported to be significantly increased despite elevated expression of GZMB and PRF1, which may originate from both $GZMH^+$

$CD8^+$ T and NK_{dim} cells. Although *GZMB* and *PRF1* have been described as markers for $CD8^+$ T cell subsets enriched in SLE (*54*), *GZMH* was higher expressed, more ubiquitous, and more differentially expressed between cases and controls. The function of granzyme-H is not well characterized, but previous work demonstrated its divergent roles in initiating caspase-dependent apoptosis in T cells while initiating caspase-independent apoptosis in NK cells (*55*, *56*). The significant clonal expansion of $GZMH^+$ $CD8^+$ T cells, specifically within the cytotoxic subpopulation, suggests a pathogenic role for these cells in SLE and is consistent with independent work (*54*). One model for the initiation and exacerbation of SLE suggested by these results is an adaptive immune response initiated by foreign and autoantigens followed by chronic exposure to antigens in damaged tissue, resulting in "epitope spreading," where new autoantigens are introduced to the immune system and become future targets of the autoimmune response (*57*). Analysis of immune repertoires of both B and T cells and matching analysis of their antigenic specificity of SLE patients longitudinally would be instructive for deciphering the role of cell-mediated immunity in pathogenesis.

Integrating measurements of cellular composition and cell type–specific expression with genotyping provided an opportunity to assess the genetic determinants of cell type– and cell context–specific gene expression and ascribe functionality to SLE-associated variants. In the presence of pleiotropic effects, mux-seq enabled the decomposition of gene expression into shared and cell type–specific components and mapping of cis-eQTLs associated with these components. Enrichment analyses of orthogonal functional genomic datasets supported the annotation of cell type–specific cis-eQTLs. Integrated analysis of GWAS data and cell type–specific cis-eQTLs provided insight into immune cell types that mediate disease associations; for individual loci, it enabled the fine-mapping and annotation of disease-associated variants. Using decomposed expression components also significantly improved our ability to identify novel disease-associated genes using TWAS compared to using pseudobulk expression profiles over PBMCs or individual cell types. Finally, using quantitative measures of interferon activation from mux-seq, we identified cis-eQTLs whose effects on gene expression could be modified by elevated interferon levels, a critical disease environment in SLE. These results highlight the importance of cellular context for the interpretation of genetic variants associated with disease risk and perhaps disease heterogeneity.

Mux-seq is a cost-effective and systematic approach for enabling cellular phenotyping of large population cohorts. Genetic analysis of cohorts across populations is important for understanding the differences in SLE risk between ancestries and the involvement of environmental triggers. Longitudinal profiling of SLE cases, particularly patients in remission or active flare, could reveal new insights into the initiation of disease, variation in disease activity, new homeostatic states in patients, and response to treatment. Although we examined and controlled for treatment-associated differences in cellular composition and cell type–specific expression between SLE and healthy controls, we did observe notable effects of treatment including the depletion of NK cells in patients treated with azathioprine. Because mux-seq leverages natural genetic variation as sample barcodes, it is compatible with multimodal single-cell profiling of chromatin state and cell surface protein abundance. The integration of richer epigenetic and cellular phenotypes along with improvements to current transcriptomic workflows will undoubtedly improve molecular subphenotyping of SLE, the power to detect cell type–specific and cell context–specific molecular QTLs, and the resolution for annotating SLE associations.

Methods summary

Detailed materials and methods can be found in the supplementary materials. Briefly, we collected PBMCs from SLE cases in the California Lupus Epidemiological Study (CLUES) cohort, matching healthy controls from the UCSF Rheumatology Clinic, and additional controls from the Immune Variation Project (ImmVar). Presence of clinical features important to SLE were recorded.

Antibody-stained or unstained PBMCs were pooled and profiled using 10x Genomics' Chromium Single Cell 3′ V2 chemistry and processed using the 10x Cell Ranger pipeline. Freemuxlet was used to assign cells to their donor of origin and, along with Scrublet (*13*), remove doublets. Platelets, megakaryocytes, and red blood cells were removed using gene markers. Technical variation was removed using COMBAT and regressing out nUMIs, and mitochondrial percent. Standard approaches in Scanpy version 1.6 were used to filter cells, perform dimensionality reduction, cluster using Louvain, and project cells using UMAP (*58*). Cell types were annotated using canonical marker genes and confirmed in cells with antibody staining.

For each cell type, percentage is calculated as the number of cells divided by the total number of cells assigned to the sample. Differences in percentages were compared using weighted least squares. UCSF electronic health record queries compared individuals with multiple heathy encounters and cases with a M32.* ICD-10 code. Mendelian randomization was performed using the GSMR package version 1.91.5beta on UK Biobank cell count QTLs and a separate SLE study (*4*). To examine changes in expression, pseudobulk expression profiles were computed for each cell type and individual using EdgeR. EdgeR was used to perform differential expression analysis (*59*). $CD8_{GZMH}$ signature scores were calculated using Scanpy score_genes on canonical markers. Module scores per individual were calculated by the mean pseudobulk expression for genes in each module. Coexpression analysis was performed on the top 300 DE genes, and clustered by Spearman correlation. Expression modules were recovered by hierarchical clustering of DE genes, revealing six modules. ToppGene was used to find enrichment of modules in pathways (*60*). Molecular clusters were defined using PCA. RNA velocity was performed on cM using the scVelo package. Sklearn's Logistic Regression function was used for all prediction models.

TCR sequencing was performed by amplifying TRA and TRB CDR3 sequences from cDNA and processed with the Cell Ranger pipeline. Only cells with paired TRA and TRB were used. TCRs were analyzed with the singleTCR package. Expanded clonotypes, defined as a TCR sequence detected in at least two cells, were identified using normalized Shannon's entropy.

Samples collected at UCSF were genotyped using the Affymetrix World LAT array. ImmVar samples were genotyped on the OmniExpressExome54 chip. Data were processed using Axiom Best Practices or by previously published methods for the ImmVar cohort. Samples were evaluated for call rate, missingness, and heterozygosity, then imputed using the Michigan Imputation Server with the Haplotype Reference Consortium version 1.1 reference set. Only SNPs with Rsq > 0.3 and minor allele frequency > 10% were retained. Heritability was calculated with the GCTA package's Bivariate GREML function. Cis-eQTLs were mapped ±100 kb of each gene using the MatrixEQTL package accounting for genotype PCs, expression PCs, age, sex, SLE status, and batch as covariates in the linear model. Cell type–specific eQTLs were mapped using the fastGxC method (*31*). CLUES Asian, CLUES European, and ImmVar samples were analyzed separately, then meta-analyzed using the METASOFT package. Empirical *P* values and FDRs were calculated with the qvalue package. LocusZoom was used to visualize loci. SLE cases were analyzed for reQTLs with MatrixEQTL using the ISG score as an interaction term and accounting for genotype PCs, age, sex, and batch.

ATAC-seq enrichment was calculated using a Mann-Whitney test and previously published ATAC-seq peaks from sorted cell types. GWAS enrichment was calculated using LDscore regression (*33*). TWAS analyses were performed using CONTENT (*47*). Colocalization analyses were performed with COLOC (*36*).

The 10x Chromium scATAC-seq kit was used to process PBMCs from five healthy individuals incubated for 8 hours with IFNB or culture media alone. Sequencing data were processed with CellRanger and demultiplexed with Freemuxlet. The ArchR package and Scanpy were used for downstream processing (*61*).

REFERENCES AND NOTES

1. E. E. Carter, S. G. Barr, A. E. Clarke, The global burden of SLE: Prevalence, health disparities and socioeconomic impact. *Nat. Rev. Rheumatol.* **12**, 605–620 (2016). doi: 10.1038/nrrheum.2016.137; pmid: 27558659
2. A. Kaul *et al*., Systemic lupus erythematosus. *Nat. Rev. Dis. Primers* **2**, 16039 (2016). doi: 10.1038/nrdp.2016.39; pmid: 27306639
3. R. Banchereau *et al*., Personalized Immunomonitoring Uncovers Molecular Networks that Stratify Lupus Patients. *Cell* **165**, 551–565 (2016). doi: 10.1016/j.cell.2016.03.008; pmid: 27040498
4. J. Bentham *et al*., Genetic association analyses implicate aberrant regulation of innate and adaptive immunity genes in the pathogenesis of systemic lupus erythematosus. *Nat. Genet.* **47**, 1457–1464 (2015). doi: 10.1038/ng.3434; pmid: 26502338
5. J. Banchereau, V. Pascual, Type I interferon in systemic lupus erythematosus and other autoimmune diseases. *Immunity* **25**, 383–392 (2006). doi: 10.1016/j.immuni.2006.08.010; pmid: 16979570
6. D. Nehar-Belaid *et al*., Mapping systemic lupus erythematosus heterogeneity at the single-cell level. *Nat. Immunol.* **21**, 1094–1106 (2020). doi: 10.1038/s41590-020-0743-0; pmid: 32747814
7. S. Sharma *et al*., Widely divergent transcriptional patterns between SLE patients of different ancestral backgrounds in sorted immune cell populations. *J. Autoimmun.* **60**, 51–58 (2015). doi: 10.1016/j.jaut.2015.04.002; pmid: 25921064
8. H. M. Kang *et al*., Multiplexed droplet single-cell RNA-sequencing using natural genetic variation. *Nat. Biotechnol.* **36**, 89–94 (2018). doi: 10.1038/nbt.4042; pmid: 29227470
9. C. M. Lanata *et al*., Genetic contributions to lupus nephritis in a multi-ethnic cohort of systemic lupus erythematous patients. *PLOS ONE* **13**, e0199003 (2018). doi: 10.1371/journal.pone.0199003; pmid: 29953444
10. T. Raj *et al*., Polarization of the effects of autoimmune and neurodegenerative risk alleles in leukocytes. *Science* **344**, 519–523 (2014). doi: 10.1126/science.1249547; pmid: 24786080
11. M. N. Lee *et al*., Common genetic variants modulate pathogen-sensing responses in human dendritic cells. *Science* **343**, 1246980 (2014). doi: 10.1126/science.1246980; pmid: 24604203
12. C. J. Ye *et al*., Intersection of population variation and autoimmunity genetics in human T cell activation. *Science* **345**, 1254665 (2014). doi: 10.1126/science.1254665; pmid: 25214635
13. S. L. Wolock, R. Lopez, A. M. Klein, Scrublet: Computational identification of cell Doublets in Single-cell transcriptomic data. *Cell Syst.* **8**, 281–291.e9 (2019). doi: 10.1016/j.cels.2018.11.005; pmid: 30954476
14. V. A. Traag, L. Waltman, N. J. van Eck, From Louvain to Leiden: Guaranteeing well-connected communities. *Sci. Rep.* **9**, 5233 (2019). doi: 10.1038/s41598-019-41695-z; pmid: 30914743
15. L. McInnes, J. Healy, J. Melville, UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction. arXiv 1802.03426 (2018).
16. Z. Zhu *et al*., Causal associations between risk factors and common diseases inferred from GWAS summary data. *Nat. Commun.* **9**, 224 (2018). doi: 10.1038/s41467-017-02317-2; pmid: 29335400
17. C. Bycroft *et al*., The UK Biobank resource with deep phenotyping and genomic data. *Nature* **562**, 203–209 (2018). doi: 10.1038/s41586-018-0579-z; pmid: 30305743
18. V. R. Moulton *et al*., Pathogenesis of human systemic lupus erythematosus: A cellular perspective. *Trends Mol. Med.* **23**, 615–635 (2017). doi: 10.1016/j.molmed.2017.05.006; pmid: 28623084
19. K. Rubtsova, A. V. Rubtsov, M. P. Cancro, P. Marrack, Age-Associated B Cells: A T-bet-Dependent Effector with Roles in Protective and Pathogenic Immunity. *J. Immunol.* **195**, 1933–1937 (2015). doi: 10.4049/jimmunol.1501209; pmid: 26297793
20. A. Ferraro *et al*., Interindividual variation in human T regulatory cells. *Proc. Natl. Acad. Sci. U.S.A.* **111**, E1111–E1120 (2014). doi: 10.1073/pnas.1401343111; pmid: 24610777
21. Y. Kotliarov *et al*., Broad immune activation underlies shared set point signatures for vaccine responsiveness in healthy individuals and disease activity in patients with lupus. *Nat. Med.* **26**, 618–629 (2020). doi: 10.1038/s41591-020-0769-8; pmid: 32094927
22. H. Shigematsu *et al*., Plasmacytoid dendritic cells activate lymphoid-specific genetic programs irrespective of their cellular origin. *Immunity* **21**, 43–53 (2004). doi: 10.1016/j.immuni.2004.06.011; pmid: 15345219
23. A.-C. Villani *et al*., Single-cell RNA-seq reveals new types of human blood dendritic cells, monocytes, and progenitors. *Science* **356**, eaah4573 (2017). doi: 10.1126/science.aah4573; pmid: 28428369
24. V. Bergen, M. Lange, S. Peidli, F. A. Wolf, F. J. Theis, Generalizing RNA velocity to transient cell states through dynamical modeling. *Nat. Biotechnol.* **38**, 1408–1414 (2020). doi: 10.1038/s41587-020-0591-3; pmid: 32747759
25. G. La Manno *et al*., RNA velocity of single cells. *Nature* **560**, 494–498 (2018). doi: 10.1038/s41586-018-0414-6; pmid: 30089906
26. J. P. Buyon *et al*., The effect of combined estrogen and progesterone hormone replacement therapy on disease activity in systemic lupus erythematosus: A randomized trial. *Ann. Intern. Med.* **142**, 953–962 (2005). doi: 10.7326/0003-4819-142-12_Part_1-200506210-00004; pmid: 15968009
27. L. R. Shiow *et al*., CD69 acts downstream of interferon-α/β to inhibit S1P1 and lymphocyte egress from lymphoid organs. *Nature* **440**, 540–544 (2006). doi: 10.1038/nature04606; pmid: 16525420
28. B. Han, E. Eskin, Random-effects model aimed at discovering associations in meta-analysis of genome-wide association studies. *Am. J. Hum. Genet.* **88**, 586–598 (2011). doi: 10.1016/j.ajhg.2011.04.014; pmid: 21565292
29. A. A. Shabalin, Matrix eQTL: Ultra fast eQTL analysis via large matrix operations. *Bioinformatics* **28**, 1353–1358 (2012). doi: 10.1093/bioinformatics/bts163; pmid: 22492648
30. J. Yang, S. H. Lee, M. E. Goddard, P. M. Visscher, GCTA: A tool for genome-wide complex trait analysis. *Am. J. Hum. Genet.* **88**, 76–82 (2011). doi: 10.1016/j.ajhg.2010.11.011; pmid: 21167468
31. A. Lu *et al*., Fast and powerful statistical method for context-specific QTL mapping in multi-context genomic studies. bioRxiv 448889 (2021). doi: 10.1101/2021.06.17.448889
32. D. Calderon *et al*., Landscape of stimulation-responsive chromatin across diverse human immune cells. *Nat. Genet.* **51**, 1494–1505 (2019). doi: 10.1038/s41588-019-0505-9; pmid: 31570894
33. H. K. Finucane *et al*., Heritability enrichment of specifically expressed genes identifies disease-relevant tissues and cell types. *Nat. Genet.* **50**, 621–629 (2018). doi: 10.1038/s41588-018-0081-4; pmid: 29632380
34. E. Nashi, Y. Wang, B. Diamond, The role of B cells in lupus pathogenesis. *Int. J. Biochem. Cell Biol.* **42**, 543–550 (2010). doi: 10.1016/j.biocel.2009.10.011; pmid: 19850148
35. X. Hu *et al*., Integrating autoimmune risk loci with gene-expression data identifies specific pathogenic immune cell subsets. *Am. J. Hum. Genet.* **89**, 496–506 (2011). doi: 10.1016/j.ajhg.2011.09.002; pmid: 21963258
36. C. Giambartolomei *et al*., 0 CommonMind Consortium, A Bayesian framework for multiple trait colocalization from summary association statistics. *Bioinformatics* **34**, 2538–2545 (2018). doi: 10.1093/bioinformatics/bty147; pmid: 29579179
37. M. F. Moffatt *et al*., Genetic variants regulating ORMDL3 expression contribute to the risk of childhood asthma. *Nature* **448**, 470–473 (2007). doi: 10.1038/nature06014; pmid: 17611496
38. L. Jostins *et al*., Host-microbe interactions have shaped the genetic architecture of inflammatory bowel disease. *Nature* **491**, 119–124 (2012). doi: 10.1038/nature11582; pmid: 23128233
39. J. C. Barrett *et al*., Genome-wide association study and meta-analysis find that over 40 loci affect risk of type 1 diabetes. *Nat. Genet.* **41**, 703–707 (2009). doi: 10.1038/ng.381; pmid: 19430480
40. B. Morgan *et al*., Aiolos, a lymphoid restricted transcription factor that interacts with Ikaros to regulate lymphocyte differentiation. *EMBO J.* **16**, 2004–2013 (1997). doi: 10.1093/emboj/16.8.2004; pmid: 9155026
41. L. Li, Y. Li, Y. Bai, Role of GSDMB in Pyroptosis and Cancer. *Cancer Manag. Res.* **12**, 3033–3043 (2020). doi: 10.2147/CMAR.S246948; pmid: 32431546
42. Y. Zhang *et al*., The ORMDL3 Asthma Gene Regulates ICAM1 and Has Multiple Effects on Cellular Inflammation. *Am. J. Respir. Crit. Care Med.* **199**, 478–488 (2019). doi: 10.1164/rccm.201803-0438OC; pmid: 30339462
43. B. James, S. Milstien, S. Spiegel, ORMDL3 and allergic asthma: From physiology to pathology. *J. Allergy Clin. Immunol.* **144**, 634–640 (2019). doi: 10.1016/j.jaci.2019.07.023; pmid: 31376405
44. J. Dang *et al*., ORMDL3 facilitates the survival of splenic B cells via an ATF6α-endoplasmic reticulum stress-Beclin1 autophagy regulatory pathway. *J. Immunol.* **199**, 1647–1659 (2017). doi: 10.4049/jimmunol.1602124; pmid: 28747345
45. J. Yang *et al*., Conditional and joint multiple-SNP analysis of GWAS summary statistics identifies additional variants influencing complex traits. *Nat. Genet.* **44**, 369–375 (2012). doi: 10.1038/ng.2213; pmid: 22426310
46. B. J. Schmiedel *et al*., 17q21 asthma-risk variants switch CTCF binding and regulate IL-2 production by T cells. *Nat. Commun.* **7**, 13426 (2016). doi: 10.1038/ncomms13426; pmid: 27848966
47. M. Thompson *et al*., Multi-context genetic modeling of transcriptional regulation resolves novel disease loci. bioRxiv 461579 (2021). doi: 10.1101/2021.09.23.461579
48. C. J. Ye *et al*., Genetic analysis of isoform usage in the human anti-viral response reveals influenza-specific regulation of *ERAP2* transcripts under balancing selection. *Genome Res.* **28**, 1812–1825 (2018). doi: 10.1101/gr.240390.118; pmid: 30446528
49. J. F. Degner *et al*., DNase I sensitivity QTLs are a major determinant of human expression variation. *Nature* **482**, 390–394 (2012). doi: 10.1038/nature10808; pmid: 22307276
50. S. J. Rivero, E. Díaz-Jouanen, D. Alarcón-Segovia, Lymphopenia in systemic lupus erythematosus: Clinical, diagnostic, and prognostic significance. *Arthritis Rheum.* **21**, 295–305 (1978). doi: 10.1002/art.1780210302; pmid: 646828
51. E. F. Morand *et al*., Trial of anifrolumab in active systemic lupus erythematosus. *N. Engl. J. Med.* **382**, 211–221 (2020). doi: 10.1056/NEJMoa1912196; pmid: 31851795
52. M. Cella *et al*., Plasmacytoid monocytes migrate to inflamed lymph nodes and produce large amounts of type I interferon. *Nat. Med.* **5**, 919–923 (1999). doi: 10.1038/11360; pmid: 10426316
53. T. B. Niewold *et al*., Age- and sex-related patterns of serum interferon-α activity in lupus families. *Arthritis Rheum.* **58**, 2113–2119 (2008). doi: 10.1002/art.23619; pmid: 18576315
54. P. Blanco *et al*., Increase in activated $CD8^+$ T lymphocytes expressing perforin and granzyme B correlates with disease activity in patients with systemic lupus erythematosus. *Arthritis Rheum.* **52**, 201–211 (2005). doi: 10.1002/art.20745; pmid: 15641052
55. L. Casciola-Rosen, F. Andrade, D. Ulanet, W. B. Wong, A. Rosen, Cleavage by granzyme B is strongly predictive of autoantigen status: Implications for initiation of autoimmunity. *J. Exp. Med.* **190**, 815–826 (1999). doi: 10.1084/jem.190.6.815; pmid: 10499920
56. M. Faroudi *et al*., Lytic versus stimulatory synapse in cytotoxic T lymphocyte/target cell interaction: Manifestation of a dual activation threshold. *Proc. Natl. Acad. Sci. U.S.A.* **100**, 14145–14150 (2003). doi: 10.1073/pnas.2334336100; pmid: 14610278
57. C. L. Vanderlugt, S. D. Miller, Epitope spreading in immune-mediated diseases: Implications for immunotherapy. *Nat. Rev. Immunol.* **2**, 85–95 (2002). doi: 10.1038/nri724; pmid: 11910899
58. F. A. Wolf, P. Angerer, F. J. Theis, SCANPY: Large-scale single-cell gene expression data analysis. *Genome Biol.* **19**, 15 (2018). doi: 10.1186/s13059-017-1382-0; pmid: 29409532
59. M. D. Robinson, D. J. McCarthy, G. K. Smyth, edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2010). doi: 10.1093/bioinformatics/btp616; pmid: 19910308
60. J. Chen, E. E. Bardes, B. J. Aronow, A. G. Jegga, ToppGene Suite for gene list enrichment analysis and candidate gene prioritization. *Nucleic Acids Res.* **37**, W305–W311 (2009). doi: 10.1093/nar/gkp427; pmid: 19465376
61. J. M. Granja *et al*., Author Correction: ArchR is a scalable software package for integrative single-cell chromatin accessibility analysis. *Nat. Genet.* **53**, 935 (2021). doi: 10.1038/s41588-021-00850-x; pmid: 33790476

62. R. K. Perez *et al*., Multiplexed scRNA-seq reveals the cellular and genetic correlates of systemic lupus erythematosus Analysis Code (2021), doi: 10.5281/zenodo.4724043.

ACKNOWLEDGMENTS

We thank all members of the Ye lab for discussions. **Funding:** Supported by NIH grant P30AR070155 (C.J.Y., L.A.C., and J.Y.); NIH grants R01AR071522, U01HG012192, R21AI133337, and CZI P0535277 (C.J.Y.); CDC grant U01DP005120 (L.A.C. and J.Y.); the Lupus Research Alliance (L.A.C.); NIH grants K25HL121295, U01HG009080, R01HG006399, R01CA227237, and R03DE025665 and DoD grant W81XWH-16-2-0018 (N.Z.); NIH grants R01CA194511 and R01CA223484 (L.F.); the Manton Foundation, Klarman Cell Observatory, and Howard Hughes Medical Institute (A.R.); NIH grant 1F31HG011007 (M.G.G.); NSF grant GRFP 1650113 (G.C.H.); and NIH grant T32HG002536 (M.T.). **Author contributions:** M.Su., G.C.H., S.T., Y.S., and L.M. performed all experiments. R.K.P., M.G.G., M.Su., and C.J.Y. wrote the manuscript. R.K.P., M.G.G., and C.J.Y. revised the manuscript. R.K.P. and C.J.Y. performed all preprocessing, cell type annotations, single-cell analysis, pseudobulk, DE analysis, and clinical predictions. M.C.K. performed the trajectory analysis and RNA velocity. N.R. performed the UCSF EHR database queries. M.G.G. performed the Mendelian randomization analysis and eQTL analysis. S.S.K. and T.L. performed the TCR sequencing experiments. R.K.P. performed the TCR analysis. M.M., M.C., L.T., C.L., M.D., J.Y., and L.A.C. provided access to CLUES samples and all patient information. M.G.G., B.B., A.L., M.T., and N.Z. developed and implemented the decomposition and modified TWAS methods. A.D. aided heritability and subtype analyses. M.Sl., J.W., D.D., and O.R. performed the Immvar sequencing experiments. M.G.G., A.O., and R.B. organized and disseminated data into appropriate repositories. M.D., L.F., A.R., J.Y., L.A.C., and N.Z. provided critical edits and feedback to the manuscript. **Competing interests:** A.R. is a co-founder and equity holder of Celsius Therapeutics, an equity holder in Immunitas, and was an SAB member of ThermoFisher Scientific, Syros Pharmaceuticals, Neogene Therapeutics, and Asimov until 31 July 2020. From 1 August 2020, A.R. is an employee of Genentech. O.R.R. is a co-inventor on patent applications filed at the Broad related to single-cell genomics. O.R.R. has given numerous lectures about single-cell genomics to a wide variety of audiences and, in some cases, has received remuneration to cover time and costs. C.J.Y. is a SAB member for and holds equity in Related Sciences and ImmunAI, is a consultant for and holds equity in Maze Therapeutics, and is a consultant for Trex Bio. C.J.Y. has received research support from Chan Zuckerberg Initiative, Chan Zuckerberg Biohub, and Genentech. **Data and materials availability:** All data are available in the Human Cell Atlas Data Coordination Platform and at GEO accession number GSE137029. Genotypes are available at dbGap accession number phs002812.v1.p1. Code is available at 10.5281/zenodo.4724043 (*62*).

SUPPLEMENTARY MATERIALS

16 October 2020; accepted 8 November 2021
10.1126/science.abf1970

RESEARCH ARTICLE SUMMARY

IMMUNOGENOMICS

Single-cell eQTL mapping identifies cell type–specific genetic control of autoimmune disease

Seyhan Yazar†, Jose Alquicira-Hernandez†, Kristof Wing†, Anne Senabouth, M. Grace Gordon, Stacey Andersen, Qinyi Lu, Antonia Rowson, Thomas R. P. Taylor, Linda Clarke, Katia Maccora, Christine Chen, Anthony L. Cook, Chun Jimmie Ye, Kirsten A. Fairfax, Alex W. Hewitt*†, Joseph E. Powell*†

INTRODUCTION: The human immune system has evolved to maintain tissue homeostasis and target exogenous pathogens by regulating specialized cell populations. It displays substantial variation between individuals, defining how people vary in susceptibility to disease and respond to pathogens or cancer.

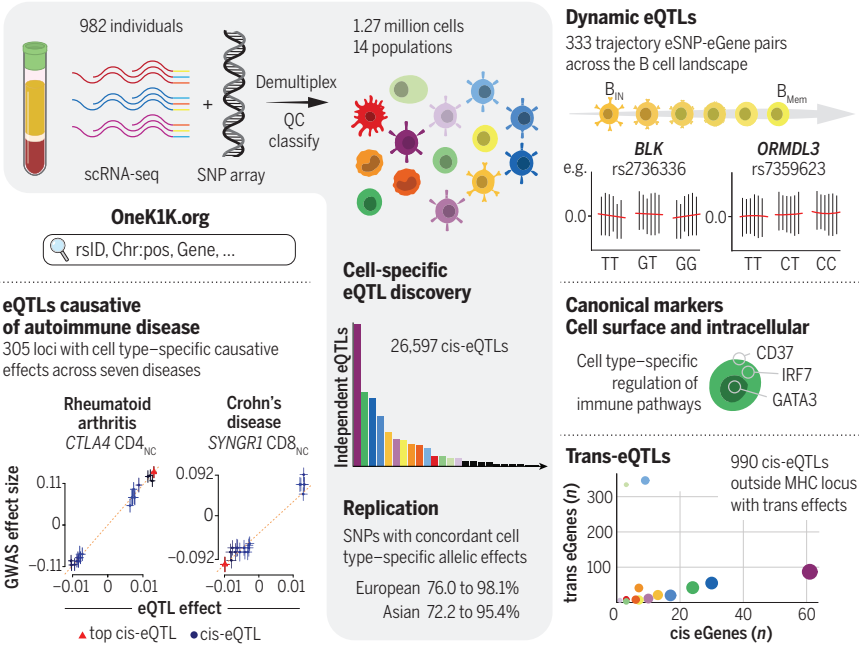
RATIONALE: Our knowledge of how genetic differences contribute to immune variation at the cellular level has been limited by two main challenges in the generation of data at single-cell resolution. One of these challenges is to sequence from many individuals and the other is to sequence a large number of cells from each individual. Addressing these challenges is necessary to dissect the genetic and molecular underpinnings of common, heterogeneous diseases.

RESULTS: We present the OneK1K cohort, which consists of single-cell RNA sequencing (scRNA-seq) data from 1.27 million peripheral blood mononuclear cells (PMBCs) collected from 982 donors. We developed a framework for the classification of individual cells, and by combining the scRNA-seq data with genotype data, we mapped the genetic effects on gene expression in each of 14 immune cell types and identified 26,597 independent cis–expression quantitative trait loci (eQTLs). We show that most of these have an allelic effect on gene expression that is cell type–specific. Our results replicated in two independent cohorts, one of which comprises individuals with a different ancestry to our discovery cohort. Over all loci, our discovery and replication cohorts have a concordance of allelic direction ranging from 72.2 to 98.1% across cell types.

Using the top associated eQTL single-nucleotide polymorphism (eSNP) at each locus outside the major histocompatibility complex (MHC) region, we identified 990 trans-acting effects, most (63.6%) of which were cell type–specific. We show how eQTLs have dynamic allelic effects in B cells that are transitioning from naïve to memory states. Overall, we identified a set of 1988 eSNP–eGene (a gene with an eQTL) pairs expressed across the B cell maturation landscape, of which 333 have a statistically significant change in their allelic effect as B cells differentiate. Of these, 66% were only identified from the dynamic eQTL analysis and were not observed when testing for effects independently in cell types, highlighting the importance of investigating cell state–specific effects that underlie immune cell function. We investigated how eQTLs affect the expression variation of essential immune genes in specific cell types and provided experimental support for established hypotheses of cellular mechanisms in complex autoimmune diseases.

Finally, we integrated genetic association data for seven common autoimmune diseases and identified significant enrichment of genetic effects operating in a cell type–specific manner. Through colocalization of single-cell eQTL and genome-wide association study (GWAS) loci, we found that 19% of cis-eQTLs share the same causal locus as a GWAS risk association. Using a Mendelian randomization approach, we uncovered the causal route by which 305 loci contribute to autoimmune disease through changes in gene expression in specific cell types and subsets. Of the shared causal loci, 38.4% are outside the MHC region and exhibit highly cell-specific effects. Highlighting multiple sclerosis, we identified the causal route underlying 57 risk loci. For example, we show that the loci at 3q12 causally acts through changes in *EAF2* expression, but only in immature and naïve B (B_{IN}) and memory B (B_{Mem}) cells, despite this gene being ubiquitously expressed in all cell types in our data.

CONCLUSION: This work brings together population genetics and scRNA-seq data to uncover drivers of interindividual variation in the immune system. Our results demonstrate how segregating genetic variation influences the expression of genes that encode proteins involved in critical immune regulatory and signaling pathways in a cell type–specific manner. Understanding the genetic underpinnings of immune system regulation will have broad implications in the treatment of autoimmune diseases and infections, transplantation, and cancers. ■



Single-cell eQTL mapping and colocalization with autoimmune disease risk loci. scRNA-seq data from 1.27 million PBMCs were used to identify 26,597 cis-eQTLs (gray box). Dynamic eQTLs were uncovered as cells move from a naïve to a memory state (top right). Genetic variation between individuals influences immune regulation in a cell type–specific manner (middle right). In this study, 990 trans-eQTL effects (bottom right) and the causal effects for 305 autoimmune disease loci were identified (bottom left). Browsable results are available at www.onek1k.org. $CD4_{NC}$, CD4 naïve and central memory T cells; $CD8_{NC}$, CD8 naïve and central memory T cells; QC, quality control.

The list of author affiliations is available in the full article online.
*Corresponding author. Email: hewitt.alex@gmail.com (A.W.H.); j.powell@garvan.org.au (J.E.P.)
†These authors contributed equally to this work.
Cite this article as S. Yazar *et al.*, *Science* **376**, eabf3041 (2022). DOI: 10.1126/science.abf3041

READ THE FULL ARTICLE AT
https://doi.org/10.1126/science.abf3041

RESEARCH ARTICLE

IMMUNOGENOMICS

Single-cell eQTL mapping identifies cell type–specific genetic control of autoimmune disease

Seyhan Yazar[1]†, Jose Alquicira-Hernandez[1,2]†, Kristof Wing[3,4]†, Anne Senabouth[1], M. Grace Gordon[5,6,7], Stacey Andersen[2], Qinyi Lu[3], Antonia Rowson[3,8], Thomas R. P. Taylor[3], Linda Clarke[9], Katia Maccora[3,8], Christine Chen[8], Anthony L. Cook[10], Chun Jimmie Ye[5,6,7,11,12,13], Kirsten A. Fairfax[3], Alex W. Hewitt[3,4,9]*†, Joseph E. Powell[1,14]*†

The human immune system displays substantial variation between individuals, leading to differences in susceptibility to autoimmune disease. We present single-cell RNA sequencing (scRNA-seq) data from 1,267,758 peripheral blood mononuclear cells from 982 healthy human subjects. For 14 cell types, we identified 26,597 independent cis-expression quantitative trait loci (eQTLs) and 990 trans-eQTLs, with most showing cell type–specific effects on gene expression. We subsequently show how eQTLs have dynamic allelic effects in B cells that are transitioning from naïve to memory states and demonstrate how commonly segregating alleles lead to interindividual variation in immune function. Finally, using a Mendelian randomization approach, we identify the causal route by which 305 risk loci contribute to autoimmune disease at the cellular level. This work brings together genetic epidemiology with scRNA-seq to uncover drivers of interindividual variation in the immune system.

The expression of genes in immune cells is highly variable between individuals (*1*–*6*), with this variation being both a cause and a consequence of differences in susceptibility to immune-related diseases (*7*, *8*). Investigations into the underlying genetic contribution to immune regulation and disease development have uncovered many associated variants (*9*). Yet the complexity of circulating immune populations has made their mechanisms of action difficult to dissect.

Coupling transcriptional profiles with genetic variation allows the direct identification of genomic regulators of gene expression. This is important because disease-associated genetic risk variants identified through genome-wide association studies (GWASs), including those linked to common immune-mediated diseases, are often mapped to regulatory regions of the genome (*10*–*13*). Both empirical results and theoretical models provide evidence that most common disease-associated variants act through changes in gene expression rather than directly influencing protein structure or function (*14*). By combining genetic information with bulk RNA sequencing (RNA-seq), the downstream effects of disease-associated genetic risk factors have been linked to expression quantitative trait loci (eQTLs). Efforts such as GTEx (*15*), eQTL-Gen (*16*), CAGE (*17*), and ImmVar (*18*) have identified eQTLs across a variety of cell types and tissues but have used bulk RNA-seq approaches, where gene expression levels represent the averaged signal over large numbers of cells. The data from these ensemble analyses are valid, but the gene expression heterogeneity between individual cells is still largely unexplored.

An important step is to define the cellular and environmental contexts in which disease-risk single-nucleotide polymorphisms (SNPs) affect gene expression levels. This will help determine the molecular and cellular mechanisms by which disease develops and inform therapeutic strategies. Beyond the ability to annotate individual disease associations, cell type–specific eQTLs are enriched for heritability across complex traits (*4*). This is important because many eQTL effects are tissue specific (*2*, *18*, *19*), and both fluorescence-activated cell sorting (FACS) and computational deconvolution of cell types from bulk samples have evidenced cell type–specific eQTLs (*20*–*22*). Although these studies have helped demonstrate variation in the role of genetic loci in cell subsets, challenges remain. For example, bulk RNA-seq of FACS cell populations is biased toward known cell types that are defined by a limited set of marker genes. It does not capture the heterogeneity within a sorted population. Likewise, computational methods that deconvolute a bulk signal into cell types struggle to identify less abundant cell types and rely on approximations to estimate cell proportions (*23*). By contrast, single-cell RNA sequencing (scRNA-seq) enables the simultaneous, unbiased determination of cellular composition and cell type–specific gene expression, capturing intraindividual cell heterogeneity.

Results

The OneK1K cohort

We characterized the transcriptional variation across circulating immune cells of a large cohort (OneK1K) to explore how allelic variation is associated with changes in gene expression in a cell type–specific manner (Fig. 1A). The OneK1K cohort consists of 982 individuals of Northern European ancestry (Fig. 1B) who reported no active infection at the time of sample collection. We generated genotype data on 759,993 SNPs (figs. S1 to S3) and imputed SNPs against the Haplotype Reference Consortium panel (*24*). After quality control, we retained 5,328,917 SNPs with a minor allele frequency greater than 0.05 (fig. S4). We generated scRNA-seq data on 1,449,385 peripheral blood mononuclear cells (PBMCs) using a pooled multiplexing strategy. After demultiplexing, removal of doublets, and quality control, we retained 1,267,758 cells for further analysis (Fig. 1C).

Classification of individual cells

We developed a framework to independently classify each cell into one of 14 different immune cell types across the myeloid and lymphoid lineages based on their transcriptional profiles. This framework, implemented in *scPred* (*25*), uses a combination of hierarchical supervised and unsupervised classification methods, using FACS-sorted PBMC scRNA-seq data as a reference (*26*) (Fig. 1D and table S1). Cell composition ranged from 0.7% dendritic cells (DCs) to 36.6% $CD4^+$ naïve and central memory T ($CD4_{NC}$) cells (Fig. 1E), with the mean and range of proportions matching those reported elsewhere (*27*, *28*) (Fig. 1F and table S2). Visualization of cell types with uniform manifold approximation and projection (UMAP) reflects the hierarchical relationship among these cell types (Fig. 1G), which is also supported by cell coordinates across the first two principal components (fig. S5). Cells were classified using their complete transcriptional profiles. Still, to aid interpretation against other studies, we show concordance with the expression patterns of canonical markers and other

[1]Garvan-Weizmann Centre for Cellular Genomics, Garvan Institute of Medical Research, Sydney, NSW, Australia. [2]Institute for Molecular Bioscience, University of Queensland, Brisbane, QLD, Australia. [3]Menzies Institute for Medical Research, University of Tasmania, Hobart, TAS, Australia. [4]Department of Ophthalmology, Royal Hobart Hospital, Hobart, TAS, Australia. [5]Division of Rheumatology, Department of Medicine, University of California, San Francisco, San Francisco, CA, USA. [6]Department of Epidemiology and Biostatistics, University of California, San Francisco, San Francisco, CA, USA. [7]Institute for Human Genetics, University of California, San Francisco, San Francisco, CA, USA. [8]Department of Surgery, School of Clinical Science at Monash Health, Monash University, VIC, Australia. [9]Centre for Eye Research Australia, University of Melbourne, East Melbourne, VIC, Australia. [10]Wicking Dementia Research and Education Centre, University of Tasmania, Hobart, TAS, Australia. [11]Institute of Computational Health Sciences, University of California, San Francisco, San Francisco, CA, USA. [12]Parker Institute for Cancer Immunotherapy, San Francisco, CA, USA. [13]Chan Zuckerberg Biohub, San Francisco, CA, USA. [14]UNSW Cellular Genomics Futures Institute, University of New South Wales, Sydney, NSW, Australia.
***Corresponding author. Email: hewitt.alex@gmail.com (A.W.H.); j.powell@garvan.org.au (J.E.P.)**
†These authors contributed equally to this work.

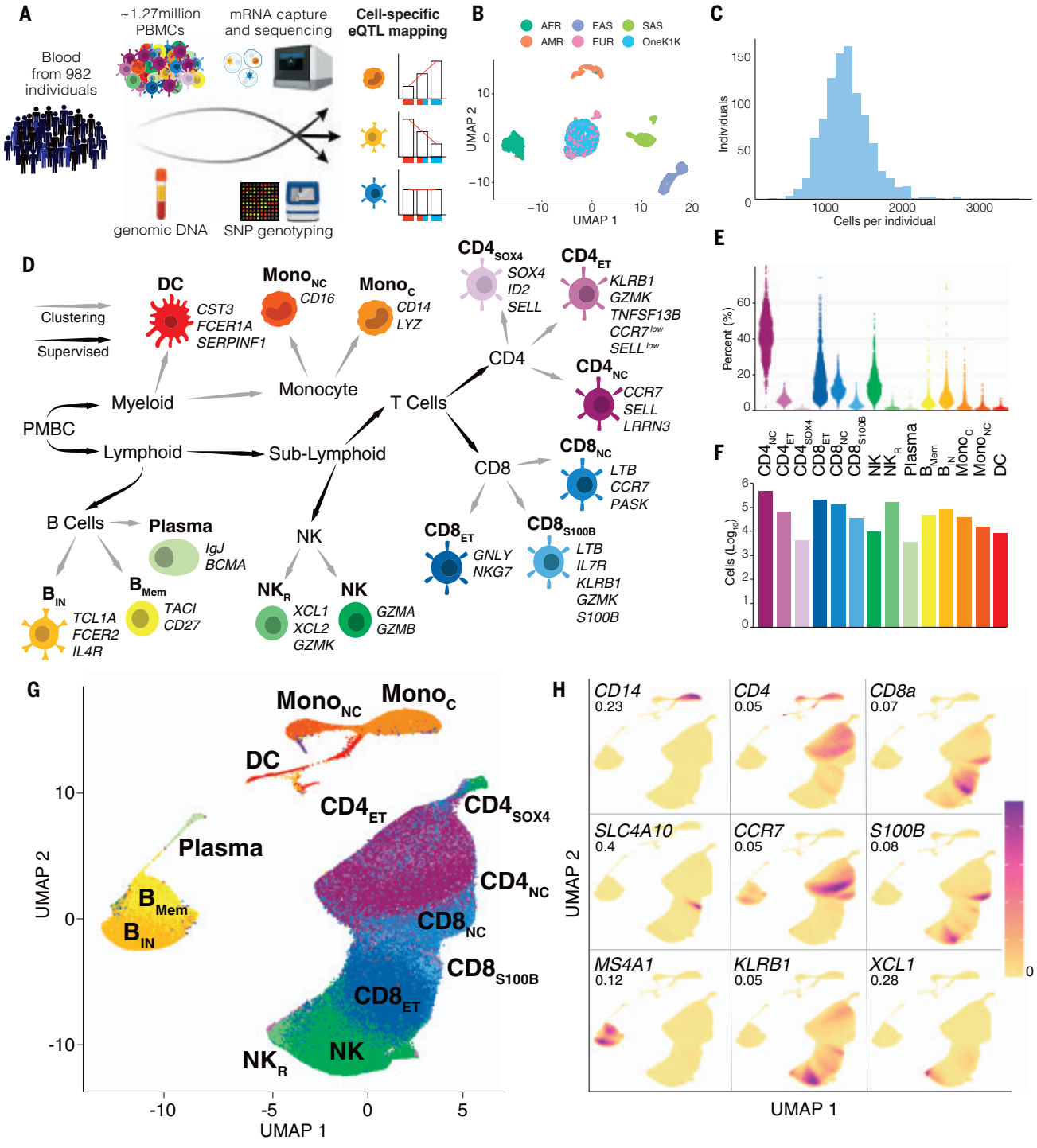


Fig. 1. Population-scale scRNA-seq identifies 14 transcriptionally distinct mononuclear populations in peripheral blood. (**A**) scRNA-seq data from PMBCs were generated using a pooled multiplexing strategy for 982 healthy individuals. Simultaneously, SNPs were genotyped, and data were integrated for single-cell eQTL analysis. (**B**) UMAP analysis shows the genetic relationship between the individuals from the OneK1K cohort and the 1000 Genomes Project (*83*). Individuals from the OneK1K cohort are embedded with individuals with Northern European genetic ancestry. AFR, African; AMR, ad-mixed American; EAS, East Asian; EUR, European; SAS, South Asian. (**C**) Mean of 1291 individual cells per donor, ranging from 62 to 3501 after scRNA-seq, demultiplexing, and quality-control filtering. (**D**) Hierarchical classification of cells. Each cell underwent up to four rounds of supervised clustering based on similarity to each node, as indicated by the black arrows. After that, unsupervised clustering by Seurat (gray arrows) yielded 14 transcriptionally distinct cell types. Classification of each cell was confirmed based on cosine similarity to FACS reference data and further assessed through the interrogation of differentially expressed and prototypical genes. (**E**) Total percentage of each cell type as a proportion of the total sequenced population for each individual. (**F**) The total number of cells per cell type after sequencing, demultiplexing, and quality-control filtering. (**G**) UMAP of 1,267,758 PBMCs across all individuals, with 14 transcriptionally distinct populations. Color coding is the same as in (D). (**H**) Density plots of nine differentially expressed canonical markers of peripheral immune cells, demonstrating robust concordance with canonical markers (see fig S11 for additional markers). Values denote maximal density. The abbreviations for each cell type are displayed in table S1. The color scale is relative, ranging from no density (0) to highest density.

single-cell sequencing studies (*26*, *29*, *30*) (Fig. 1H).

After batch correction, we found no evidence for variation in cell identity, transcriptional signatures, or cell proportions across the capture pools (figs. S6 to S8). Across individuals, we sequenced an average of 1291 cells per donor (Fig. 1C). Although most of the individuals had scRNA-seq data for all 14 cell types, because of sampling variance, some cell types [predominantly $CD4^+$ T cells expressing SOX4 ($CD4_{SOX4}$ cells), plasma cells, and nonclassical monocytes ($Mono_{NC}$)] were not sequenced for some individuals (fig. S9 and table S7). Therefore, for subsequent analyses, the sample size for eQTL analysis varied by cell type, although 12 out of the 14 populations had $n > 930$.

Single-cell eQTL analysis reveals cell-type specificity of transcriptional changes that occur because of common variants

To understand how genetic variation between individuals influences gene expression in a cell type–specific manner, we tested for the association between the genotypes of SNPs within a 1-Mb cis region of either end of a gene including the gene body and the expression of genes in each of the 14 cell types. This approach identifies eQTLs in each cell type, enabling us to assess the degree to which the genetic effects on gene expression are shared across PBMCs. Multiple SNPs within a cis region can be associated with gene expression because of the correlation between genotypes induced by linkage disequilibrium and numerous independent loci associated with the expression levels of the gene. To differentiate between these scenarios, we performed a conditional analysis for each identified eQTL, fitting the lead eQTL SNP(s) [eSNP(s)] as conditional covariates in subsequent rounds of analysis.

In total, we identified 26,597 eQTLs for 39.7% of the genes tested, with 16,597 ($eSNP_1$) in the first round of analysis and a further 10,000 ($eSNP_2$ to $eSNP_5$) from the four rounds of conditional tests (Fig. 2A and tables S9 and S10). The number of independent eQTLs varied between cell types, with 6473 identified in $CD4_{NC}$ cells and 399 in plasma cells (Fig. 2B). This variation in the number of eQTLs determined per cell type is likely a function of statistical power. There is a strong relationship between both cell proportions (Fig. 1E and fig. S17) and the number of individuals with identifiable cells (table S7). The conditional eQTL analysis identified secondary loci influencing expression in 8.1 to 19.2% of genes with an initial eQTL and more than three independent eQTLs for 10.6 to 40.6% of genes (Fig. 2A and table S9).

These conditional eQTLs identify instances where there are multiple independent loci within the cis region whose genotypes are associated with the expression levels of a gene. For example, in $CD4_{NC}$ cells, we identified a primary eQTL for *PADI4*. This gene encodes an enzyme that is responsible for converting arginine residues to citrulline residues (*31*), thereby regulating the activity of histone H1 and consequently the maintenance of stem cells (*32*). *PADI4* has been implicated in the pathogenesis of rheumatoid arthritis (RA) at both a genetic and cellular level (*33*). The top $eSNP_1$ for this eQTL is rs10788663, where each copy of the T allele causes a decrease of an average of 0.28 mRNA transcript molecules per cell (fig. S12). In a subsequent round of conditional analysis, we fitted rs10788663 as a covariate and tested for associations again across the cis region, identifying a secondary independent eQTL marked by the top $eSNP_2$ rs1612843. On average, individuals carrying each copy of the C allele of rs1612843 have a decrease of 0.24 mRNA transcript molecules per cell. rs10788663 is located in the first intron, whereas rs1612843 is located in the intron between exons 15 and 16 of *PADI4*, suggesting that independent transcription factors likely regulate multiple independent sites and are required for the regulation of the expression of *PADI4*. In the OneK1K cohort, the linkage disequilibrium between rs10788663 and rs1612843 is 0.0678, providing further evidence that multiple independent eQTLs influence the expression of *PADI4* in $CD4_{NC}$ cells. Indeed, after confirming the expected additive effect of two independent loci, we observed a mean difference of 1.04 mRNA transcripts per cell for individuals carrying homozygous T/T and C/C compared with C/C and G/G for rs10788663 and rs1612843, respectively (fig. S12). Both rs10788663 and rs1612843 associations were replicated in eQTL-Gen data (*34*).

The allelic effect of genetic loci on gene expression may be distinctive to a particular cell type and absent in other cell types—a relationship we define as "cell type–specific." We explored its prevalence by investigating the deviation of test statistics from a null distribution for cis-eQTLs in other cell types where they did not initially meet study-wide significance (Fig. 2B). The mean proportion of cis-eQTLs identified in one cell type that showed inflation of their test statistics in one other cell type was $\pi_1 = 0.53$ (0.19 to 0.96) (fig. S13). This is evidence that with larger sample sizes, cis-eQTLs currently identified in a single cell type should reach study-wide significance in one or more other cell types. However, the magnitude of their allelic effect is likely to vary between cell types. For 3060 genes with an eQTL (eGenes) identified in only a single cell type, we do not find any evidence for allelic effects in other cell types, suggesting that these are indeed cell type–specific (fig. S14). The observation of cell type–specific eQTLs has multiple possible explanations: The gene may only be detectably expressed in one cell type, there may be low statistical power to detect eQTLs in multiple cell types, or there is true regulatory heterogeneity across cell types.

To evaluate these different scenarios, we performed a series of analyses for each of the genes with at least one eQTL (eGene $n = 6469$). Only 43 (0.7%) of these eGenes are expressed in a single cell type. The remaining 6426 are expressed in multiple cell types, with these genes expressed in an average of 11 cell types, in addition to the one with a significant eQTL (fig S15). Indeed, when we tested for the correlation in the expression levels of each of these 6426 eGenes between a pair of cell types, we identified a high overall concordance in coexpression (Fig. 2C). The pattern of average correlation in eGene expression levels between a couple of cell types followed the hematopoietic lineage relationship. For example, of the 6473 eGenes with an eQTL found only in $CD4_{NC}$ cells, 1392 were expressed in $CD8^+$ naïve and central memory T ($CD8_{NC}$) cells and the mean correlation in gene expression between the cells was 0.97 (Fig. 2C). By contrast, in classical monocytes ($Mono_C$), only 168 of the plasma cell eGenes were expressed, but the mean correlation of expression with plasma cells was 0.79. From these results, we can conclude that most of the eGenes with an eQTL identified in just one cell type are not due to cell type–specific expression of the eGene in most instances but rather may be due to cell type–specific expression of regulatory factors.

Having identified that these eGenes are expressed in multiple cell types, we next sought to evaluate if the observation of cell type–specific eQTLs was due to low statistical power to detect allelic effects in more than one cell type. To assess this hypothesis, we implemented an empirical framework to test the rank of the test statistics for eGene allelic effects across the nonsignificant cell types. In almost all instances, we observed none or minimal enrichment of the test statistic across cell types, suggesting that in most cases, cell type–specific eQTLs are due to specific cell regulatory mechanisms (fig. S15). In instances where we identified a marked enrichment, cell types closely related in the hematopoietic lineage existed. However, for most eGenes, we did not identify an enrichment in the test statistics, again suggesting that effects are cell type–specific. These results collectively demonstrate that most of the eQTLs identified for the 2367 eGenes are specific to just a single cell type.

For the remaining 4102 eGenes, we identified a total of 14,230 eQTLs across two or more cell types, although, for 1386 of these eGenes, we observed different lead eSNPs between cell types (Fig. 2B). Under this scenario, one hypothesis is that the same variant underlies eQTLs in multiple cell types, with differences in top eSNPs being due to variation in gene expression patterns. An alternative hypothesis is

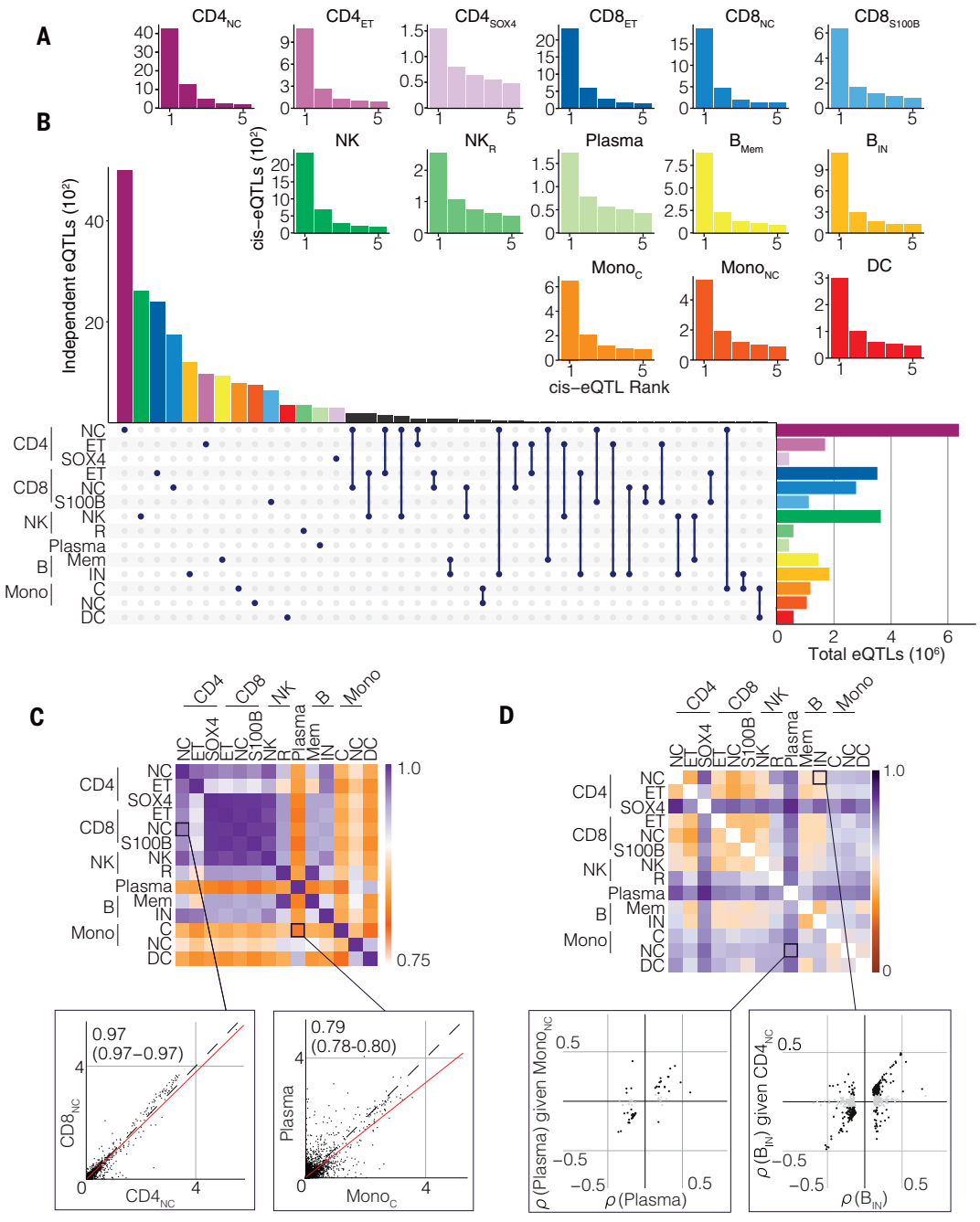


Fig. 2. Conditional eQTL analysis reveals cell-type specificity of transcriptional changes due to common variants. (**A**) Up to five rounds of conditional eQTL analysis (*x* axis) are shown, with the number of cis-eQTLs (*y*-axis value times 10^2) detected at a study-wide FDR less than 0.05 in each round of analysis. (**B**) Using $eSNP_1$ to $eSNP_5$, we identify most of the eQTLs in only a single cell type. eQTLs identified in multiple cell types are connected by lines. The *x* axis is truncated at 30 eQTLs in a given category. The total number of eQTLs detected per cell type is shown on the right. (**C**) For each eQTL identified in a single cell type, we tested whether it was present in other cell types and compared eGene expression levels. Each unit of the heatmap (top) shows the correlation of expression levels of genes associated with an eQTL only in cell type A against cell type B. Cell type–specific eQTL genes remain consistently expressed across other cell types, with examples shown for the expression levels of $CD4_{NC}$ eQTL genes in $CD8_{NC}$ cells (bottom left) and for $Mono_C$ eQTL genes in plasma cells (bottom right), with *x*- and *y*-axis units representing mean UMIs per cell. The scale of the correlation coefficients between pairwise combinations of cells ranges from 0.75 to 1.0. (**D**) To investigate the independence of eQTLs for genes with eQTLs in more than one cell type (but tagged by different eSNPs), we tested for the change in allelic effect in cell type A, after conditioning on the eSNP from cell type B. Significant changes ($p < 0.05$) imply the same eQTL in both cell types (or linkage disequilibrium between eSNPs). A lack of change provides evidence that the gene has independent eQTLs in each cell type. For example, the allelic effects of eSNPs from plasma cells after conditioning on the lead eSNPs from $Mono_{NC}$ cells (bottom left) and of B_{IN} eSNPs after conditioning on the lead eSNPs from $CD4_{NC}$ (bottom right) cells are shown. The heatmap (top) shows the pairwise correlations in allelic effects. ρ represents the original correlation coefficient. The abbreviations for each cell type are displayed in table S1.

that the eQTLs result from independent variants that influence expression in different cell types. To test between these hypotheses, we performed a regression strategy to evaluate the change in the test statistic of an eSNP after regressing out the effects of an eSNP from another cell type. Under this strategy, if the eSNPs tag the same causal variant for that gene or are in linkage disequilibrium with one another, then the allelic effect size of the original eSNP will decrease in the conditional analysis. Similarly, if they tag independent variants, the allelic effect will remain relatively unchanged. We performed this strategy for each pairwise combination of eQTLs where different top eSNPs were identified in different cell types.

We tested whether each eGene was tagged by two distinct variants by conditioning the lead eSNP from the first cell type on the lead eSNP from the second cell type for every pair of cell types (182 pairs). The correlation coefficients of significant independent eSNPs from shared eGenes pre- and postconditioning are shown in Fig. 2D and fig. S16. Whereas most lymphoid immune cell eQTLs had a considerable change in the correlation coefficients after conditioning, among the myeloid immune cells, the eQTL correlation coefficients remained similar (fig. S16). This finding suggests that lymphoid cell types are more likely to share genetic control of gene expression between cell types compared with myeloid cells.

Evidence suggests that cell type–specific chromatin accessibility underlies a proportion of cell type–specific cis-eQTLs

To explore the functional regulation underlying cis-eQTLs, we tested for the overlap of eSNP locations and regions of open chromatin generated from single-cell assay for transposase accessible chromatin sequencing (scATAC-seq) data from 8876 cells. Cells were classified into each of the 14 cell types, and open chromatin peaks were called for each cell type that had more than five classified cells. This filtering retained 11 cell types, comprising the most abundant populations [except $CD4^+$ T cells with an effector memory or central memory phenotype ($CD4_{ET}$), $CD4_{SOX4}$, and plasma cells] (fig. S18). On average, we identified 52,048 peaks per cell type, with the mean distance between an eSNP and the nearest peak ranging from 7485 to 31,383 base pairs. To determine whether the location of cis-eQTLs was significantly closer to open chromatin regions, we compared the distances between the cis-eQTLs. We randomly sampled SNPs that were selected based on the same distance distribution from the transcript to the nearest peaks per cell type using a bootstrapping technique. We observed a significant difference between the cis-eQTL distances across all cell types except $CD4_{SOX4}$ cells [false discovery rate (FDR) < 0.05] (fig. S19). We conclude from these results that cell type–specific chromatin accessibility is likely to contribute to variation in allelic effects on gene expression between cell types.

Single-cell eQTLs replicate in multiethnic cohorts and bulk eQTL studies

To verify cell-specific eQTL findings, we replicated our lead eSNP results in two independent cohorts of European and Asian ancestry, consisting of 113 and 89 individuals, respectively. Of the 16,597 $eSNP_1$-eGene pairs, 10,071 were present with a minor allele frequency greater than 0.05 in both cohorts. Of these, 3198 (26%) in the European cohort and 2243 (22%) in the Asian cohort replicated at the FDR threshold of 5%, which is encouraging given the differences between the sample sizes of these cohorts and the sample size of the OneK1K discovery cohort (tables S5, S12, and S13).

Indeed, correcting the FDR distributions under the assumptions of equal sample size in the discovery and replication cohorts leads to 87 and 78% replication rates in the European and Asian cohorts, respectively. Similarly, the concordance of allelic direction over all tested loci was 76.0 to 98.1% in the European cohort and 72.2 to 95.4% in the Asian cohort. This concordance increases to 99.3 to 100% and 96.9 to 99.8%, respectively, for eQTLs replicating at an FDR less than 0.05 (Fig. 3, A and B). The discrepancy in replication rates between cohorts likely reflects differences in the allele frequencies of eSNPs between population groups. However, the results indicate that cell type–specific eQTLs are likely to be largely shared among populations. The discovery of OneK1K eQTLs was tested for replication in all cell types in the replication cohorts. At an FDR less than 0.05, replicating eQTLs and eGenes are predominantly identified in a single cell type (Fig. 3, C and D), providing further evidence for cell type–specific effects of loci on gene expression in PBMCs. The concordance of correlation coefficients between the OneK1K and replication cohorts are shown in Fig. 3E for both the European and Asian samples. We were able to replicate 62.5 and 40.4% of cis-eQTLs identified in bulk RNA-seq studies of blood samples from the eQTL-Gen Consortium (*34*) and GTEx Consortium (*5*), respectively (fig. S20 and table S14).

Identification of dynamic eQTL allelic effects across the B cell landscape

We investigated the dynamic effects of eQTLs across the pseudotime landscape of immature and naïve B (B_{IN}) cells through to memory B (B_{Mem}) cells. Cells were categorized into six quantiles (Q1 to Q6) based on their relative position on the pseudotime curve (Fig. 4, A and B). Overlaying the expression of classical markers revealed a graded change across the derived trajectory from B_{IN} (Q1) to B_{Mem} cells (Q6). For example, *TCL1A* and *IL4R* are highly expressed in naïve B cells (*35*, *36*) and were found to be down-regulated across the transition to B_{Mem} cells (Fig. 4C). Conversely, the expression of *CD27*, a canonical B_{Mem} cell marker (*37*), increased as the cells transitioned to a memory state. *IgJ* expression, a marker of immunoglobulin M (IgM) and IgA production, was up-regulated in the higher quantiles, indicating that they contain cells poised to become plasma cells (*38*) (Fig. 4C).

We sought to identify instances where eQTL allelic effects exhibited either linear or nonlinear changes across the trajectory of naïve to memory B cell transition. Dynamic B cell eQTLs were determined by testing the interaction between the genotype and quantile ranks using both linear and quadratic models. Of the 3074 cis-eQTLs identified in B_{IN} and B_{Mem} cells, 1988 were expressed in at least three pseudotime quantiles and tested for dynamic effects. Of these, we identified significant changes in the allelic effect across the trajectory for 333 of them (FDR < 0.05) (Fig. 4D and fig. S21).

Many of the genes with dynamic eQTL effects have a role in fine-tuning B cell migration, activation, survival, or function. For example, *SELL* is involved in integrin-mediated migration to and within tissues (*39*, *40*). Migration to and organization of B cells within the germinal center is a critical component in generating appropriate memory and humoral outputs. The allelic effect of the intronic variant rs4987360-G on *SELL* expression is largest in immature cells, decreasing over each of the subsequent quantiles (Fig. 4E). The opposite trend is identified for SNPs that influence the expression of the Src family tyrosine kinase B lymphocyte kinase (*BLK*), a gene responsible for regulating the amplitude of signaling downstream of the B cell receptor. Both rs2736336 and rs2409780 show the greatest allelic effects in Q5 and Q6 (Fig. 4E and table S15). Interestingly, rs2736336, a variant in the promoter of *BLK*, is associated with systemic lupus erythematosus (SLE) (*41*), whereas rs2409780, an intronic variant, is in high linkage disequilibrium with variants associated with SLE and RA [coefficient of determination (R^2) = 0.99, and coefficient of linkage disequilibrium (D') = 0.99] (*13*, *42*). Another gene responsible for interpreting the signaling downstream of B cell surface receptors and influencing subsequent B cell proliferation and survival is c-Rel, encoded by the transcription factor *REL* (*43*). rs12989427 is in high linkage disequilibrium with variants associated with SLE (R^2 = 0.88, and D' = 0.98), and the allelic effect follows a nonlinear relationship, peaking at the medium point of the B cell trajectory (Fig. 4E). *ORMDL3* promotes mature B cell survival by suppressing apoptosis and promoting autophagy (*44*). rs7359623 and rs8067378 are in high linkage disequilibrium with risk variants (R^2 > 0.8, and D' > 0.9) implicated in a range of autoimmune

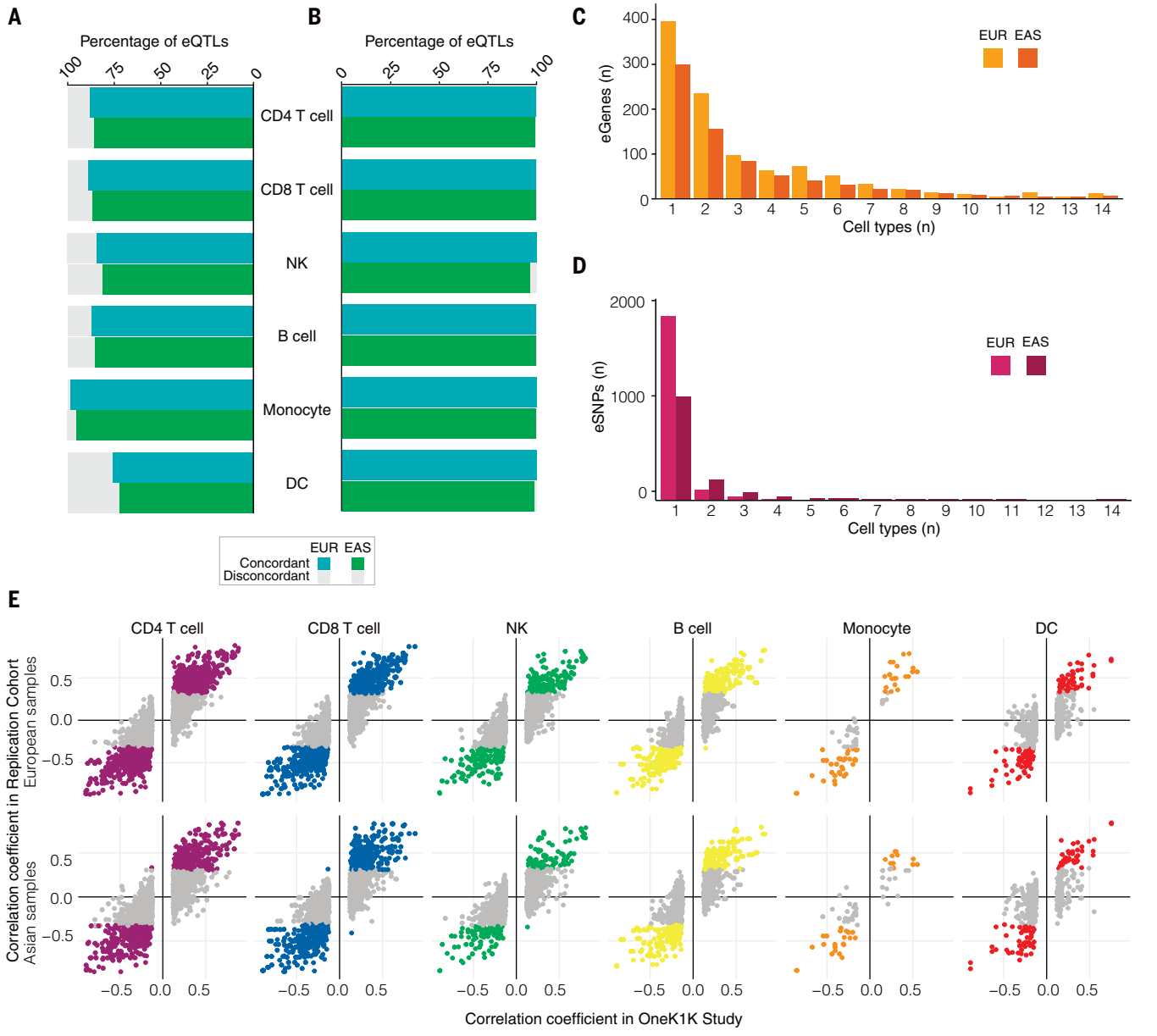


Fig. 3. Replication cohort shows concordance of eQTLs across European and Asian samples. (**A**) Percentage of concordance of the allelic direction of effect for all single-cell eQTLs tested for replication across cell types. Concordances for single-cell RNA of European (EUR) and Asian (EAS) populations are shown. (**B**) Percentage of concordance for eQTLs that replicate at a study-wide FDR significance threshold of 0.05. (**C**) The numbers of cell types in which the eGene is observed for eQTLs that replicate in the European or Asian cohorts. (**D**) The number of cell types in which eSNPs are identified in the replication cohorts. Most replicating eQTLs are cell type–specific. (**E**) The correlation coefficient for eQTLs in the OneK1K and replication cohorts per cell type indicated. The direction of the correlation coefficients denotes the direction of the allelic effect with respect to the reference allele. European samples are shown in the first row, and Asian samples are shown in the second row. Colored dots denote eQTLs that replicate at a study-wide FDR less than 0.05.

diseases (*45*–*48*) and have a dynamic eQTL effect on *ORMDL3* in B cells across the trajectory.

Genetic variation controls transcriptional regulation in a cell type–specific manner to regulate immune pathways

Although it is widely accepted that immune regulation is variable between individuals (*1*), the factors that cause this variation are poorly understood. By selecting genes described in the literature that affect immune regulation, we demonstrate how genetic loci contribute to variation in the expression of immune regulatory genes in a cell type–specific manner (Fig. 5 and table S10).

Leukocyte recirculation between the blood and lymph nodes is an essential property of the immune system. It depends on the lymph node homing receptor CD62L (L-selectin) encoded by the *SELL* gene (*49*). We observed opposing regulation of *SELL* mRNA between the innate and adaptive immune systems under the influence of rs4987360, a common polymorphism in linkage disequilibrium with rs4987353 (R^2 = 1, and D' = 1) that is associated with monocyte blood cell counts (*50*). The rs4987360-G allele decreased *SELL* mRNA in $Mono_C$ but increased *SELL* mRNA in B_{IN} cells (Fig. 5A), illustrating how a single inherited allele can act through different cell types to influence gene expression. The dynamic eQTL analysis identified that the allelic effect of rs4987360 varied across the B cell–state landscape (Fig. 4E).

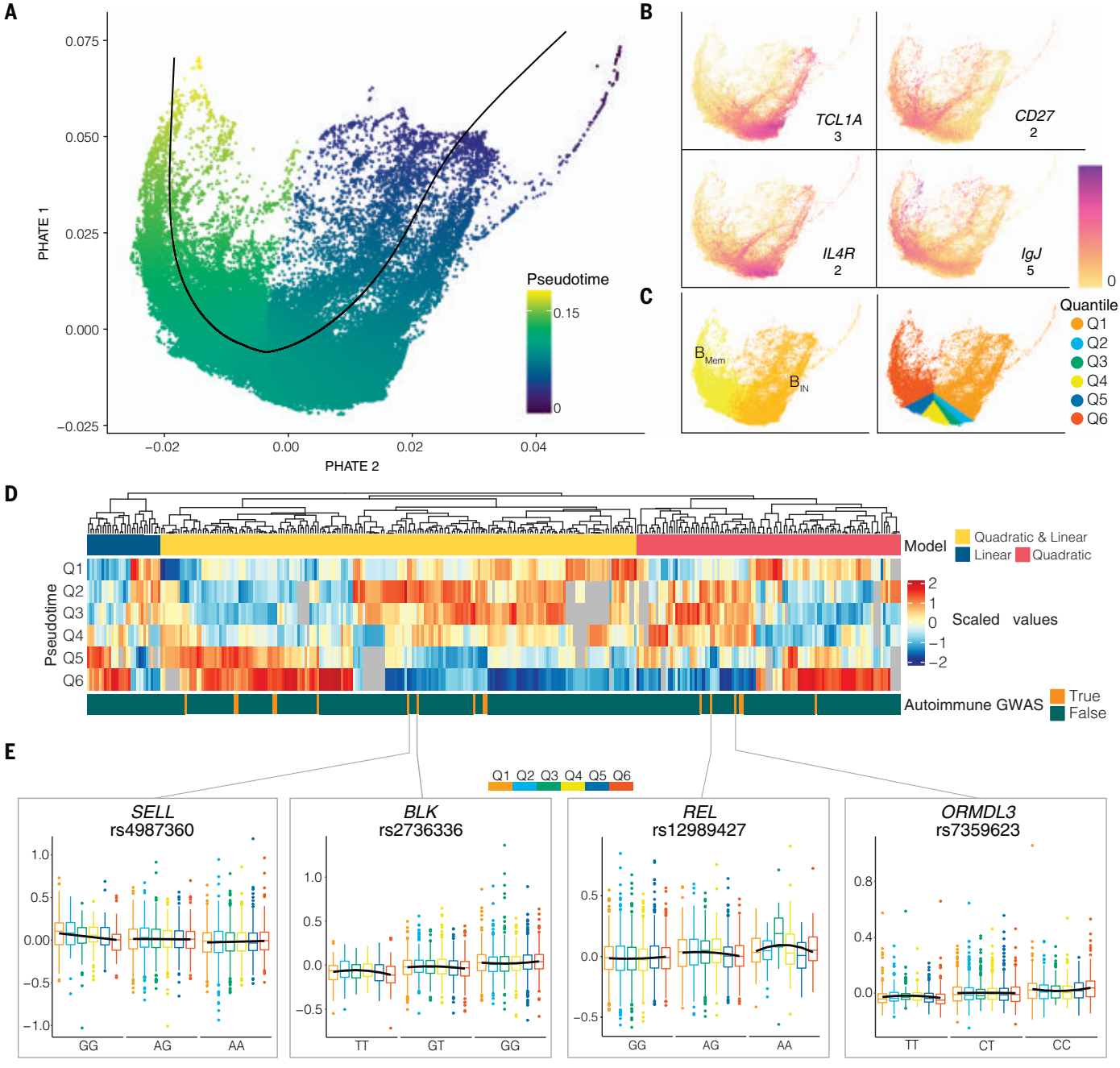


Fig. 4. Dynamic eQTLs across B cell trajectories. (**A**) The pseudotime projection of 124,968 B cells was derived from their progression from immature or naïve to memory cells. The pseudotime curve is represented by the solid black line. The pseudotime is represented with a color scale from 0 (the earliest pseudotime) to 1 (the latest pseudotime). (**B**) Mapping of B_{IN} and B_{Mem} cells and the division of landscape into six quantiles across the pseudotime trajectory. The color scale shows expression, ranging from lowest (0) to highest (maximum for each gene). (**C**) Density plots of canonical markers highlight B cell profile changes from immature or naïve to memory B cells across the derived pseudotime trajectory. (**D**) eSNP-eGene pairs with a statistically significant difference in eQTL effect size across the B cell landscape. Both linear and quadratic models were applied to SNP-gene pairs across the pseudotime quantiles. SNPs known to be in high linkage disequilibrium ($R^2 > 0.8$) with variants identified through GWAS of autoimmune diseases are displayed. Instances where the eGene was not expressed in a given quantile are shown in gray. The entire heat map is available in fig S21. β, estimate of an eQTL effect size. (**E**) Examples of allele-specific changes in expression profiles across cell quantiles in the B cell pseudotime landscape. The scaled β values are shown for each eSNP-eGene pair, with the box plots colored by cell quantile with the same color coding used in (C).

The rs4987360 association replicates in bulk RNA-seq eQTL data from eQTL-Gen (*34*) and GTEx (*5*) and has allelic effects with the opposite direction in bulk B cells and monocytes (given rs2223286 with a $R^2 = 1$ and $D' = 1$) (*51*).

CTLA4 is a gene-dosage sensitive, essential inhibitory receptor on T cells (*52*–*55*). In contrast to the example of *SELL*, the rs3087243-G allele downstream of *CTLA4*, which is associated with susceptibility to type 1 diabetes mellitus (T1DM) and RA (*47*, *56*–*58*), acts in multiple cell types in the same allelic direction by decreasing *CTLA4* mRNA expression in four T cell subsets (Fig. 5B). The polymorphism rs231770 is located less than 10 kb away from

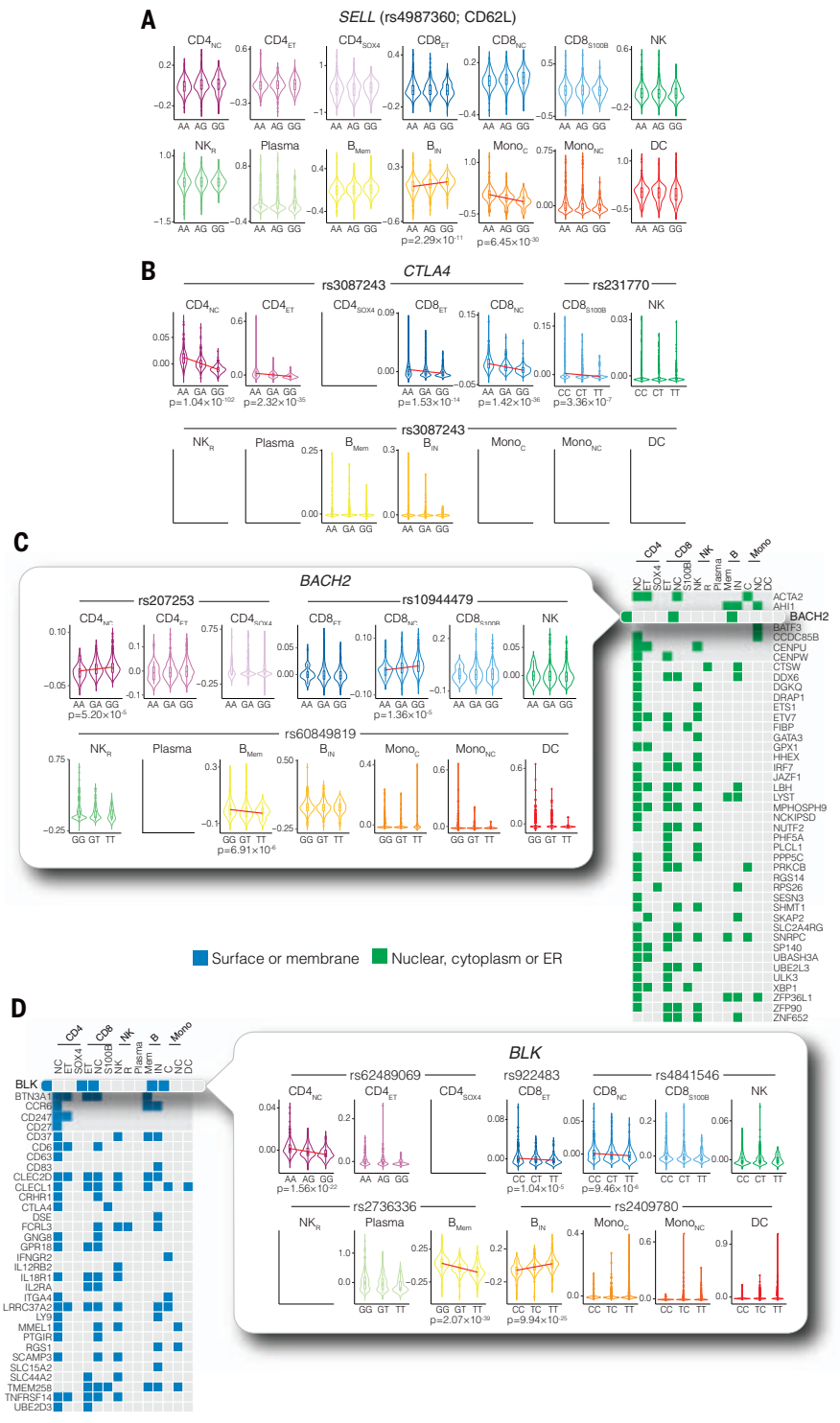


Fig. 5. Genetic variation leads to cell type–specific immune regulation. Cell type–specific eQTLs for genes known to play a role in immune function and common autoimmune diseases were identified. (**A**) The eQTL for *SELL* (CD62L) exhibits different allelic directions between the lymphoid and myeloid lineages; the effects for the eSNP rs4987360 are shown. (**B**) Allelic plots for the inhibitory receptor *CTLA4*. (**C**) With regard to nuclear, cytoplasm, or ER genes, we highlight *BACH2*, which showed significant and independent eQTLs in three cell types tagged by eSNPs rs207253, rs10944479, and rs60849819. (**D**) Cell type–specific eQTLs for cell surface receptors or membrane-associated proteins implicated in common autoimmune diseases. Focusing on *BLK*, we identified cell type–specific eQTLs and expression patterns across cell types. *p* values are from Spearman's rank correlation testing. Red lines indicate the allelic effect of significant eQTLs identified at an FDR less than 0.05.

rs3087243 but is in linkage equilibrium (R^2 = 0.5). rs231770-T is similarly associated with decreased *CTLA4* mRNA expression in $CD8^+$ T cells with expression of S100B ($CD8_{S100B}$) T cells and is associated with the autoimmune condition myasthenia gravis (*59*).

By linking allelic effects to changes in the expression of genes known to be implicated in autoimmune disease, we can support established hypotheses and identify previously uncharacterized examples of cellular mechanisms that underlie conditions and control immune regulation. By focusing on genes involved in autoimmune diseases, we evaluate how allelic effects vary across cell types, highlighting genes that encode membrane, nuclear, cytoplasmic, or endoplasmic reticulum (ER) proteins (Fig. 5, C and D). One example is *BACH2*, an essential transcription factor involved in differentiating memory B and T cells (*60*). We identify rs10944479, which was previously associated with thyroid peroxidase antibody positivity and hyperthyroidism (*61*) and has an eQTL effect on *BACH2* in $CD8_{NC}$ cells. We identify eQTLs for *BACH2* in $CD4_{NC}$, $CD8_{NC}$, and B_{Mem} cells, although the loci controlling expression in each cell type are independent of one another (R^2 = 0 to 0.11; Fig. 5C). We demonstrate that rs60849819-T is associated with a significant down-regulation of *BACH2* in individuals who are homozygous for the T allele in B_{Mem} cells, and rs207253-A has a similar effect in $CD4_{NC}$ cells (Fig. 5C).

Another example that provides insight into autoimmune disease is *BLK*. Five eSNPs were identified as being associated with *BLK* expression in $CD4_{NC}$, $CD8^+$ T cells with an effector memory phenotype ($CD8_{ET}$), $CD8_{NC}$, B_{Mem}, and B_{IN} cells (Fig. 5D) and are associated with RA, SLE, Sjögren's syndrome, and systemic scleroderma (*41*, *58*, *62*, *63*) (Fig. 5D). One of these loci, rs2736336, results in the differential expression of *BLK* in B_{Mem} cells. The dynamic eQTL analysis shows that allelic effects vary significantly across the B cell lineage, with the largest genetic effects observed in the quantiles of the memory B cells. rs2736336 is associated with SLE (*41*), and carrying copies of the autoimmune risk allele has been implicated in hyperactivation of B cells, with enhanced T cell costimulatory capacities (*64*). These results suggest that an allelic variation at rs2736336 contributes to interindividual variation in maintaining tolerance of B lymphocytes. Src family tyrosine kinases, such as *BLK*, are critical components of the signaling pathways that act downstream of the antigen receptor and determine the strength of the signal that a cell receives as a consequence of antigen engagement.

Finally, we sought to evaluate the impact of eQTLs on cellular composition within the OneK1K cohort. For each $eSNP_1$, we tested for the association between an individual's

genotype and cell-type proportion. At a study-wide significance threshold ($p < 3.0 \times 10^{-6}$), we identified five associations, all of which affect the proportion of $CD8_{S100B}$ cells (table S11). The eGenes—*LSS*, *S100B*, *PRMT2*, *DIP2A*—and *PCNT* are all located within a 1-Mb region on chromosome 21q22, and the SNPs are in modest to high linkage disequilibrium with one another (R^2 = 0.31 to 0.97), suggesting that a single variant influences the proportion of $CD8_{S100B}$ cells.

Identification of cell type–specific trans-eQTLs suggests that distal genome regulation is highly cell type–specific

We performed trans-eQTL analysis, testing the top eSNPs from each cis-eQTL against the gene expression levels of all other genes, excluding those within ±2 Mb of the cis-eGene and the major histocompatibility complex (MHC) locus. At a study-wide FDR of 0.01, we identified 990 trans-eQTL (median of one per cis-eSNP) (table S16). The number of trans-eGenes identified in each cell type was weakly correlated with the total number of cis-eQTLs (Spearman's ρ = 0.37) (Fig. 6A). Compared with cis-eGenes, most trans-eGenes were specific for a cell type, and none were found ubiquitously across cell types (Fig. 6B and fig. S14).

A total of 630 cell type–specific trans-eQTL effects were identified. For example, rs2077041 has a cis effect on *ERN1* expression in $CD8_{ET}$ cells, with the C allele decreasing expression. This locus has the same allelic direction of effect in seven trans-eGenes (Fig. 6C). *ERN1* is an unfolded protein response stress sensor with dual roles as a protein kinase and ribonuclease (*65*) and can catalyze the splicing of *XBP1* in a spliceosome-independent manner (*66*). Up-regulation of the master transcriptional regulator of the unfolded protein response, *XBP1*, promotes protein maturation. Individuals carrying copies of the C allele of rs2077041 have down-regulation of *XBP1* and *SEC61G*, *SEC61B*, and *SEC11C*, which are involved in the translocation, signal peptide removal, and integration of proteins across the ER membrane (*67*, *68*). Interestingly, rs74787440 was found to also have a significant cis effect

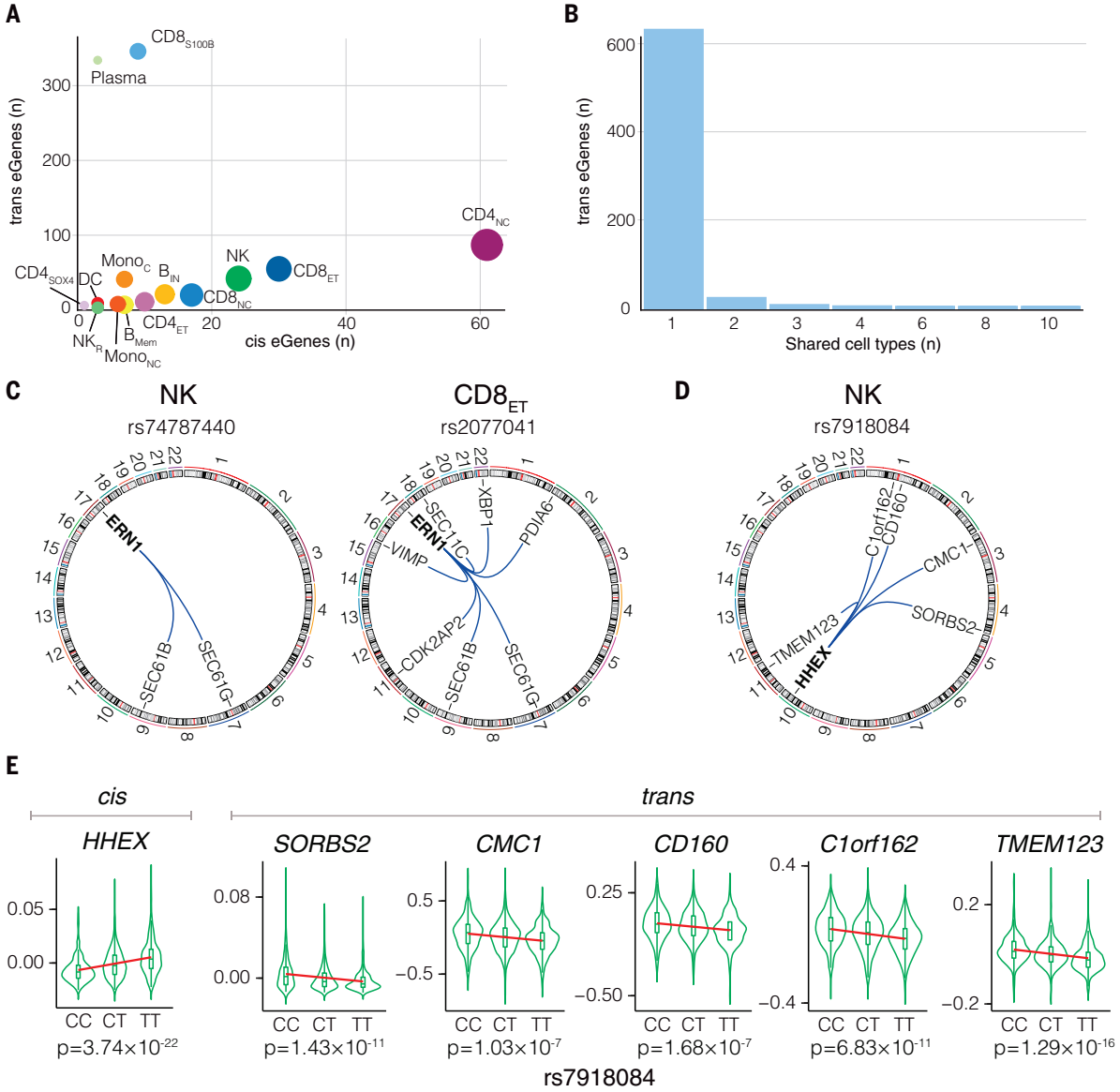


Fig. 6. Trans-acting eQTL mapping at single-cell resolution. (**A**) The number of genes with a trans-eQTL (trans-eGenes) as a function of genes with a cis-eQTL (cis-eGenes). Bubble size corresponds to the relative number of cis-eSNPs identified as per Fig. 2A. (**B**) The number of trans-eGenes identified across the corresponding number of cell types. (**C**) eQTL associations of rs74787440 and rs2077041 exert cis effects on *ERN1* in NK and $CD8_{ET}$ cells, respectively. (**D**) rs7918084 exerts a cis effect on *HHEX* expression and a trans effect in the opposite direction on *CD160*, *CMC1*, *SORBS2*, *TMEM123*, and *C1orf162*. (**E**) Change in the expression profile associated with rs7918084 genotypes on cis- and trans-genes. *p* values are from Spearman's rank correlation testing.

on *ERN1* expression in natural killer (NK) cells. Yet this same variant has a trans effect on *SEC61G* and *SEC61B* but not on the other genes associated with rs2077041.

When the locus on chromosome 21q22 that contains eQTLs associated with cellular composition was inspected, we identified many trans-eQTLs in this region and found that the expression levels of 118 genes throughout the genome were associated with these eSNPs (table S16). The route by which genetic variation affects $CD8_{S100B}$ frequency is unclear, and we find no evidence for the enrichment of functional pathways from the trans-eGenes. Across the tests, we observe a genomic inflation factor (λ) of 1.05, suggesting a limited impact of single cell eQTLs on cell composition, although other significant associations would be uncovered with larger sample sizes.

Trans-eQTLs were identified at established autoimmune risk loci, including rs7918084-T (Fig. 6D), which is a cis-eQTL for *HHEX* in NK cells and is associated with atopic asthma and eosinophil counts in peripheral blood (*50*, *69*). *HHEX* binds and represses the proapoptotic factor *BIM* (*70*), increasing the number of NK cells. In NK cells, rs7918084-T yields trans-eQTL effects across four chromosomes, decreasing the expression of *CD160*, *CMC1*, *SORBS2*, *TMEM123*, and *C1orf162* (Fig. 6E). *CD160* is a stimulatory receptor that is important in facilitating NK cell interferon-γ (IFN-γ) production (*71*), with NK cell recruitment being pivotal in the development of the airway eosinophilia typical of asthma in murine models (*72*) and IFN-γ secretion from NK cells in animal models of asthma being associated with reduced airway inflammation (*73*).

The production of IFN-γ within airway inflammation models plays a complex role in regulating inflammation, and it has been shown that IFN-γ acting on the airway epithelium will limit inflammation, such that lower IFN-γ levels may lead to more asthma-related airway inflammation and obstruction (*74*). Mechanistically, the rs7918084-T risk allele for asthma may combine derepression of NK cell proliferation in an *HHEX*-dependent cis-acting mechanism with reduced IFN-γ production by NK cells through *CD160* down-regulation, yielding the hallmarks of asthma.

Colocalization of genetic risk variants and single-cell cis-eQTLs identified cell type–specific mechanisms for autoimmune diseases

We applied an integrative approach to identify the relationship between cell type–specific eQTLs and genetic risk loci for seven common autoimmune diseases. We tested the extent to which cis-eQTLs (using $eSNP_1$) from each cell type were enriched for 2335 trait-associated SNPs for the seven autoimmune diseases selected for cis-QTL exploration in Fig. 5, C and D, using both colocalization and mendelian randomization approaches. Colocalization identified that 19% of cis-eQTLs have the same causal loci as GWAS risk variants (table S19). The overlap in eQTLs with GWAS loci shows significant enrichment for all diseases (Bonferroni adjusted $p < 5.1 \times 10^{-4}$) and in all cell types (fig. S22). The overlap was highest in $CD4_{NC}$ and NK cells. Similarly, in NK recruiting (NK_R) cells, there are high enrichments of overlap for inflammatory bowel disease (IBD), RA, ankylosing spondylitis (AS), and Crohn's disease (CD), which are low for multiple sclerosis (MS), SLE, and T1DM (fig. S22). These results highlight the complexity with which the polygenetic effects of genetic risk for these common autoimmune diseases act at the cellular level.

Focusing on MS as an example, we identify overlapping cis-eQTL for 108 risk genes (table S17). Colocalization identified 530 gene–cell type pairs with a shared causal effect through eQTLs (Fig. 7A). The eQTL overlap for MS disease risk loci is highly cell type–specific: Of the 108 genes, 69 show eQTL overlap in just a single cell type. There are an additional 20 genes where eQTLs are identified in two cell types, 10 with eQTLs in three cell types, and five with eQTLs in four cell types. For example, for *RMI2*, which is a gene expressed in all PBMC types, we identify an overlapping eQTL and MS association in $CD4_{NC}$ cells only.

By contrast, for *METTL21B*, overlapping eQTLs are observed in $CD4_{NC}$, $CD4_{ET}$, and $CD8_{NC}$ cells. These results are concordant with our observations of cell type–specific eQTLs and provide further evidence for the genetic risk of common autoimmune diseases acting in a highly cell type–specific manner, where each locus contributes through changes to the function of a limited number of cell types. Still, collectively, genetic risk is endowed through the immune system.

Although overlapping GWAS SNPs and eQTLs imply that altered gene expression is involved in disease pathogenesis, there are two alternative hypotheses. One is that both the GWAS loci and eQTL have the same causal variant, but the effects on the two phenotypes are independent—that is, pleiotropy. A second explanation is that there are two independent causal loci, one for the GWAS association and the other for the eQTL. Still, they are in linkage disequilibrium with one another. To distinguish between these two hypotheses, we implemented a Mendelian randomization approach to identify evidence for the direction of causation by which risk loci for autoimmune diseases act (*75*). We tested for the causal relationship between all disease-associated variants ($p < 1 \times 10^{-8}$) and OneK1K eQTLs across each of the 14 cell types using GWAS data from the seven autoimmune diseases previously introduced. In total, we identified 305 loci (study-wide FDR < 0.05) where the associated risk loci are identified as having a causal effect of disease risk through changes in the expression of a specific gene in one or more cell types, ranging from 4 (T1DM) to 47 (IBD) (Table 1 and table S18). Of the 305 loci, 188 are located in the MHC region, where causal effects display largely ubiquitous effects across cell types. The remaining 117 loci show patterns of highly cell type–specific causal effects, with 76 loci identified as having a causal effect in only one cell type (Table 1).

Again, using MS as an example, we evaluated the causal genes and the cell types in which they act for 90 risk loci (*13*). Of these, we were able to test for the causal direction of 57 risk loci based on the overlap of eQTLs in one or more cell types in OneK1K data. Our analysis identified significant (study-wide FDR < 0.05) effects for 39 genes (Fig. 7B and table S18). In the MHC region, we identified 73 loci whose causal effects on MS risk predominantly act through changes in the expression of genes in multiple cell types. For example, rs9264579 is identified as working through changes in human lymphocyte antigen class B (HLA-B) expression in all 14 cell types, whereas rs9501393 has a causal effect by changing the expression levels of *SKIV2L* in $CD4_{NC}$ cells only. Outside of the MHC region, we identified an additional 17 loci with causal effects that act in a more cell type–specific manner. For example, SNPs in the 1q23 region have previously been identified as associated with MS, with FCRL3 tagged by rs7528684 ($p = 8.9 \times 10^{-9}$) located within a promoter element. Our analysis identified the proximally located *FCRL3* as the causal gene for MS risk in $CD8_{ET}$ ($p = 5.0 \times 10^{-7}$) and B_{IN} ($p = 6.6 \times 10^{-7}$) cells (Fig. 7C).

Another example is the MS risk locus at 3q12, which is tagged by lead SNP rs9882971 ($p = 6.5 \times 10^{-9}$), where Mendelian randomization analysis identified *EAF2* as the causal gene in B_{IN} ($p = 1.7 \times 10^{-8}$) and B_{Mem} ($p = 2.8 \times 10^{-8}$) cells. Because *EAF2* is universally expressed, our results provide a clear example of the ability to identify cell-type genetic effects on gene expression and pinpoint the cells in which genetic risk factors are acting. A final example is the risk locus at 19p13, which is tagged by top SNP rs12984330 ($p = 2.8 \times 10^{-9}$) located in the intronic region of *PIK3R2*. Our analyses identify the causal gene as *MAST3* in $CD8_{ET}$ and NK cells, which is located about 65 kb from the lead SNP. *MAST3* is also universally expressed, although there is known evidence of the risk locus overlapping with regulatory elements, which presents an interesting case for further exploration.

Discussion

This study reveals the allelic architecture of cell type–specific eQTLs in circulating immune cells. We mapped genetic effects of 14 cell types

on gene expression and identified more than 26,000 independent cis-acting eQTLs and 990 trans-eQTLs outside the MHC locus. Summary statistics for our cis-eQTLs can be found in table S10 and are available for browsing from www.onek1k.org/. We show that most of these eQTLs have an allelic effect on gene expression, which is largely cell type–specific, yet replicate in two independent cohorts. We identify examples of how genetic loci contribute to key immune function pathways. Lineage-dynamic analyses applied to B cells demonstrated expected changes in markers of B cell maturation. They facilitated the identification of dynamic eQTLs, many of which had not been identified through our primary cis-eQTL analysis. By integrating scRNA-seq eQTL data with autoimmune risk loci identified through GWASs, we uncovered both the causal gene at these loci and resolved the cells through which these genes exert their pathogenetic effects. Mendelian randomization and colocalization analysis of our eQTLs and disease-associated SNPs were performed, providing complementary insights into the relationship between eQTLs and disease risk loci. The colocalization analysis provides evidence that the same

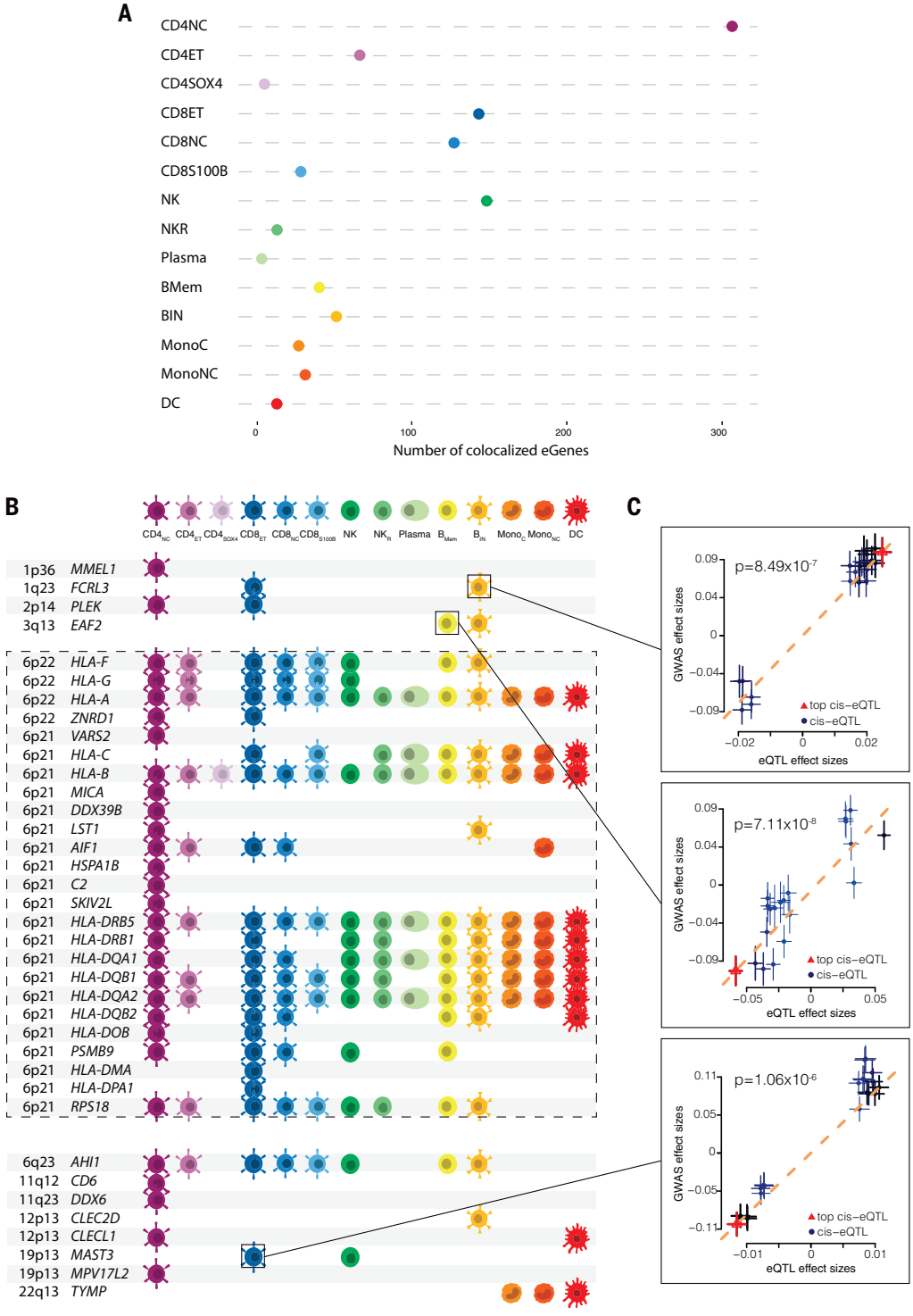


Fig. 7. Dissection of autoimmune disease loci using eQTL mapping at single-cell resolution.
(**A**) Breakdown of cis-eGenes colocalized with GWAS associations for MS using Bayes factors. (**B**) Mendelian randomization was used to establish causation between overlapping GWAS loci for MS and identified eSNPs. Significant results were identified for 39 MS-related genes (FDR < 0.05), with the 12 outside the MHC locus (dashed box) displaying highly cell type–specific effects. Colored symbols depict cell types. Differences were identified between the direct NHGRI-EBI GWAS Catalog (9) overlap and Mendelian randomization analysis for eGene and cell-specific profiles. (**C**) The effect sizes of OneK1K eQTL SNPs plotted against the allelic effects from the MS GWAS for *FCRL3*, *EAF2*, and *MAST3* in B_{IN}, B_{Mem}, and $CD8_{ET}$ cells, respectively, are displayed. *p* values are from the heterogeneity in dependent instruments (HEIDI) test.

Table 1. A summary of significant evidence of causation between overlapping GWAS loci and identified eQTLs for autoimmune diseases. The number of cell types in which causal effect is identified is given in parentheses. For loci with causal effects acting in multiple cell types, multiple independent eQTLs are often present (table S18). Significance threshold of FDR < 0.5.

Disease	Loci	Genes
Systemic lupus erythematosus	19	*AIF1* (3), *BLK* (3), *BTN2A1* (3), *BTN2A2* (1), *BTN3A2* (9), *C6orf48* (4), *FAM167A* (1), *HLA-B* (2), *HLA-C* (4), *HLA-DMA* (1), *HLA-DQA1* (2), *HLA-DQB1* (6), *HLA-DQB2* (2), *HLA-DRB1* (3), *HLA-DRB5* (9), *MICB* (1), *UBE2L3* (3), *XXbac-BPG181B23.7* (3), *ZFP57* (5)
Rheumatoid arthritis	37	*AIF1* (3), *ANKRD55* (3), *B3GALT4* (1), *C2* (1), *C6orf48* (1), *CTLA4* (5), *DDX6* (3), *HLA-A* (9), *HLA-B* (6), *HLA-C* (12), *HLA-DMA* (1), *HLA-DOB* (2), *HLA-DPA1* (2), *HLA-DPB1* (5), *HLA-DQA1* (11), *HLA-DQA2* (13), *HLA-DQB1* (12), *HLA-DQB2* (4), *HLA-DRB1* (9), *HLA-DRB5* (12), *HSD17B8* (1), *HSPA1B* (1), *IL6ST* (1), *LST1* (4), *MDC1* (1), *MICA* (1), *MMEL1* (1), *RP11-279F6.3* (1), *RP11-973H7.4* (1), *SKIV2L* (2), *SYNGR1* (6), *TAP1* (1), *TAPBP* (4), *UQCC2* (1), *XXbac-BPG181B23.7* (2), *XXbac-BPG299F13.17* (14), *ZFP57* (5)
Crohn's disease	18	*ADCY7* (2), *BRD7* (3), *C5orf56* (1), *CCDC101* (1), *CTD-2260A17.2* (1), *CYLD* (1), *ERAP2* (13), *GSDMB* (1), *IP6K2* (1), *IRF1* (4), *ORMDL3* (1), *RNASET2* (9), *SLC22A5* (1), *SLC2A4RG* (1), *SNX20* (1), *SPNS1* (3), *TUFM* (2), *UQCRQ* (1)
Inflammatory bowel disease	47	*ADCY7* (3), *BRD7* (3), *C5orf56* (1), *CCDC101* (2), *CYLD* (1), *EGR2* (1), *ETS2* (1), *FCGR3B* (1), *FYB* (1), *GMEB2* (1), *GPANK1* (1), *GPX1* (2), *GSDMB* (1), *HCG23* (1), *HLA-DOB* (3), *HLA-DQA1* (8), *HLA-DQA2* (13), *HLA-DQB1* (7), *HLA-DQB2* (3), *HLA-DRB1* (4), *HLA-DRB5* (8), *LAMB1* (8), *LST1* (4), *MICB* (1), *NDUFS2* (1), *ORMDL3* (5), *PAPD5* (1), *PEX13* (1), *PNMT* (1), *RBM6* (1), *RNASET2* (7), *RP11-229P13.20* (1), *RP11-324I22.4* (1), *RP11-94L15.2* (1), *SLC22A5* (1), *SLC2A4RG* (2), *STMN3* (2), *TCTA* (1), *TNFRSF9* (1), *TUFM* (8), *UBE2L3* (4), *UQCRQ* (1)
Multiple sclerosis	39	*AHI1* (8), *AIF1* (5), *C2* (1), *CD6* (1), *CLEC2D* (1), *CLECL1* (2), *DDX39B* (1), *DDX6* (1), *EAF2* (2), *FCRL3* (2), *HLA-A* (13), *HLA-B* (14), *HLA-C* (9), *HLA-DMA* (1), *HLA-DOB* (2), *HLA-DPA1* (1), *HLA-DQA1* (11), *HLA-DQA2* (13), *HLA-DQB1* (13), *HLA-DQB2* (6), *HLA-DRB1* (9), *HLA-DRB5* (13), *HLA-F* (8), *HLA-G* (6), *HSPA1B* (1), *LST1* (2), *MAST3* (2), *MICA* (1), *MMEL1* (1), *MPV17L2* (1), *PLEK* (2), *PSMB9* (5), *RPS18* (9), *SKIV2L* (1), *TYMP* (3), *VARS2* (1), *XXbac-BPG181B23.7* (3), *XXbac-BPG299F13.17* (9), *ZNRD1* (2)
Ankylosing spondylitis	10	*AIF1* (2), *C6orf48* (3), *HLA-A* (1), *HLA-B* (9), *HLA-C* (9), *HLA-DQA1* (1), *LST1* (3), *MICB* (1), *NCR3* (2), *XXbac-BPG181B23.7* (4)
Type 1 diabetes mellitus	4	*HLA-DQA1* (9), *HLA-DQA2* (13), *HLA-DQB1* (9), *HLA-DQB2* (3)

causal loci are shared between an eQTL and GWAS risk loci. This observation can be explained by either a causal effect or pleiotropy. Mendelian randomization takes this one step further, addressing these alternative hypotheses to provide evidence of the direction of causal effect (i.e., DNA to RNA to disease).

Single-cell eQTL analyses have several advantages over alternative methods that are used to map the allelic architecture of transcriptional regulation, such as cellular deconvolution from bulk RNA-seq data. For example, scRNA-seq–based approaches can identify previously uncharacterized and rare cell types, which are challenging to detect using deconvolution methods (*22*, *76*). scRNA-seq also accurately quantifies transcriptional abundance, because amplified libraries can be collapsed back to the level of individual transcript molecules using unique molecular identifier (UMI) barcodes. Nevertheless, ongoing work investigating trans-acting variants and gene-environment interactions at single-cell resolution is required, particularly in the immune system, where exposure to antigens or cytokines can trigger changes in the transcriptional profile of cells.

This work brings together genetic epidemiology with scRNA-seq to uncover drivers of interindividual variation in the immune system. Our results demonstrate how segregating genetic variation influences the expression of genes that encode proteins involved in critical immune regulatory and signaling pathways in a cell type–specific manner. Understanding the genetic underpinnings of immune system regulation will have broad implications in the treatment of autoimmune diseases and infections, transplantation, and cancer.

Materials and methods summary

We collected peripheral blood from 1104 individuals. After DNA extraction, samples were genotyped using the Illumina Infinium Global Screening Array. Poor genotyping quality, cryptic relatedness, and ethnic outliers were removed, yielding 1034 participants. Imputation was performed using the Michigan Imputation Server (*24*). PMBCs were isolated through density-gradient centrifugation from heparinized whole blood (8-ml cell preparation tubes; BD Biosciences Australia; catalog no. 362753), with live cells isolated with the Miltenyi Dead Cell Removal Kit (Miltenyi; catalog no. 130-090-101). Live cells were subsequently pooled with 12 to 14 participant samples per pool, which underwent single-cell RNA capture and barcoding with the Single Cell 3′ Library and Gel Bead Kit (10x Genomics) to target the capture of 20,000 cells per well. Library preparation and multiplex sequencing using an Illumina NovaSeq 2000 generated 49 billion reads. Reads underwent processing using the Cell Ranger Single Cell Software Suite (v 2.2.0; 10x Genomics) into FASTQ files, followed by demultiplexing into their respective pools, and

were mapped to GRCh37/hg19 (release 84) using STAR (*77*).

Cells were assigned using genotype data to individual participants using Dexmuxlet (*78*), with droplets containing two or more cells excluded using Demuxlet and Scrublet (*79*), yielding 982 individuals in the final cohort. Cells were classified using supervised clustering into major immune populations using reference data from Zheng *et al.* (*26*) and then underwent unsupervised clustering using Seurat v3.0 (*80*). Expression values for genes were first normalized by the pool for the distribution of the total number of UMIs, the number of genes, and the percentage of mitochondrial gene expression and were subsequently adjusted for sex, age, six genotyping principal components, and two probabilistic estimation of expression residuals (PEER) factors. Subsequent single-cell cis-eQTL mapping was undertaken through five rounds of iterative conditional analysis to yield cell type–specific $eSNP_1$ to $eSNP_5$. Lead cis-eQTLs were replicated in two independent cohorts of participants by creating pseudo-bulk populations, and trans-eQTL mapping was performed. Lineage-dynamic analysis was undertaken using SCTransform (*81*) to identify 500 differentially expressed genes and filter out contaminating cells. A two-dimensional space was created using PHATE (*82*) and slingshot (*83*). Six quantiles were analyzed for the presence of dynamic eQTLs.

REFERENCES AND NOTES

1. P. Brodin, M. M. Davis, Human immune system variation. *Nat. Rev. Immunol.* **17**, 21–29 (2017). doi: 10.1038/nri.2016.125; pmid: 27916977
2. L. Chen *et al.*, Genetic drivers of epigenetic and transcriptional variation in human immune cells. *Cell* **167**, 1398–1414.e24 (2016). doi: 10.1016/j.cell.2016.10.026; pmid: 27863251
3. Y. I. Li *et al.*, RNA splicing is a primary link between genetic variation and disease. *Science* **352**, 600–604 (2016). doi: 10.1126/science.aad9417; pmid: 27126046
4. F. Hormozdiari *et al.*, Leveraging molecular quantitative trait loci to understand the genetic architecture of diseases and complex traits. *Nat. Genet.* **50**, 1041–1047 (2018). doi: 10.1038/s41588-018-0148-2; pmid: 29942083
5. GTEx Consortium, The GTEx Consortium atlas of genetic regulatory effects across human tissues. *Science* **369**, 1318–1330 (2020). doi: 10.1126/science.aaz1776; pmid: 32913098
6. W. J. Astle *et al.*, The allelic landscape of human blood cell trait variation and links to common complex disease. *Cell* **167**, 1415–1429.e19 (2016). doi: 10.1016/j.cell.2016.10.042; pmid: 27863252
7. C. C. Whitacre, Sex differences in autoimmune disease. *Nat. Immunol.* **2**, 777–780 (2001). doi: 10.1038/ni0901-777; pmid: 11526384
8. G. S. Cooper, B. C. Stroehla, The epidemiology of autoimmune diseases. *Autoimmun. Rev.* **2**, 119–125 (2003). doi: 10.1016/S1568-9972(03)00006-5; pmid: 12848952
9. A. Buniello *et al.*, The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. *Nucleic Acids Res.* **47**, D1005–D1012 (2019). doi: 10.1093/nar/gky1120; pmid: 30445434
10. L. A. Hindorff *et al.*, Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 9362–9367 (2009). doi: 10.1073/pnas.0903103106; pmid: 19474294
11. M. Boyd *et al.*, Characterization of the enhancer and promoter landscape of inflammatory bowel disease from human colon biopsies. *Nat. Commun.* **9**, 1661 (2018). doi: 10.1038/s41467-018-03766-z; pmid: 29695774
12. Y. Okada, S. Eyre, A. Suzuki, Y. Kochi, K. Yamamoto, Genetics of rheumatoid arthritis: 2018 status. *Ann. Rheum. Dis.* **78**, 446–453 (2019). doi: 10.1136/annrheumdis-2018-213678; pmid: 30530827
13. J. Bentham *et al.*, Genetic association analyses implicate aberrant regulation of innate and adaptive immunity genes in the pathogenesis of systemic lupus erythematosus. *Nat. Genet.* **47**, 1457–1464 (2015). doi: 10.1038/ng.3434; pmid: 26502338
14. B. D. Umans, A. Battle, Y. Gilad, Where are the disease-associated eQTLs? *Trends Genet.* **37**, 109–124 (2021). doi: 10.1016/j.tig.2020.08.009; pmid: 32912663
15. GTEx Consortium, Genetic effects on gene expression across human tissues. *Nature* **550**, 204–213 (2017). doi: 10.1038/nature24277; pmid: 29022597
16. H.-J. Westra *et al.*, Systematic identification of trans eQTLs as putative drivers of known disease associations. *Nat. Genet.* **45**, 1238–1243 (2013). doi: 10.1038/ng.2756; pmid: 24013639
17. L. R. Lloyd-Jones *et al.*, The genetic architecture of gene expression in peripheral blood. *Am. J. Hum. Genet.* **100**, 228–237 (2017). doi: 10.1016/j.ajhg.2016.12.008; pmid: 28065468
18. C. J. Ye *et al.*, Intersection of population variation and autoimmunity genetics in human T cell activation. *Science* **345**, 1254665 (2014). doi: 10.1126/science.1254665; pmid: 25214635
19. J. E. Powell *et al.*, Genetic control of gene expression in whole blood and lymphoblastoid cell lines is largely independent. *Genome Res.* **22**, 456–466 (2012). doi: 10.1101/gr.126540.111; pmid: 22183966
20. K. Ishigaki *et al.*, Polygenic burdens on cell-specific pathways underlie the risk of rheumatoid arthritis. *Nat. Genet.* **49**, 1120–1125 (2017). doi: 10.1038/ng.3885; pmid: 28553958
21. B. J. Schmiedel *et al.*, Impact of genetic polymorphisms on human immune cell gene expression. *Cell* **175**, 1701–1715.e16 (2018). doi: 10.1016/j.cell.2018.10.022; pmid: 30449622
22. S. Kim-Hellmuth *et al.*, Cell type–specific genetic regulation of gene expression across human tissues. *Science* **369**, eaaz8528 (2020). doi: 10.1126/science.aaz8528; pmid: 32913075
23. F. Avila Cobos, J. Alquicira-Hernandez, J. E. Powell, P. Mestdagh, K. De Preter, Benchmarking of cell type deconvolution pipelines for transcriptomics data. *Nat. Commun.* **11**, 5650 (2020). doi: 10.1038/s41467-020-19015-1; pmid: 33159064
24. S. McCarthy *et al.*, A reference panel of 64,976 haplotypes for genotype imputation. *Nat. Genet.* **48**, 1279–1283 (2016). doi: 10.1038/ng.3643; pmid: 27548312
25. J. Alquicira-Hernandez, A. Sathe, H. P. Ji, Q. Nguyen, J. E. Powell, scPred: Accurate supervised method for cell-type classification from single-cell RNA-seq data. *Genome Biol.* **20**, 264 (2019). doi: 10.1186/s13059-019-1862-5; pmid: 31829268
26. G. X. Y. Zheng *et al.*, Massively parallel digital transcriptional profiling of single cells. *Nat. Commun.* **8**, 14049 (2017). doi: 10.1038/ncomms14049; pmid: 28091601
27. V. Orrù *et al.*, Genetic variants regulating immune cell levels in health and disease. *Cell* **155**, 242–256 (2013). doi: 10.1016/j.cell.2013.08.041; pmid: 24074872
28. P. Brodin *et al.*, Variation in the human immune system is largely driven by non-heritable influences. *Cell* **160**, 37–47 (2015). doi: 10.1016/j.cell.2014.12.020; pmid: 25594173
29. E. Cano-Gamez *et al.*, Single-cell transcriptomics identifies an effectorness gradient shaping the response of $CD4^+$ T cells to cytokines. *Nat. Commun.* **11**, 1801 (2020). doi: 10.1038/s41467-020-15543-y; pmid: 32286271
30. A.-C. Villani *et al.*, Single-cell RNA-seq reveals new types of human blood dendritic cells, monocytes, and progenitors. *Science* **356**, eaah4573 (2017). doi: 10.1126/science.aah4573; pmid: 28428369
31. B. P. Fairfax *et al.*, Innate immune activity conditions the effect of regulatory variants upon monocyte gene expression. *Science* **343**, 1246949 (2014). doi: 10.1126/science.1246949; pmid: 24604202
32. M. A. Christophorou *et al.*, Citrullination regulates pluripotency and histone H1 binding to chromatin. *Nature* **507**, 104–108 (2014). doi: 10.1038/nature12942; pmid: 24463520
33. F. Cornélis *et al.*, New susceptibility locus for rheumatoid arthritis suggested by a genome-wide linkage study. *Proc. Natl. Acad. Sci. U.S.A.* **95**, 10746–10750 (1998). doi: 10.1073/pnas.95.18.10746; pmid: 9724775
34. U. Võsa *et al.*, Large-scale cis- and trans-eQTL analyses identify thousands of genetic loci and polygenic scores that regulate blood gene expression. *Nat. Genet.* **53**, 1300–1310 (2021). doi: 10.1038/s41588-021-00913-z; pmid: 34475573
35. J. W. Said *et al.*, TCL1 oncogene expression in B cell subsets from lymphoid hyperplasia and distinct classes of B cell lymphoma. *Lab. Invest.* **81**, 555–564 (2001). doi: 10.1038/labinvest.3780264; pmid: 11304575
36. E. F. Wagner *et al.*, Novel diversity in IL-4-mediated responses in resting human naive B cells versus germinal center/memory B cells. *J. Immunol.* **165**, 5573–5579 (2000). doi: 10.4049/jimmunol.165.10.5573; pmid: 11067912
37. S. G. Tangye, Y. J. Liu, G. Aversa, J. H. Phillips, J. E. de Vries, Identification of functional human splenic memory B cells by expression of CD148 and CD27. *J. Exp. Med.* **188**, 1691–1703 (1998). doi: 10.1084/jem.188.9.1691; pmid: 9802981
38. A. Kallies *et al.*, Initiation of plasma-cell differentiation is independent of the transcription factor Blimp-1. *Immunity* **26**, 555–566 (2007). doi: 10.1016/j.immuni.2007.04.007; pmid: 17509907
39. M. L. Tang, D. A. Steeber, X. Q. Zhang, T. F. Tedder, Intrinsic differences in L-selectin expression levels affect T and B lymphocyte subset-specific recirculation pathways. *J. Immunol.* **160**, 5113–5121 (1998). pmid: 9590263
40. M. Togni *et al.*, Regulation of in vitro and in vivo immune functions by the cytosolic adaptor protein SKAP-HOM. *Mol. Cell. Biol.* **25**, 8052–8063 (2005). doi: 10.1128/MCB.25.18.8052-8063.2005; pmid: 16135797
41. C. D. Langefeld *et al.*, Transancestral mapping and genetic load in systemic lupus erythematosus. *Nat. Commun.* **8**, 16021 (2017). doi: 10.1038/ncomms16021; pmid: 28714469
42. P. K. Gregersen *et al.*, REL, encoding a member of the NF-κB family of transcription factors, is a newly defined risk locus for rheumatoid arthritis. *Nat. Genet.* **41**, 820–823 (2009). doi: 10.1038/ng.395; pmid: 19503088
43. F. Köntgen *et al.*, Mice lacking the c-*rel* proto-oncogene exhibit defects in lymphocyte proliferation, humoral immunity, and interleukin-2 expression. *Genes Dev.* **9**, 1965–1977 (1995). doi: 10.1101/gad.9.16.1965; pmid: 7649478
44. J. Dang *et al.*, ORMDL3 facilitates the survival of splenic B cells via an ATF6α-endoplasmic reticulum stress-Beclin1 autophagy regulatory pathway. *J. Immunol.* **199**, 1647–1659 (2017). doi: 10.4049/jimmunol.1602124; pmid: 28747345
45. A. H. Beecham *et al.*, Analysis of immune-related loci identifies 48 new susceptibility variants for multiple sclerosis. *Nat. Genet.* **45**, 1353–1360 (2013). doi: 10.1038/ng.2770; pmid: 24076602
46. K. M. de Lange *et al.*, Genome-wide association study implicates immune activation of multiple integrin genes in inflammatory bowel disease. *Nat. Genet.* **49**, 256–261 (2017). doi: 10.1038/ng.3760; pmid: 28067908
47. S. Onengut-Gumuscu *et al.*, Fine mapping of type 1 diabetes susceptibility loci and evidence for colocalization of causal variants with lymphoid gene enhancers. *Nat. Genet.* **47**, 381–386 (2015). doi: 10.1038/ng.3245; pmid: 25751624
48. J. C. Barrett *et al.*, Genome-wide association study and meta-analysis find that over 40 loci affect risk of type 1 diabetes. *Nat. Genet.* **41**, 703–707 (2009). doi: 10.1038/ng.381; pmid: 19430480
49. W. M. Gallatin, I. L. Weissman, E. C. Butcher, A cell- surface molecule involved in organ-specific homing of lymphocytes. *Nature* **304**, 30–34 (1983). doi: 10.1038/304030a0; pmid: 6866086
50. M.-H. Chen *et al.*, Trans-ethnic and ancestry-specific blood-cell genetics in 746,667 individuals from 5 global populations. *Cell* **182**, 1198–1213.e14 (2020). doi: 10.1016/j.cell.2020.06.045; pmid: 32888493
51. B. P. Fairfax *et al.*, Genetics of gene expression in primary immune cells identifies cell type-specific master regulators and roles of HLA alleles. *Nat. Genet.* **44**, 502–510 (2012). doi: 10.1038/ng.2205; pmid: 22446964
52. P. Waterhouse *et al.*, Lymphoproliferative disorders with early lethality in mice deficient in Ctla-4. *Science* **270**, 985–988 (1995). doi: 10.1126/science.270.5238.985; pmid: 7481803
53. E. A. Tivol *et al.*, Loss of CTLA-4 leads to massive lymphoproliferation and fatal multiorgan tissue destruction, revealing a critical negative regulatory role of CTLA-4. *Immunity* **3**, 541–547 (1995). doi: 10.1016/1074-7613(95)90125-6; pmid: 7584144
54. D. Schubert *et al.*, Autosomal dominant immune dysregulation syndrome in humans with *CTLA4* mutations. *Nat. Med.* **20**, 1410–1416 (2014). doi: 10.1038/nm.3746; pmid: 25329329
55. H. S. Kuehn *et al.*, Immune dysregulation in human subjects with heterozygous germline mutations in *CTLA4*. *Science* **345**, 1623–1627 (2014). doi: 10.1126/science.1255904; pmid: 25213377

56. Y.-F. Wang *et al*., Identification of 38 novel loci for systemic lupus erythematosus and genetic heterogeneity between ancestral groups. *Nat. Commun.* **12**, 772 (2021). doi: 10.1038/s41467-021-21049-y; pmid: 33536424
57. J. D. Cooper *et al*., Meta-analysis of genome-wide association study data identifies additional type 1 diabetes risk loci. *Nat. Genet.* **40**, 1399–1401 (2008). doi: 10.1038/ng.249; pmid: 18978792
58. K. Kim *et al*., High-density genotyping of immune loci in Koreans and Europeans identifies eight new rheumatoid arthritis risk loci. *Ann. Rheum. Dis.* **74**, e13 (2015). doi: 10.1136/annrheumdis-2013-204749; pmid: 24532676
59. A. E. Renton *et al*., A genome-wide association study of myasthenia gravis. *JAMA Neurol.* **72**, 396–404 (2015). doi: 10.1001/jamaneurol.2014.4103; pmid: 25643325
60. R. Roychoudhuri *et al*., BACH2 represses effector programmes to stabilize Treg-mediated immune homeostasis - a new target in tumor immunotherapy? *J. Immunother. Cancer* **1**, O14 (2013). doi: 10.1186/2051-1426-1-S1-O14
61. M. Medici *et al*., Identification of novel genetic loci associated with thyroid peroxidase antibodies and clinical thyroid disease. *PLOS Genet.* **10**, e1004123 (2014). doi: 10.1371/journal.pgen.1004123; pmid: 24586183
62. C. J. Lessard *et al*., Variants at multiple loci implicated in both innate and adaptive immune responses are associated with Sjögren's syndrome. *Nat. Genet.* **45**, 1284–1292 (2013). doi: 10.1038/ng.2792; pmid: 24097067
63. A. Márquez *et al*., Meta-analysis of Immunochip data of four autoimmune diseases reveals novel single-disease and cross-phenotype associations. *Genome Med.* **10**, 97 (2018). doi: 10.1186/s13073-018-0604-8; pmid: 30572963
64. K. R. Simpfendorfer *et al*., Autoimmune disease-associated haplotypes of BLK exhibit lowered thresholds for B cell activation and expansion of Ig class-switched B cells. *Arthritis Rheumatol.* **67**, 2866–2876 (2015). doi: 10.1002/art.39301; pmid: 26246128
65. A. V. Korennykh *et al*., The unfolded protein response signals through high-order assembly of Ire1. *Nature* **457**, 687–693 (2009). doi: 10.1038/nature07661; pmid: 19079236
66. K. P. K. Lee *et al*., Structure of the dual enzyme Ire1 reveals the basis for catalysis and regulation in nonconventional RNA splicing. *Cell* **132**, 89–100 (2008). doi: 10.1016/j.cell.2007.10.057; pmid: 18191223
67. T. Becker *et al*., Structure of monomeric yeast and mammalian Sec61 complexes interacting with the translating ribosome. *Science* **326**, 1369–1373 (2009). doi: 10.1126/science.1178535; pmid: 19933108
68. K. U. Kalies, E. Hartmann, Membrane topology of the 12- and the 25-kDa subunits of the mammalian signal peptidase complex. *J. Biol. Chem.* **271**, 3925–3929 (1996). doi: 10.1074/jbc.271.7.3925; pmid: 8632014
69. Z. Zhu *et al*., Shared genetic and experimental links between obesity-related traits and asthma subtypes in UK Biobank. *J. Allergy Clin. Immunol.* **145**, 537–549 (2020). doi: 10.1016/j.jaci.2019.09.035; pmid: 31669095
70. S. W. Jang *et al*., Homeobox protein Hhex negatively regulates Treg cells by inhibiting Foxp3 expression and function. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 25790–25799 (2019). doi: 10.1073/pnas.1907224116; pmid: 31792183
71. T. C. Tu *et al*., CD160 is essential for NK-mediated IFN-γ production. *J. Exp. Med.* **212**, 415–429 (2015). doi: 10.1084/jem.20131601; pmid: 25711213
72. M. Korsgren *et al*., Natural killer cells determine development of allergen-induced eosinophilic airway inflammation in mice. *J. Exp. Med.* **189**, 553–562 (1999). doi: 10.1084/jem.189.3.553; pmid: 9927517
73. S. Matsubara *et al*., IL-2 and IL-18 attenuation of airway hyperresponsiveness requires STAT4, IFN-γ, and natural killer cells. *Am. J. Respir. Cell Mol. Biol.* **36**, 324–332 (2007). doi: 10.1165/rcmb.2006-0231OC; pmid: 17038663
74. C. Mitchell, K. Provost, N. Niu, R. Homer, L. Cohn, IFN-γ acts on the airway epithelium to inhibit local and systemic pathology in allergic airway disease. *J. Immunol.* **187**, 3815–3820 (2011). doi: 10.4049/jimmunol.1100436; pmid: 21873527
75. Z. Zhu *et al*., Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets. *Nat. Genet.* **48**, 481–487 (2016). doi: 10.1038/ng.3538; pmid: 27019110
76. A. Dobin *et al*., STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013). doi: 10.1093/bioinformatics/bts635; pmid: 23104886
77. H. M. Kang *et al*., Multiplexed droplet single-cell RNA-sequencing using natural genetic variation. *Nat. Biotechnol.* **36**, 89–94 (2018). doi: 10.1038/nbt.4042; pmid: 29227470
78. S. L. Wolock, R. Lopez, A. M. Klein, Scrublet: Computational identification of cell doublets in single-cell transcriptomic data. *Cell Syst.* **8**, 281–291.e9 (2019). doi: 10.1016/j.cels.2018.11.005; pmid: 30954476
79. T. Stuart *et al*., Comprehensive integration of single-cell data. *Cell* **177**, 1888–1902.e21 (2019). doi: 10.1016/j.cell.2019.05.031; pmid: 31178118
80. C. Hafemeister, R. Satija, Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression. *Genome Biol.* **20**, 296 (2019). doi: 10.1186/s13059-019-1874-1; pmid: 31870423
81. K. R. Moon *et al*., Visualizing structure and transitions in high-dimensional biological data. *Nat. Biotechnol.* **37**, 1482–1492 (2019). doi: 10.1038/s41587-019-0336-3; pmid: 31796933
82. K. Street *et al*., Slingshot: Cell lineage and pseudotime inference for single-cell transcriptomics. *BMC Genomics* **19**, 477 (2018). doi: 10.1186/s12864-018-4772-0; pmid: 29914354
83. A. Auton *et al*., A global reference for human genetic variation. *Nature* **526**, 68–74 (2015). doi: 10.1038/nature15393; pmid: 26432245
84. S. Yazar, J. A. Hernandez, Population-scale single-cell eQTL mapping identifies cell type specific genetic control of autoimmune disease. Zenodo (2021); .doi: 10.5281/zenodo.5084395

ACKNOWLEDGMENTS

We thank the participants of this study and are grateful to D. MacArthur and C. Goodnow, who provided feedback about this work. **Funding:** This research was supported by a National Health and Medical Research Council Research Fellowship (S.Y.), Practitioner Fellowship (A.W.H.), Career Development Fellowship (J.E.P., 1107599), and Investigator Fellowship (J.E.P., 1175781). K.A.F. is supported by the Alex Gadomski Fellowship, funded by Maddie Riewoldt's Vision. Additional grant support was provided by the National Health and Medical Research Council (1150144 and 1143163), the Australian Research Council (180101405), and the Royal Hobart Hospital Research Foundation. **Author contributions:** Conceptualization: J.E.P., A.W.H.; Methodology: S.Y., J.A.-H., K.W., A.S., S.A., K.A.F., A.W.H., J.E.P.; Resources: K.W., M.G.G., Q.L., A.R., T.R.P.T., L.C., K.M., C.C., A.L.C., C.J.Y., K.A.F., A.W.H., J.E.P.; Data analysis: S.Y., J.A.-H., K.W., A.S., M.G.G., K.A.F., A.W.H., J.E.P.; Writing – original draft: S.Y., J.A.-H., K.W., K.A.F., A.W.H., J.E.P.; Writing – review and editing: All other authors; Supervision and project administration: A.W.H., J.E.P.; Funding acquisition: C.J.Y., A.W.H., J.E.P. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** OneK1K single-cell gene expression and genotype data are available via Gene Expression Omnibus (GSE196830). The cell by gene data are available at Human Cell Atlas (HCA) (https://cellxgene.cziscience.com/collections/dde06e0f-ab3b-46be-96a2-a8082383c4a1). GWAS summary statistics for MS are available from the Multiple Sclerosis Genetics Consortium (https://imsgc.net/). 1000 Genome Phase 3 (02.05.2013 release) data are available from (http://ftp.ebi.ac.uk/1000g/ftp/). Summary statistics for IBD, CD, RA, AS, and T1DM GWASs are available for public access in the following repositories: (i) for IBD and CD from the IBD Genetics Consortium website (www.ibdgenetics.org), (ii) for RA from the Japanese Encyclopedia of Genetic associations by RIKEN (http://jenger.riken.jp/en/result), and (iii) for AS and T1DM from Pan-UK Biobank (https://pan.ukbb.broadinstitute.org/). The SLE GWAS summary statistics data can be downloaded from the following link: http://urr.cat/data/GWAS_SLE_summaryStats.zip. There are no known limitations or data transfer agreements regarding access to the summary data described above. Data used to replicate the single-cell eQTLs are available via Gene Expression Omnibus (GSE174188), and genotype data are available via dbGap (46198). Cell by gene data are available at HCA (https://cellxgene.cziscience.com/collections/436154da-bcf1-4130-9c8b-120ff9a888f2). Analysis code is available on Zenodo (*84*).

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abf3041
Materials and Methods
Figs. S1 to S23
Tables S1 to S19
References (*85*–*108*)
MDAR Reproducibility Checklist

16 October 2020; accepted 22 February 2022
10.1126/science.abf3041

RESEARCH ARTICLE SUMMARY

PROTEIN ENGINEERING

Tuning T cell receptor sensitivity through catch bond engineering

Xiang Zhao, Elizabeth M. Kolawole, Waipan Chan, Yinnian Feng, Xinbo Yang, Marvin H. Gee, Kevin M. Jude, Leah V. Sibener, Polly M. Fordyce, Ronald N. Germain, Brian D. Evavold, K. Christopher Garcia*

INTRODUCTION: The T cell receptor (TCR) controls T cell antigen specificity and helps determine response sensitivity upon recognizing peptide–major histocompatibility complexes (pMHCs). In immunotherapy, TCRs that react with tumor antigens are used in adoptive cell therapy (ACT) to eradicate tumors, but most endogenous tumor-specific TCRs elicit weak functional responses. To overcome this limitation, tumor-reactive TCRs have been affinity matured to enhance their killing potency. However, high-affinity TCRs can exhibit off-target toxicity in clinical trials, which suggests that new approaches are needed. Engineering TCRs to display high potency toward tumor targets while retaining low physiological affinities could potentially enhance the efficacy of T cell therapies without increasing the risk of off-target side effects. Catch bonds prolong the bond lifetime between proteins under increasing applied force, triggering TCR activation upon pMHC engagement. However, whether catch bonds can be engineered to enhance TCR potency and whether such TCRs would preserve their natural specificities and affinities is not known.

RATIONALE: We hypothesized that an alternative strategy to affinity maturation was needed to endow clinically useful TCRs with high potency yet low affinity [i.e., three-dimensional (3D) binding affinity (K_D) of ~5 to 50 μM]. We therefore devised an engineering strategy called catch bond fishing that relies on a functional selection to recruit catch bonds between poorly reactive TCRs and pMHCs. We surmised that new catch bonds could be acquired by mutating certain TCR residues into small libraries composed of charged or polar amino acids followed by, paradoxically, screening for high-potency, low-affinity TCR variants.

RESULTS: We first applied this engineering strategy to an HIV peptide–specific human TCR (TCR55), which binds the human lymphocyte antigen B35 (HLA-B35)–HIV complex with a physiological 3D binding affinity but fails to activate downstream signaling because of an apparent lack of catch bond formation on cells, as measured by biomembrane force probe (BFP). Our functional selection isolated CD69-high and pMHC tetramer staining–low T cells, thereby enriching for catch bond–engineered TCRs that trigger in a low-affinity regime. Single amino acid positions on TCR55 α and β chains were catch bond hotspots, and several amino acid substitutions at those sites resulted in potent signaling despite retaining physiological 3D binding affinities. These signaling-active TCR mutants had acquired catch bonds based on a BFP assay on cells, and those longer bond lifetimes correlated with signal strength.

We next applied this catch bond engineering strategy to a melanoma antigen MAGE-A3–specific TCR. An affinity-matured version of this TCR, TCR-A3A, which has previously been used in clinical trials, resulted in patient deaths as a result of off-target toxicity elicited by HLA-A2 presenting a peptide from the cardiovascular tissue–derived TITIN molecule. We isolated several high-potency, low-affinity variants of the parental TCR that could facilitate the killing of MAGE-A3–positive cancer cell lines with physiological affinities (K_D ~ 10 to 50 μM). Furthermore, the catch bond–engineered TCR variants did not appreciably cross-react with TITIN peptide–pulsed cells. We used a yeast-displayed HLA-A1 peptide library to screen for cross-reactivity of the catch bond–engineered TCR variants. We found negligible cross-reactivity for predicted human self-antigens compared with their affinity-matured TCR-A3A counterparts.

CONCLUSION: We have shown that catch bond acquisition between TCRs and pMHCs is an engineerable parameter that can directly enhance TCR sensitivity while marginally affecting the 3D binding affinity. Furthermore, TCR sensitivity can be precisely fine-tuned by different levels of peak bond lifetime. Catch bond engineering of clinically useful, tumor-reactive TCRs is a viable alternative to affinity maturation for generating high-potency, low-affinity TCRs with reduced likelihoods of off-target toxicity for immunotherapy. ■

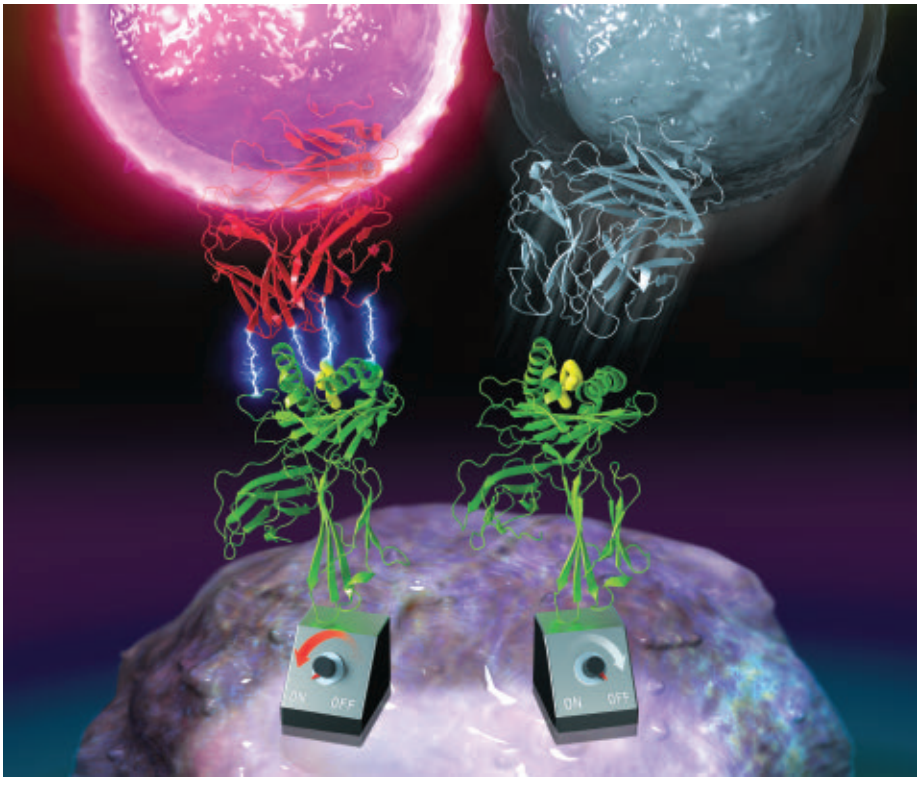


IMAGE CREATED BY ERIC SMITH AND CHRIS GARCIA

TCR catch bond engineering. An engineered TCR (left, red), with enriched catch bonds depicted as lightning bolts between pMHC and TCR, could trigger stronger T cell signaling compared with the signaling-off wild-type TCR (right, blue).

The list of author affiliations is available in the full article online.
*Corresponding author. Email: kcgarcia@stanford.edu
Cite this article as X. Zhao *et al.*, *Science* 376, eabl5282 (2022). DOI: 10.1126/science.abl5282

READ THE FULL ARTICLE AT
https://doi.org/10.1126/science.abl5282

RESEARCH ARTICLE

PROTEIN ENGINEERING

Tuning T cell receptor sensitivity through catch bond engineering

Xiang Zhao[1], Elizabeth M. Kolawole[2], Waipan Chan[3], Yinnian Feng[4], Xinbo Yang[1], Marvin H. Gee[1,5], Kevin M. Jude[1], Leah V. Sibener[1,5], Polly M. Fordyce[4,6,7,8], Ronald N. Germain[3], Brian D. Evavold[2], K. Christopher Garcia[1,9]*

Adoptive cell therapy using engineered T cell receptors (TCRs) is a promising approach for targeting cancer antigens, but tumor-reactive TCRs are often weakly responsive to their target ligands, peptide–major histocompatibility complexes (pMHCs). Affinity-matured TCRs can enhance the efficacy of TCR–T cell therapy but can also cross-react with off-target antigens, resulting in organ immunopathology. We developed an alternative strategy to isolate TCR mutants that exhibited high activation signals coupled with low-affinity pMHC binding through the acquisition of catch bonds. Engineered analogs of a tumor antigen MAGE-A3–specific TCR maintained physiological affinities while exhibiting enhanced target killing potency and undetectable cross-reactivity, compared with a high-affinity clinically tested TCR that exhibited lethal cross-reactivity with a cardiac antigen. Catch bond engineering is a biophysically based strategy to tune high-sensitivity TCRs for T cell therapy with reduced potential for adverse cross-reactivity.

T cells mediate many important aspects of cellular immunity, including the elimination of cells expressing cancer-related self-antigens. T cells express clonotypic T cell receptors (TCRs) that interact with specific peptides that are bound to and presented on the cell surface by major histocompatibility complex (MHC) molecules, known as pMHCs. Recognition of pMHCs by the TCR leads to activation of downstream signaling and effector functions in T cells, including cytokine secretion and target cell killing. The molecular and structural parameters that determine TCR sensitivity in response to pMHCs have been extensively studied but remain incompletely defined (*1*). TCR activation potency is often correlated with pMHC binding affinity, and TCR affinity maturation can result in TCRs with enhanced responsiveness to pMHC targets. However, the three-dimensional (3D) binding affinity generally fails to predict sensitivity, which suggests that additional mechanisms modulate TCR-pMHC interactions that result in functional intracellular signaling (*2*–*4*).

[1]Departments of Molecular and Cellular Physiology and Structural Biology, Stanford University School of Medicine, Stanford, CA 94305, USA. [2]Department of Pathology, University of Utah School of Medicine, Salt Lake City, UT 84132, USA. [3]Lymphocyte Biology Section, Laboratory of Immune System Biology, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD 20892, USA. [4]Department of Genetics, Stanford University, Stanford, CA 94305, USA. [5]Program in Immunology, Stanford University School of Medicine, Stanford, CA 94305, USA. [6]Department of Bioengineering, Stanford University, Stanford, CA 94305, USA. [7]ChEM-H Institute, Stanford University, Stanford, CA 94305, USA. [8]Chan Zuckerberg BioHub, San Francisco, CA 94158, USA. [9]Howard Hughes Medical Institute, Stanford University School of Medicine, Stanford, CA 94305, USA.
*Corresponding author. Email: kcgarcia@stanford.edu

Mechanical force has recently been shown to play a key role as a biophysical determinant of TCR triggering and signaling (*5*–*7*), with the TCR transforming cellular shear forces into biochemical signals when binding to agonist pMHC (*5*–*8*). Single-molecule force measurements on cells have shown that there is extended bond lifetime during productive antigenic pMHC-TCR interactions, referred to as catch bonds (*6*, *9*, *10*). There is a close correlation between the detection of catch bonds with a given TCR on a T cell and the agonist potency of a particular pMHC (*6*). Nonstimulatory pMHC ligands have also been identified that do not exhibit catch bonds but bind TCRs with solution affinities characteristic of many agonist TCR-pMHC interactions (*6*). Mutants of these nonstimulatory pMHC ligands that show agonist activity were found to have acquired catch bonds with the TCR, but they do not have substantially higher 3D affinities (*11*). Thus, in the environment of the T cell membrane, the presence or absence of catch bonds can act as a switch for TCR signaling and is not coupled to pMHC binding affinity (*11*). We aimed to take advantage of this cellular TCR triggering mechanism to address the limitations of current clinical TCRs used for cancer immunotherapy.

Adoptive T cell transfer [known as adoptive cell therapy (ACT)] with engineered T cells (TCR-T) [or chimeric antigen receptor (CAR)–T] is currently being used for cancer treatment (*12*, *13*). In this regimen, T cells are transduced with a tumor antigen–specific TCR or CAR, respectively, and then, after in vitro expansion of cell number, are administered into cancer patients (*14*). One advantage of TCR-T ACT over CAR-T is the natural sensitivity of TCRs to very low antigen densities on tumors. However, a drawback is that many tumor antigen–specific TCRs have low affinity for tumor-associated pMHCs that only weakly activate the TCR-T cells they bind to. To overcome this problem, a common strategy is to increase the affinity of the TCR for the tumor pMHC (*15*). However, in some cases, affinity-matured TCRs have shown substantial off-target toxicities (*14*, *16*, *17*). In fact, an affinity-matured TCR recognizing MAGE-A3, a promising tumor antigen, showed lethal off-target cross-reactivity with a cardiac peptide from the TITIN protein. High-affinity TCRs likely have a higher propensity to engage off-target pMHC ligands, so alternative approaches that bypass affinity maturation will be valuable for improving ACT with TCR-T cells. Here, we report an alternative TCR engineering strategy, which we call catch bond fishing, that harnesses a biophysical parameter mediating many adhesive cell surface protein-protein interactions.

Results

Design of catch bond fishing libraries

Our previous studies showed that TCR55 does not produce measurable T cell activation although it binds to an HIV peptide ($Pol_{448\text{-}456}$) presented by the human lymphocyte antigen (HLA)–B35 MHC molecule with physiological affinities. This TCR-pMHC interaction does not form catch bonds during the binding event (*11*). However, HIV peptide mutants isolated from HLA-B35 yeast pMHC libraries, such as pep20, gained the capacity to form catch bonds with TCR55 and potently activated T cells bearing this receptor while maintaining comparable affinity to the nonstimulatory parent pMHC (Fig. 1, A and B, and figs. S1 and S2) (*11*). We then investigated whether, in a reciprocal manner, a functional screen could isolate mutants of TCR55 that acquire catch bond capacity and enable functional T cell responses evoked by the nonstimulatory HIV peptide.

Although the source of catch bonds in force-dependent triggering has been attributed to multiple structural elements of the TCR (*18*), we focused our library design on the TCR-pMHC interface. Our TCR library design was guided by the biophysical characteristics of catch bonds, which are mediated by the transient formation of hydrogen bonds or salt bridges encountered during the TCR-pMHC shearing step that precedes disengagement. This leads to extended bond lifetimes that manifest as a transient resistive force before unbinding (*19*). Thus, our strategy was to lightly mutate the complementarity determining region (CDR) residues of TCR55 to encode polar or charged amino acids that would act as fishhooks (bait) to probe for H bonding and/or salt bridging residues (prey) on the pMHC binding surface during disengagement. We chose TCR CDR residue positions for the libraries that were

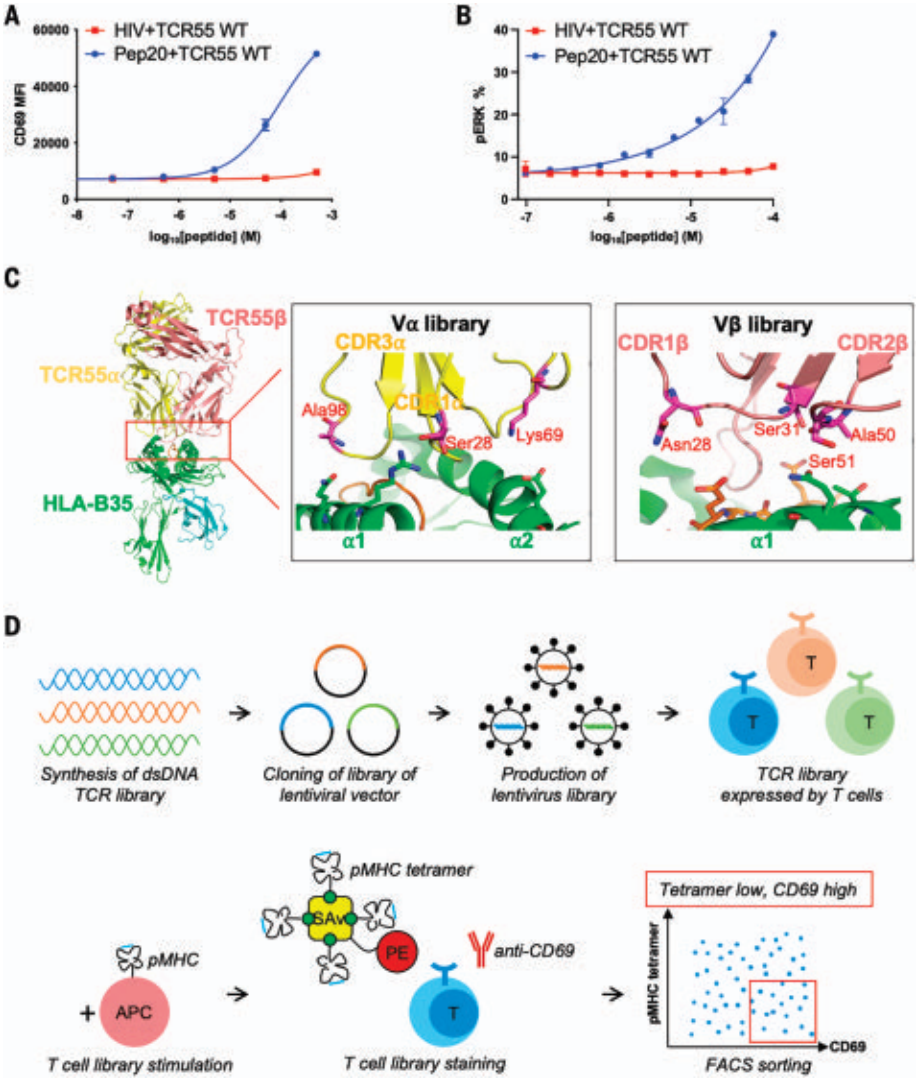


Fig. 1. The design of catch bond fishing libraries and selection strategy. (**A**) TCR55-transduced SKW3 T cells were stimulated by KG-1 cells pulsed with titrated HIV or Pep20 peptides for 14 hours. Anti-CD69 staining was performed on the SKW3 T cells and analyzed by flow cytometry. (**B**) TCR55-transduced SKW3 T cells were stimulated by KG-1 cells pulsed with titrated HIV or Pep20 peptides for 15 min. Anti–phospho-ERK staining was performed on the SKW3 T cells and analyzed by flow cytometry. (**C**) The design of TCR55 libraries. Each library has three or four residues selected to be randomized. The side chains of the residues selected for mutation on TCR55 are shown as sticks in the figure. (**D**) Workflow of catch bond engineering of TCR. [(A) and (B)] Data are representative of three independent experiments. Data are shown as means ± SDs of technical triplicates. APC, antigen-presenting cell; SAv, streptavidin; PE, phycoerythrin.

too distant from the pMHC to form direct contacts in the bound state to mitigate selecting for affinity-matured TCRs.

On the basis of the structure of the TCR55-HIV-B35 complex (*11*), three residues on the TCR55 α chain and four residues on the TCR55 β chain were selected for the library positions (Fig. 1C). Our library consisted of mainly charged and polar residues including glutamine, glutamate, asparagine, aspartate, arginine, lysine, serine, and histidine to increase the chances of forming adventitious polar interactions. The three randomized residues on the TCR55 α chain were combined as one library with a diversity of 1728 muteins (Vα library), and the four randomized residues on TCR55 β chain were combined as a second library with diversity of 20,736 muteins (Vβ library). Full-length TCR55 libraries were synthesized and cloned into a lentiviral backbone vector. Lentivirus libraries were constructed and used to infect the SKW3 T cell line at low multiplicity of infection (MOI), and TCR libraries were expressed on the surface of T cells. The Vα library was paired with the wild-type (WT) TCR55 β chain, and the Vβ library was paired with the WT TCR55 α chain in the transduced SKW3 cells. The libraries were stimulated with 10 μM HIV peptides and sorted for pMHC tetramer staining–low (no higher than the pMHC tetramer staining of WT TCR55) together with costaining for activation antigen CD69-high [top 5% population based on anti-CD69 mean fluorescence intensity (MFI)] populations to enrich for low-affinity, high-potency TCR mutants (Fig. 1D).

Single–amino acid substitutions in TCR55 trigger activation through catch bond formation

We carried out three rounds of fluorescence-activated cell sorting (FACS) selections on the TCR55α CDR library (diversity: 1728) and enriched a population with a tetramer-low, CD69-high staining phenotype (Fig. 2A and fig. S3, A and B). Approximately 100 single-cell clones were recovered and individually tested for activation by the HIV(Pol) peptide. The two clones (clone 8 and clone 17) that showed the most-potent responses to this pMHC ligand (fig. S3C) encoded identical TCR mutations on the TCR55 α chain—S28G and A98H. To directly examine whether the identified mutations conferred increased potency, SKW3 T cells were transduced with the TCR55α–S28G A98H and WT TCR55 β chain and stimulated by B35-associated HIV peptide (fig. S3D). To deconvolute which mutation was responsible for the activation, we tested the mutations individually (fig. S3, D and E) and found that the single mutation of alanine to histidine in the TCR55α CDR3 was sufficient to endow the nonresponsive TCR55 with the ability to be activated upon exposure to the B35-HIV pMHC (Fig. 2B and fig. S4).

The 3D affinity of TCR55α-A98H binding to the B35-HIV pMHC was measured by surface plasmon resonance (SPR) as K_D (3D binding affinity) = 5.9 μM, which is approximately threefold lower than the WT TCR55 binding to B35-HIV (K_D = 17 μM) but is still in the physiological affinity range for TCR-pMHC interactions and is higher than that measured for the binding of TCR589 to B35-HIV (K_D = 4 μM), a receptor-ligand pair with agonist qualities (Fig. 2C and fig. S3F) (*11*). Biomembrane force probe (BFP) experiments were conducted to determine whether TCR55α-A98H forms catch bonds with B35-HIV. The nonresponsive WT TCR55 showed progressively shorter bond lifetime with increasing force, which is consistent with slip bond formation. By contrast, application of force increased bond lifetime between TCR55α-A98H and B35-HIV, indicating catch bond formation (Fig. 2D). Analysis of the previously published structure of TCR55 bound to B35-HIV (*11*) suggests that the residues Q65 and T69 on the B35 MHC heavy chain molecule might form bonds with H98 on TCR55α (fig. S3G). Q65 or T69 was mutated to alanine, and only the Q65A mutation substantially abrogated the activation of TCR55α-A98H, which suggests that the triggering catch bond may involve an interaction between

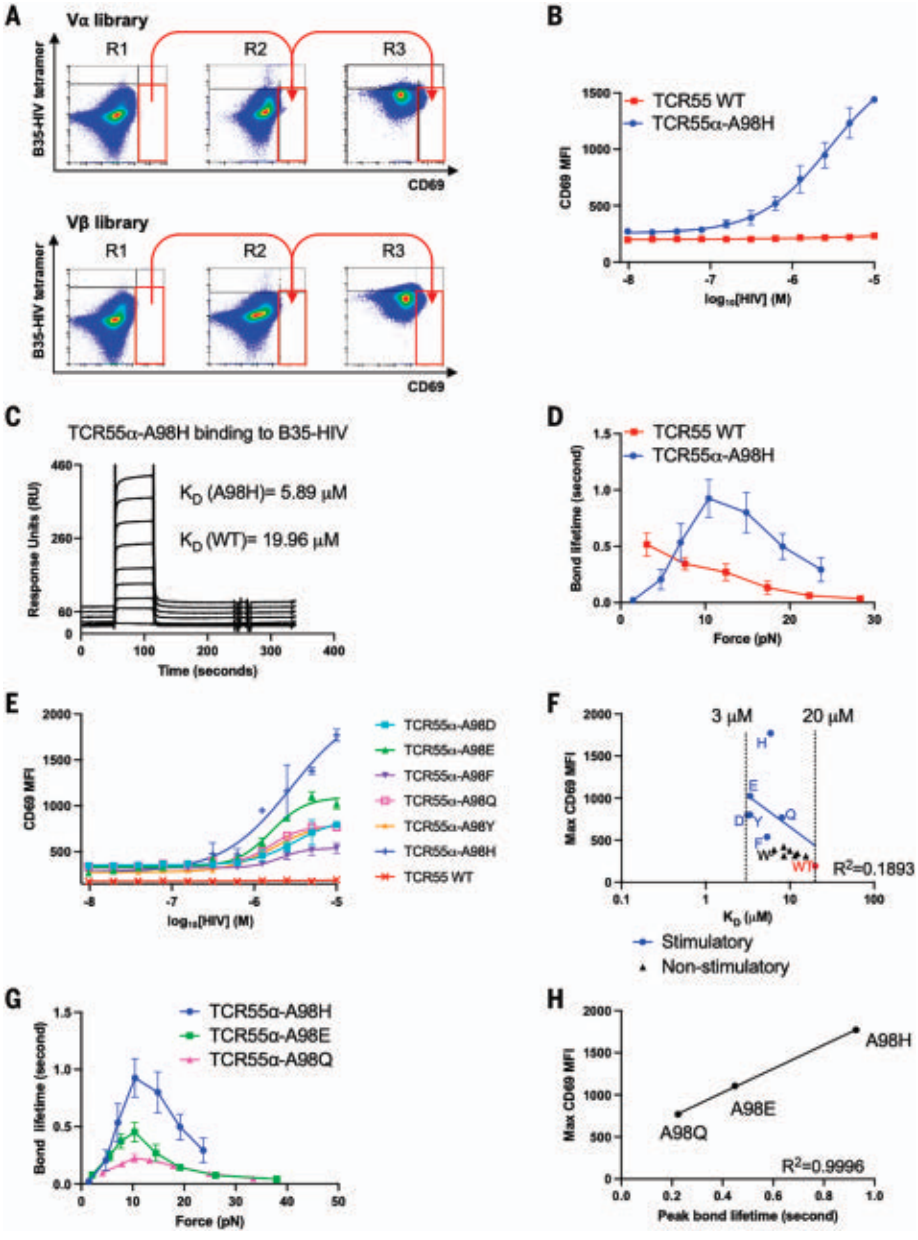


Fig. 2. A hotspot on the TCR can tune TCR signaling strength. (**A**) B35-HIV tetramer staining and anti-CD69 staining of cells transduced with library clones in each round of selection. The gate is based on the staining of WT TCR55. (**B**) A stimulatory clone, TCR55α-A98H, was selected from the library and was stimulated by KG-1 cells pulsed with titrated HIV peptides for 14 hours. Anti-CD69 staining was performed on the transduced SKW3 T cells and analyzed by flow cytometry. (**C**) SPR experiments of TCR55α-A98H protein binding to B35-HIV. Biotinylated B35-HIV monomer was immobilized on the streptavidin chip, and the TCR55α-A98H protein was flowed through the chip. (**D**) BFP experiments to measure bond lifetime force curves for TCR55α-A98H or TCR55 WT binding to B35-HIV. (**E**) TCR55α-A98 was mutated to D, E, F, Q, Y, and H and used to transduce SKW3 T cells with WT TCR55β. The transfectants were stimulated by KG-1 cells pulsed with titrated HIV peptides for 14 hours. Anti-CD69 staining was performed on the transduced SKW3 T cells and analyzed by flow cytometry. (**F**) Mean value of maximal anti-CD69 MFI versus 3D binding affinity K_D of TCR55α-A98 mutants transfectants. The linear correlation analysis was performed for stimulatory mutants and TCR55 WT. (**G**) BFP experiments to measure bond lifetime force curves for TCR55α-A98H, TCR55α-A98E, or TCR55α-A98Q T cell transfectants binding to B35-HIV. (**H**) Mean value of maximal anti-CD69 MFI versus peak bond lifetime of TCR55α-A98 mutants transfectants. [(B) and (E)] Data are representative of three independent experiments. Data are shown as means ± SDs of technical triplicates. [(D) and (G)] Data are shown as means ± SEMs of 500+ individual bond lifetimes per force curve. Single-letter abbreviations for the amino acid residues are as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.

B35-Q65 and TCR55α-A98H (fig. S3H). BFP showed that B35-Q65A–HIV formed catch bonds with TCR55α-A98H but exhibited shorter peak bond lifetimes than the B35-HIV–TCR55α-A98H interaction (fig. S3I).

Calibrating TCR55 signaling strength by bond lifetime

The acquisition of T cell activation by B35-HIV(Pol) coincident with catch bond formation by a single-point mutant of TCR55 provided an opportunity to investigate structure-function relationships between amino acid substitutions and activation strength. We mutated the TCR55α-A98 to 12 different amino acids to investigate how residue identity at this position affected the strength of TCR signaling. In addition to histidine, mutations to aspartate, glutamate, phenylalanine, glutamine, and tyrosine also enabled TCR55 signaling through B35-HIV(Pol) engagement for lymphocyte activation, albeit to different extents (Fig. 2E). By contrast, mutations to cysteine, lysine, asparagine, arginine, serine, threonine, and tryptophan did not activate TCR55 (fig. S5A). Therefore, only select polar, aromatic, and charged amino acids replacing residue TCR55α-A98 enabled effective signaling in response to B35-HIV. To investigate whether there was a correlation between signaling capacity and binding strength, we measured the 3D affinity by SPR for each of the different TCR55α-A98 mutants binding to B35-HIV pMHC. Most mutants have a 3D affinity in a narrow range between K_D = 3 μM and K_D = 20 μM (fig. S6 and table S1). Neither the maximum CD69 MFI [coefficient of determination (R^2) = 0.1893] nor the median effective concentration (EC_{50}) (R^2 = 0.02855) of stimulatory mutants was correlated to the SPR affinity of stimulatory mutants, which suggests that 3D affinity could not explain the gain of function exhibited by the stimulatory mutants (Fig. 2F and fig. S5B). TCR55α-A98W (K_D = 6.5 μM), a variant that exhibited higher affinity than WT-TCR55 (K_D = 19 μM), did not enable TCR-dependent activation in response to B35-HIV(Pol). Furthermore, the most ligand-sensitive of the TCR mutants, TCR55α-A98H (K_D = 5.9 μM), did not have the highest affinity (Fig. 2F and table S1). Based on BFP measurements for two B35-HIV responsive mutants—TCR55α-A98E and TCR55α-A98Q (Fig. 2G)—we found that the maximal effect (E_{max}) was correlated with the peak bond lifetime (R^2 = 0.996) rather than affinity (Fig. 2H). Thus, the strength of the catch bonds is a key parameter for the discrimination between agonist and nonagonist TCR-pMHC interactions.

We carried out a parallel screen on the TCR55β CDR library (diversity: 20,736) and identified a TCR55 variant, clone 36, that exhibited a high level of T cell activation by B35-HIV(Pol) (fig. S7, A and B). Clone 36 contained two mutations: a CDR1 mutation, TCR55β-N28Q,

and a CDR2 mutation, TCR55β-A50D. We identified the isolated TCR55β-A50D mutation as necessary and sufficient to enable T cell activation by B35-HIV (fig. S7C). Replacing the TCR55β-A50 position with alternative amino acids showed that aspartate, glutamate, phenylalanine, histidine, asparagine, glutamine, serine, threonine, and tyrosine supported TCR55 mutant responses to B35-HIV to different degrees, whereas cysteine, lysine, arginine, and tryptophan did not support effective signaling (fig. S7, D and E). The SPR 3D affinities of TCR55β-A50 mutants exhibited a range of K_D = 2 to 20 μM, similar to those of the TCR55α mutants and falling within the natural physiological range of TCR affinities (fig. S7, F and G; fig. S8; and table S2). There was a better correlation between maximal CD69 MFI versus K_D (R^2 = 0.7558) among the TCR55β-A50 mutants than among the TCR55α-A98 mutants (fig. S9A). However, the EC_{50} was not correlated with the 3D affinity (R^2 = 0.3543) (fig. S9B), which again suggests that affinity alone was not sufficient to explain the gain of function with these mutant TCRs. BFP experiments with the TCR55β-A50E, TCR55β-A50D, TCR55β-A50H, and TCR55β-A50T mutants (fig. S9C) again showed that peak bond lifetime correlated with E_{max} for TCR55β-A50 mutants stimulated by the B35-HIV pMHC ligand (R^2 = 0.8644) (fig. S9D). Analysis of the crystal structure of the TCR55-HIV-B35 complex (*11*) shows that residues T69 and Q72 on the B35-HIV pMHC potentially mediate the formation of new hydrogen bonds with TCR55β-A50E (fig. S9E). K562 cells transduced with B35-T69A prevented the activation of T cells bearing TCR55β-A50E, whereas the B35-Q72A mutation had no effect (fig. S9F). Consistent with these results, BFP measurements showed that B35-T69A–HIV only formed slip bonds with TCR55β-A50E (fig. S9G).

Signaling landscape of catch bond–engineered TCR

To assess how the catch bond–engineered TCR55 mutants affect intracellular signaling in T cells in response to B35-HIV pMHC ligand, we used a live cell imaging reporter system to measure the activation dynamics of the extracellular signal–regulated kinase (ERK), p38, and NFAT2 signaling pathways (fig. S10, A to E). In this system, translocation of fluorescent reporter molecules can be visualized in real time and quantified on a single-cell basis. Upon engagement with HIV peptide–pulsed B35-expressing antigen-presenting cells, reporter Jurkat T cells expressing the catch bond–engineered TCR variants displayed enhanced pathway activation when compared with the nonresponding parent TCR55, using the signaling-responsive TCR589 as a positive control (Fig. 3, A to C). Although both TCR55α-A98H and TCR55β-A50E mutants were able to activate the ERK and p38 signaling pathways for a similar duration at the population level, substantial differences in NFAT2 activation dynamics were observed (Fig. 3C). These results were quantified by single-cell area under the curve (AUC) analysis (Fig. 3, D to F, and tables S3 and S4), which demonstrated significant differences in both ERK and NFAT2 signaling responses for all the tested TCR variants. Because of the substantially lower signal-to-noise ratio of the p38–kinase translocation reporter (KTR), we observed more-subtle p38 signaling differences that follow the same hierarchy of mean AUC distribution compared with ERK or NFAT2 activation (Fig. 3, D to G). We find a strong correlation between mean ERK (R^2 = 0.9370) or NFAT2 (R^2 = 0.9415) AUC distribution and peak bond lifetime, which further supports the idea that catch bond strength plays a critical role in TCR-ligand engagements that result in functional intracellular signaling. (Fig. 3H).

Applied force activation of TCR at physiological pMHC density

To investigate the triggering of catch bond–engineered TCR55 at extremely low but physiologically relevant levels of pMHC (HIV–HLA-B35), we used the BATTLES (biomechanically-assisted T cell triggering for large-scale exogenous-pMHC screening) technique (*20*). The BATTLES technique uses temperature-sensitive polymer beads coated with pMHC proteins displayed at physiological densities (3 to 4.5 pMHCs per cell) to apply ramping forces (estimated maximum magnitude = 20 to 27.5 pN/s) to T cells interacting with bead surfaces (Fig. 3I) (*21*). Upon activation of force, we monitored Ca^{2+} signaling (which is correlated with initial T cell triggering) for >1000 SKW3 T cells transduced with engineered TCR55s containing either TCR55α-A98H, TCR55α-A98E, TCR55α-A98Q, TCR55β-A50E, TCR55β-A50H, TCR55β-A50D, or TCR55β-A50T substitutions interacting with HIV peptides (Fig. 3J). Although some T cells exhibited sustained increases in cellular Ca^{2+} flux (fig. S10F, top and middle rows), most cells showed decreasing fluorescence intensities and resulted in negative accumulated signals, indicating no triggering (fig. S10F, bottom row). This is consistent with prior literature showing that only a small fraction of T cells is activated at low pMHC densities, even with optimal force (*21*). All tested substitutions except TCR55β-A50T yielded higher integrated per-cell Ca^{2+} signals as compared with WT, with the magnitude of the integrated signal showing a strong correlation with measured peak bond lifetimes (Fig. 3K). These results, using force-induced activation of single T cells, provide evidence that engineered TCRs can drive efficient activation under the low-density pMHC conditions encountered in vivo.

Application of TCR catch bond engineering to TCR–T cell therapy

Catch bond engineering has implications for ACT with TCR–T cells because many WT tumor-reactive TCRs have low-affinity binding to tumor pMHC and low sensitivity to signaling in response to relevant tumor-associated antigens, which results in inefficient tumor killing (*22*–*24*). The melanoma antigen MAGE-A3–specific TCR (WT) was chosen for catch bond engineering. The antigen is HLA-A1 restricted with a reported K_D = 500 μM to the WT TCR (*16*, *25*). This TCR shows extremely poor T cell activation in response to the tumor antigen MAGE-A3, whereas an affinity-matured mutant of the WT MAGE-A3 TCR, A3A TCR, mediates greatly enhanced T cell activation by the same ligand (Fig. 4A). However, in clinical trials for melanoma, the A3A TCR was found to cross-react with HLA-A1–presented TITIN peptide, which is expressed mainly in cardiovascular tissue, leading to a high level of cardiotoxicity (Fig. 4A) (*16*, *17*). We explored whether we could use catch bond engineering to improve the sensitivity of the poorly responsive parental WT TCR to the MAGE-A3 ligand while maintaining low affinity to avoid cross-reactivity with TITIN.

We did not have a crystal structure of the low-affinity WT TCR complex with HLA-A1–MAGE-A3, but a structure of the affinity-matured version of the TCR with the HLA-A1–MAGE-A3 complex was available (*25*). We thus modeled the WT TCR binding to HLA-A1–MAGE-A3 and designed a library on the TCR α chain (Fig. 4B). Following the design strategy for TCR55, the residues chosen for the library (CDR1α positions 28 and 30 and CDR2α positions 52 and 54) fall within the CDR loops and are relatively close to the pMHC but do not directly contact the pMHC (Fig. 4B). The SKW3 T cell line was transduced with the library at low MOI, and CD69-high, tetramer-low clones were selected as described earlier (Fig. 4C and fig. S11A). After three rounds of selection, 96 single-cell clones were selected from the enriched population and tested for TCR-dependent activation. We isolated 13 distinct mutant-transduced SKW3 clones that showed enhanced responsiveness to the MAGE-A3 peptide at a concentration unable to trigger T cells expressing the parental WT TCR (Fig. 4, D and E, and table S5). By comparing the E_{max} of the TCR mutants, we defined eight clones as high-sensitivity mutants compared with the A3A TCR (Fig. 4D and fig. S11B) and five clones as intermediate-sensitivity mutants (Fig. 4E). We measured K_D for six high-sensitivity mutants and two intermediate-potency mutants binding to HLA-A1–MAGE-A3. The affinities ranged from K_D = 10 to 50 μM, substantially lower affinities than that of A3A (K_D = 1.24 μM) (fig. S12 and table S6). We did not observe a correlation between E_{max} versus 3D affinity (R^2 = 0.3718) (Fig. 4F) but observed

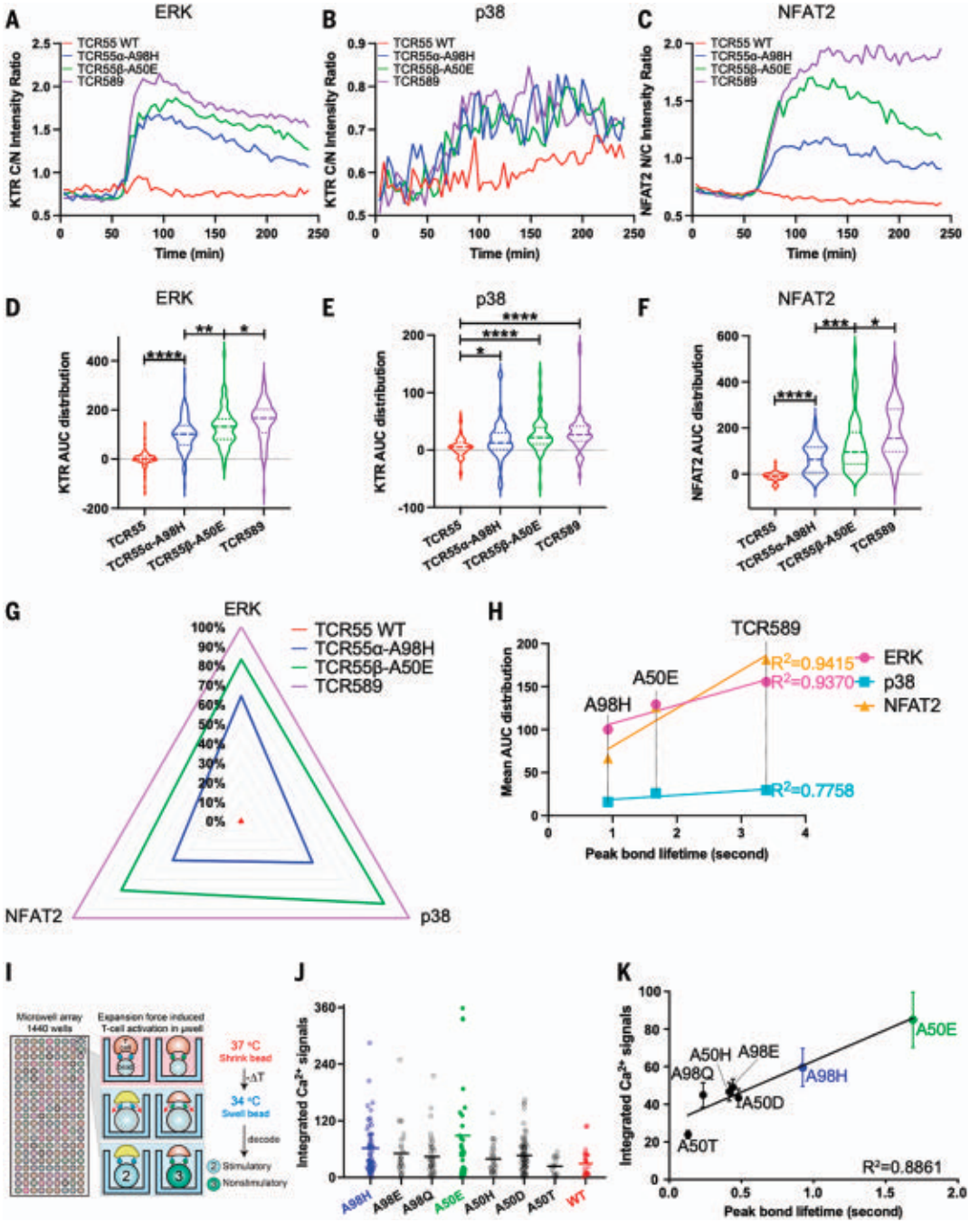


Fig. 3. Signaling landscape of catch bond–engineered TCRs. (**A**) ERK activation dynamics induced by B35-HIV engagement with the indicated TCR55 variant or TCR589, measured by ERK-KTR-mScarlet cytoplasmic/nuclear (C/N) intensity ratio over imaging time. (**B**) p38 activation dynamics measured by p38-KTR-mScarlet cytoplasmic/nuclear intensity ratio over imaging time. (**C**) NFAT2 activation dynamics measured by GFP1-11-NFAT2 nuclear/cytoplasmic intensity ratio over imaging time. (**D**) AUC distribution of single-cell ERK activation dynamics. (**E**) AUC distribution of single-cell p38 activation dynamics. (**F**) AUC distribution of single-cell NFAT2 activation dynamics. (**G**) Radar summary plot with normalized mean AUC values to illustrate the signaling landscape of TCR55 variant or TCR589 in response to B35-HIV engagement. (**H**) Mean ERK, p38, and NFAT2 AUC distributions versus peak bond lifetime measurements. (**I**) Schematic illustration of bead–T cell interaction in BATTLES. (**J**) Calcium flux signaling strength of different TCR55 mutant transfectants. Individual cell signals are shown as circular markers, and lines represent the mean values. (**K**) The correlation between calcium flux signaling strength and peak bond lifetime of different TCR55 mutant transfectants. Errors represent standard errors of the mean. [(A) to (F) and (I)] Data are representative of two independent experiments. *P < 0.05; **P < 0.01; ***P < 0.001; ****P < 0.0001.

a weak correlation between EC_{50} and affinity (R^2 = 0.5998) (fig. S11C). We tested whether the eight high-sensitivity mutants showed cross-reactive functional responses to the TITIN peptide. The A3A-transduced SKW3 cells were strongly activated by the TITIN pMHC ligand (Fig. 4G). Four mutants (20a-18, 20a-new 12, 94a-14, and 94a-30) exhibited no cross-reactivity with the TITIN peptide, whereas the remaining four displayed very weak activation by TITIN only at high peptide concentrations (Fig. 4G).

We also measured the binding affinity of all catch bond–engineered TCR mutants to HLA-A1–TITIN, and they had very low or unmeasurable 3D binding affinities (K_D > 100 μM), whereas the A3A affinity for TITIN was K_D = 7.7 μM (table S7 and fig. S13). BFP experiments were performed for WT TCR, A3A TCR, and TCR mutants 94a-14 and 20a-18, and all formed catch bonds with HLA-A1–MAGE-A3, with the mutant 94a-14 having a higher peak bond lifetime than A3A and WT TCR (Fig. 4H). The peak bond lifetimes of WT, A3A, 94a-14, and 20a-18 TCR were well correlated to the maximal CD69 MFI measured in Fig. 4D (R^2 = 0.9781) (Fig. 4I). A force of ~10 pN for a CD8-TCR-agonist has been demonstrated to promote optimal effector signaling (*6*, *8*–*10*). At ~10 pN of force, 94a-14 TCR has a significantly higher peak bond lifetime than both WT and A3A TCRs (Fig. 4J). BFP experiments for 94a-14 or 20a-18 TCR with HLA-A1–TITIN indicate that only slip bond formation was observed for both TCRs (fig. S14A), consistent with the loss of TITIN cross-reactivity by 94a-14 and 20a-18 TCRs.

To test whether the MAGE-A3 TCR mutants could efficiently kill HLA-A1–MAGE-A3$^+$ tumor cells, human primary T cells were transduced with the WT, A3A, and TCR mutants and cocultured with the HLA-A1–MAGE-A3$^+$ melanoma cell line A375 (Fig. 5, A to E) or the

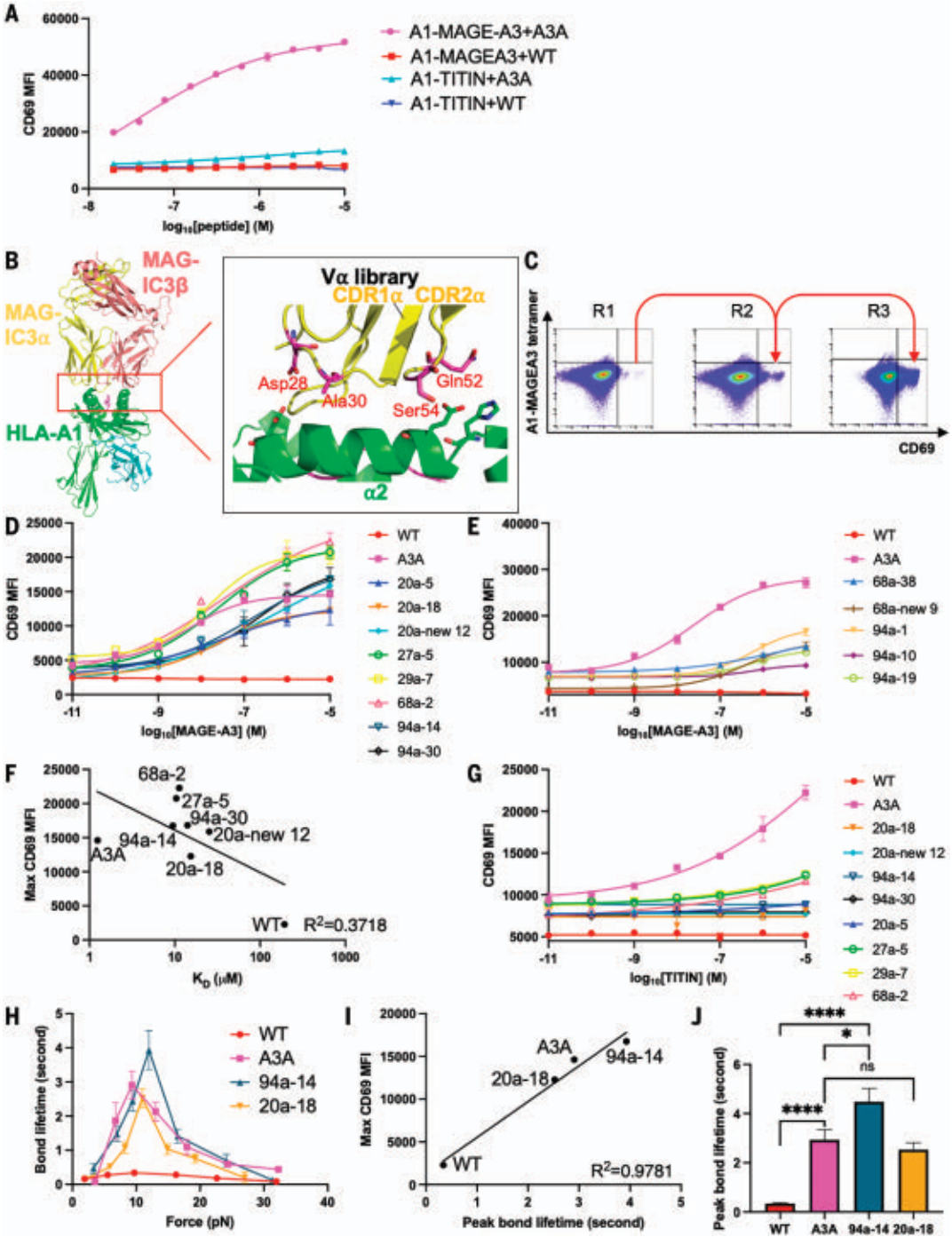


Fig. 4. Catch bond engineering of MAGE-A3–specific TCRs. (**A**) The WT TCR or A3A TCR chains were transduced in SKW3 T cells. The transfectants were stimulated by HLA-A1$^+$ 293T cells pulsed with titrated MAGE-A3 peptide or TITIN peptide. Anti-CD69 staining was performed on the T cells and analyzed by flow cytometry. (**B**) The design of the MAGE-A3 TCR Vα library. The library has four residues picked to be randomized. The side chains of the selected residues on the TCR are shown as sticks in the figure. (**C**) Three rounds of selection of the MAGE-A3 TCR Vα library on tetramer staining–low and anti-CD69 staining–high gate. The gate is based on the staining of MAGE-A3 WT TCR. (**D**) The eight high-potency MAGE-A3 TCR mutants were transduced into SKW3 T cells. The transfectants were stimulated by HLA-A1$^+$ 293T cells pulsed with titrated MAGE-A3 peptide. Anti-CD69 staining was performed on the T cells and analyzed by flow cytometry. (**E**) The five intermediate-potency MAGE-A3 TCR mutants were transduced into SKW3 T cells. The transfectants were stimulated by HLA-A1$^+$ 293T cells pulsed with titrated MAGE-A3 peptide. Anti-CD69 staining was performed on the T cells and analyzed by flow cytometry. (**F**) The correlation between mean value of maximal anti-CD69 MFI and 3D affinity of selected MAGE-A3 TCR mutants binding to HLA-A1–MAGE-A3. (**G**) The eight high-potency MAGE-A3 TCR mutants were transduced in SKW3 T cells. The transfectants were stimulated by HLA-A1$^+$ 293T cells pulsed with titrated TITIN peptide. Anti-CD69 staining was performed on the T cells and analyzed by flow cytometry. (**H**) BFP experiments to measure bond lifetime force curves for WT, A3A, 94a-14, or 20a-18 TCR binding to HLA-A1–MAGE-A3. Data are shown as means ± SEMs of 500+ individual bond lifetimes per force curve. (**I**) Mean value of maximal anti-CD69 MFI versus peak bond lifetime of MAGE-A3 TCR mutants transfectants. (**J**) Multiple measurements of bond lifetime at 10 pN for WT, A3A, 94a-14, and 20a-18 TCR. ns, not significant; *P < 0.05; **P < 0.01; ***P < 0.001; ****P < 0.0001. [(A), (D), (E), and (G)] Data are representative of three independent experiments. Data are shown as means ± SDs of technical triplicates.

HLA-A1–MAGE-A3$^+$ colon cancer cell line HCT-116 (Fig. 5, F to J, and fig. S15). In response to A375 cells, the engineered TCRs 94a-14 and 20a-18 were superior in target killing compared with the WT TCR and were at least comparable to—and in some cases superior to—A3A in target stimulated effector activity depending on the metric analyzed [interferon-γ (IFN-γ), tumor necrosis factor (TNF), or degranulation] (Fig. 5, A to E). Similar trends were seen in response to HCT-116 cells, which express lower levels of the MAGE-A3 antigen (Fig. 5, F to J). The mutants 20a-5 and 27a-5 were also tested in human primary T cells and showed a high level of cytotoxicity against A375 melanoma cells (fig. S11, D to H) and HCT-116 colon cancer cells (fig. S11, I to M).

To examine whether TCR clones 94a-14 and 20a-18 exhibited cross-reactivity to TITIN, primary human T cells transduced with the respective

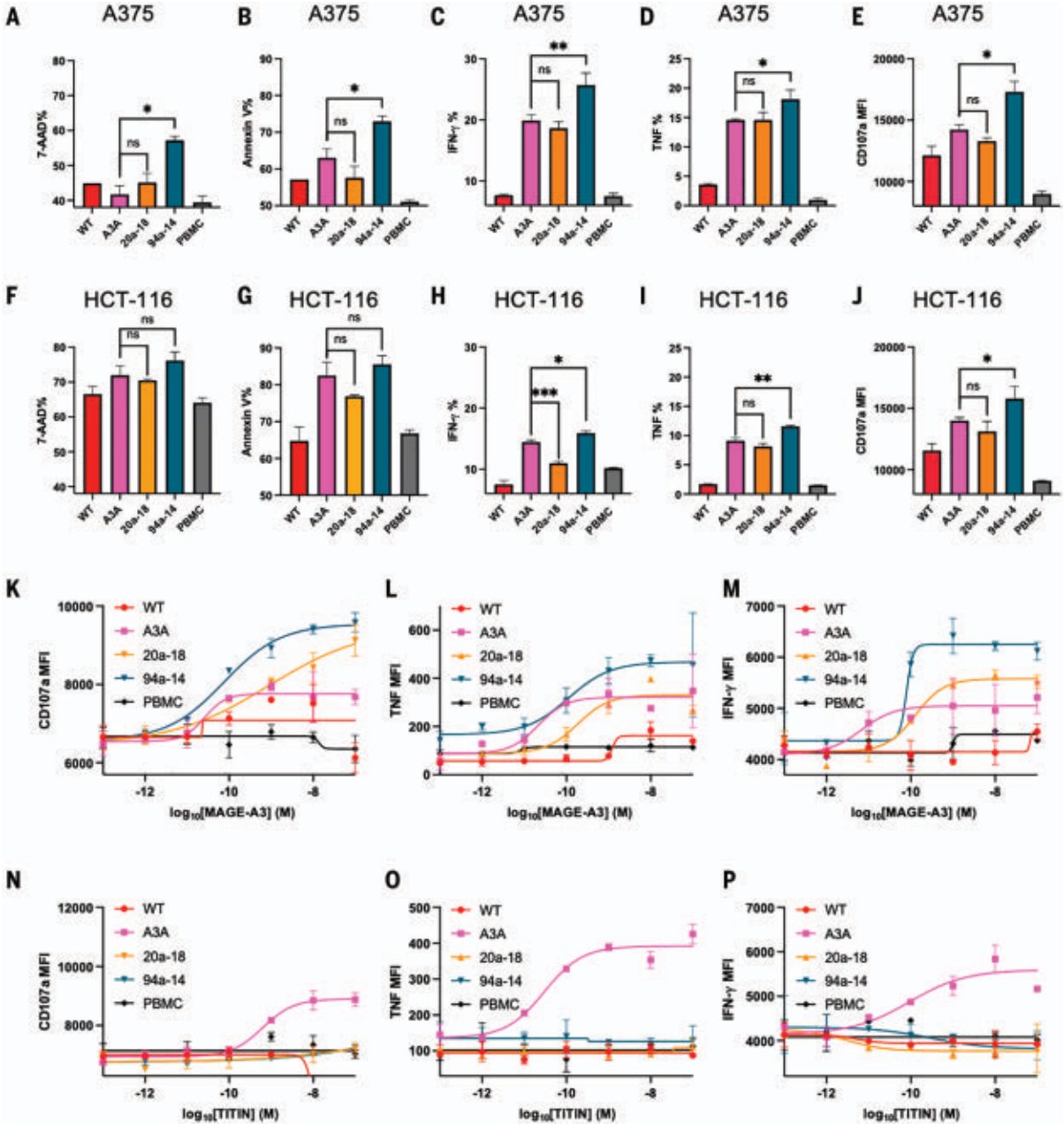


Fig. 5. Cytotoxicity and specificity of engineered MAGE-A3–specific TCR. (**A** and **B**) Killing of A375 melanoma cell line by different MAGE-A3–specific TCR transduced human primary T cells. (**C** to **E**) IFN-γ, TNF, and cytotoxic granule release (CD107a staining) by different MAGE-A3–specific TCR transduced human primary T cells, induced by the A375 melanoma cell line. (**F** and **G**) Killing of HCT-116 colon cancer cell line by different MAGE-A3–specific TCR transduced human primary T cells. (**H** to **J**) IFN-γ, TNF, and cytotoxic granule release (CD107a staining) by different MAGE-A3–specific TCR transduced human primary T cells, induced by the HCT-116 colon cancer cell line. (**K** to **M**) Cytotoxic granule release (CD107a staining), TNF, and IFN-γ by different MAGE-A3–specific TCR transduced human primary T cells, induced by HLA-A1^{+} 293T cells pulsed with a titration of MAGE-3 peptide. (**N** to **P**) Cytotoxic granule release (CD107a staining), TNF, and IFN-γ by different MAGE-A3–specific TCR transduced human primary T cells, induced by HLA-A1^{+} 293T cells pulsed with a titration of TITIN peptide. [(A) to (P)] Data are representative of three independent experiments. Data are shown as means ± SDs of technical duplicates. ns, not significant; *P < 0.05; **P < 0.01; ***P < 0.001; ****P < 0.0001.

TCRs were cocultured with MAGE-A3 or TITIN peptide–pulsed antigen-presenting cells. Although 20a-18 or 94a-14 showed enhanced cytotoxicity, degranulation, and cytokine secretion (Fig. 5, K to M) after coculturing with MAGE-A3–pulsed cells, none of these TCR clones responded to the presented TITIN peptide (Fig. 5, N to P). Similarly, the 20a-5 and 27a-5 clones mediated potent cytotoxic responses to MAGE-A3 (fig. S11, N to P) but only minimal cross-reactivity to TITIN at high concentrations of peptide (fig. S11, Q to S).

Profiling the cross-reactivity of engineered MAGE-A3 TCRs

Although the engineered TCRs lacked substantial reactivity with the TITIN peptide, we asked whether the engineered TCRs had acquired new, other off-target peptide reactivities. We turned

to a yeast-display pMHC library system originally used to characterize the cross-reactivity of TCRs (*26*, *27*) and to uncover the specificities of TCRs derived from tumor-resident T cells (*28*). We first generated an HLA-A*01 9–amino acid peptide library to survey the cross-reactive landscape of the WT, affinity-matured A3A and three catch bond–engineered MAGE-A3 TCR variants (Fig. 6A). The library was designed based on peptide sequences known to bind HLA-A*01, fixing anchor residues in positions P3 to aspartate and glutamate and P9 to tyrosine to ensure proper presentation of the peptides in the HLA groove. All remaining positions allowed flexibility to all 20 amino acids for a library diversity of 1.8×10^8.

We performed selections following established methods with soluble, recombinant forms of the WT MAGE-A3 TCR, A3A, 94a-14, 20a-18, or 94a-30. Although the WT TCR failed to enrich any yeast clones, presumably because of its very low 3D binding affinity (K_D > 500 μM) for MAGE-A3 (*16*), the high-affinity A3A and the engineered mutants strongly enriched populations of yeast clones (Fig. 6B). The selected library pools were sequenced to isolate individual sequences. The selected peptides showed strong convergence at the N-terminal end for all the TCR variants, with a lack of C-terminal specificity, as previously described for A3A (*29*) (Fig. 6C). Aside from the fixed anchor residues, P1 GLU, P4 PRO, and P5 ISO showed strong conservation and notably exist in both MAGE-A3 and TITIN peptides. The three catch bond–engineered TCR variants showed very similar sequence preferences, indicating that the specificities of the TCRs were minimally changed by catch bond engineering. The deep sequencing data were used to make off-target predictions using previously developed statistical methods (tables S8 to S11). For the A3A TCR, both TITIN and MAGE-A3 were top-ranked predictions, ranking as 1 and 7 respectively (table S8). However, for the three catch bond–engineered TCRs, TITIN was not predicted in the top 35 peptides, whereas the MAGE-A3 peptide was predicted to bind to all three catch bond–engineered TCRs—ranking as first for TCR 94a-14 (table S9), ranking as second for TCR 20a-18 (table S10), and ranking as 34th for TCR 94a-30 (table S11).

We tested the top 20 putative off-target predictions for the A3A TCR and catch bond–engineered TCRs with T cell activation assays. The top 20 predicted peptides for each TCR were synthesized and used for screening each TCR (60 peptides in total after removing repetitive peptides, listed in table S12). For the A3A TCR, we found that, in addition to MAGE-A3 and TITIN, it was also activated by two previously discovered epitopes, MAGE-A6 and FAT2 (*30*) (Fig. 6, D and E). For the three catch bond–engineered TCRs (94a-14, 20a-18, and 94a-30), only the MAGE-A3 peptide activated the T cells over baseline (Fig. 6, D and E). For the WT TCR, none of the peptides substantially stimulated the T cells compared with the dimethyl sulfoxide (DMSO) control (fig. S16A). The collective results of these cross-reactivity profiling experiments show that the screen could identify both known on- and off-target specificities for the high-affinity A3A TCR and that catch bond engineering did not introduce off-target specificities corresponding to known sequences in the human proteome. Although we cannot formally rule out the possibility that different types of cross-reactivity screens could identify off-target specificities that we did not find, the yeast-display pMHC screen represents a stringent test that shows the absence of unanticipated human antigen cross-reactivity while clearly identifying the source of cardiac toxicity seen with the A3A TCR.

Discussion

In environments where cell-cell interactions are subject to shear stresses, mechanical force plays an important role in signal transduction by a variety of receptor-ligand systems. Catch bonds have been observed as a natural signal-potentiating mechanism in various low-affinity cell surface adhesion systems, such as those involving cadherins; selectins; Notch; and, more recently, the TCR (*31*–*33*). Effective TCR signaling upon T cell engagement with an agonist pMHC ligand on an antigen-presenting cell involves the formation of catch bonds that extend receptor-ligand interaction lifetime upon application of a pulling force (*5*, *6*, *8*–*10*, *27*, *28*). The presence or absence of catch bonding residues in peptide antigens can decouple TCR triggering from conventional measurements of pMHC binding strength (*11*). In this work, by screening for mutant TCRs with a combination of modest solution affinity but high sensitivity to ligand-induced signaling, we show that TCRs with increased catch bond formation, as measured by BFP on T cells, dominate among the effective mutant TCRs isolated. These newly acquired catch bonds have not obviously predisposed the TCRs to increased human antigen cross-reactivity, as evident from screening pMHC libraries. This suggests that although a slow off rate, per se, can enable effective TCR signaling upon pMHC binding (*34*, *35*), catch bonds can play a deterministic role for antigen-responsive TCRs expressed on T cells. The degree to which catch bonds are contributed to by cellular factors such as membrane fluidity remains unknown (*36*).

The ease with which we identified such TCRs in the screen suggests that catch bonds may play a substantial role in the overall operational TCR repertoire and helps explain the existing discrepancies in the literature between measured solution binding affinities for specific pMHCs and the capacity of those pMHCs to show agonist properties in terms of T cell activation (*2*). The motility of T lymphocytes when scanning for ligand on antigen-presenting or target cells, along with the activity of cellular filipodia (*37*), provide tugging or shear forces that would favor prolongation of TCR-ligand interactions by catch bond formation to enable effective phosphatase exclusion as compared with intrinsic slow–off-rate binding that could be disrupted by such forces. This finding has direct implications for the emerging field of TCR-T therapy (*12*, *13*, *38*, *39*), where the inherently weak self-tumor reactivity of TCRs presents limitations to clinical activity.

Our selection strategy was critical to the successful isolation of ligand-sensitive yet low-affinity clones for several reasons. First, we focused our libraries on polar and charged residues that can maximize the likelihood of mutant substitutions engaging in adventitious polar interactions during TCR-pMHC disengagement. Second, we designed the libraries to focus on residues that were not in direct contact with the pMHC so that the selection did not simply isolate high-affinity (especially slow–off-rate) TCRs. We chose residues that were in the second shell, as it were, of TCR CDR residues—in close proximity to the pMHC surface but too distant to form direct interactions in the ground state complex. These residues would be ideally positioned to act as hooks during shearing of the TCR-pMHC interface. Third, our functional selection strategy directly isolated signaling active (CD69-high) but low-affinity (tetramer-low) clones. Although the 3D SPR K_D of the isolated clones does trend to slightly higher affinities than those of the the WT TCRs, the affinities remain firmly in the physiological regime, and K_D does not correlate with activity, validating the screening principles.

For our proof-of-concept studies, we used the TCR55-B35-HIV system because of the physiological binding affinity (K_D = 17 μM) of this TCR with the B35-HIV pMHC and the undetectable TCR activation after ligand binding (*11*, *40*). All the stimulatory single-site TCR mutants had affinities within the physiological regime (K_D ~ 2 μM to 20 μM), comparable to the WT TCR55, and showed different degrees of bond lifetime extension that correlated with activation strength. These results show that catch bond–engineered TCRs can be tuned for sensitivity through scanning different amino acid substitutions at hotspot positions. Such tunability allows for careful curation of clones with the desired balance of activation versus affinity. We emphasize that TCR signaling can be affected by both TCR affinity maturation and catch bond engineering. There was a weak positive correlation between the TCR mutants' sensitivity and affinity. However, catch bond engineering enables potency enhancement while

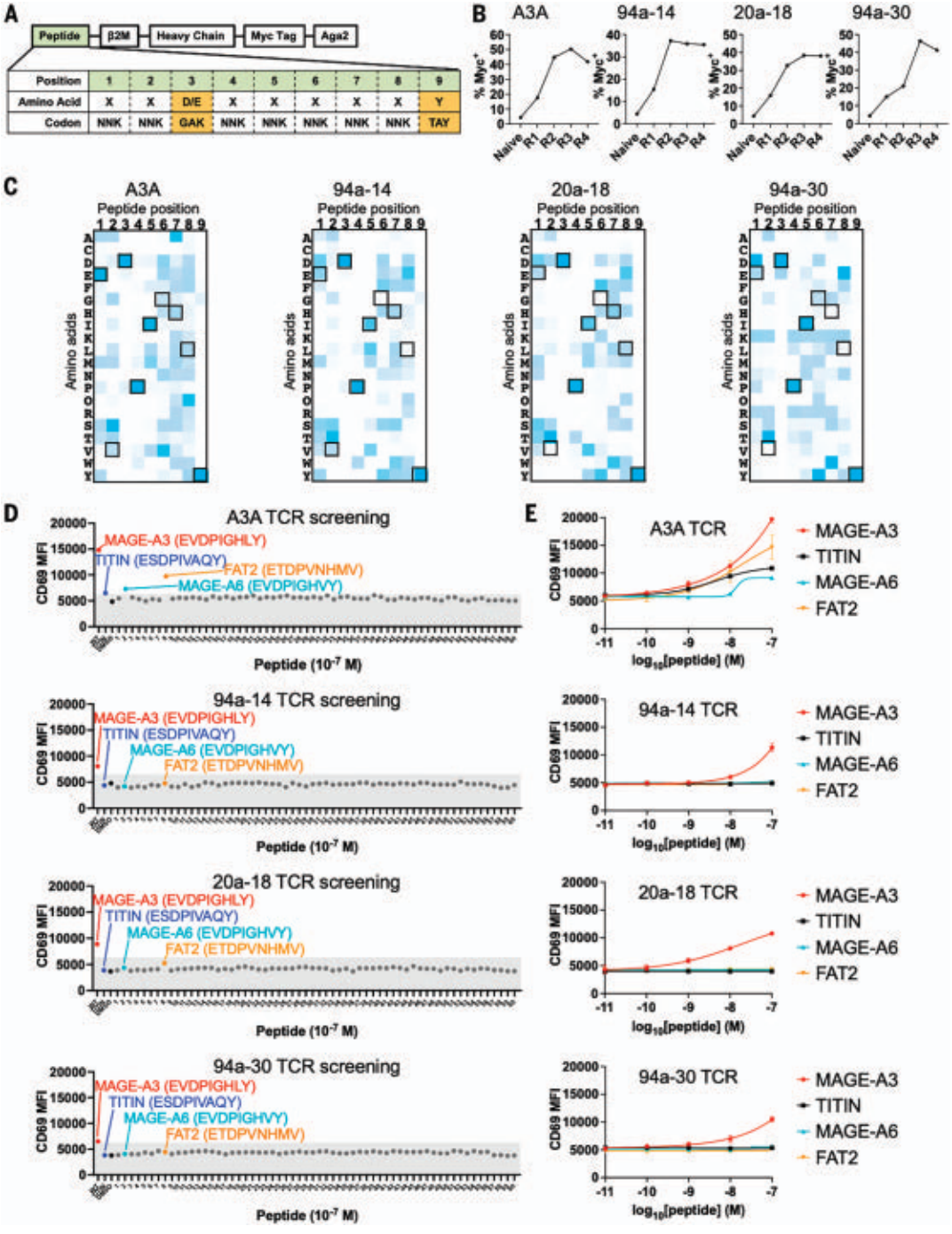


Fig. 6. Cross-reactivity screening of MAGE-A3 TCR variants with pMHC libraries. (**A**) Design of the single-chain HLA-A*01 yeast-display peptide library. The DNA peptide library design shows an NNK codon library for all positions except anchor positions P3 (GAK) and P9 (TAY) to maximize peptides displayed by HLA-A*01. The single-chain trimer construct is N-terminal to the Myc tag fused to Aga2 for expression on yeast. (**B**) Increasing myc tag expression on yeast over rounds of selection represents enrichment of peptide HLA-A*01 and positive selection of the library. (**C**) Heatmap of round 4 selected peptides showing peptide position by amino acid accounting for the number of reads detected per peptide. Boxed amino acids represent the MAGE-A3 peptide EVDPIGHLY. Dark blue represents a more enriched amino acid in that position. (**D**) MAGE-A3 (red dot), TITIN (blue dot), DMSO (black dot), and 60 predicted peptides (MAGE-A6, cyan dot; FAT2, orange dot) were used to pulse 293T–HLA-A1 cells to stimulate SKW3 T cells expressing different TCRs for 14 hours. Anti-CD69-APC staining was performed and analyzed on flow cytometry. (**E**) 293T–HLA-A1 cells were pulsed with titrated MAGE-A3, TITIN, MAGE-A6, or FAT2 peptides to stimulate SKW3 T cells expressing MAGE-A3 TCR variants for 14 hours. Anti-CD69-APC staining was performed and analyzed on flow cytometry. (D) Data are representative of two independent experiments. (E) Data are representative of two independent experiments. Data are shown as means ± SDs of technical duplicates.

maintaining physiologic affinity, reducing the predisposition toward off-target cross-reactivity compared with affinity-matured TCRs.

Although the safety of engineered T cell therapy will ultimately depend on the degree of preferential expression of the target tumor antigen versus healthy tissue, the strategy of catch bond engineering to maintain physiological affinity yet strong agonist signaling responses may help to reduce the chance of unwanted cross-reactivity with other pMHCs for clinically directed TCRs. Enhancing the efficacy of clinical TCRs has generally involved affinity maturation (*15*, *17*, *41*, *42*). However, some affinity-matured TCRs have displayed off-target toxicity (*17*, *43*). The extreme peptide selectivity of catch bond–engineered TCRs may even be helpful in mitigating on-target–off-tumor reactivities—for example, by enhancing therapeutic indices based on relative expression levels of unmutated self-tumor antigens, or neoantigens, with very close sequence similarity to a self-antigen expressed in healthy versus cancerous tissues (*43*). Given the relative ease with which we isolated such mutants and the simplicity of the screen, this lends itself well to a general approach in the TCR-T clinical development pipeline without requiring specialized structural insights.

Materials and methods

Cell lines

SKW3 T cells (DSMZ) were cultured in RPMI-1640+GluMax (Thermo Fisher Scientific) complemented with 10% fetal bovine serum (FBS) (Sigma-Aldrich), 10 mM HEPES, and 50 U/mL

Pen-Strep (Thermo Fisher Scientific) at 37°C and 5% CO_2.

LentiX cells and 293T cells were cultured in DMEM (Thermo Fisher Scientific) supplemented with 10% FBS, 2 mM L-Glutamine, 10 mM HEPES, and 50 U/mL Pen-Strep (Thermo Fisher Scientific) at 37°C and 5% CO_2.

KG-1 cells (ATCC) were cultured in IMDM (Thermo Fisher Scientific) supplemented with 10% FBS and 50 U/mL Pen-Strep (Thermo Fisher Scientific) at 37°C and 5% CO_2.

SF9 cells were cultured in SF900-III media (Thermo Fisher) supplemented with 10% FBS and 10 mg/mL gentamicin sulfate (Thermo Fisher) at 27°C and atmospheric CO_2.

Hi5 cells were grown in insect cell culture medium (Expression Systems) supplemented with 10 mg/mL gentamicin sulfate (Thermo Fisher) at 27°C and atmospheric CO_2.

Jurkat cell lines were cultured in RPMI 1640 supplemented with 10% FBS, 2 mM L-Glutamine, 50 U/mL Penicillin, 50 μg/mL Streptomycin, and 50 μM β-mercaptoethanol at 37°C and 5% CO_2.

Packaging of lentivirus

HEK293T-derived LentiX cells were seed in 6-well plate at a density of 3×10^5 cells/mL (2 mL in total). On the next day, for each well of cells, 750 ng plasmid of interest, 500 ng psPAX, 260 ng pMD2.G were mixed with 4.5 μL Fugene transfection reagent (Promega) in 100 μL Opti-MEM and rested for 20 min. Fresh cRPMI media were added to each well. Then, the DNA/Fugene mixture was added to each well. Optionally, 12 hours after the transfection, the supernatant of each well was replaced with 2 mL fresh cRPMI. 48 hours after the transfection, the supernatant was ready to infect 10^6 cells.

Cloning of TCR library

The double-stranded DNA (dsDNA) of the TCR library was synthesized commercially by GeneArt technology (Thermo Fisher Scientific) and was cloned into pHR lentiviral vector by HiFi assembly (New England Biolabs). Specifically, 20 ng dsDNA of TCR library, 100 ng linearized pHR vector, and 10 μL HiFi assembly mastermix were mixed and incubated at 50°C for 1 hour (eight replicates). 10 μL assembly product was analyzed on agarose gel to check the success of assembly. The remaining assembly product was purified by polymerase chain reaction (PCR) product clean up kit (Qiagen) and eluted in 30 μL water. The electrocompetent cells MegaX DH10B T1R Electrocomp Cells (Thermo Fisher Scientific) was defrosted on ice for 30 min. Then, 50 μL MegaX cells were mixed with 5 μL (>100 ng) HiFi assembly product. The tube was tapped three times and incubated on ice for 30 min. The bacteria-DNA mix was then transferred to chilled electroporation cuvette. The electroporation was conducted at 2.0 kV, 200 Ω, 25 μF. The cells were immediately recovered in 1000 μL SOC media. The competent cells culture was then recovered at 37°C, 225 rpm for 1 hour. After the recovery, 10 μL and 1000 μL cell culture was plated on the square bioassay dish (Corning) and cultured at 37°C overnight. The square bioassay dish plated with 10 μL culture was used for calculating the colony forming unit (cfu). All the colonies were scraped from the square bioassay dish and the plasmids were extracted by maxiprep (Qiagen).

TCR library display by T cells

Lentivirus of the TCR library was packaged by the method above. Lentivirus of TCR55 Vα library was titrated and coinfected SKW3 T cells with WT TCR55β lentivirus. Lentivirus of TCR55 Vβ library was titrated and coinfected SKW3 T cells with WT TCR55α lentivirus. Lentivirus of MAGE library was titrated and coinfected SKW3 T cells with WT MAGE-A3 TCRβ lentivirus. 48 hours after the infection, the percentage of TCR-positive population was determined by anti-CD3 (clone OKT3, BioLegend) staining and analyzed by flow cytometry. The titration of lentivirus that led to 20% infection efficiency was used to infect 100 to 200 million SKW3 T cells to have a low MOI. TCR-positive cells were sorted (Sony SH800S) and used for further sorting selection.

TCR library selection

Ten million KG-1 cells were labeled with carboxyfluorescein diacetate succinimidyl ester (CFSE) according to manufacturer's protocol (Thermo Fisher Scientific). The KG-1 cells were then pulsed with 10 μM HIV peptide for 3 hours at 37°C, 5% CO_2. The KG-1 cells were resuspended at 5×10^5 cells/mL and aliquoted into 96-well plate at 200 μL per well. The KG-1 cells were washed once to remove excess peptides. The library of 10 million T cells were resuspended at 5×10^5 cells/mL and aliquoted into the 96-well plate with KG-1 cells at 200 μL per well. After 14-hour activation, the cells were stained with anti-CD69-APC (clone FN50, BioLegend) and B35-HIV-PE tetramer (the method of making pMHC tetramer is described below) on ice for 30 min. Cells were sorted to select tetramer staining–low (comparable to TCR55 WT T cell's tetramer staining), anti-CD69 staining–high (top 5% in terms of anti-CD69 MFI) population. Cells were sorted into FBS to maintain cell health. Sorted cells were cultured in cRPMI. It took 2 weeks to grow enough cells to continue the next round of selection. After three to five rounds of selection, single-cell clones were obtained by diluting cells to 2.5 cells/mL and aliquoting 200 μL cell dilution to each well of 96-well U-bottom plate (Corning). It took 2 to 4 weeks to grow enough number of cells from single-cell clone. Each single-cell clone was tested by TCR55 signaling assay described below.

Sequencing of TCR mutants

Single-cell clones of SKW3 T cells with expected phenotype were used to extract genomic DNA according to the manufacturer's protocol. The TCR mutant DNA fragment was cloned by PCR and ligated into the pHR vector. The product of ligation was used to transform competent *Escherichia coli* cells and 30 single colonies was picked for sequencing the TCR mutants. More than one TCR sequence might be found in each single-cell clone (each T cell might still be transduced with more than one lentiviral particle at the beginning), and each TCR sequence should be tested individually by transducing SKW3 T cells for further TCR activation signaling assay.

TCR55 signaling assay

Peptide was dissolved and titrated in DMSO. KG-1 cells were labeled with CFSE and then resuspended at 5×10^5 cells/mL. 200 μL KG-1 cells were aliquoted to each well of 96-well U-bottom plate. KG-1 cells were pulsed with titrated peptides for 3 hours at 37°C, 5% CO_2. After that, KG-1 cells were washed once to remove excess peptides. SKW3 T cell transfectants were resuspended at 5×10^5 cells/mL and 200 μL T cells were added to each well with peptide-pulsed KG-1 cells. The stimulation was performed at 37°C, 5% CO_2 for 14 hours. After the stimulation, the cells were stained with anti-CD69-APC and anti-αβTCR-BV421 (clone IP26, BioLegend) on ice for 30 min and analyzed by CytoFLEX flow cytometer (Beckman). For phosphor-ERK staining, the stimulation was performed for only 15 min at 37°C, 5% CO_2. After the stimulation, the cells were immediately fixed with 4% PFA and shake for 15 min. The cells were then washed with PBS (0.5% BSA) and permeabilized in ice cold methanol for 30 min on ice. The cells were then washed with PBS (0.5% BSA) two times and stained with 1:50 dilution of anti-pERK1/2 (clone 197G2, Cell Signaling Technology) for 1 hour at room temperature with shaking. The cells were washed once and analyzed by CytoFLEX.

MAGE-A3–specific TCR signaling assay

MAGE-A3 (EVDPIGHLY) or TITIN (ESDPIVAQY) peptide (80% purity, Elim peptide) was dissolved and titrated in DMSO. HLA-A1–P2A–EGFP lentiviral vector was used to transfect HEK293T cells and green fluorescent protein (GFP)–positive cells were sorted and used as antigen-presenting cells (293T-A1). The 293T-A1 cells were resuspended at 5×10^5 cells/mL and pulsed with titrated peptide for 3 hours at 37°C, 5% CO_2. 200 μL KG-1 cells were aliquoted to each well of 96-well U-bottom plate. After the pulsing, the 293T-A1 cells were washed

once to remove excess peptides. MAGE-A3–specific TCR mutants-transduced SKW3 cells were resuspended at 5×10^5 cells/mL and 200 μL T cells were added to each well with peptide-pulsed 293T-A1 cells. The stimulation was performed at 37°C, 5% CO_2 for 14 hours. After the stimulation, the cells were stained with anti-CD69-APC and anti-Vβ5.1-BV421 (clone LC4, ThermoFisher Scientific) on ice for 30 min and analyzed by CytoFLEX flow cytometer (Beckman).

Transduction of human primary T cells with TCR

Human whole blood from healthy anonymous volunteer donors was purchased from Stanford Blook Bank under the approved protocol of APB-2749-KG1018. 6-well plate was coated with 2 mL of 2.5 μg/mL anti-CD3 (OKT3 clone) overnight. The next day, human peripheral blood mononuclear cells (PBMCs) were added to the plate with 5 μg/mL anti-CD28 and cultured at 37°C, 5% CO_2 for 3 days. 4 million LentiX cells were seed in 10-cm dish and transfected with lentiviral vector of MAGE-A3–specific TCR α chain or β chain. The lentivirus was made as described above. In total, 40 mL of TCR virus were concentrated to 500 μL using 100-kDa cutoff filter. 5 million preactivated human PBMCs were resuspended in 500 μL media and mixed with 500 μL concentrated TCR virus and 5 μg/mL Polybrene and 100 U/mL human IL-2. The virus-cells mixture was processed with spin infection under 2800 rpm, 32°C for 2 hours.

Killing assay of tumor cells

Twenty thousand A375 or HCT-116 cells were seed in each well of 96-well plate. 60,000 MAGE-A3–specific TCR-transduced human primary cells were added to each well with tumor cells and cocultured for 24 hours. The plate was washed in EDTA-free buffer and stained with 7-AAD (ThermoFisher Scientific) and Annexin V-APC (BioLegend) for 10 min. The plate was analyzed by CytoFLEX.

Cytotoxicity, cytokine, and granule release assays

Two hundred thousand tumor cells or peptide-pulsed 293T-A1 cells were seeded in each well of 96-well plate overnight. The next day, 200,000 MAGE-A3–specific TCR-transduced human primary cells were mixed with 1:100 anti-CD107a-PE (clone H4A3, BioLegend) and 1:1000 brefeldin A, and then added to each well. Coculture was done for 6 hours at 37°C, 5% CO_2. After 6 hours, the plate was stained with anti-CD8-BV421 (clone RPA-T8, BD Biosciences), anti-Vβ5.1-APC. Then the plate was fixed with IC fixation and permeabilized by permeabilization buffer. The plate was further stained with anti–IFN-γ–BV605 (clone B27, BioLegend) and anti-TNF-PE-Cy7 clone MAb11, BioLegend) on ice for 30 min. The plate was then washed and analyzed by CytoFLEX.

Production of MHC and β-2-microglobulin inclusion body

The protein of B35 MHC heavy chain and human β-2-microglobulin were made in *E. coli* as inclusion body. Specifically, B35 MHC heavy chain or human β-2-microglobulin was cloned into pET28a vector and transformed into BL21 (DE3) *E. coli* strain. Single colony was picked and resuspended in 10 mL LB media containing 50 μg/mL kanamycin and shake at 250 rpm, 37°C for 12 to 16 hours. Then the 10 mL culture was added into 1 L LB media containing 50 μg/mL kanamycin and shake at 250 rpm, 37°C for ~3 hours until the optical density (OD) = 0.5 to 0.6. Isopropyl-β-D-thiogalactopyranoside (IPTG) was added into the culture at final concentration of 1 mM and continued to shake for another 3 hours. The bacteria culture was spin down at 6000 rpm for 20 min. The bacteria pellet was resuspended in 50 mL buffer 1 [50 mM Tris-HCl, pH 8.0, 100 mM NaCl, 1 mM dithiothreitol (DTT), 5% Triton X-100, 1 mM EDTA, 0.2 mM phenylmethylsulfonyl fluoride (PMSF)]. Then the bacteria were sonicated under the program of 2 min sonication plus 2 min of rest. The sonication program was repeated four times continuously. After that, bacteria were spin 7500 rpm for 15 min. It was repeated two more times to resuspend the bacteria pellet in buffer 1 and do the sonication. The bacteria pellet was then resuspended in 50 mL buffer 2 (50 mM Tris-HCl, pH 8.0, 100 mM NaCl, 1 mM EDTA). Then the bacteria were sonicated under the program of 2 min sonication plus 2 min rest. The sonication program was repeated four times continuously. After that, bacteria were spin 7500 rpm for 15 min. It was repeated one more time to resuspend the bacteria pellet in buffer 2 and do the sonication. The inclusion body was pelleted and solubilized in 25 mL buffer (8 M urea, 50 mM Tris-HCl pH 8.0, 10 mM EDTA, 10 mM DTT).

Refolding of pMHC

Refolding buffer was prepared as 100 mM Tris-HCl pH 8, 400 mM arginine, 5 M urea, 0.5 mM oxidized glutathione, 5 mM reduced glutathione, 2 mM EDTA. 30 mg peptide was dissolved in DMSO and added to each liter of refolding buffer. For each liter of refolding buffer, 30 mg MHC heavy chain inclusion body and 30 mg human β-2-microglobulin inclusion body were mixed in a syringe and added into each liter of refolding buffer drop by drop. Then, the refold buffer/protein were poured into dialysis tubing (Spectrum Labs) and dialyzed into 10 L 10 mM Tris pH 8.0. The 10 L 10 mM Tris pH 8.0 buffer was changed every 12 hours and repeated four times in total. Then, the protein was purified by using weak anion exchange resin (DEAE Cellulose, Santa Cruz Biotechnologies). Specifically, DEAE-Cellulose was equilibrated with 10 mM Tris-HCl, pH 8.0 in a column. Then the dialyzed refolded protein solution flowed through the cellulose column drop by drop and repeated the flowing one more time. The refolded protein was eluted in 30 mL 10 mM Tris-HCl, pH 8.0 plus 0.5 M NaCl. The protein was buffer exchanged into 10 mM Tris-HCl, pH 8.0 and concentrated to 500 μL and biotinylated overnight. Biotinylated refolded protein was analyzed by size exclusion chromatography (Superdex 200, GE Healthcare) and ion exchange (MonoQ, GE Healthcare) on AKTAPurifier (GE Healthcare).

pMHC tetramer

For staining each 10 million cells, 20 μg biotinylated pMHC protein and 30 μg streptavidin-PE (Thermo Fisher Scientific) were aliquoted. 20% of total amount of streptavidin-PE were added into biotinylated pMHC each time at an interval time of 1 hour and repeated five times. During the interval time, the tetramer was incubated on ice. The pMHC tetramer was stored at 4°C overnight before using.

Production of TCR protein by Expi293

The TCR protein used for SPR was produced in Expi293 cells (Thermo Fisher Scientific). Specifically, TCR α chain was cloned into pD649 vector with basic zipper, and TCR β chain was cloned into pD649 vector with acid zipper. 15 μg TCR α chain constructs and 15 μg TCR β chain constructs were transfected into 75 million Expi293 cells according to the manufacturer's protocol. Four days after the transfection, the cell culture was spin down at 400 g for 5 min and the supernatant was saved. A onefold volume of PBS was added to the supernatant and final concentration of 20 mM Tris-HCl pH 8.0 buffer was added. 2 mL nickel-NTA was added to the supernatant and the solution was rotated overnight at 4°C. Then, the solution was flowed through a column to collect the Ni-NTA and bounded protein. 1× HBS pH 7.2 containing 10 mM imidazole was used to wash the Ni-NTA and protein once, and the protein was eluted by 1× HBS pH 7.2 containing 300 mM imidazole. The protein was concentrated in a 30-kDa filter (Millipore) and buffer exchanged in 1× HBS pH 7.2. The protein was purified by size-exclusion chromatography using Superdex200 column on AKTAPurifier (GE Healthcare). The purified protein was collected from the according fraction based on the size and run on SDS–polyacrylamide gel electrophoresis (SDS-PAGE) to check the size and 1:1 stoichiometry.

Production of TCR protein by insect cells

The TCR α chain was cloned into pAcGP67a vector with basic zipper, and the TCR β chain

was cloned into pAcGP67a vector with acid zipper. 2 μL baculovirus linear DNA and 2 μg TCR constructs were mixed with 100 μL Opti-MEM (Thermo Fisher Scientific) and 6.6 μL Fugene (Promega), and rest for 15 min. The mixture was added into 2 million SF9 cells and wait for 6 to 7 days. The cell culture was spin down at 2000 rpm for 8 min. The supernatant was saved as P0 virus. The P1 virus was made by adding 25 μL P0 virus to 25 mL SF9 cells at 2 million cells/mL. 25 mL media was added to the culture after 24 hours. Six to 7 days later, the P1 virus was collected by spinning down the cell culture at 2000 rpm for 8 min and saving the supernatant. The P1 virus of TCR α chain and TCR β chain was used and titrated to coinfect 2 million Hi5 cells to determine the optimal amount of P1 virus used to get the highest amount of 1:1 expression. Usually, 1 to 4 mL P1 virus for each chain was used for 1 L Hi5 cells (2 million cells/mL). Optimal amount of P1 virus of TCR α chain and TCR β chain was added to Hi5 cells. 72 hours after the coinfection, the cell culture was spin down at 1500 rpm for 15 min. The supernatant was collected, and for each liter of supernatant, 100 mL 1 M Tris pH 8.0, 1 mL 1 M $NiCl_2$, and 1 mL 5 M $CaCl_2$ was added and stirred for 30 min. After that, the solution was spin down at 6000 rpm for 15 min. The supernatant was collected and 3 mL Ni-NTA was added to each liter of the solution. The solution was stirred for 5 hours or overnight. Then, the solution was filtered through Buchner funnel and the Ni-NTA was transferred to a filter column. The protein-bound Ni-NTA was washed with 500 mL 1× HBS pH 7.2 containing 20 mM imidazole. Then, the protein was eluted with 15 mL 1× HBS pH 7.2 containing 300 mM imidazole. The protein was concentrated in a 30-kDa filter and washed once with 1xHBS pH 7.2. The protein was purified by size-exclusion chromatography using Superdex200 column on AKTAPurifier (GE Healthcare). The purified protein was collected from the according fraction based on the size and run on SDS-PAGE to check the size and 1:1 stoichiometry.

SPR

The affinity of TCR binding to the specific pMHC was measured by SPR on Biacore T100 (GE Healthcare). The refolded pMHC protein was biotinylated and immobilized on streptavidin chip (GE Healthcare). The TCR protein was treated with 3C protease to remove the basic/acid zipper. The pMHC protein was immobilized until a 100-200 RU increase, and the titrated TCR protein was flowed through the flow cell at 25°C. The affinity of the steady state was determined by the Biacore software. No surface regeneration was required because the sample completely returned to the baseline after the dissociation.

BFP assay

The BFP force clamp assay has previously been described in detail (*6*, *44*, *45*). In brief, a T cell of interest were aspirated onto a piezo driven micropipette controlled by Labview (National Instrument) programs. An opposing micropipette as an aspirated RBC biotinylated with EZ-link NHS-PEG-Biotin (Thermo Fisher Scientific). At the apex of this RBC was a streptavidin-maleimide (Sigma-Aldrich) bound glass bead coated with the pMHCs of interest [HLA B35–HIV($Pol_{448-456}$), B35-Pep20, A1-MAGE-A3 or A1-TITIN]. This RBC:bead complex served as a force probe sensor. Each T cell was repetitively brought into contact, held and then retracted to the distance controlled by the piezo actuator. The retraction and hold phase generated a force on the TCR:MHC bond, which could be altered, based on the distance the T cell was retracted. The position of the edge of the bead was tracked by the high-resolution camera (1600 frames per second) with <3 nm displacement precision. The camera then recorded the time it took for the T cell to disengage the glass bead, which can visually be seen by the RBC retracting and the bead returning to its starting position. Multiple repeated cycles (known as force-clamp cycles) could be carried at a single force to generate an average bond lifetime between the TCR and peptide:MHC complex. Varying the level of force and recording lifetimes allowed for the determination of the average bond lifetime and the type of bond formation.

Molecular cloning of TCR signaling reporter plasmids

LCAG-HBG and LEG11-NFAT2 lentiviral expression plasmids were created by Gibson Assembly cloning based on a split-GFP system described previously (*46*, *47*). EF1α-ERK-KTR-mScarlet or EF1α-p38-KTR-mScarlet lentiviral expression vector was generated by Gibson Assembly cloning based on an ERK-KTR-Clover or a p38-KTR-mCerulean3 plasmid from the Markus Covert laboratory (Addgene no. 59150 or no. 59155) (*48*).

Jurkat ERK and p38-NFAT2 reporter cell lines

To create a live cell nuclear marker with GFP1-10 expression, Jurkat cell line was transduced with the LCAG-HBG lentiviral expression vector. Stable H2B-tBFP+ Jurkat cells were isolated by FACS and transduced with the LE-EKS lentiviral expression vector. Stable ERK-KTR-mScarlet+ Jurkat cells were then isolated by FACS to create the ERK reporter cell line. To create the p38-NFAT2 reporter cell line, H2B-tBFP+ Jurkat cells were transduced with the LE-38KS and the LEG11-NFAT2 lentiviral expression vectors. Stable p38-KTR-mScarlet+ and GFP1-11-NFAT2+ Jurkat cells were isolated by FACS.

Live cell confocal microscopy

Live cell fluorescence time-lapse imaging data were collected using a Leica SP8 microscope with a 63× NA 1.4 oil objective (Biological Imaging Section, Research Technologies Branch, NIAID). Glass-bottom 8-well imaging chambers were coated with poly-D-lysine overnight at 4°C and washed twice with PBS. Cells were imaged in a heated 37°C environment with 5% CO_2. Imaging data were processed by Imaris Cell module, customized Batch analysis, and TranslocQ pipelines.

BATTLES

To produce thermo-responsive smart beads (~47 μm in diameter), we generated a mixture of N-isopropylacrylamide (NIAPM, 9.2% w/v), poly(ethylene glycol) diacrylate (PEGDA, MW = 700, 2.8437% v/v), lanthanide nanophosphors, sodium acrylate (1M, 5.5% v/v) and lithium phenyl-2,4,6-trimethylbenzoylphosphinate (LAP, 39.2 mg/mL, 2.5% v/v). We then injected this mixture and a fluorinated HFE7500 oil suspension with 2% ionic Krytox 157 FSH surfactant and 0.05% v/v acrylic acid into a microfluidic droplet generator to produce water-in-oil droplets that were subsequently polymerized into solid beads under flood UV light (IntelliRay, UV0338) at 100% amplitude (17.78 cm away from the lamp, power = ~50 to 60 mW/cm^2) for 2 min (*49*). After polymerization, carboxylated smart beads were washed with 2 mL dimethylformamide for 20 s; 2 mL dichloromethane for 10 s; and 2 mL methanol for 20 s before being resuspended in 1 mL PBST buffer. To coat smart beads with streptavidin, we preactivated ~200,000 beads with 1%w/v the N-(3-dimethylaminopropyl)-N′-ethylcarbodiimide hydrochloride (EDC) in 400 μL 0.1 M MES buffer (pH = 4.5) supplemented with 0.01% (v/v) Tween-20 for 3.5 hours at RT on an end-over-end rotator (10 rpm). The beads were spined down, washed with 1 mL 0.1 M borate buffer (pH = 8.5) supplemented with 0.01% (v/v) Tween-20 and subsequently resuspended in 400 μL of the same buffer. We then added 16 μL of streptavidin solution (dissolved in 1× PBS at 1 mg/mL) into the mixture and rotated the mixture overnight at 4°C. The next day, we quenched the conjugation reaction by adding 10 μL of 0.25 M ethanolamine in 0.1 M borate buffer (pH = 8.5) to the mixture and rotating for 30 min at 4°C. The final product was washed three times with PBST buffer, resuspended in 200 μL of the same buffer and stored at 4°C for further use. pMHC functionalized smart beads were generated by mixing 0.5 μL of 10 nM biotin-pMHCs with ~20,000 streptavidin smart beads in 50 μL PBST buffer. A PDMS microwell array (1440 wells) was then used to colocalized the pMHC coated beads and the calcium dye (Cal-250, 2μM) stained T cells. To exert mechanical load on bead-associated

T cells, the chip was heated to and maintained at 37°C for 1 min and then cooled to and kept at 34°C for 2 min. Immediately after cooling, we acquired a total of 150 Ca^{2+} fluorescence images at 4-s intervals. Integrated Ca^{2+} signals for single T cells were analyzed by ImageJ and a custom-written MATLAB code.

Yeast-display HLA-A1 peptide library

The yeast-display HLA-A1 peptide library was generated similarly to previously described protocol (*11*, *27*, *28*). To express the HLA-A1 peptide, a single-chain format of peptide library, β-2-microglobulin (β2M) and A1 heavy chain connected by linkers was fused N-terminal to Aga2. The A1 heavy chain contains a Y84A mutation to allow an opening at the terminal of MHC groove and a linker can connect the peptide with β2M. For the peptide library, P3 and P9 were set as anchoring residues with limited diversity: P3 as asparate or glutamate, P9 as tyrosine only. For other positions of peptide library, NNK codon was used to allow all 20 amino acids. The peptide library was synthesized as short nucleotide primers which were amplified via PCR to generate the single chain of pMHC-Aga2 inserts. To generate yeast-display HLA-A1 peptide library, competent EBY-100 yeast cells were electroporated with pMHC-Aga2 library inserts and linear pYAL vector. The pMHC-Aga2 library inserts were ligated to pYAL vector inside yeast cells via homologous recombination. By plating the initial yeast library at 1:10,000, 1:1,000, 1:100, and 1:10, the library size was calculated to have 1.8×10^8 functional diversity. The yeast library was grown in SDCAA pH 4.5 media. The yeast library was then induced to express the pMHC library protein by growing in SGCAA pH 4.5 media.

Selection of yeast-displayed HLA-A1 peptide library

Yeast-display HLA-A1 peptide library was selected with streptavidin-coated magnetic beads coated with biotinylated TCR proteins. The number of yeast cells used for each round of selection should be 10 times as high as the diversity of the last selection step (round 1 should use yeast cells number of 10 times of naïve library diversity). The yeast library was first incubated with 250 μL streptavidin magnetic beads in 10 mL PBE buffer (PBS +0.5% FBS+1 mM EDTA) and rotated at 4°C for 1 hour to do negative selection and remove unspecific binding to streptavidin magnetic beads. After incubation, the yeast-beads mixture was passed through an LS column (Miltenyi) and washed with PBE buffer three times, and all the flow-through was collected. Streptavidin magnetic beads coated with TCR protein was prepared by mixing 400 nM biotinylated TCR monomer with 250 μL streptavidin beads in 4.7 mL PBE buffer for 15 min at 4°C. The flow-through was incubated with TCR-beads for 3 hours at 4°C on a rotator. The yeast cells were washed and pelleted down at 5000 g for 1 min. The yeast cells were resuspended in 5 mL PBE buffer and passed through an LS column and washed with PBE buffer three times. The flow-through was discarded. The cells in the column were eluted by 5 mL PBE buffer and pelleted down. The pellet was washed one time with SDCAA media and resuspend again in 3 mL SDCAA media to grow overnight. When the OD is >2, yeast cells were induced in SGCAA for 2 to 3 days before the next round of selection. The yeast library was stained with specific TCR tetramer and anti-Myc antibody after each round of selection. The TCR tetramer was prepared at the final concentration of 400 nM by mixing TCR monomer and streptavidin-A647 at the ratio of 5:1. 100,000 yeast cells were stained with TCR tetramer and 2 μL anti-c-Myc-488 antibody (9402S, Cell Signaling) in 200 μL buffer. FACS plots were gated based on the yeast cells induced by SGCAA and stained with streptavidin-A647. Further rounds of selection were repeated with 10×10^8 yeast with only a modification done to the negative and positive selection using only 50 μL of streptavidin-coated beads with or without TCR in 500 μL of PBE.

Deep sequencing

Yeast DNA was extracted by Zymoprep II Kit (Zymo Research) for each round of selection from 50 million yeast cells. Barcoding PCR was first done for each DNA sample. The barcoding primes were designed as: Forward barcoding primer 5′ CTACACGACGCTCTTCCGATCTNNNNNNNN6 nucleotide barcode of your choice beginning of your sequence Tm (annealing) = 60 3′; Reverse barcoding primer 5′ end of your sequence Tm annealing = 60NNNNNNNNAGATCGGAAGAGCGGTTCAGCAGGAAT 3′. The barcoding PCR product was run on agarose gel and gel purified. Illumina PCR was then done by using the barcoding PCR product as template and specific Illumina PCR primers: Illumina F 5′AATGATACGGCGACCACACGAGTCTACACTCTTTCCCTACACGACGCTCTTCCGA 3′; Illumina R (order the reverse complement)- 5′GAAGAGCGGTTCAGCAGGAATGCCGAGACCGATCTCGTATGCCGTCTTCTGCTTG 3′. The PCR product was purified by gel extraction. The Illumina PCR product was quantified by nanodrop. The amount of each Illumina PCR product and water needed to obtain 40 μL 8 nM solution was calculated, aliquoted, and mixed together. We used the Illumina V2 2x300 cycle kit following the manufacturer's protocol for a low-diversity library.

Analysis of deep sequencing data and prediction of WT peptides from yeast selection

The sequencing results were first paired by PANDASEQ. The paired sequences were then imported into Geneious software to parse barcodes for each round of selection. Peptides were trimmed from the sequences and frequencies of amino acids were counted by custom Perl scripts used prior (*27*, *28*, *50*). To predict WT peptides for each TCR, a positional frequency matrix was determined based on peptides from round 3 selection. To score 9–amino acid peptides in the human proteome data, unique peptides counted more than 10 were used to generate position weight matrices (PWM). Each PWM from individual TCR selections were then used to predicted WT peptides from human proteome. The *Homo sapiens* proteome used was from UniProtKB (Proteome ID UP000005640; June 2020 update). Python was used for algorithm for weighted positional frequency matrix and ranking a reference proteome (*28*).

Screening of predicted WT peptides

The top 20 predicted WT peptides for TCR A3A, 94a-14, 20a-18, and 94a-30 were synthesized, and there were 59 different peptides all together after removing repetitive peptides. Because MAGE-A12 was shown to be cross-reactive in a previous study (*43*), the HLA-A1–restricted MAGE-A12 peptide was also synthesized and tested. In total, 60 different WT peptides were used to screen activity of different TCRs. Briefly, 100,000 293-A1 cells were pulsed with different WT peptides in each well of 96-well plate for 3 hours at 37°C, 5% CO_2. The 293-A1 cells were then washed with completed RPMI to remove excess peptides. 100,000 SKW3 cells expressing different TCRs were added to each well and cocultured for 14 hours at 37°C, 5% CO_2. Anti-CD69-APC and anti-TCR-BV421 staining of cells were done on ice and analyzed on flow cytometer. To do dose response of MAGE-A3, TITIN, MAGE-A6, and FAT2 peptides, 100,000 HLA-A1 cells were pulsed with titrated peptides in each well of 96-well plate for 3 hours at 37°C, 5% CO_2. The 293-A1 cells were then washed one time with completed RPMI to remove excess peptides. 100,000 SKW3 cells expressing different TCRs were added to each well and cocultured for 14 hours at 37°C, 5% CO_2. Anti-CD69-APC and anti-TCR-BV421 staining of cells were done on ice and analyzed on flow cytometer.

REFERENCES AND NOTES

1. J. D. Stone, A. S. Chervin, D. M. Kranz, T-cell receptor binding affinities and kinetics: Impact on T-cell activity and specificity. *Immunology* **126**, 165–176 (2009). doi: 10.1111/j.1365-2567.2008.03015.x; pmid: 19125887
2. M. Degano *et al*., A functional hot spot for antigen recognition in a superagonist TCR/MHC complex. *Immunity* **12**, 251–261 (2000). doi: 10.1016/S1074-7613(00)80178-8; pmid: 10755612
3. A. M. Kalergis *et al*., Efficient T cell activation requires an optimal dwell-time of interaction between the TCR and the pMHC complex. *Nat. Immunol.* **2**, 229–234 (2001). doi: 10.1038/85286; pmid: 11224522
4. E. M. Kolawole, T. J. Lamb, B. D. Evavold, Relationship of 2D Affinity to T Cell Functional Outcomes. *Int. J. Mol. Sci.* **21**, 7969 (2020). doi: 10.3390/ijms21217969; pmid: 33120989

5. S. T. Kim *et al*., The αβ T cell receptor is an anisotropic mechanosensor. *J. Biol. Chem.* **284**, 31028–31037 (2009). doi: 10.1074/jbc.M109.052712; pmid: 19755427
6. B. Liu, W. Chen, B. D. Evavold, C. Zhu, Accumulation of dynamic catch bonds between TCR and agonist peptide-MHC triggers T cell signaling. *Cell* **157**, 357–368 (2014). doi: 10.1016/j.cell.2014.02.053; pmid: 24725404
7. Y. Liu *et al*., DNA-based nanoparticle tension sensors reveal that T-cell receptors transmit defined pN forces to their antigens for enhanced fidelity. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 5610–5615 (2016). doi: 10.1073/pnas.1600163113; pmid: 27140637
8. Y. Feng *et al*., Mechanosensing drives acuity of *αβ* T-cell recognition. *Proc. Natl. Acad. Sci. U.S.A.* **114**, E8204–E8213 (2017). doi: 10.1073/pnas.1703559114; pmid: 28811364
9. D. K. Das *et al*., Force-dependent transition in the T-cell receptor β-subunit allosterically regulates peptide discrimination and pMHC bond lifetime. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 1517–1522 (2015). doi: 10.1073/pnas.1424829112; pmid: 25605925
10. B. Liu, E. M. Kolawole, B. D. Evavold, Mechanobiology of T Cell Activation: To Catch a Bond. *Annu. Rev. Cell Dev. Biol.* **37**, 65–87 (2021). doi: 10.1146/annurev-cellbio-120219-055100; pmid: 34213954
11. L. V. Sibener *et al*., Isolation of a Structural Mechanism for Uncoupling T Cell Receptor Signaling from Peptide-MHC Binding. *Cell* **174**, 672–687.e27 (2018). doi: 10.1016/j.cell.2018.06.017; pmid: 30053426
12. A. D. Fesnak, C. H. June, B. L. Levine, Engineered T cells: The promise and challenges of cancer immunotherapy. *Nat. Rev. Cancer* **16**, 566–581 (2016). doi: 10.1038/nrc.2016.97; pmid: 27550819
13. F. Manfredi *et al*., TCR Redirected T Cells for Cancer Treatment: Achievements, Hurdles, and Goals. *Front. Immunol.* **11**, 1689 (2020). doi: 10.3389/fimmu.2020.01689; pmid: 33013822
14. R. A. Morgan *et al*., Cancer regression in patients after transfer of genetically engineered lymphocytes. *Science* **314**, 126–129 (2006). doi: 10.1126/science.1129003; pmid: 16946036
15. A. P. Rapoport *et al*., NY-ESO-1–specific TCR–engineered T cells mediate sustained antigen-specific antitumor effects in myeloma. *Nat. Med.* **21**, 914–921 (2015). doi: 10.1038/nm.3910; pmid: 26193344
16. B. J. Cameron *et al*., Identification of a Titin-derived HLA-A1–presented peptide as a cross-reactive target for engineered MAGE A3–directed T cells. *Sci. Transl. Med.* **5**, 197ra103 (2013). doi: 10.1126/scitranslmed.3006034; pmid: 23926201
17. G. P. Linette *et al*., Cardiovascular toxicity and titin cross-reactivity of affinity-enhanced T cells in myeloma and melanoma. *Blood* **122**, 863–871 (2013). doi: 10.1182/blood-2013-03-490565; pmid: 23770775
18. W. Hwang, R. J. Mallis, M. J. Lang, E. L. Reinherz, The *αβ*TCR mechanosensor exploits dynamic ectodomain allostery to optimize its ligand recognition site. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 21336–21345 (2020). doi: 10.1073/pnas.2005899117; pmid: 32796106
19. P. Wu *et al*., Mechano-regulation of Peptide-MHC Class I Conformations Determines TCR Antigen Recognition. *Mol. Cell* **73**, 1015–1027.e7 (2019). doi: 10.1016/j.molcel.2018.12.018; pmid: 30711376
20. Y. Feng, X. Zhao, A. K. White, K. C. Garcia, P. M. Fordyce, Structure-activity mapping of the peptide- and force-dependent landscape of T-cell activation. *bioRxiv*2021.04.24.441194 [Preprint] (2021). .doi: 10.1101/2021.04.24.441194
21. K. L. Hui, L. Balagopalan, L. E. Samelson, A. Upadhyaya, Cytoskeletal forces during signaling activation in Jurkat T-cells. *Mol. Biol. Cell* **26**, 685–695 (2015). doi: 10.1091/mbc.E14-03-0830; pmid: 25518938
22. S. Zhong *et al*., T-cell receptor affinity and avidity defines antitumor response and autoimmunity in T-cell immunotherapy. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 6973–6978 (2013). doi: 10.1073/pnas.1221609110; pmid: 23576742
23. J. E. Thaxton, Z. Li, To affinity and beyond: Harnessing the T cell receptor for cancer immunotherapy. *Hum. Vaccin. Immunother.* **10**, 3313–3321 (2014). doi: 10.4161/21645515.2014.973314; pmid: 25483644
24. M. M. Hoffmann, J. E. Slansky, T-cell receptor affinity in the age of cancer immunotherapy. *Mol. Carcinog.* **59**, 862–870 (2020). doi: 10.1002/mc.23212; pmid: 32386086
25. M. C. C. Raman *et al*., Direct molecular mimicry enables off-target cardiovascular toxicity by an enhanced affinity TCR designed for cancer immunotherapy. *Sci. Rep.* **6**, 18851 (2016). doi: 10.1038/srep18851; pmid: 26758806
26. J. J. Adams *et al*., T cell receptor signaling is limited by docking geometry to peptide-major histocompatibility complex. *Immunity* **35**, 681–693 (2011). doi: 10.1016/j.immuni.2011.09.013; pmid: 22101157
27. M. E. Birnbaum *et al*., Deconstructing the peptide-MHC specificity of T cell recognition. *Cell* **157**, 1073–1087 (2014). doi: 10.1016/j.cell.2014.03.047; pmid: 24855945
28. M. H. Gee *et al*., Antigen Identification for Orphan T Cell Receptors Expressed on Tumor-Infiltrating Lymphocytes. *Cell* **172**, 549–563.e16 (2018). doi: 10.1016/j.cell.2017.11.043; pmid: 29275860
29. M. H. Gee, X. Yang, K. C. Garcia, Facile method for screening clinical T cell receptors for off-target peptide-HLA reactivity. *bioRxiv*472480 [Preprint] (2018). .doi: 10.1101/472480
30. T. Kula *et al*., T-Scan: A Genome-wide Method for the Systematic Discovery of T Cell Epitopes. *Cell* **178**, 1016–1028.e13 (2019). doi: 10.1016/j.cell.2019.07.009; pmid: 31398327
31. C. D. Buckley *et al*., The minimal cadherin-catenin complex binds to actin filaments under force. *Science* **346**, 1254211 (2014). doi: 10.1126/science.1254211; pmid: 25359979
32. B. T. Marshall *et al*., Direct observation of catch bonds involving cell-adhesion molecules. *Nature* **423**, 190–193 (2003). doi: 10.1038/nature01605; pmid: 12736689
33. V. C. Luca *et al*., Notch-Jagged complex structure implicates a catch bond in tuning ligand sensitivity. *Science* **355**, 1320–1324 (2017). doi: 10.1126/science.aaf9739; pmid: 28254785
34. R. M. Pielak *et al*., Early T cell receptor signals globally modulate ligand:receptor affinities during antigen discrimination. *Proc. Natl. Acad. Sci. U.S.A.* **114**, 12190–12195 (2017). doi: 10.1073/pnas.1613140114; pmid: 29087297
35. P. A. van der Merwe, O. Dushek, Mechanisms for T cell receptor triggering. *Nat. Rev. Immunol.* **11**, 47–55 (2011). doi: 10.1038/nri2887; pmid: 21127503
36. L. Limozin *et al*., TCR-pMHC kinetics under force in a cell-free system show no intrinsic catch bond, but a minimal encounter duration before binding. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 16943–16948 (2019). doi: 10.1073/pnas.1902141116; pmid: 31315981
37. E. Cai *et al*., Visualizing dynamic microvillar search and stabilization during ligand detection by T cells. *Science* **356**, eaal3118 (2017). doi: 10.1126/science.aal3118; pmid: 28495700
38. S. A. Rosenberg, N. P. Restifo, J. C. Yang, R. A. Morgan, M. E. Dudley, Adoptive cell transfer: A clinical path to effective cancer immunotherapy. *Nat. Rev. Cancer* **8**, 299–308 (2008). doi: 10.1038/nrc2355; pmid: 18354418
39. L. Zhao, Y. J. Cao, Engineered T Cell Therapy for Cancer in the Clinic. *Front. Immunol.* **10**, 2250 (2019). doi: 10.3389/fimmu.2019.02250; pmid: 31681259
40. T. Ueno, H. Tomiyama, M. Fujiwara, S. Oka, M. Takiguchi, Functionally impaired HIV-specific CD8 T cells show high affinity TCR-ligand interactions. *J. Immunol.* **173**, 5451–5457 (2004). doi: 10.4049/jimmunol.173.9.5451; pmid: 15494492
41. L. Poncette, X. Chen, F. K. M. Lorenz, T. Blankenstein, Effective NY-ESO-1–specific MHC II–restricted T cell receptors from antigen-negative hosts enhance tumor regression. *J. Clin. Invest.* **129**, 324–335 (2019). doi: 10.1172/JCI120391; pmid: 30530988
42. A. Isser, J. P. Schneck, High-affinity T cell receptors for adoptive cell transfer. *J. Clin. Invest.* **129**, 69–71 (2019). doi: 10.1172/JCI125471; pmid: 30530992
43. R. A. Morgan *et al*., Cancer regression and neurological toxicity following anti-MAGE-A3 TCR gene therapy. *J. Immunother.* **36**, 133–151 (2013). doi: 10.1097/CJI.0b013e3182829903; pmid: 23377668
44. Y. Chen *et al*., Fluorescence Biomembrane Force Probe: Concurrent Quantitation of Receptor-ligand Kinetics and Binding-induced Intracellular Signaling on a Single Cell. *J. Vis. Exp.* **102**, e52975 (2015). doi: 10.3791/52975; pmid: 26274371
45. W. Chen, V. I. Zarnitsyna, K. K. Sarangapani, J. Huang, C. Zhu, Measuring Receptor-Ligand Binding Kinetics on Cell Surfaces: From Adhesion Frequency to Thermal Fluctuation Methods. *Cell. Mol. Bioeng.* **1**, 276–288 (2008). doi: 10.1007/s12195-008-0024-8; pmid: 19890486
46. S. Cabantous, T. C. Terwilliger, G. S. Waldo, Protein tagging and detection with engineered self-assembling fragments of green fluorescent protein. *Nat. Biotechnol.* **23**, 102–107 (2005). doi: 10.1038/nbt1044; pmid: 15580262
47. S. Cabantous *et al*., A new protein-protein interaction sensor based on tripartite split-GFP association. *Sci. Rep.* **3**, 2854–2859 (2013). doi: 10.1038/srep02854; pmid: 24092409
48. S. Regot, J. J. Hughey, B. T. Bajar, S. Carrasco, M. W. Covert, High-sensitivity measurements of multiple kinase activities in live single cells. *Cell* **157**, 1724–1734 (2014). doi: 10.1016/j.cell.2014.04.039; pmid: 24949979
49. Y. Feng, A. K. White, J. B. Hein, E. A. Appel, P. M. Fordyce, MRBLES 2.0: High-throughput generation of chemically functionalized spectrally and magnetically encoded hydrogel beads using a simple single-layer microfluidic device. *Microsyst. Nanoeng.* **6**, 109–113 (2020). doi: 10.1038/s41378-020-00220-3; pmid: 33299601
50. M. H. Gee *et al*., Stress-testing the relationship between T cell receptor/peptide-MHC affinity and cross-reactivity using peptide velcro. *Proc. Natl. Acad. Sci. U.S.A.* **115**, E7369–E7378 (2018). doi: 10.1073/pnas.1802746115; pmid: 30021852

ACKNOWLEDGMENTS

We thank M. Yen, R. Fernandes, C. Glassman, L. Su, J. Rodrigues, and F. Liu for reading the manuscript, helpful discussions, and/or reagents. **Funding:** K.C.G. is supported by NIH grant 5R01AI103867, the Howard Hughes Medical Institute, the Parker Foundation for Cancer Immunotherapy, the Mathers Foundation, and a Bio-X seed grant. W.C. and R.N.G. are supported by the Intramural Research Program of the National Institute of Allergy and Infectious Diseases, National Institutes of Health. B.D.E. is supported by NIH grants R01 AI147641 and R01 NS071518. P.M.F. was partially supported by NIH grants 1DP2GM123641 and R01GM107132 and a Stanford Bio-X Interdisciplinary Initiatives seed grant. P.M.F. is a Chan Zuckerberg Biohub investigator and acknowledges the support of a Sloan Research Foundation Fellowship. Y.F. is a Cancer Research Institute Postdoctoral Fellow. Part of this work was performed at the Stanford Nano Shared Facilities (SNSF), supported by the National Science Foundation under award ECCS-1542152. **Author contributions:** K.C.G. conceived of the project. X.Z. and K.C.G. designed the overall experimental strategy. K.C.G. and X.Z. wrote the manuscript. X.Z., K.M.J., X.Y., K.C.G., and L.V.S. designed the TCR libraries. X.Z. performed lentivirus production, transduction of TCR libraries, selection of TCR libraries, screening of single-cell clones, and validation of activation of deconvoluted TCRs. X.Z. performed all the TCR activation flow cytometry assays. X.Z. performed protein expression, protein purification, and SPR experiments. X.Z. performed human primary T cells transduction, killing assays, and cytotoxicity assays. E.M.K. performed BFP experiments. R.N.G. and W.C. designed and W.C. performed the Jurkat signaling reporter microscopy imaging experiments. Y.F. performed BATTLES experiments. M.H.G. and X.Y. performed yeast peptide–MHC selection. X.Z. and X.Y. did deep sequencing and analyzed yeast selection data. X.Z. did predicted peptides screening. P.M.F., R.N.G., B.D.E., and K.C.G. supervised the research. All authors edited the manuscript. **Competing interests:** X.Z. and K.C.G. are coinventors of a patent (serial no. US 63/158, 131) covering the use of engineered MAGE-A3 TCR sequences for T cell immunotherapy. M.H.G., L.V.S., and K.C.G. are cofounders of 3T Biosciences. Y.F., P.M.F., X.Z., and K.C.G. are coinventors of a patent (serial no. US 63/108.162) covering the BATTLES microfluidics platform. The authors declare no other competing interests. **Data and materials availability:** All data are available in the main text or the supplementary materials. Requests for resources and reagents should be directed to the corresponding author.

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abl5282
Figs. S1 to S16
Tables S1 to S12
MDAR Reproducibility Checklist

22 July 2021; resubmitted 19 December 2021
Accepted 8 March 2022
10.1126/science.abl5282

RESEARCH ARTICLES

VIROME

Cryptic and abundant marine viruses at the evolutionary origins of Earth's RNA virome

Ahmed A. Zayed[1,2,3]†, James M. Wainaina[1,3]†, Guillermo Dominguez-Huerta[1,2,3]†, Eric Pelletier[4,5], Jiarong Guo[1,2,3], Mohamed Mohssen[1,3,6], Funing Tian[1,3], Akbar Adjie Pratama[1,2], Benjamin Bolduc[1,2,3], Olivier Zablocki[1,2,3], Dylan Cronin[1,2,3], Lindsey Solden[1], Erwan Delage[5,7], Adriana Alberti[4,5]§, Jean-Marc Aury[4,5], Quentin Carradec[4,5], Corinne da Silva[4,5], Karine Labadie[4,5], Julie Poulain[4,5], Hans-Joachim Ruscheweyh[8], Guillem Salazar[8], Elan Shatoff[9], Tara Oceans Coordinators‡, Ralf Bundschuh[6,9,10,11], Kurt Fredrick[1], Laura S. Kubatko[12,13], Samuel Chaffron[5,7], Alexander I. Culley[14], Shinichi Sunagawa[8], Jens H. Kuhn[15], Patrick Wincker[4,5], Matthew B. Sullivan[1,2,3,6,12,16]*

Whereas DNA viruses are known to be abundant, diverse, and commonly key ecosystem players, RNA viruses are insufficiently studied outside disease settings. In this study, we analyzed ≈28 terabases of Global Ocean RNA sequences to expand Earth's RNA virus catalogs and their taxonomy, investigate their evolutionary origins, and assess their marine biogeography from pole to pole. Using new approaches to optimize discovery and classification, we identified RNA viruses that necessitate substantive revisions of taxonomy (doubling phyla and adding >50% new classes) and evolutionary understanding. "Species"-rank abundance determination revealed that viruses of the new phyla "*Taraviricota*," a missing link in early RNA virus evolution, and "*Arctiviricota*" are widespread and dominant in the oceans. These efforts provide foundational knowledge critical to integrating RNA viruses into ecological and epidemiological models.

RNA viruses of 47 of 103 established families included in the riboviriad (with RNA genomes) kingdom *Orthornavirae* [orthornavirans; encoding an RNA-directed RNA polymerase (RdRp) for replication] have been studied deeply and mechanistically for their roles in human, livestock, and plant diseases (*1*–*3*). The remaining viruses are less well studied because they infect less economically critical but nevertheless ecologically essential organisms, such as invertebrates, fungi, protists, and bacteria. Not surprisingly, virus discovery efforts, largely by using environmental RNA sequencing, have recently forced drastic changes in our understanding of orthornaviran diversity and evolution (*4*–*7*). Specifically, these studies have expanded diversity within known orthornaviran groups (*4*–*6*), revealed altered genome architecture among viruses with broad host ranges (*4*), and posited large host range jumps as driving much of orthornaviran evolution (*8*, *9*).

Because the gene encoding RdRp is ancient, thought to be among the first genes of the peptide-RNA world (*10*–*12*), it serves as a deep evolutionary gene marker and is often used to understand orthornaviran origins and more generally to explore the origins of life (*7*, *12*–*15*). Recently, RdRp-inferred orthornaviran evolutionary relationships resolved five major branches (*7*), which were subsequently recognized by the International Committee on Taxonomy of Viruses (ICTV) as five phyla (*16*, *17*). This five-branch phylogenetic structure that underpins current orthornaviran megataxonomy was hypothesized to be stable, and the question of whether phylum-rank diversity was saturated was opened (*5*, *17*). Beyond taxonomy, the evolutionary origins of orthornavirans, because of challenges in deep phylogenetic inferences (*18*), remain contentious, puzzling, and complex (*19*–*21*). Also problematic is that environmental surveys lack scalable and systematic approaches to taxonomically classify new data and assess their impact on our understanding of orthornaviran evolution.

In this study, we update several key analytics and apply these to ≈28 terabases (Tb) of Global Ocean RNA metatranscriptome sequences to identify and characterize previously unknown RNA viruses and use them to (i) test hypotheses about orthornaviran megataxonomy stability and evolutionary origins and (ii) establish baseline planetary-scale ocean biogeographic context.

Marine RNA viruses double known orthornaviran phyla from 5 to 10

Given how little RNA virus diversity is explored in the Global Ocean (tables S1 and S2), we sought to leverage systematically collected and globally distributed *Tara* Oceans resources (table S3). These include RNA-sequencing data from 771 metatranscriptomes (table S4 for sample metadata) that span 10 organismal size fractions (fig. S1), three ocean layers, and 121 locations distributed throughout the world's five oceans and include ≈6 Tb of new sequencing data from 143 metatranscriptomes obtained throughout the Arctic Ocean (Fig. 1A and table S4). To maximize our inferences from these metatranscriptomes, we developed and/or improved and benchmarked methods for the identification, classification, and organization of the orthornaviran genome–derived sequence space.

We first searched our Global Ocean data for nucleic acids that encode RdRps, which are specific to orthornavirans and have no known relationship to cellular RdRps (*22*) or DNA-directed RNA polymerases (*23*). Given notoriously divergent RdRp sequences, we maximized RdRp identification by means of an iterative search-and-update hidden Markov model (HMM) approach that we improved and automated in our work (supplementary materials, materials and methods, and fig. S2). This approach identified 44,779 RdRp-encoding contigs (after removing 134 false positives) (materials and methods and fig. S2C) (details per contig are available in table S5), a ≈26-fold improvement over standard BLAST (Basic Local Alignment Search Tool)–based approaches (fig. S2G). Of these 44,779 contigs, 6686 encoded complete or near-complete RdRp domain sequences (≥90% completeness) (materials and methods).

Because the oceans are vastly undersampled for orthornavirans, we sought to assess how these new data compared with the current five-branch understanding of orthornaviran megataxonomy (*7*). This introduced our second major analytical challenge because although this phylogeny-based unified framework

[1]Department of Microbiology, Ohio State University, Columbus, OH 43210, USA. [2]EMERGE Biology Integration Institute, Ohio State University, Columbus, OH 43210, USA. [3]Center of Microbiome Science, Ohio State University, Columbus, OH 43210, USA. [4]Génomique Métabolique, Genoscope, Institut François-Jacob, CEA, CNRS, Univ Evry, Université Paris-Saclay, 91000 Evry, France. [5]Research Federation for the Study of Global Ocean Systems Ecology and Evolution, FR2022/Tara Oceans GOSEE, 75016 Paris, France. [6]The Interdisciplinary Biophysics Graduate Program, Ohio State University, Columbus, OH 43210, USA. [7]Nantes Université, CNRS UMR 6004, LS2N, F-44000 Nantes, France. [8]Department of Biology, Institute of Microbiology and Swiss Institute of Bioinformatics, ETH Zurich, Zurich, Switzerland. [9]Department of Physics, Ohio State University, Columbus, OH 43210, USA. [10]Department of Chemistry and Biochemistry, Ohio State University, Columbus, OH 43210, USA. [11]Division of Hematology, Department of Internal Medicine, Ohio State University, Columbus, OH 43210, USA. [12]Department of Evolution, Ecology, and Organismal Biology, Ohio State University, Columbus, OH 43210, USA. [13]Department of Statistics, Ohio State University, Columbus, OH 43210, USA. [14]Département de Biochimie, Microbiologie et Bio-informatique, Université Laval, Québec, Québec G1V 0A6, Canada. [15]Integrated Research Facility at Fort Detrick, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Fort Detrick, Frederick, MD 21702, USA. [16]Department of Civil, Environmental, and Geodetic Engineering, Ohio State University, Columbus, OH 43210, USA.
*Corresponding author. Email: sullivan.948@osu.edu
†These authors contributed equally to this work.
‡The *Tara* Oceans Coordinators are listed in the supplementary materials.
§Present address: Université Paris-Saclay, CEA, CNRS, Institute for Integrative Biology of the Cell (I2BC), 91198, Gif-sur-Yvette, France.

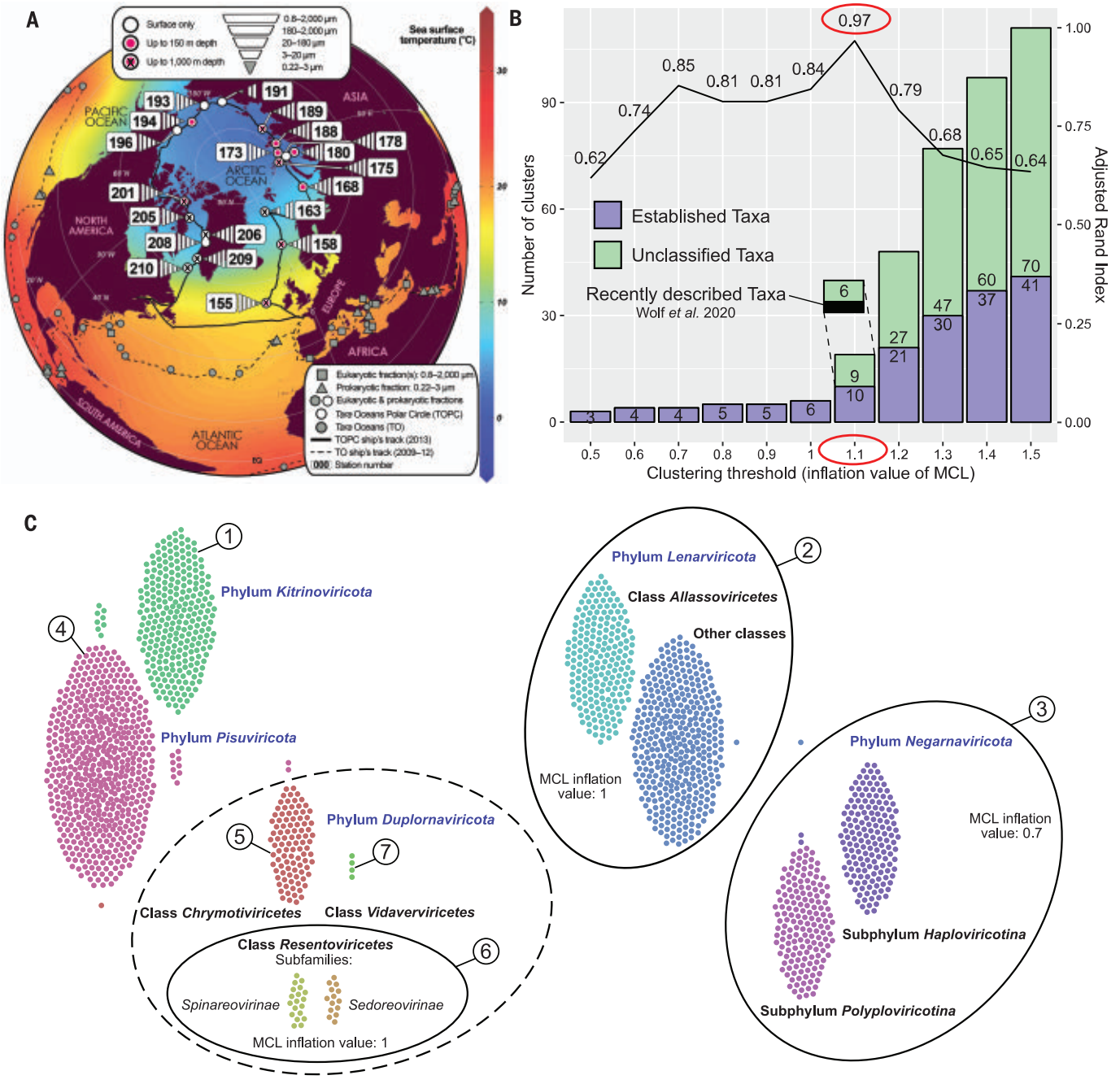


Fig. 1. Establishment of RdRp domain megaclusters. (**A**) Arctic projection of the Global Ocean highlighting the new size-fractionated metatranscriptomes described here (white polygons). Gray symbols indicate previously published metatranscriptomes, whereas numbered stations indicate circumpolar Arctic Ocean data. Sea surface temperature gridding was done by using the weighted-average method in Ocean Data View (*43*) from the in situ temperature measurements collected during *Tara* expeditions. TO, *Tara* Oceans; TOPC, *Tara* Oceans Polar Circle. (**B**) Percent agreement (line) of our network-guided and phylogeny-based mega-taxonomy at different clustering thresholds (materials and methods). Stacked bars represent the number of taxonomic clusters of near-complete RdRp domains (at least 90% of the domain) (materials and methods) at these different clustering thresholds. Only sequences representing established taxa (violet) were used for calculating the agreement percentage. At an inflation value of 1.1, three (black box) of the nine unclassified clusters have been recently described by Wolf *et al*. (*5*), bringing the number of new major taxa in our study to six. (**C**) Swarm plot of the 10 ICTV-established taxa emerging at an inflation value 1.1 in the Markov Clustering Algorithm (MCL) analysis [from (A)]. Solid lines encompass taxa that were exclusively joined at a lower inflation value, as indicated within each ellipse. The dashed line encompasses the three established duplornaviricot classes, which were never exclusively joined at lower inflation values. Dots that have the same color but are not part of their swarm represent discrepancies from GenBank taxonomy (aligned vertically with the cluster that recruited them in the network). The resultant seven clusters (numbered) along with the six new clusters from our study (A) were used to build the 13 individual phylogenetic trees in Fig. 2A. Phylum *Kitrinoviricota* encompasses two of the three recently described unclassified megaclusters (A) at an MCL inflation value of 1. The third megacluster represents viruses with permuted motifs in the RdRp domain ("permutotetra-like" and "birna-like" viruses) and hence was excluded from phylogenetic analyses.

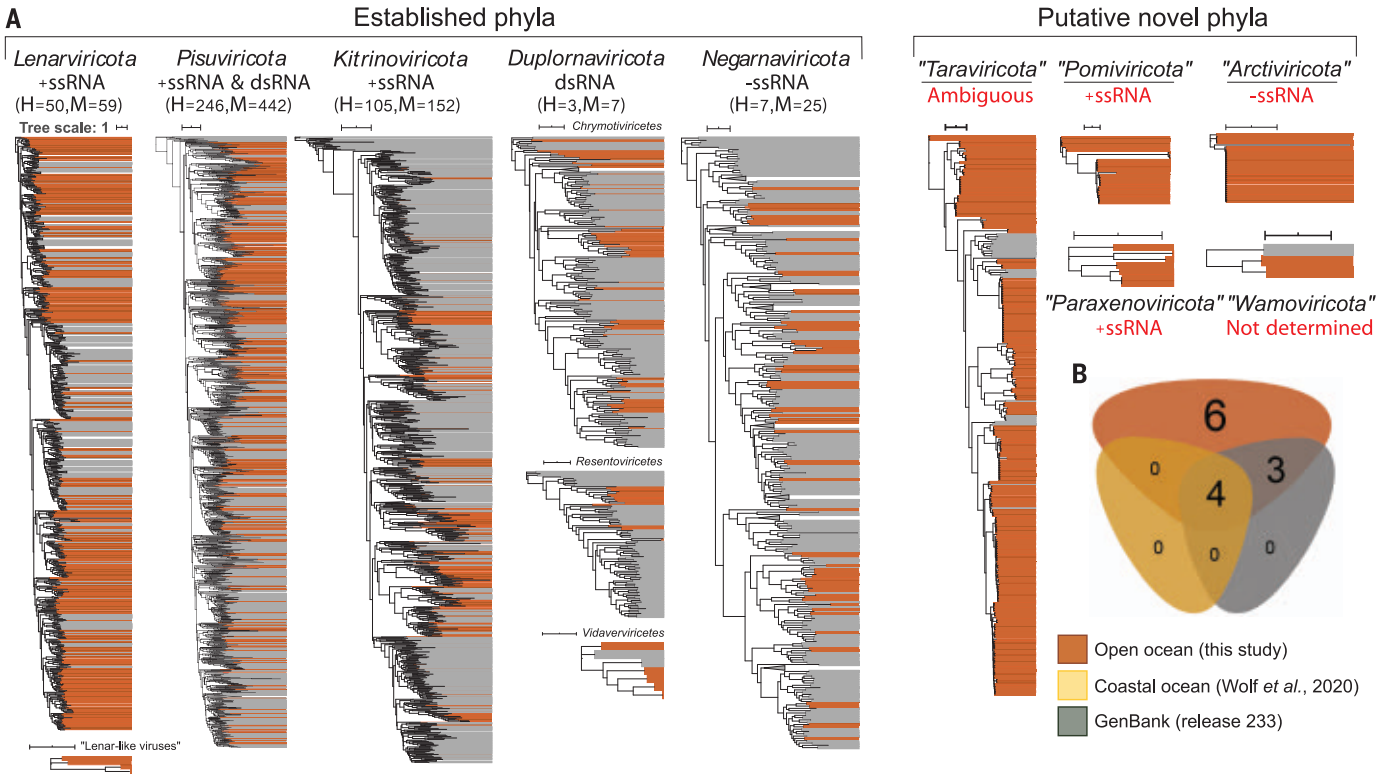


Fig. 2. Phylum- and class-rank RdRp-based phylogenetic analyses showing the taxonomic diversity of Global Ocean orthornavirans. (**A**) Thirteen maximum-likelihood phylogenetic trees encompassing the 19 megaclusters that emerged from network analyses of near-complete RdRp sequences (details in Fig. 1). Brown color indicates virus sequences discovered in this study, whereas gray indicates previously known reference sequences. The scale bar indicates one amino acid substitution per site. Classes were merged into a unified phylum-ranked tree only if the results from both phylogeny and network-guided clustering analysis were in agreement (materials and methods). Sequences were preclustered at 50% identity, and clades supported by 100% bootstrap values were collapsed. Genome strandedness (red text) for the new phyla was inferred in this study (as described in fig. S8 and materials and methods). A conservative estimate of the number of new complete or high-quality (H) and medium-quality (M) genomes retrieved in this study is indicated with parentheses. Underlined new phyla are supported by long- and short-read assemblies, whereas the remainder were supported by multiple independent assemblies from short-read assemblies (domain motifs are available in table S10). (**B**) Euler diagram of the shared, well-resolved phylum- or class-rank clusters of the near-complete RdRp domains across all available data from GenBank, a prior coastal ocean survey, and this study. Established megataxa represented in all datasets are *Lenarviricota*, *Pisuviricota*, *Kitrinoviricota*, and *Duplornaviricota*; *Chrymotiviricetes*. Established megataxa represented in our dataset and GenBank are *Duplornaviricota*; *Vidaverviricetes*, *Duplornaviricota*; *Resentoviricetes*, and *Negarnaviricota*. Unestablished megataxa inferred in this study are "*Taraviricota*," "*Pomiviricota*," "*Paraxenoviricota*," "*Arctiviricota*," "*Wamoviricota*," and "lenar-like viruses." In all analyses, RdRp domain clusters with permuted motifs ("permutotetra-like" and "birna-like" viruses) were excluded.

was groundbreaking, RdRp phylogenies are complex and require a manual and stepwise approach for construction, including a laborious iterative process of multiple sequence alignments, manual refinement, tree building, and representative selections to establish the global phylogeny. We worried that as seen in the literature (*7*, *24*), subjectivity in the iterative manual curation step could lead to varied perspectives on orthornaviran evolutionary inferences. Thus, to mitigate these concerns, we developed and benchmarked a scalable, network-based, iterative clustering approach to assess RdRp diversity; once performed, it nearly completely recapitulated the previously established phylogeny-based ICTV-accepted taxonomy (*7*, *17*) at the phylum and class ranks (97% agreement) (Fig. 1, B and C, and materials and methods).

With this approach, we then evaluated the Global Ocean data to classify the subset with complete or nearly complete RdRp domains and assess their novelty. Joint analysis of 111,760 complete or nearly complete RdRp domain sequences from all available (terrestrial and oceanic) viruses—6686 from our dataset, 101,819 from GenBank [release 233; only 3850 established species (*25*), indicating high species-rank redundancy] (materials and methods), and 3255 from coastal ocean RNA viromes (*5*)—revealed 19 "megaclusters" (Fig. 1B and table S6). Whereas our dataset represents only ≈6% of the total sequences in this analysis, our data covered vast diversity across the RNA orthovirosphere as follows (Fig. 2 and fig. S3): 13 of the 19 megaclusters from our analysis were known previously; together they compose the five ICTV-recognized phyla of the orthornaviran megataxonomy (*17*), with ocean-representative viruses for all five established phyla, all 20 established classes, and 49 of 103 established families (Fig. 2 and figs. S3 and S4). Although "known" at these taxon ranks, virtually all (99.7%) of the ocean viruses that could be evaluated represent new species (determined from whole-genome or contig information as described later) (table S5) that substantially augment undersampled taxa, because as much as 70% of sequences for some families were ocean derived (fig. S4A and table S7).

Beyond these more established taxa of the five-phylum system, 6 of the 19 megaclusters from our analysis were new (hereafter indicated with double quotation marks) and dominated by Global Ocean RdRps (Fig. 2A and data S1 and S2) (explanations for the

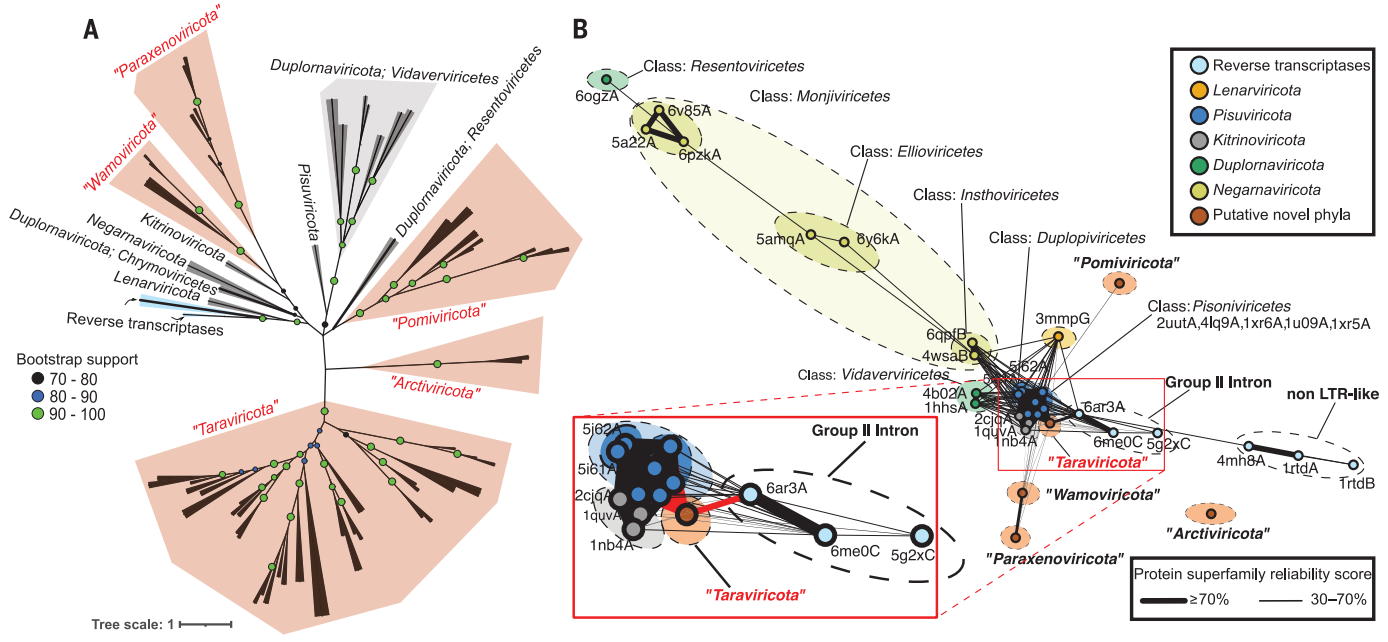


Fig. 3. Global RdRp-based phylogeny and network analyses inferring the early evolutionary history of orthornavirans. (**A**) Maximum-likelihood phylogenetic tree of RdRp domain sequences with RT sequences (cyan). The gray branches and polygons represent established megataxa, whereas the brown polygons represent megataxa inferred here. Each branch represents either a consensus or an individual sequence from a megataxon (materials and methods). Nodes in each branch represent bootstrap support. The scale bar indicates one amino acid substitution per site. (**B**) Three-dimensional structure similarity network of predicted (brown) and experimentally resolved (other colors; labeled with accession numbers) RdRp and RT protein domain structures. Each node represents a different structure, and the edges represent the reliability scores, for each connected pair, that they belong to the same protein superfamily (materials and methods). (Inset) The probability of "taraviricot" RdRps belonging to the same superfamily as group II–intron RTs and pisuviricot RdRps is 75 and 98%, respectively. In all analyses, RdRp domain clusters with permuted motifs ("permutotetra-like" and "birna-like" viruses) were excluded. LTR, long terminal repeat.

suggested names are provided in the supplementary materials, materials and methods). In the current orthornaviran megataxonomic framework (*17*), these six clusters would correspond to five new phyla, which we suggest to call "*Arctiviricota*," "*Paraxenoviricota*," "*Pomiviricota*," "*Taraviricota*" [includes the 22 previously identified "quenyaviruses" (*24*) with near-complete RdRp domains], and "*Wamoviricota*," as well as a new lenarviricot class, which we refer to here as "lenar-like viruses." Manual sequence inspection revealed that three of seven canonical RdRp motifs (*26*) are missing from members of this class-rank megataxon. Cluster-specific phylogenetic analyses (data S3) revealed that some virus groups were well represented in the oceans and elsewhere (such as ICTV-recognized pisuviricots), whereas others were primarily ("taraviricots") or exclusively ("pomiviricots," "paraxenoviricots," "arctiviricots," and "lenar-like viruses") oceanic (Fig. 2A).

To further assess the validity of our RdRp-inferred five new phyla, we evaluated phylogenetic (primary sequences) (Fig. 3A) and three-dimensional (3D) alignment (predicted and resolved tertiary structures) (Fig. 3B, fig. S5, and table S8) analyses of the RdRp domain, as well as other genomic features for which data were available (such as domain enrichments outside the RdRp, available for 7 of the 10 phyla) (table S9). In all cases, the network-derived clusters were supported by the phylogenetic and 3D-structure network information and contained features (statistically significant enrichment of domains outside the RdRp) (complete list is provided in table S9) that are consistent with variation observed at the established phylum rank. Marine representatives from established families have genome organizations similar to those from nonmarine taxa, whereas virus contigs of new phyla and classes were poorly annotated beyond the RdRp domains (figs. S6 and S7 and table S9). Together, these findings further suggest that the Global Ocean sequences add five phyla to the five already established as well as increase the number of known orthornaviran classes >50% by adding at least 11 classes (figs. S3 and S7) within previously established phyla. This expands the current megataxonomic framework beyond a stable five-phylum structure (*5*, *17*) and invites further exploration of its sequence space.

Marine RNA viruses revise the early evolution of orthornaviran megataxa

RdRp domain–based phylogeny has been used to infer deep orthornaviran evolutionary history (*7*), with different opinions on its robustness for this purpose (*21*, *24*, *27*) owing to the challenges of assigning homology in highly divergent primary sequences (*28*, *29*). The deepest parts of the RdRp phylogenetic tree are controversial (*21*, *27*) because only 55 of 441 sites showed an alignment homogeneity score ≥0.3 (as compared with 128 or more such sites for more broadly accepted phyla) (*27*). Although controversial and challenging, we interpret current literature to suggest that RdRp primary-sequence inferences lack confidence for interphyla relationships (*7*, *21*, *24*, *27*) but do suggest most phyla appear monophyletic (*27*). Given the extensive, new orthornaviran diversity, we revisited these deep evolutionary inferences using primary sequence–inferred phylogeny but also other features such as RdRp 3D structures and network-based clusters, other genomic domains, and whole-genome characteristics.

First, we assessed the monophyletic origin of double-stranded RNA (dsRNA) viruses of *Duplornaviricota*, which is one of the five orthornaviran phyla thought to have more recently evolved from positive-sense single-stranded RNA (+ssRNA) viruses (*7*). Previously, all viruses in *Duplornaviricota* were placed in a single phylum with three classes because *Duplornaviricota* and *Negarnaviricota* were

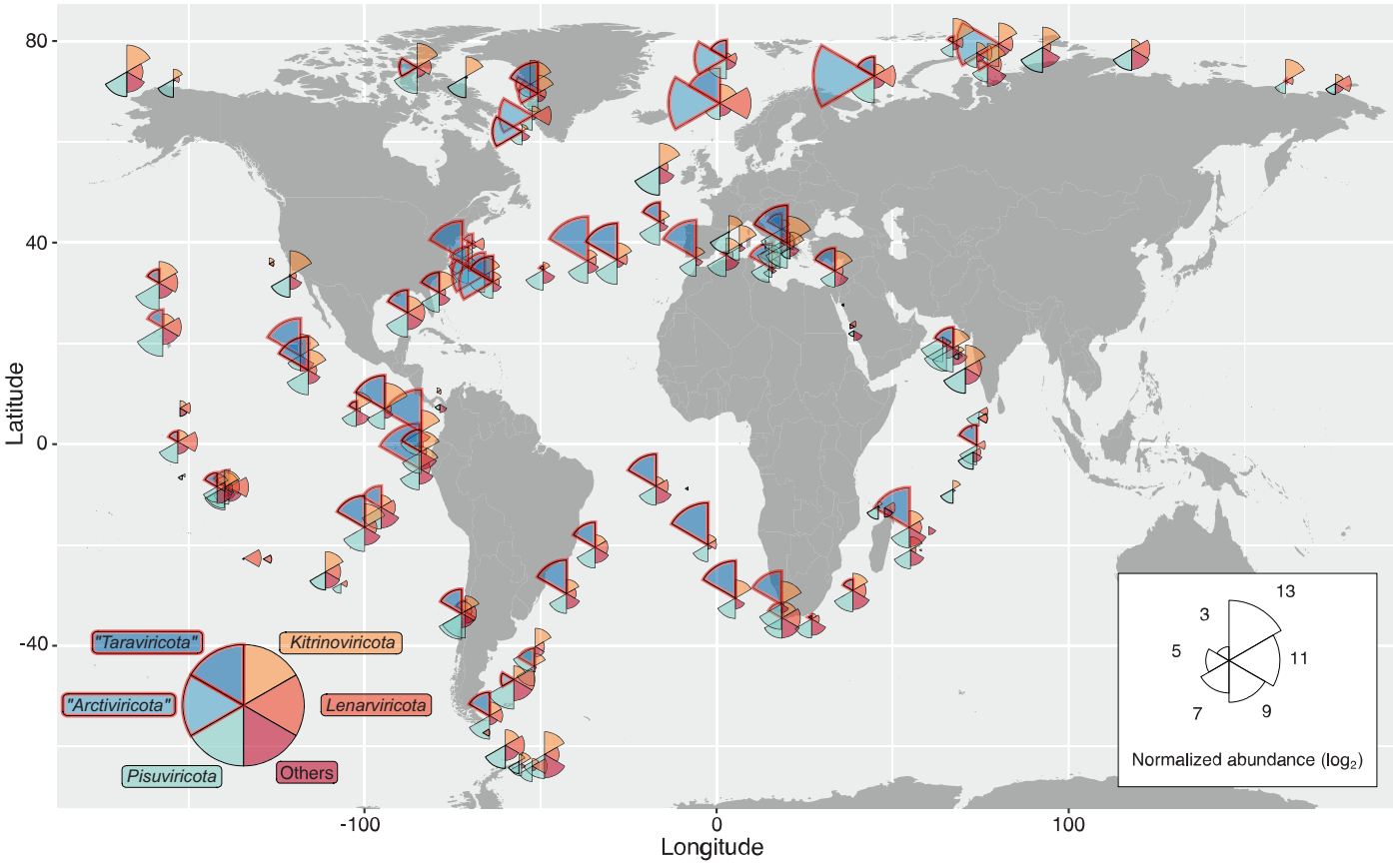


Fig. 4. Biogeography of orthornaviran megataxa. Global map showing the distribution and average relative abundance (on a $\log_2$ scale) of vOTUs inferred in this study per phylum. The position and color of the wedges are fixed for the same megataxon across the Global Ocean. Wedge lengths are proportional to the average abundance in the sample as well as across the global dataset. Biogeography per size fraction is provided in fig. S11.

strongly monophyletic [*Duplornaviricota* and *Negarnaviricota* are labeled as branches 4 and 5, respectively, in (*7*, *17*)]. However, reexamination of alignment homogeneity from previous work (*27*) suggests that these taxa are polyphyletic because (i) only 72 sites within the duplornaviricot sequence alignment showed homogeneity ≥0.3 as compared with at least 128 sites for sequences from the other phyla and (ii) *Duplornaviricota* showed a paraphyletic relationship with respect to *Negarnaviricota* (*7*), which hinted toward accommodating *Duplornaviricota* taxonomically by at least three phyla (*7*, *17*). Our global phylogenetic tree also suggests, with strong support, that these dsRNA viruses are polyphyletic (Fig. 3A). The *Duplornaviricota* polyphyly we observed is further supported by (i) the lack of strong duplornaviricot intertaxon connections in our 3D structure network (Fig. 3B), (ii) the absence of a homogeneous cluster encompassing these taxa that are emerging from our iterative clustering approach (Fig. 1), and (iii) differential extraneous-to-RdRp domain enrichment across these taxa (table S9). Hence, the grouping of all dsRNA viruses (apart from the class *Duplopiviricetes*) into one phylum (*Duplornaviricota*), as established currently (*7*), appears incorrect. Instead, we suggest—as the ICTV has done for +ssRNA viruses that were recently split into three phyla [*Lenarviricota*, *Pisuviricota*, and *Kitrinoviricota*; also supported by our data (Figs. 2 and 3)] (*7*)—that *Duplornaviricota* represent three different phyla along the lines of the currently recognized classes. If ultimately ICTV approved, this would expand currently known diversity to a total of 12 phyla.

The second deep evolutionary orthornaviran inference we assessed was the proposition that negative-sense single-stranded RNA (–ssRNA) viruses (phylum *Negarnaviricota*) evolved from the dsRNA duplornaviricots, which is considered a low-confidence link in the literature (*7*, *17*, *27*). Our global phylogenetic tree also indicates a last common ancestor of negarnaviricots and one of the dsRNA virus "classes," but we found the well-supported sister taxon to be the dsRNA "class" *Chrymotiviricetes* (Fig. 3A), as opposed to the prior observed "class" *Resentoviricetes* (*7*). Because such deep evolutionary phylogenetic inferences are prone to long branch attraction artefacts, we evaluated other lines of evidence. This revealed that these prior proposed relationships were not supported in (i) our 3D structure network (only *Resentoviricetes* was connected, and only weakly, to *Negarnaviricota*) (Fig. 3B) or (ii) our iterative primary sequence–based clustering approach (the two taxa never formed a homogeneous cluster) (Fig. 1). Additionally, domain enrichment analysis (table S9, section B) showed that negarnaviricots did not share any domains with dsDNA viruses but did share a virus-capping methyltransferase domain (Pfam: PF14314) with >50 viruses classified in *Pisuviricota* and *Kitrinoviricota* (table S9). When we examined the suggested phyla for their "strandedness" (materials and methods and fig. S8), which helps identify the virus genome type (+ssRNA, –ssRNA, or dsRNA), "*Arctiviricota*" emerged as –ssRNA. Both phylogenetic (Fig. 3A) and 3D structure network (Fig. 3B) analyses suggest that "arctiviricots" evolved independently from negarnaviricots (and dsRNA viruses) and represent a second –ssRNA phylum and further polyphyly within the orthornavirans. These findings argue that all orthornaviran

genome types (+ssRNA, −ssRNA, and dsRNA viruses) have multiple evolutionary origins.

Third, we revisited the RdRp primary sequence–inferred hypothesis that considers orthornavirans monophyletic and assumes reverse transcriptases (RTs) of retroelements as the root of the global RdRp tree (*7*). In that scenario, lenarviricots (some of which infect bacteria and carry capsid proteins) are a sister group to the remaining orthornavirans, and retroelements appear more likely (and parsimoniously) to be ancestral to orthornavirans (*7*), arguing against the emergence of virus RdRp in the peptide-RNA world (*12*, *30*). Instead, our RdRp phylogeny revealed lenarviricot RdRps sharing ancestry with RTs (well supported) (Fig. 3A and data S4), which (assuming a monophyletic origin of orthornavirans) suggests a capsidless RNA replicon as the ancestor of both retroelements and RNA viruses and agrees with the thinking that virus RdRps were part of the earlier peptide-RNA world. *Lenarviricota* harbors the short (<5 kb) capsidless RNA replicons (mitovirids that carry only an RdRp, infect eukaryotes, and replicate in host mitochondria).

An alternative scenario, however, was inferred from 3D structure analyses, which are often considered more informative than primary-sequence information for deep evolutionary inferences (*31*). These analyses suggest, with high calculated probability (materials and methods), that viruses from our suggested phylum "*Taraviricota*" represent a missing link between retroelements (riboviriad pararnavirans) and orthornavirans (Fig. 3B). If true, this implies that "*Taraviricota*" RdRp represents the capsidless RNA replicon ancestor of retroelements and orthornaviran RdRps—potentially the RdRp replicon postulated to have originated from junctions of proto-tRNAs (11, 12). To evaluate this scenario further, we examined genomic information of "taraviricots" as follows.

First, similar to mitovirids (phylum *Lenarviricota*), all but four of the marine "taraviricots" that were recovered from short- (n = 220) or long-read (n = 32) assemblies (Fig. 2A) have short genomes (<3.4 kb) (fig. S7) and encode only RdRp. No other well-sampled (>10 viruses) phylum in our dataset showed such a feature, which we interpret to be due to either short virus genome length or consistent genome segmentation ["quenyaviruses" always encode RdRp on its own segment (*24*)]. If the former is true—that most "taraviricots" have short genomes—it implies that orthornavirans evolved from an RdRp-only ancestor through gene gains (and potential later losses) (*7*). If the latter is true, then genome segmentation in orthornavirans evolved early and potentially contributed to an accelerated early diversification of orthornavirans (Fig. 3A, "*Taraviricota*"). Genome segmentation is not common among lenarviricots, and many of its non-segmented lineages encode single jelly-roll capsid proteins that were hypothesized (although, notably, unparsimoniously) to be horizontally transferred from viruses of other phyla (*7*). Both of these observations support our alternative 3D structure–inferred scenario presented here.

Second, of the four marine "taraviricots" encoding more than just RdRp, two encoded only a putative phospholipase [Pfam, PF11618 (CL14603) or PF02230 (CL0028); not found in any other orthornaviran (table S9)]. This observation suggests that at least some "taraviricots" ancestrally or currently infect a cell wall–deficient prokaryotic host or the mitochondria of eukaryotes (sensu mitovirids). Although this link is still speculative, we interpret this finding—together with "taraviricots" overwhelmingly encoding just the RdRp on very short genomes and/or potential consistent genome segmentation and their 3D structure resemblance to multiple orthornaviran types (+ssRNA and dsRNA) and RTs—to provide a parsimonious scenario for "*Taraviricota*" as an early basal lineage from which other orthornaviran phyla have subsequently evolved.

Collectively, we sought to reevaluate deep evolutionary inferences using multiple data types beyond primary sequence, and these analyses suggest (i) polyphyletic origins of dsRNA "phylum" *Duplornaviricota* (splitting it into three different phyla) and −ssRNA phyla (*Negarnaviricota* and "*Arctiviricota*") and (ii) an ancient presence of "taraviricots" on Earth, with a potential important role in the orthornaviran and pararnaviran evolution.

Abundance and biogeography of orthornaviran "species"

Given this extensive, new orthornaviran diversity, we next sought to biogeographically contextualize it globally, at least for the oceans. Such analyses are possible because of two major advances: (i) systematic *Tara* Oceans' global sampling (table S4) and (ii) a recent consensus approach (*32*) that establishes virus operational taxonomic units (vOTUs; a species-rank approximation) by evaluating genomic sequence space for discontinuities. Applying this approach to our whole-genome and contig data revealed such a discontinuity, although at different cutoffs supported by our sensitivity analyses (fig. S9 and materials and methods). The empirically derived vOTU definition suggested from these analyses was 90% average nucleotide identity over 80% coverage of the smaller contig and ≥1 kb in length. Dereplicating our 44,779 virus contigs at this cutoff revealed 5504 vOTUs (vOTU contig length range of 1001 to 25,584 nucleotides, with a median of 1958) (table S5). Of these 5504 vOTUs, a subset (n = 624) is related enough to known complete virus genomes that we can estimate their completeness—433 high-quality or complete genomes (belonging to 188 vOTUs), 719 medium-quality genomes (belonging to 246 additional vOTUs), and 807 low-quality genomes (belonging to 190 additional vOTUs)—whereas the remainder (n = 4880) are so divergent from reference genomes that their completeness cannot be estimated by using available approaches (table S5). Virtually all of these vOTUs (n = 5485; 99.7%), including those with at least medium-quality genomes (n = 430; 99.6%), belong to new species (table S5). Additionally, to compare our methods with those that rely on just the RdRp domain sequences for vOTU construction [for example, (*33*)], we examined a range of clustering and contig length cutoffs (materials and methods) and found general and robust agreement for contigs ≥1 kb in length (at least 93% agreement) (fig. S9 and materials and methods). Hence, our vOTU definition both respects RdRp-inferred relationships among individual contigs in a cluster and expands on them by including genomic information to resolve ambiguity in RdRp-based identity cutoffs (fig. S9).

Given this robustness, we quantified vOTUs by means of read mapping to assess abundance and global biogeography across the 771 Global Ocean metatranscriptomes (materials and methods). This revealed three phyla—*Pisuviricota*, *Kirinoviricota*, and "*Taraviricota*"—as collectively abundant and widespread (fig. S10). The first two phyla include "picorna-like" and "tombus-like" viruses commonly found in site-focused surveys (*34*, *35*), whereas the third phylum ("*Taraviricota*") consists of at least 220 previously unknown viruses (with near-complete RdRp domain sequences) described here. This phylum's vOTUs were, on average, the most abundant across most temperate and tropical waters (Fig. 4). This finding suggests ecological importance for these previously overlooked viruses and provides broader context for previously described viruses ("quenyaviruses") that were found to be abundant in some arthropods and other animals (*24*) and are now more clearly recognized as members of the most abundant ocean orthornaviran phylum. Although with more restricted geographic range, vOTUs belonging to the −ssRNA phylum "*Arctiviricota*" were, on average, the most abundant across most of the Atlantic Arctic waters (Fig. 4). None of the other −ssRNA viruses (negarnaviricots) showed similar patterns in any area of the ocean, suggesting a specific ecological footprint for the "arctiviricots" described here. Although the biogeographic data shown here represent relative abundances of a mixture of abundances derived from genomes and transcripts, the relative abundances of

"*Taraviricota*" and "*Arctiviricota*" are likely mostly derived from their genomes (fig. S8). Together, these data provide an orthornaviran-wide, systematically sampled, and large-scale complement to prior RNA virus diversity studies in the ocean (*24*, *33*–*35*).

Last, having established this environmental context and vast ocean-derived orthornaviran diversity, we sought to identify their hosts. Unfortunately, host identification for environmental RNA virus contigs is challenging, which limits us to reporting only domain-rank hosts for the new megataxa from multiple analytical approaches that include preestablished host linkages to previously known RNA virus taxa, abundance-based co-occurrence networks, and screening of endogenous virus elements (materials and methods). Results from this effort revealed that viruses of "*Taraviricota*," "*Arctiviricota*," "*Pomiviricota*," "*Wamoviricota*," and eight of the new classes are associated with eukaryotes (table S11), whereas only pisuviricot class 27 viruses likely infect prokaryotes (table S12). The latter finding of infecting prokaryotes is rare but not unknown for RNA viruses and is supported by a statistically significant signal of Shine-Dalgarno motifs (table S12 and materials and methods) and one of the representative virus genomes encoding a putative preprotein translocase subunit SecY of the bacterial type II secretion system (fig. S7). The remaining new megataxa (one phylum and two classes) could not be associated with hosts. Together, these findings suggest that eukaryotes remain the main hosts of orthornavirans but suggest addition of our new pisuviricot class 27 to known RNA phage groups alongside levivirids (phylum *Lenarviricota*), cystovirids (phylum *Duplornaviricota*), and potentially (*36*) picobirnavirids (phylum *Pisuviricota*).

Conclusions

Although clear population- and genome-resolved approaches have been developed for dsDNA viruses and revealed the existence of hundreds of thousands of distinct dsDNA virus species in the oceans alone (*37*), few parallel studies for RNA viruses exist—despite urgent needs (*38*) and suggestions that our understanding of the virosphere will increase with the study of microbial eukaryotes (*4*, *5*). Our study and several prior studies (*4*, *5*, *39*) confirm this suggestion and are now reshaping our understanding of RNA virus diversity and evolution, with thousands of previously unknown RNA virus species presented in this study alone. Although documentation of such RNA virus diversity might now be scalable to that observed in nature, several challenges need to be addressed. These include (i) identifying hosts for previously undiscovered viruses, (ii) scalably improving genome completeness in survey approaches, and (iii) directly capturing RNA virus particles from environmental samples to assess their diversity in a targeted manner and complement the host metatranscriptomic sequence space–based abundance calculations presented in this study. Although challenges remain, the global and systematic effort presented here provides critical information and resources, an analytical roadmap, and foundational advances to feed the predictive models that are needed to assess RNA virus ecosystem, eco-evolutionary, and epidemiological impacts.

REFERENCES AND NOTES

1. M. E. J. Woolhouse, L. Brierley, *Sci. Data* **5**, 180017 (2018).
2. K.-B. G. Scholthof *et al*., *Mol. Plant Pathol.* **12**, 938–954 (2011).
3. A. Brun, *Methods Mol. Biol.* **1349**, 1–24 (2016).
4. M. Shi *et al*., *Nature* **540**, 539–543 (2016).
5. Y. I. Wolf *et al*., *Nat. Microbiol.* **5**, 1262–1270 (2020).
6. M. Shi *et al*., *Nature* **556**, 197–202 (2018).
7. Y. I. Wolf *et al*., *mBio* **9**, e02329–e18 (2018).
8. M. Krupovic, V. V. Dolja, E. V. Koonin, *Biol. Direct* **10**, 12 (2015).
9. V. V. Dolja, E. V. Koonin, *Virus Res.* **244**, 36–52 (2018).
10. C. W. Carter Jr., *Life* **5**, 294–320 (2015).
11. S. Chatterjee, S. Yadav, *Life* **9**, 25 (2019).
12. A. Pereira Dos Santos Jr., M. V. José, S. Torres de Farias, *Biosystems* **206**, 104442 (2021).
13. S. T. de Farias, A. P. Dos Santos Jr., T. G. Rêgo, M. V. José, *Front. Genet.* **8**, 125 (2017).
14. P. Forterre, *Virus Res.* **117**, 5–16 (2006).
15. P. Forterre, D. Prangishvili, *Res. Microbiol.* **160**, 466–472 (2009).
16. J. H. Kuhn *et al*., *Nature* **566**, 318–320 (2019).
17. E. V. Koonin *et al*., *Microbiol. Mol. Biol. Rev.* **84**, e00061–e19 (2020).
18. J. B. Whitfield, P. J. Lockhart, *Trends Ecol. Evol.* **22**, 258–265 (2007).
19. M. Krupovic, V. V. Dolja, E. V. Koonin, *Nat. Rev. Microbiol.* **18**, 661–670 (2020).
20. D. M. Kristensen, A. R. Mushegian, V. V. Dolja, E. V. Koonin, *Trends Microbiol.* **18**, 11–19 (2010).
21. E. C. Holmes, S. Duchêne, *mBio* **10**, 1–2 (2019).
22. A. M. Burroughs, Y. Ando, L. Aravind, *Wiley Interdiscip. Rev. RNA* **5**, 141–181 (2014).
23. L. M. Iyer, E. V. Koonin, L. Aravind, *BMC Struct. Biol.* **3**, 1 (2003).
24. D. J. Obbard, M. Shi, K. E. Roberts, B. Longdon, A. B. Dennis, *Virus Evol.* **6**, vez061 (2020).
25. A. Dance, *Nature* **595**, 22–25 (2021).
26. A. J. W. te Velthuis, *Cell. Mol. Life Sci.* **71**, 4403–4420 (2014).
27. Y. I. Wolf *et al*., *mBio* **10**, e00542–e19 (2019).
28. B. Rost, *Protein Eng.* **12**, 85–94 (1999).
29. P. M. Zanotto, M. J. Gibbs, E. A. Gould, E. C. Holmes, *J. Virol.* **70**, 6083–6096 (1996).
30. S. T. de Farias, T. G. Rêgo, M. V. José, *Life (Basel)* **6**, 15 (2016).
31. K. Illergård, D. H. Ardell, A. Elofsson, *Proteins* **77**, 499–508 (2009).
32. S. Roux *et al*., *Nat. Biotechnol.* **37**, 29–37 (2019).
33. J. A. Gustavsen, D. M. Winget, X. Tian, C. A. Suttle, *Front. Microbiol.* **5**, 703 (2014).
34. A. I. Culley, A. S. Lang, C. A. Suttle, *Science* **312**, 1795–1798 (2006).
35. A. Culley, *Virus Res.* **244**, 84–89 (2018).
36. S. Ghosh, Y. S. Malik, *Front. Vet. Sci.* **7**, 615293 (2021).
37. A. C. Gregory *et al*., *Cell* **177**, 1109–1123.e14 (2019).
38. R. K. French, E. C. Holmes, *Trends Microbiol.* **28**, 165–175 (2020).
39. C.-X. Li *et al*., *eLife* **4**, e05378 (2015).
40. A. A. Zayed, J. M. Wainaina, G. Dominguez-Huerta, Cryptic and abundant marine viruses at the evolutionary origins of Earth's RNA virome. *CyVerse Data Commons* (2021).
41. J. M. Wainaina, A. A. Zayed, G. Dominguez-Huerta, B. Bolduc, M. B. Sullivan, Supporting trees and alignments for the publication: Cryptic and abundant marine viruses at the evolutionary origins of Earth's RNA virome. *DRYAD* (2022).
42. J. Guo, rdrpsearch, Version 0.1.*Zenodo* (2021).
43. R. Schlitzer, Ocean Data View (2018); https://odv.awi.de/.

ACKNOWLEDGMENTS

We thank Y. I. Wolf (National Center for Biotechnology Information, U.S. National Library of Medicine, National Institutes of Health) for advice and guidance in analyzing RdRp sequences and A. Crane (Integrated Research Facility at Fort Detrick, National Institute of Allergy and Infectious Diseases, National Institutes of Health) for critically editing the manuscript. *Tara* Oceans would not exist without the leadership of the *Tara* Expeditions Foundation and the continuous support of 23 institutes. The extensive *Tara* Oceans expeditionary support is detailed in the supplementary text. **Funding:** The virus-specific work presented here was supported in part through the following: Gordon and Betty Moore Foundation (award #3790); US National Science Foundation (awards OCE#1829831, ABI#1759874, and DBI#2022070); The Ohio Supercomputer and Ohio State University's Center of Microbiome Science; Ramon-Areces Foundation Postdoctoral Fellowship to G.D.-H.; Laulima Government Solutions, LLC prime contract with the U.S. National Institute of Allergy and Infectious Diseases (NIAID—Contract No. HHSN272201800013C); SNSF project (grant 205321_184955 to S.S.); and France Génomique for funding for the sequencing (ANR-10-INBS-09) (P.W.). **Author contributions:** A.A.Z., G.D.-H., J.M.W., and M.B.S. planned and supervised the work, interpreted the results, and wrote the manuscript with inputs from all authors. A.A.Z., J.M.W., G.D.-H., E.P., J.G., M.M., F.T., B.B., O.Z., A.A.P., S.C., D.C., L.S., E.D., E.S., R.B., and K.F. developed and/or implemented the informatic analyses. A.A., J.-M.A., Q.C., C.d.S., K.L., E.P., J.P., H.-J.R., G.S., A.A.Z., S.S., P.W., and *Tara* Oceans coordinators all contributed to expeditionary infrastructure needed for global ocean sampling, sample processing, and/or previously published data resource development. L.S.K., A.I.C., and J.H.K. provided domain expertise on phylogenetics, RNA virus ecology, and taxonomy, respectively. All authors read and commented on the manuscript and approved it in its final form. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** The authors declare that all data reported here are fully and freely available from the date of publication without restrictions and that all of the analyses, publications, and ownership of data are free from legal entanglement or restriction by the various nations whose waters were sampled during the *Tara* Oceans expeditions. This article is contribution number 129 of *Tara* Oceans. Newly generated raw sequence reads for the 143 eukaryote-size fraction metatranscriptomes from the Arctic Ocean are available at ENA/SRA under BioProjectID PRJEB9738 and PRJEB9739. Processed data are publicly available through iVirus (*40*), including all metatranscriptome assemblies, RNA virus contigs and vOTUs, RdRp sequences and clusters, and HMM profiles. Multiple sequence alignments and phylogenetic trees are publicly available through DRYAD (*41*), whereas the HMM pipeline developed in this work can be accessed through Zenodo (*42*). In addition, scripts used to generate figures are uploaded to the MAVERICKlab bitbucket page (https://bitbucket.org/MAVERICLab/global-rna-virus-evolution-2021).

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abm5847
Materials and Methods
Supplementary Text
Figs. S1 to S11
Tables S1 to S12
References (*44*–*153*)
Data S1 to S4

28 September 2021; accepted 7 February 2022
10.1126/science.abm5847

STRUCTURAL BIOLOGY

Structure of a Janus kinase cytokine receptor complex reveals the basis for dimeric activation

Caleb R. Glassman[1]†, Naotaka Tsutsumi[1,2]†, Robert A. Saxton[1,2], Patrick J. Lupardus[1]‡, Kevin M. Jude[1,2], K. Christopher Garcia[1,2,3]*

Cytokines signal through cell surface receptor dimers to initiate activation of intracellular Janus kinases (JAKs). We report the 3.6-angstrom–resolution cryo–electron microscopy structure of full-length JAK1 complexed with a cytokine receptor intracellular domain Box1 and Box2 regions captured as an activated homodimer bearing the valine→phenylalanine (VF) mutation prevalent in myeloproliferative neoplasms. The seven domains of JAK1 form an extended structural unit, the dimerization of which is mediated by close-packing of the pseudokinase (PK) domains from the monomeric subunits. The oncogenic VF mutation lies within the core of the JAK1 PK interdimer interface, enhancing packing complementarity to facilitate ligand-independent activation. The carboxy-terminal tyrosine kinase domains are poised for transactivation and to phosphorylate the receptor STAT (signal transducer and activator of transcription)–recruiting motifs projecting from the overhanging FERM (four-point-one, ezrin, radixin, moesin)–SH2 (Src homology 2)-domains. Mapping of constitutively active JAK mutants supports a two-step allosteric activation mechanism and reveals opportunities for selective therapeutic targeting of oncogenic JAK signaling.

Cytokines are a multifarious family of secreted proteins that have broad and pleiotropic effects on cell growth, hematopoiesis, immunity, and inflammation (*1*, *2*). Cytokines initiate signaling by binding to the extracellular domains of Type I single-pass transmembrane receptors to facilitate receptor dimerization which is required to initiate transduction (*3*–*5*). This extracellular dimerization event is structurally conveyed to the intracellular domains (ICDs), resulting in the activation and transphosphorylation of noncovalently associated Janus kinases (JAKs) (*6*–*8*). All four members of the JAK family (JAK1, JAK2, JAK3, and TYK2) associate with the membrane-proximal regions of cytokine receptor ICDs through two distinct conserved motifs in the receptor: a proline-rich segment termed "Box1" and a hydrophobic segment called "Box2" (*9*). Once activated, JAKs phosphorylate tyrosine residues within the cytokine receptor (ICDs), which subsequently serve as docking sites for the STAT (signal transducer and activator of transcription) transcription factors (*10*). Recruitment of STATs to the receptor–JAK complex enables STAT phosphorylation by the activated JAKs, leading to STAT dimerization and translocation to the nucleus to initiate transcription of cytokine-responsive genes.

All JAK family members are composed of seven JAK homology (JH) domains that comprise a four-point-one, ezrin, radixin, moesin (FERM) domain (JH5, JH6, and JH7), an Src homology 2 (SH2) domain (JH3 and JH4), and tandem kinase domains JH2 and JH1 which encode a pseudokinase (PK) and tyrosine kinase (TK), respectively (Fig. 1A) (*11*, *12*). The FERM and SH2 domains at the N-terminal end of JAK associate with the intracellular juxtamembrane segment from the paired cytokine receptor (*13*). Our current understanding of full-length JAK structure and activation mechanisms is derived from extrapolations of structures of monomeric JAK fragments. Crystal structures of the human JAK1, JAK2, and TYK2 FERM-SH2 fragments have revealed that these domains are tightly associated and thus form a single receptor-binding module that accommodates the Box1/Box2 peptide at multiple interaction sites (*14*–*17*). In addition, structural models for the PK-TK modules from TYK2 and JAK2 have suggested a mechanism of negative regulation by the pseudokinase (*18*, *19*). Furthermore, numerous structures of cytokines complexed with their receptor extracellular domains (ECDs) in homo- or heterodimeric complexes have shown common features and structural diversity in the overall architectures of the extracellular assemblies that are presumably communicated to the inside of the cell for JAK activation (*20*). However, how ECD dimerization brings two intracellular JAKs into proper orientation and proximity for activation remains unresolved as a result of the absence of structural information on full-length JAK proteins in activated states (*8*).

Naturally occurring mutations in cytokine receptors, JAKs, and STATs lead to immunodeficiency and myeloproliferative disorders in humans (*10*, *21*). Disruption of JAK1 and JAK2 genes is lethal (*22*–*24*), whereas loss-of-function (LOF) mutations in JAK3 cause severe combined immunodeficiency (SCID) (*25*–*27*). On the other hand, gain-of-function (GOF) mutations in JAK genes are responsible for a family of blood disorders known as myeloproliferative neoplasms (MPNs), which include polycythemia vera, primary myelofibrosis, and essential thrombocythemia, as well as leukemias (*28*). In a classic series of papers reported in 2005 (*29*–*32*), a point mutation in the PK domain of JAK2—Val^{617}→Phe (V617F), which results in constitutive activity—was shown to be present in >90% of patients with polycythemia vera and in ~50% of patients with essential thrombocythemia and primary myelofibrosis. Analogous mutations in human JAK paralogs also result in constitutive activity, suggesting a shared activation mechanism across JAK family members, likely involving ligand-independent dimerization at the cell surface (*3*, *34*, *35*). Ruxolitinib is a small-molecule inhibitor of JAK2 (and JAK1) kinase activity and targets both wild-type (WT) JAK2 and JAK2-V617F, resulting in side effects such as thrombocytopenia and anemia (*21*). A better understanding of how mutations in JAK—particularly JAK2-V617F—result in constitutive activity is needed to guide drug design to target mutant JAK2. Here we report the cryo–electron microscopy (cryo-EM) structure at 3.6-Å resolution of full-length mouse JAK1 complexed with the interferon λ receptor 1 (IFNλR1) intracellular Box1/Box2 segment, which provides a structural blueprint to understand both cytokine and oncogenic mutant-driven signal activation.

Engineering an active JAK1-IFNλR1 complex for cryo-EM imaging

Full-length JAKs have been recalcitrant to structural analysis by x-ray crystallography and electron microscopy (*8*). Imaging a JAK1 complex with cytokine receptor ICD required several protein engineering steps to produce an activated, stable, nonaggregated complex suitable for cryo-EM imaging. First, we determined that full-length mouse JAK1 has better expression and solubility properties when produced from insect cells, compared with other JAK paralogs and orthologs. Second, we introduced the V657F mutation into mouse JAK1 (analogous to hJAK2 V617F) to stabilize the activated state. Third, so that we could affinity purify full-length JAK1 with the receptor ICDs, we focused on the JAK1 binding Box1/Box2 domains from interferon λ receptor 1 (IFNλR1) on the basis of a screen that identified this ICD as among the highest affinity JAK1-ICD interactions (*14*). Fourth, we replaced the transmembrane domains of the receptor with the homodimeric GCN4 leucine zipper fused to the IFNλR1 Box1/Box2 to create a soluble mimic of a dimerized receptor

[1]Department of Molecular and Cellular Physiology, Stanford University School of Medicine, Stanford, CA 94305, USA. [2]Howard Hughes Medical Institute, Stanford University School of Medicine, Stanford, CA 94305, USA. [3]Department of Structural Biology, Stanford University School of Medicine, Stanford, CA 94305, USA.
*Corresponding author. Email: kcgarcia@stanford.edu
†These authors contributed equally to this work.
‡Present address: Synthekine, Menlo Park, CA 94025, USA.

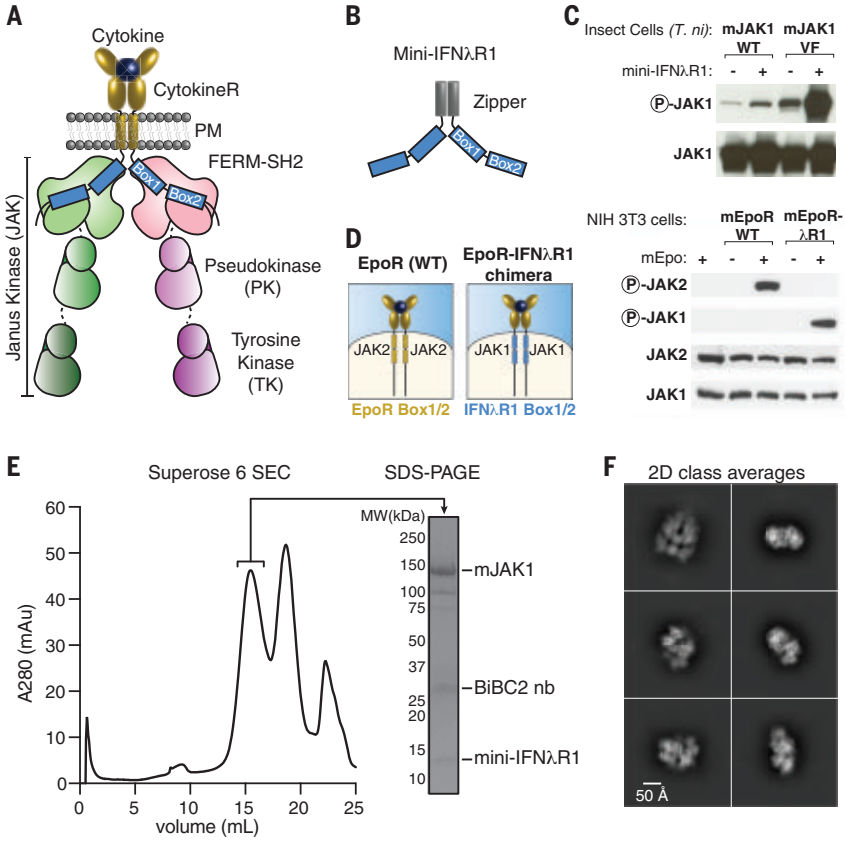


Fig. 1. Purification and biochemical characterization of an active JAK1-IFNλR1 complex. (**A**) Schematic of a cell-surface, ligand-induced cytokine receptor-JAK dimer. (**B**) Schematic of a soluble cytokine receptor dimer mimetic ("mini-IFNλR1") in which the transmembrane domains of the receptor have been replaced with a GCN4-zipper and the intracellular tail has been trucated after Box1/Box2. (**C**) Mini-IFNλR1 expression enhances JAK1 phosphorylation when coexpressed in insect cells. Wild type (WT) or Val^{657}→Phe (VF) JAK1 was coexpressed with mini-IFNλR1 in *Trichoplusia ni* (*T. ni*) cells by baculovirus transduction. JAK phosphorylation and total expression were measured 2 days after infection by immunoblot of whole-cell lysate. Results are representative of more than two independent experiments. (**D**) Schematic of the WT EpoR/Epo complex (left) and EpoR-IFNλR1 chimera (right) showing substitution of Box1/Box2 motifs. Cytokine-mediated dimerization of IFNλR1 Box1/Box2 results in JAK1 phosphorylation in mammalian cells. NIH 3T3 cells transiently expressing mEpoR or the mEpoR-IFNλR1 chimera were stimulated with Epo for 20 min before analysis of JAK phosphorylation by immunoblot. Results are representative of two independent experiments. (**E**) Affinity purification of JAK1 using mini-IFNλR1 yields a stable, nonaggregated complex. Superose 6 size exclusion chromatography (SEC, left) and sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE, right) of the JAK1-IFNλR1 complex. (**F**) Representative 2D class averages from single-particle cryo-EM imaging of the JAK1-IFNλR1 complex. PM, plasma membrane; R, receptor; mAu, milliabsorbance units; BiBC2 nb, tandem BC2 nanobody.

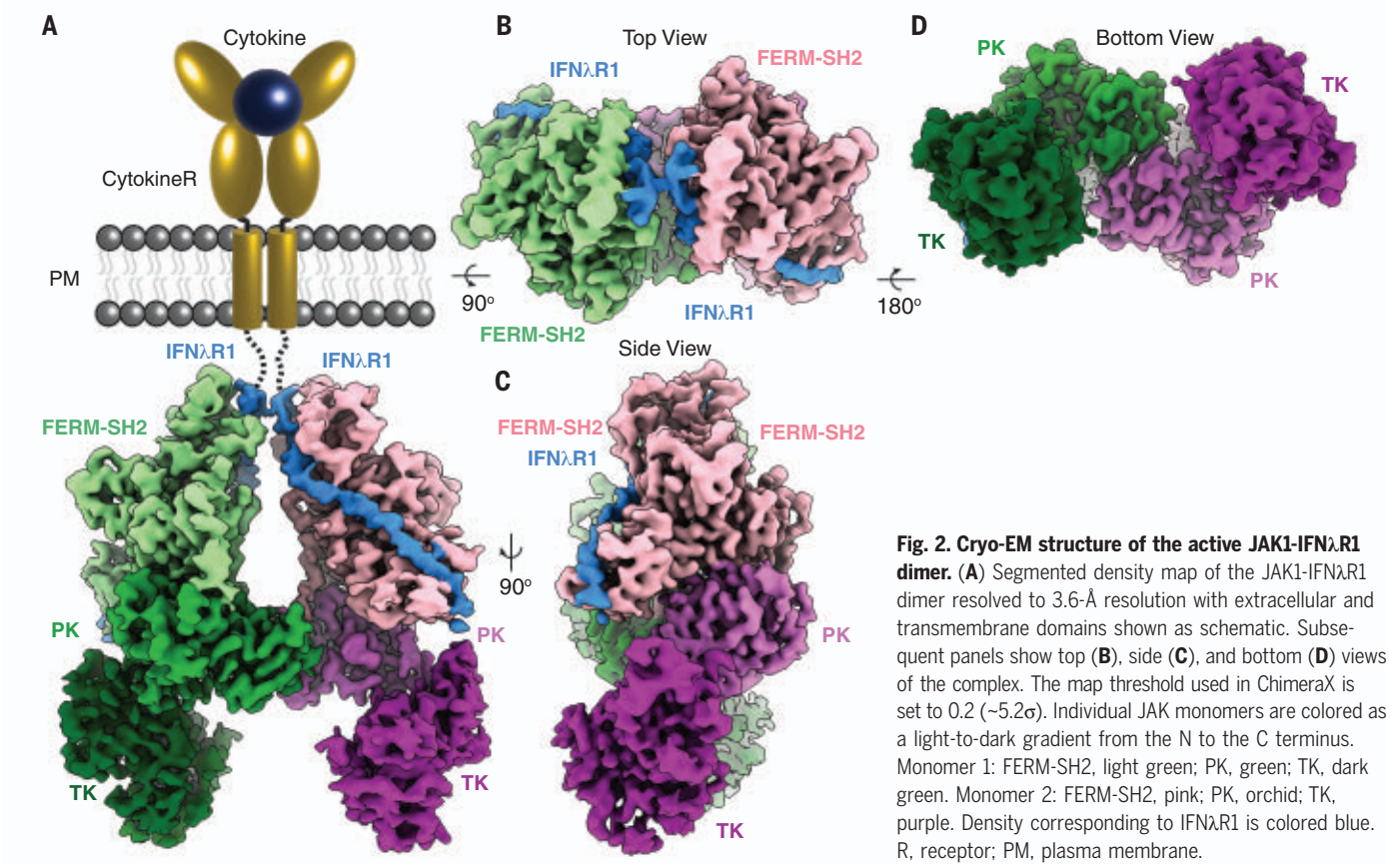


Fig. 2. Cryo-EM structure of the active JAK1-IFNλR1 dimer. (**A**) Segmented density map of the JAK1-IFNλR1 dimer resolved to 3.6-Å resolution with extracellular and transmembrane domains shown as schematic. Subsequent panels show top (**B**), side (**C**), and bottom (**D**) views of the complex. The map threshold used in ChimeraX is set to 0.2 (~5.2σ). Individual JAK monomers are colored as a light-to-dark gradient from the N to the C terminus. Monomer 1: FERM-SH2, light green; PK, green; TK, dark green. Monomer 2: FERM-SH2, pink; PK, orchid; TK, purple. Density corresponding to IFNλR1 is colored blue. R, receptor; PM, plasma membrane.

in which the zippers approximated the spatial constraints of dimerized TM domains (Fig. 1B) (*36*, *37*).

Coexpression of the soluble zippered IFNλR1 ICD peptide ("mini-IFNλR1") with either full-length WT JAK1 or JAK1-V657F protein resulted in increased activation loop phosphorylation as measured by Western blot, validating the construct design strategy (Fig. 1C). In the context of IFNλ signaling on cells, JAK1-IFNλR1 normally heterodimerizes with TYK2–IL-10Rβ to initiate downstream signaling. To test whether our engineered JAK1-IFNλR1 homodimer is capable of signaling in response to cytokine stimulation, we generated chimeric receptors in which the Box1/Box2 motif from IFNλR1 was substituted into the analogous position in erythropoietin receptor (EpoR), which forms an EpoR-JAK2 homodimer in response to stimulation with erythropoietin (Epo). As expected, Epo stimulation selectively induced JAK2 phosphorylation in cells expressing WT EpoR. In cells expressing the EpoR-IFNλR1 Box1/Box2 chimera, Epo stimulation resulted in phosphorylation of JAK1, indicating that the JAK1-IFNλR1 dimer is signaling-competent, and recapitulates natural JAK1-cytokine receptor dimers such as the IL-6/gp130 homodimer (Fig. 1D). On the basis of these results, we used mini-IFNλR1 to purify an active JAK1-IFNλR1 complex following coexpression in insect cells by two-step affinity-based purification (fig. S1, A to C). To further stabilize the complex, JAK1 was expressed with a C-terminal nanobody epitope tag (BC2T) which binds to the BC2 nanobody with high affinity (*38*). Dimeric BC2 nanobody was added with the logic that it might reduce conformational heterogeneity of the complex. The components coeluted as a single peak during size exclusion chromatography and were cross-linked with bis(sulfosuccinimidyl)suberate (BS3), which modifies solvent-exposed lysine residues. The cross-linked complex was vitrified on grids for cryo-EM analysis (Fig. 1, E to F, and fig. S1D).

Structure of the JAK1-IFNλR1 dimeric complex

Three-dimensional reconstruction of selected particles generated a 3.6-Å nominal resolution map of the 2:2 JAK1-IFNλR1 complex with C2 symmetry (figs. S2 and S3). Docking of individual domain crystal structures (PDB IDs: 5IXD, 4L00, and 3EYG) (*14*, *39*, *40*) was used to generate an initial model which was subject to multiple rounds of manual building and refinement, culminating in an atomic model of full-length JAK1 (Pro^{32} to Lys^{1153}) and a segment of IFNλR1 Box1/Box2 with 37 amino acids (Pro^{255} to Leu^{291}).

The JAK1-IFNλR1 complex associates into a C2 symmetric dimer (Fig. 2). At the membrane-proximal region, the N-terminal JAK1 FERM-SH2 domains are poised to receive the IFNλR1

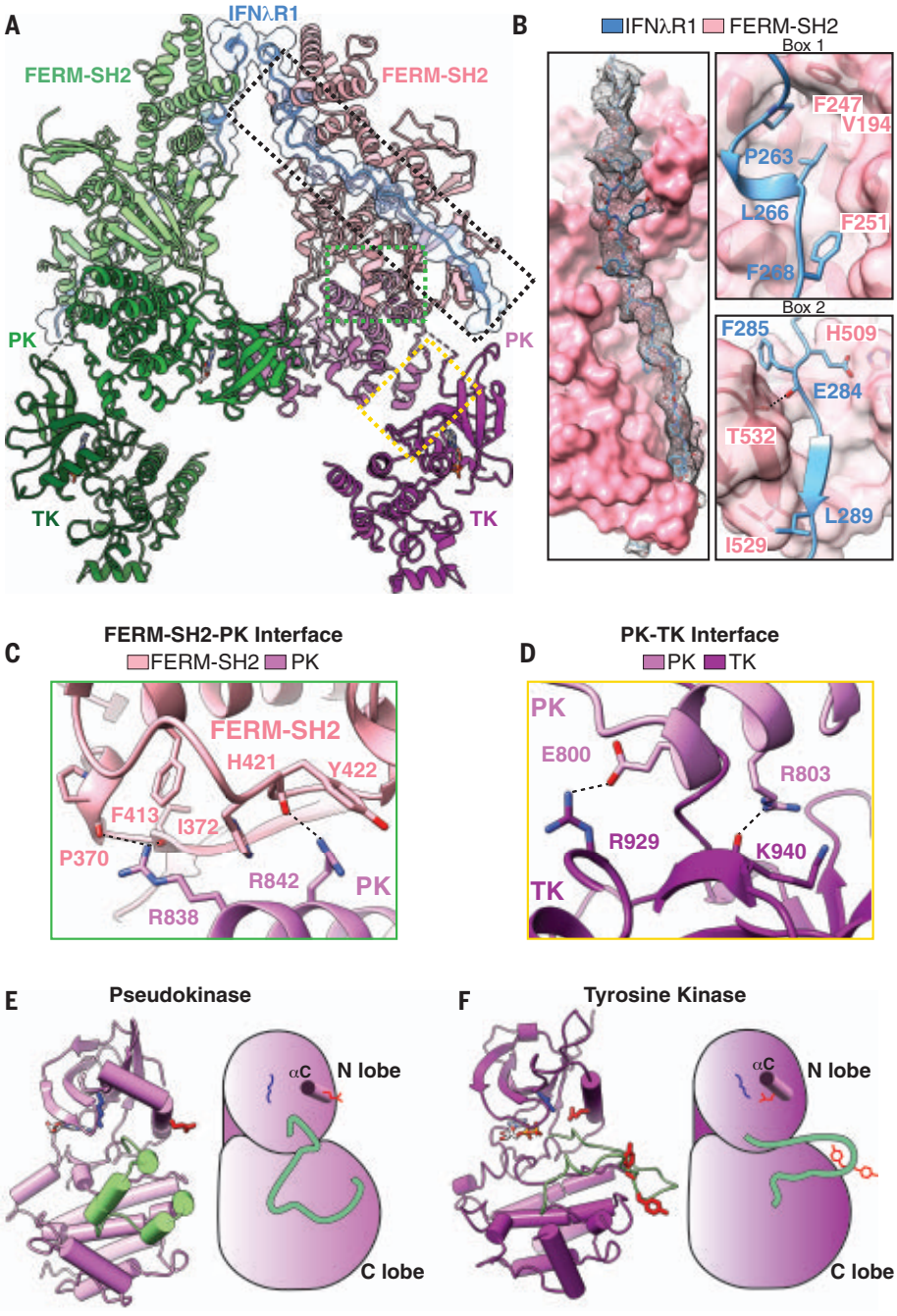


Fig. 3. Atomic model of the full-length JAK1-IFNλR1 signaling complex. (**A**) Ribbon diagram of the 2:2 JAK1-IFNλR1 complex. Dashed boxes indicate magnified views in the subsequent panels. (**B**) IFNλR1 binds JAK1 FERM and SH2 domains through N-terminal Box1 and C-terminal Box2 motifs within the receptor intracellular domain. (Left) overall interaction between IFNλR1 and FERM-SH2 shown in surface representation with peptide density from the cryo-EM map shown as black mesh contoured at ~6.1σ. (Upper right) IFNλR1 Box1 motif binds the JAK1 FERM domain via a conserved PXXLXF motif. (Lower right) IFNλR1 Box2 motif forms an antiparallel β sheet with βG1 in the JAK1 SH2 domain. Hydrogen bonds and salt bridges are shown as black dashed lines. (**C**) Interface view of the FERM-SH2-PK domains. (**D**) Closeup view of the PK-TK interaction. (**E**) Ribbon diagram (left) and schematic (right) of the PK domain in standard view. Residues corresponding to the activation loop in a functional tyrosine kinase are shown in pale green. Active site Lys^{621} is shown in blue and catalytic Glu^{636} on αC helix is shown in red. (**F**) Ribbon diagram (left) and schematic (right) of the TK domain in standard view. The TK activation loop is colored pale green with tyrosine residues Tyr^{1033} and Tyr^{1034} colored red. The catalytic Glu^{924} (red) facing inward toward Lys^{907} (blue) in the kinase active site. Amino acid abbreviations: F, Phe; V, Val; P, Pro; L, Leu; H, His; E, Glu; I, Ile; Y, Tyr; R, Arg; K, Lys; T, Thr; X, unspecified amino acid.

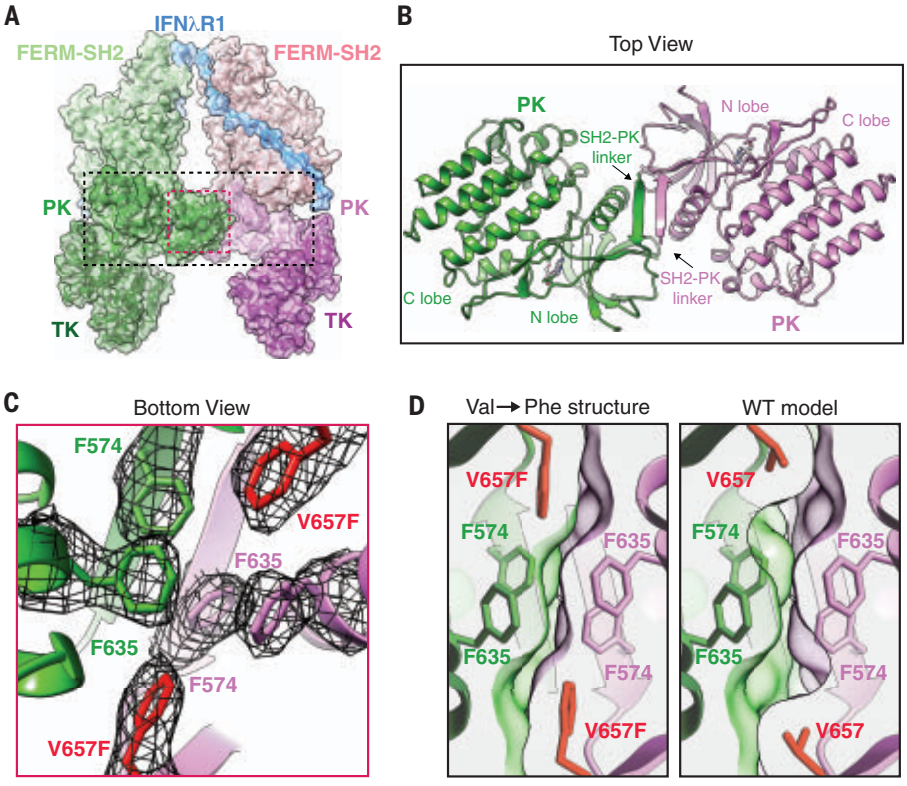


Fig. 4. JAK1 dimerization is mediated by the pseudokinase domain and enhanced by the oncogenic Val→Phe mutation. (**A**) Ribbon diagram of the JAK1-IFNλR1 complex with semi-transparent surface. Dashed boxes indicate magnified views in the subsequent panels. (**B**) Top view of the PK dimer at the center of the active JAK1 complex. The structure is shown as a ribbon diagram with nucleotides shown as sticks. Labels indicate the PK N lobe, C lobe, and SH2-PK linker. (**C**) Bottom view of the Phe triad with the cryo-EM density shown as black mesh contoured at ~9σ. The oncogenic V657F mutation is highlighted in red. (**D**) V657F enhances shape complementarity of the PK dimerization interface. Cross-section view of the PK-PK interface as seen from the bottom with the V657F cryo-EM structure compared with a model of WT Val^{657}. The WT model was created using Coot and surface clipping was set at Phe/Val^{657} Cβ for both panels to facilitate comparison. Amino acid abbreviations: F, Phe; V, Val.

ICDs as they would extend from the TM regions mimicked by the GCN4 zippers. The FERM-SH2 modules sit above inward-facing PK domains, which form a head-to-head dimer at the center of the complex. The close association between the FERM-SH2 and PK domains positions the C-terminal TK domains at the base of the JAK1 dimer, facing outwards with their catalytic clefts accessible for phosphotransferase activity. The relative positions of the kinase domains may be stabilized by the tandem BC2 nanobody bound at their C termini for imaging.

Each JAK1-IFNλR1 unit consists of four interacting modules: (i) IFNλR1 binding to JAK1 FERM-SH2, (ii) FERM domain packing against the PK C lobe, (iii) the PK domain interacting with the N lobe of TK, and (iv) the central PK dimer interface (Fig. 3A and table S1). At the membrane-proximal region of the complex, continuous density is observed for the IFNλR1 peptide, which binds along an extended groove on the surface of the FERM-SH2 through its Box1 and Box2 motifs, burying ~1650 Å^2 of surface (Fig. 3B). The IFNλR1 Box1 PXXLXF motif required for JAK1 binding forms a short 3_{10} helix which positions Leu^{266} and Phe^{268} of the peptide into a hydrophobic pocket in the JAK1 FERM domain consisting of Val^{194}, Phe^{247}, and Phe^{251} (Fig. 3B, top right), similar to a crystal structure of human JAK1 FERM-SH2 bound to IFNλR1 (*14*). We also observe density for 22 amino acids (Glu^{270} to Leu^{291}) constituting the C-terminal portion of the peptide where IFNλR1 is held in the SH2 peptide binding groove by a salt bridge interaction between Glu^{284} in IFNλR1 and His^{509} in the JAK1 SH2 domain, and a hydrogen bonding interaction between Thr^{532} in SH2 βG1 and the backbone carbonyl of IFNλR1 Phe^{285}. Beneath these specific interactions, IFNλR1 Box2 Asp^{287} to Leu^{289} forms an antiparallel β sheet with βG1 of JAK1 SH2 before the ICD exits the FERM-SH2 module and adopts a molten globule disordered state in the cytosol (*41*) that can freely interact with the kinase domains (Fig. 3B, bottom right).

Below the peptide binding region, the JAK1 FERM domain forms a broad interface with the C lobe and catalytic loop of the PK domain, burying ~1100 Å^2. At the core of this interface, the base of FERM-SH2 interacts with tandem arginine residues on successive helical turns of the PK αI helix. Arg^{838} in PK αI contacts residues P^{370} and I^{372} in FERM, whereas Arg^{842} interacts with the backbone and side chain of Tyr^{422} in the FERM-SH2 linker (Fig. 3C and fig. S4A). At the opposite face of the PK C lobe, the PK-αG helix forms a limited interaction with the N lobe of the TK domain and the PK-TK linker which buries 580 Å^2 of surface area (Fig. 3D). This site consists of a salt bridge between PK-αG Glu^{800} and Arg^{929} in TK-αC and is stabilized by a hydrogen bond between PK-Arg^{803} and the backbone carbonyl group of Lys^{940} in TK-β4.

The PK domains adopt an inactive conformation as evidenced by a closed activation loop and an outward rotation of the catalytic glutamate on the C helix (Fig. 3E). Although we observe adenosine within the nucleotide binding site, the PK domain lacks the canonical DFG motif necessary for catalytic activity, consistent with a regulatory—as opposed to catalytic—role in JAK signaling. By contrast, the TK domains adopt an active conformation with an open activation loop, catalytic glutamate facing inward toward the active site, and ADP bound at the nucleotide binding site (Fig. 3F).

Pseudokinase dimerization and stabilization by oncogenic Val→Phe mutation

The central fulcrum of the JAK1 homodimer is formed by the SH2-PK linker and PK N lobes from individual JAK1 monomers, which interact through a tightly packed hydrophobic cluster of six phenylalanine residues, in addition to an antiparallel β sheet (Fig. 4, A and B). At the membrane-proximal region of this interaction module, antiparallel β sheets from SH2-PK linkers form a lid that projects Phe^{574} into the hydrophobic interface (Fig. 4C and fig. S4B). Below the lid, Phe^{635} from the αC helix abuts oncogenic V657F (corresponding to JAK2 V617F) on β4, completing the phenylalanine triad in the JAK1 monomer. Mutation of surrounding phenylalanine residues [JAK2 Phe^{537}→Ala (mJAK1 Phe^{574}) and JAK2 Phe^{595}→Ala (mJAK1 Phe^{635})] disrupts the ability of VF to activate JAK2 (*39*, *42*). Furthermore, mutation of JAK2 Phe^{595} (mJAK1 Phe^{635})—which is central to the PK interface and packs against VF—also suppresses constitutive activation of JAK2 by a range of other clinical mutants across the PK domain (*43*). Thus the PK interface is key to ligand-independent activity of many clinically relevant MPN mutations. To better

understand the structural influence of the V→F mutation, we modeled the WT Val[657] into the structure. The smaller Val side chain results in an unfilled pocket within the dimer interface and correspondingly poorer shape complementarity (VF: 0.53, WT: 0.51), decreasing buried surface area of the side chain by ~20% (from 67 to 55 Å^2) (Fig. 4D and table S2). The hydrophobic Phe triad may also favor desolvation of the JAK monomer, further favoring dimer formation. It is well established that the corresponding V617F JAK2 mutations in all JAK family members result in constitutive activity (*34*, *35*). We generated a homology model of the JAK2 PK dimer on the basis of a previously published structure of JAK2 PK monomer (*42*). Consistent with a shared mechanism for activation by V617F, the conserved Phe side chains play a similar structural role similar to those seen in JAK1, indicating that the JAK1 structural results are generalizable to the JAK/TYK family (figs. S5 and S6).

The PK dimer interface that we visualize for the JAK1 VF mutant is likely "on pathway," stabilizing the same dimerization mode formed by cytokine-mediated activation of nonmutated JAKs. We suggest that the VF mutation simply enhances the tendency of the PK domains to naturally dimerize by improving structural and hydrophobic complementarity of the WT dimer interface. Previous structure-function data have shown that WT JAK2 requires the PK domain to enhance ligand-induced dimerization (*3*), and mutation of JAK2 Phe[595]→Ala (mJAK1 Phe[635]) in the context of WT JAK negatively influences cytokine-mediated signaling (*43*). Thus, we surmise that the WT PK interface is "detuned" relative to the VF mutant, in order to dimerize only under conditions of ligand-mediated receptor activation—an effect exploited by the Val→Phe mutation.

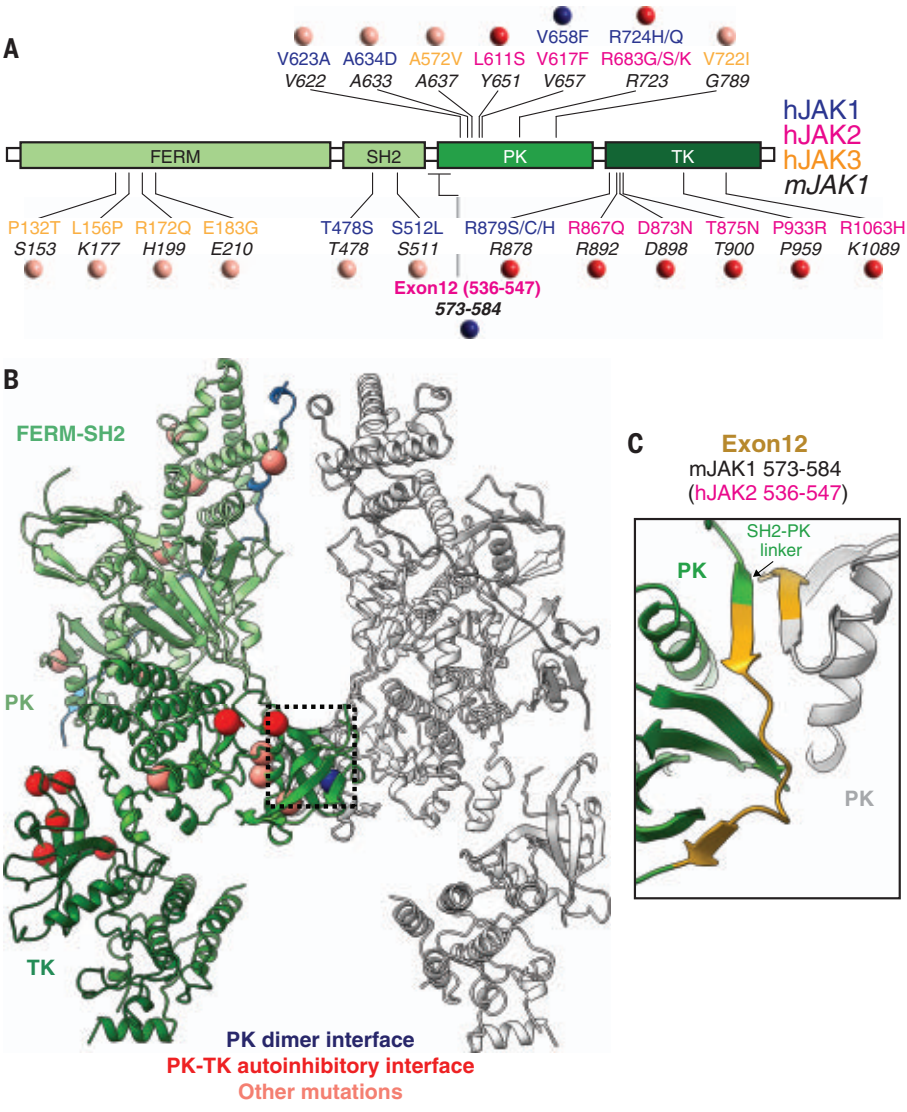


Fig. 5. Mapping human gain-of-function mutations on JAK1 suggests multiple mechanisms of oncogenic activation. (**A**) Linear diagram of JAK domains showing the location of human gain-of-function mutations. Location of patient mutations in hJAK1 (blue), hJAK2 (pink), and hJAK3 (yellow) are shown above the analogous position in mJAK1 (*44*). Colored circles indicate classification of mutations on the basis of their locations at the active PK dimer interface (blue), in the autoinhibitory PK-TK interface according to a previously reported crystal packing structure of TYK2 (red) (*18*), or at sites with unknown function (salmon). (**B**) Structure of the active JAK1-IFNλR1 complex with the position of oncogenic mutations shown as balls colored according to the proposed mechanism of action as described above. (**C**) Closeup of the PK dimer interface highlighting the residues in mJAK1 corresponding to hJAK2 exon 12, which has previously been identified as a hotspot for oncogenic mutations. Amino acid abbreviations: P, Pro; T, Thr; S, Ser; L, Leu; K, Lys; R, Arg; Q, Gln; E, Glu; G, Gly; C, Cys; D, Asp; N, Asn; V, Val; A, Ala; H, His; Y, Tyr; F, Phe.

Human gain-of-function mutations suggest a two-step mechanism for JAK activation

GOF mutations in JAK family members result in a diverse set of hematological malignancies including acute myeloid leukemia (AML), B and T cell acute lymphoblastic leukemia (B-ALL and T-ALL), and MPN. Although the Val→Phe mutation in JAK2 is best characterized, a wide variety of JAK mutations have been identified with distinct phenotypic outcomes (Fig. 5, A and B) (*44*) and many of these mutations map to the PK domain, including the JAK2 Arg[683]→Gly mutation associated with familial thrombocytosis (*45*). Previous work has identified exon12 within JAK2 to be a hotspot for oncogenic mutation (*46*, *47*). Notably, the exon 12 region of JAK2 maps to the SH2-PK linker in our JAK1 structure, which contributes to the PK dimer interface through the formation of antiparallel β sheets (Fig. 5C). However, another set of mutations that map to the N lobe of the TK, including JAK2 Thr[875]→Asn, are solvent exposed in the active JAK structure, suggesting that their mechanism of action may be distinct from Val→Phe.

To better understand how TK mutations activate JAK signaling, we aligned a previously reported structure of the autoinhibited TYK2 PK-TK domain fragment to the FERM-SH2-PK module from the full-length JAK1 dimer complex (Fig. 6A, left) (*18*). This model suggests a compact JAK monomer in which the TK is folded back on the FERM-SH2 domain, thereby occluding the activation loop and kinase active site to mediate autoinhibition. This closed state is incompatible with JAK dimerization, as a result of a steric clash between the PK domain and opposing FERM domain within the JAK dimer. However, previous negative-stain EM imaging of a JAK1

monomer suggests that it adopts a dynamic range of conformational states from a compact "closed" state to an extended "open" state, which may be compatible with dimerization (*8*). In the absence of receptor dimerization, this open state is likely transient. However, activation by cytokine-mediated receptor dimerization or VF mutation results in formation of a JAK dimer that may shift the equilibrium away from the autoinhibited state to an open state, thus releasing the TK for full activity while also driving close proximity of apposing TKs to facilitate transphosphorylation.

This two-step model of JAK activation predicts that oncogenic mutations may act by one of two possible mechanisms: (i) destabilizing the autoinhibited state (Fig. 6A, left) or (ii) stabilizing the dimeric active state (Fig. 6A, right). Consistent with this model, paired analysis of JAK2 phosphorylation and receptor dimerization as measured by single-molecule receptor tracking at physiological expression levels identified two classes of activating mutations: those that enhance JAK2 phosphorylation without affecting receptor dimerization and those that increase both JAK2 phosphorylation and receptor dimerization (*3*). Mapping these two classes of mutations onto the active JAK1-IFNλR1 structure indicates that mutations which increase JAK phosphorylation without inducing dimerization cluster at the FERM-PK-TK interface in the autoinhibited model, suggesting that destabilizing this interdomain interaction releases JAK from autoinhibition (Fig. 6B). Conversely, those mutations that induce JAK phosphorylation and receptor dimerization reside at the PK-PK dimerization interface, favoring JAK dimerization (Fig. 6C). Once the autoinhibition is relieved by JAK dimerization, the TK is ideally positioned at the base of the structure to receive the receptor intracellular domain peptide exiting the bottom of the FERM-SH2 groove to phosphorylate tyrosine residues that serve as STAT recruitment sites (Fig. 6D).

Discussion

The cryo-EM structure of activated full-length JAK1 associated with IFNλR1 ICD provides a snapshot of a complete intracellular signaling assembly at the initiating step of both cytokine-induced and oncogenic JAK-STAT signaling. Collectively, the active JAK1-IFNλR1 dimer structure—along with a wealth of previously reported biochemical and patient mutational data—suggests a mechanism for ligand-mediated JAK activation, which is then exploited by "on-pathway" pathogenic mutations found in blood cancers.

The two-step allosteric model we propose for JAK activation is supported by an abundance of previously reported structure-function data. For example, unlike other tyrosine kinases which require activation through phosphorylation by an upstream kinase, JAK TK domains show constitutive catalytic activity when expressed in isolation (*18*, *40*, *48*, *49*). The constitutive catalytic activity of the kinase domain is suppressed by expression of the tandem PK-TK domains, suggesting an autoinhibitory role for the PK domain (*18*, *49*). This autoinhibition has been rationalized in part by structural

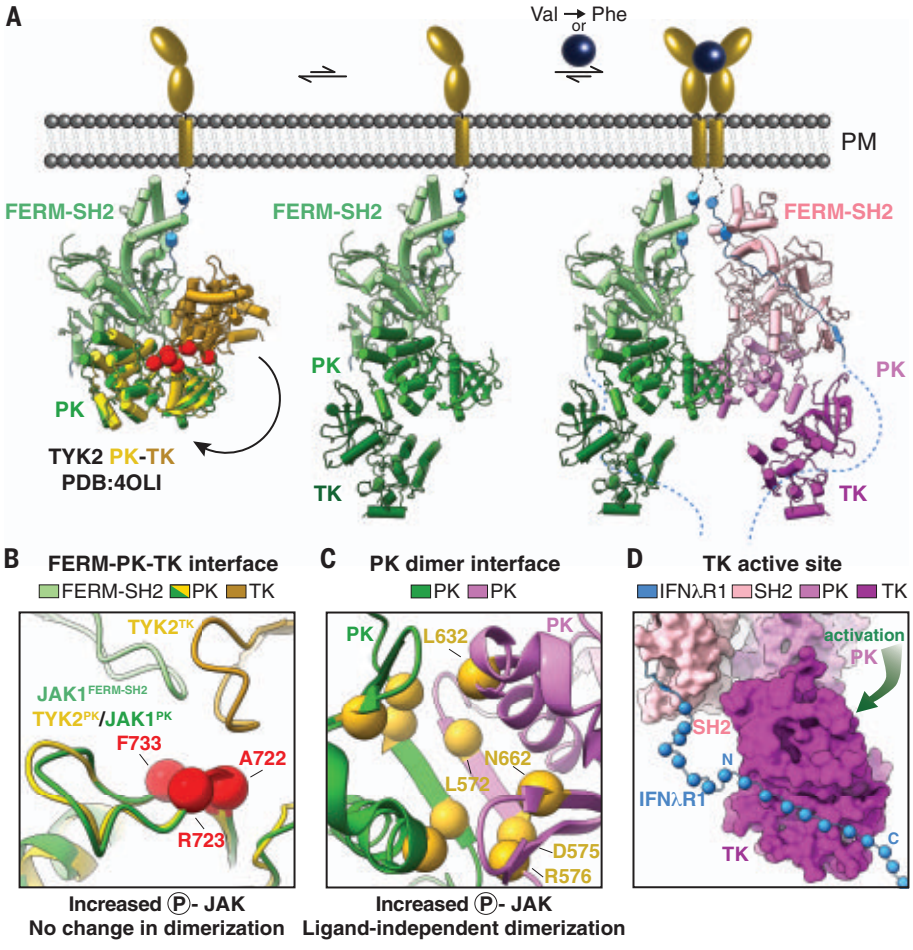


Fig. 6. Mechanistic model for JAK activation by both cytokine and oncogenic mutation. (**A**) Proposed mechanism of JAK activation by ligand-induced dimerization and Val→Phe oncogenic mutation. An autoinhibited model of full-length JAK (left) was generated by docking a crystal structure of the PK-TK domains from hTYK2 (PDB ID: 4OLI; PK, yellow; TK, gold) (*18*) into the FERM-SH2-PK from the mJAK1 cryo-EM structure. Red balls indicate the position of activating mutations in the proposed autoinhibitory interface (*44*). A dynamic equilibrium between the autoinhibited "closed" state and a partially active "open" state (middle) exposes the PK domain and SH2-PK linker to allow for JAK dimerization. Cytokine-mediated receptor dimerization or oncogenic Val→Phe mutation facilitates formation of the PK dimer, sterically preventing autoinhibition and liberating the kinase domains for phosphotransferase activity (right). (**B** and **C**) Mechanistic mutations tracking receptor dimerization and JAK2 phosphorylation support a two-step model for activation. (B) Mutations at the proposed autoinhibitory interface enhance JAK2 phosphorylation but do not affect dimerization. Closeup view of the autoinhibitory model in (A) with red balls indicating the positions of mutations previously found to increase JAK2 phosphorylation without inducing receptor dimerization (*3*). Residues are labeled according to their position in mJAK1: Ala^{722} (JAK2 Ile^{682}→Phe), Arg^{723} (JAK2 Arg^{683}→Gly), Phe^{733} (JAK2 Phe^{694}→Leu). (C) Mutations at the PK dimerization interface increase both JAK2 phosphorylation and dimerization. Closeup view of the JAK1-IFNλR1 PK dimer interface as viewed from the bottom. Yellow balls indicate the positions of mutations previously found to increase both JAK2 phosphorylation and receptor dimerization (*3*). Residues are numbered according to their position in mJAK1: Leu^{572} (JAK2 M^{535}→Ile), Asp^{575} (JAK2 His^{538}→Leu), Arg^{576} (JAK2 Lys^{539}→Leu), Leu^{632} (JAK2 Glu^{592}→Trp), Asn^{662} (JAK2 Asn^{622}→Ile). (**D**) Model of receptor phosphorylation by the JAK1 dimer. Cryo-EM structure of the JAK1-IFNλR1 dimer is shown with the TK domain in standard view. JAK1 is shown as a surface with additional residues of IFNλR1 modeled as Cα balls for every other residue exiting the JAK1 SH2 domain and projecting toward the kinase active site. Amino acid abbreviations: F, Phe; A, Ala; R, Arg; L, Leu; N, Asn; D, Asp.

models for the PK-TK domains that form head-to-head dimers through interactions between kinase N lobes, which could potentially sterically occlude substrate binding and catalytic activity (*18*, *19*). Our finding that activating Val→Phe mutations are positioned at a central PK-PK dimer interface within the active JAK1-IFNλR1 complex suggest a simple mechanism for oncogenic activation in which improved shape complementarity and hydrophobicity drive ligand-independent dimerization.

Recent discovery of highly selective TYK2 PK inhibitors, which allosterically stabilize an autoinhibited conformation underscore that the JAK family is amenable to development of allosteric rather than active-site–directed inhibitors of kinase function (*50*, *51*). One current challenge in the treatment of JAK2 V617F patients is resistance to kinase inhibitors as a result of heterodimerization and activation of JAK1 and TYK2 (*52*). Thus, new therapies could be designed to directly target the Val→Phe homodimer interface to increase specificity and reduce possibility for escape through activation of other JAKs. More generally, classification of oncogenic mutations by their mechanism of action, either through disruption of autoinhibition or increased dimerization, may provide a differential diagnostic criterion to inform therapeutic strategies.

The homodimeric JAK structure visualized here gives insight into the mechanisms underlying the "tuneability" of cytokine receptor signaling. Previous studies using genetically engineered chimeric receptors or engineered ligands have shown that the geometric variation of the cytokine receptor dimer can influence the nature of downstream signaling (*53*–*56*). The JAK PK dimer interface could potentially act as an intracellular fulcrum in the manner of a ball and socket joint to reposition the relative orientations and proximities of the C-terminal TK domains resulting in differential phosphorylation of the receptor ICDs and downstream STATs. In addition, our structure begins to rationalize how engineered cytokine ligands can elicit partial agonism. Partial agonists of cytokines have been engineered through mutational disruption of the low-affinity "site 2" cytokine receptor binding site that lowers the efficiency of receptor dimerization (*57*–*61*). The low affinity of the JAK PK dimerization interface might allow small changes in extracellular affinity to be sensitively transmitted to the downstream signaling apparatus to regulate the level of STAT activation. Indeed, the increased affinity of the JAK2 V617F mutant exploits this natural dimerization interface to drive ligand-independent signaling.

Many questions remain to refine our understanding of the cytokine receptor and JAK activation process. For example, the conformational transition from the presumed closed state of the monomeric JAK to the activated open state in the dimer is largely speculative, but resolution of this question could provide new mechanism-based opportunities to modulate cytokine receptor signaling. The resolution of the JAK1 homodimeric complex now allows for the design of small-molecule inhibitors of VF dimerization by in silico and experimental screening approaches based on the newly resolved PK dimer. Additionally, the structural basis for how cytokine receptor intracellular domains are phosphorylated at specific STAT docking sites, followed by binding, activation, and release of activated phospho-STATs, are the next frontier in the structural biology of cytokine receptor signaling.

REFERENCES AND NOTES

1. J. J. O'Shea, S. M. Holland, L. M. Staudt, *N. Engl. J. Med.* **368**, 161–170 (2013).
2. X. Wang, P. Lupardus, S. L. Laporte, K. C. Garcia, *Annu. Rev. Immunol.* **27**, 29–60 (2009).
3. S. Wilmes *et al.*, *Science* **367**, 643–652 (2020).
4. R. M. Stroud, J. A. Wells, *Sci. STKE* **2004**, re7 (2004).
5. S. S. Watowich *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **89**, 2140–2144 (1992).
6. S. R. Hubbard, *Front. Endocrinol.* **8**, 361 (2018).
7. E. Bousoik, H. Montazeri Aliabadi, *Front. Oncol.* **8**, 287 (2018).
8. P. J. Lupardus *et al.*, *Structure* **19**, 45–55 (2011).
9. M. Murakami *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **88**, 11349–11353 (1991).
10. G. R. Stark, J. E. Darnell Jr., *Immunity* **36**, 503–514 (2012).
11. L. M. LaFave, R. L. Levine, *Trends Pharmacol. Sci.* **33**, 574–582 (2012).
12. J. J. Babon, I. S. Lucet, J. M. Murphy, N. A. Nicola, L. N. Varghese, *Biochem. J.* **462**, 1–13 (2014).
13. R. Ferrao, P. J. Lupardus, *Front. Endocrinol.* **8**, 71 (2017).
14. R. Ferrao *et al.*, *Structure* **24**, 897–905 (2016).
15. D. Zhang, A. Wlodawer, J. Lubkowski, *J. Mol. Biol.* **428**, 4651–4668 (2016).
16. R. D. Ferrao, H. J. Wallweber, P. J. Lupardus, *eLife* **7**, e38089 (2018).
17. H. J. Wallweber, C. Tam, Y. Franke, M. A. Starovasnik, P. J. Lupardus, *Nat. Struct. Mol. Biol.* **21**, 443–448 (2014).
18. P. J. Lupardus *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **111**, 8025–8030 (2014).
19. Y. Shan *et al.*, *Nat. Struct. Mol. Biol.* **21**, 579–584 (2014).
20. J. B. Spangler, I. Moraga, J. L. Mendoza, K. C. Garcia, *Annu. Rev. Immunol.* **33**, 139–167 (2015).
21. Y. Luo *et al.*, *J. Allergy Clin. Immunol.* **148**, 911–925 (2021).
22. S. J. Rodig *et al.*, *Cell* **93**, 373–383 (1998).
23. H. Neubauer *et al.*, *Cell* **93**, 397–409 (1998).
24. E. Parganas *et al.*, *Cell* **93**, 385–395 (1998).
25. D. C. Thomis, C. B. Gurniak, E. Tivol, A. H. Sharpe, L. J. Berg, *Science* **270**, 794–797 (1995).
26. T. Nosaka *et al.*, *Science* **270**, 800–802 (1995).
27. P. Macchi *et al.*, *Nature* **377**, 65–68 (1995).
28. O. Kilpivaara, R. L. Levine, *Leukemia* **22**, 1813–1817 (2008).
29. E. J. Baxter *et al.*, *Lancet* **365**, 1054–1061 (2005).
30. C. James *et al.*, *Nature* **434**, 1144–1148 (2005).
31. R. Kralovics *et al.*, *N. Engl. J. Med.* **352**, 1779–1790 (2005).
32. R. L. Levine *et al.*, *Cancer Cell* **7**, 387–397 (2005).
33. D. A. Harrison, R. Binari, T. S. Nahreini, M. Gilman, N. Perrimon, *EMBO J.* **14**, 2857–2865 (1995).
34. C. Haan *et al.*, *Chem. Biol.* **18**, 314–323 (2011).
35. J. Staerk, A. Kallin, J. B. Demoulin, W. Vainchenker, S. N. Constantinescu, *J. Biol. Chem.* **280**, 41893–41899 (2005).
36. E. K. O'Shea, J. D. Klemm, P. S. Kim, T. Alber, *Science* **254**, 539–544 (1991).
37. X. Lu, A. W. Gross, H. F. Lodish, *J. Biol. Chem.* **281**, 7002–7011 (2006).
38. M. B. Braun *et al.*, *Sci. Rep.* **6**, 19211 (2016).
39. A. V. Toms *et al.*, *Nat. Struct. Mol. Biol.* **20**, 1221–1223 (2013).
40. N. K. Williams *et al.*, *J. Mol. Biol.* **387**, 219–232 (2009).
41. G. Skiniotis, P. J. Lupardus, M. Martick, T. Walz, K. C. Garcia, *Mol. Cell* **31**, 737–748 (2008).
42. R. M. Bandaranayake *et al.*, *Nat. Struct. Mol. Biol.* **19**, 754–759 (2012).
43. A. Dusa, C. Mouton, C. Pecquet, M. Herman, S. N. Constantinescu, *PLOS ONE* **5**, e11157 (2010).
44. E. Chen, L. M. Staudt, A. R. Green, *Immunity* **36**, 529–541 (2012).
45. G. Carreño-Tarragona *et al.*, *Leukemia* **35**, 3295–3298 (2021).
46. L. M. Scott, *Am. J. Hematol.* **86**, 668–676 (2011).
47. L. M. Scott *et al.*, *N. Engl. J. Med.* **356**, 459–468 (2007).
48. A. Sanz Sanz *et al.*, *Biochim. Biophys. Acta* **1844**, 1835–1841 (2014).
49. P. Saharinen, O. Silvennoinen, *J. Biol. Chem.* **277**, 47954–47963 (2002).
50. J. R. Burke *et al.*, *Sci Transl. Med.* **11**, eaaw1736 (2019).
51. J. S. Tokarski *et al.*, *J. Biol. Chem.* **290**, 11061–11074 (2015).
52. P. Koppikar *et al.*, *Nature* **489**, 155–159 (2012).
53. I. Moraga *et al.*, *Cell* **160**, 1196–1208 (2015).
54. K. Mohan *et al.*, *Science* **364**, eaav7532 (2019).
55. J. Staerk *et al.*, *EMBO J.* **30**, 4398–4413 (2011).
56. A. J. Brooks *et al.*, *Science* **344**, 1249783 (2014).
57. C. R. Glassman *et al.*, *eLife* **10**, e65777 (2021).
58. F. Mo *et al.*, *Nature* **597**, 544–548 (2021).
59. C. R. Glassman *et al.*, *Cell* **184**, 983–999.e24 (2021).
60. R. A. Saxton *et al.*, *Immunity* **54**, 660–672.e9 (2021).
61. R. A. Saxton *et al.*, *Science* **371**, eabc8433 (2021).

ACKNOWLEDGMENTS

We thank R. Fernandes and other members of the Garcia Laboratory for thoughtful discussion and helpful feedback. Cryo-EM data were collected at the Stanford cryo-EM center (cEMc). We thank E. Montabana and Y.-T. Li for generous support. **Funding:** This work was supported by National Institutes of Health grant R37AI51321 (to K.C.G.); Howard Hughes Medical Institute (to K.C.G.); Ludwig Institute for Cancer Research (to K.C.G.); Helen Hay Whitney Foundation (to R.A.S.); National Science Foundation Graduate Research Fellowship DGE-1656518 (to C.R.G.); Human Frontier Science Program Organization Fellowship LT000011/2016-L (to N.T.). **Author contributions:** Conceptualization: C.R.G., N.T., and K.C.G. Methodology: K.C.G., C.R.G. and N.T. Investigation: C.R.G., N.T., R.A.S., and K.M.J. Funding acquisition: K.C.G. Project administration: K.C.G. Supervision: K.C.G. Writing – original draft: K.C.G., C.R.G., N.T., and P.J.L. Writing – review and editing: K.C.G., C.R.G., N.T., K.M.J., R.A.S., and P.J.L. **Competing interests:** K.C.G. is the founder of Synthekine. All other authors declare no competing interests. **Data and materials availability:** The cryo-EM maps have been deposited in the Electron Microscopy Data Bank (EMDB) under accession code EMD-25715. The model coordinates have been deposited in the Protein Data Bank (PDB) under accession code 7T6F.

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abn8933
Materials and Methods
Figs. S1 to S6
Tables S1 and S2
References (*62*–*73*)
MDAR Reproducibility Checklist

29 December 2021; accepted 24 February 2022
Published online 10 March 2022
10.1126/science.abn8933

PARTICLE PHYSICS

High-precision measurement of the *W* boson mass with the CDF II detector

CDF Collaboration†‡, T. Aaltonen[1,2], S. Amerio[3,4], D. Amidei[5], A. Anastassov[6], A. Annovi[7], J. Antos[8,9], G. Apollinari[6], J. A. Appel[6], T. Arisawa[10], A. Artikov[11], J. Asaadi[12], W. Ashmanskas[6], B. Auerbach[13], A. Aurisano[12], F. Azfar[14], W. Badgett[6], T. Bae[15,16,17,18,19,20,21], A. Barbaro-Galtieri[22], V. E. Barnes[23], B. A. Barnett[24], P. Barria[25,26], P. Bartos[8,9], M. Bauce[3,4], F. Bedeschi[25], S. Behari[6], G. Bellettini[25,27], J. Bellinger[28], D. Benjamin[29], A. Beretvas[6], A. Bhatti[30], K. R. Bland[31], B. Blumenfeld[24], A. Bocci[29], A. Bodek[32], D. Bortoletto[23], J. Boudreau[33], A. Boveia[34], L. Brigliadori[35,36], C. Bromberg[37], E. Brucken[1,2], J. Budagov[11]§, H. S. Budd[32], K. Burkett[6], G. Busetto[3,4], P. Bussey[38], P. Butti[25,27], A. Buzatu[38], A. Calamba[39], S. Camarda[40], M. Campanelli[41], B. Carls[42], D. Carlsmith[28], R. Carosi[25], S. Carrillo[43]§, B. Casal[44], M. Casarsa[45], A. Castro[35,36], P. Catastini[46], D. Cauz[45,47,48], V. Cavaliere[42], A. Cerri[22], L. Cerrito[41], Y. C. Chen[49], M. Chertok[50], G. Chiarelli[25], G. Chlachidze[6], K. Cho[15,16,17,18,19,20,21], D. Chokheli[11], A. Clark[51], C. Clarke[52], M. E. Convery[6], J. Conway[50], M. Corbo[6], M. Cordelli[7], C. A. Cox[50], D. J. Cox[50], M. Cremonesi[25], D. Cruz[12], J. Cuevas[44], R. Culbertson[6], N. d'Ascenzo[6], M. Datta[6], P. de Barbaro[32], L. Demortier[30], M. Deninno[35]§, M. D'Errico[3,4], F. Devoto[1,2], A. Di Canto[25,27], B. Di Ruzza[6], J. R. Dittmann[31], S. Donati[25,27], M. D'Onofrio[53], M. Dorigo[45,54], A. Driutti[45,47,48], K. Ebina[10], R. Edgar[5], A. Elagin[34], R. Erbacher[50], S. Errede[42], B. Esham[42], S. Farrington[14], J. P. Fernández Ramos[55], R. Field[43], G. Flanagan[6], R. Forrest[50], M. Franklin[46], J. C. Freeman[6], H. Frisch[34], Y. Funakoshi[10], C. Galloni[25,27], A. F. Garfinkel[23], P. Garosi[25,26], H. Gerberich[42], E. Gerchtein[6], S. Giagu[56], V. Giakoumopoulou[57], K. Gibson[33], C. M. Ginsburg[6], N. Giokaris[57]§, P. Giromini[7], V. Glagolev[11], D. Glenzinski[6], M. Gold[58], D. Goldin[12], A. Golossanov[6], G. Gomez[44], G. Gomez-Ceballos[59], M. Goncharov[59], O. González López[55], I. Gorelov[58], A. T. Goshaw[29], K. Goulianos[30], E. Gramellini[35], C. Grosso-Pilcher[34], J. Guimaraes da Costa[46], S. R. Hahn[6], J. Y. Han[32], F. Happacher[7], K. Hara[60], M. Hare[61], R. F. Harr[52], T. Harrington-Taber[6], K. Hatakeyama[31], C. Hays[14], J. Heinrich[62], M. Herndon[28], A. Hocker[6], Z. Hong[12], W. Hopkins[6], S. Hou[49], R. E. Hughes[63], U. Husemann[64], M. Hussein[37], J. Huston[37], G. Introzzi[25,65,66], M. Iori[56,67], A. Ivanov[50], E. James[6], D. Jang[39], B. Jayatilaka[6], E. J. Jeon[15,16,17,18,19,20,21], S. Jindariani[6], M. Jones[23], K. K. Joo[15,16,17,18,19,20,21], S. Y. Jun[39], T. R. Junk[6], M. Kambeitz[68], T. Kamon[15,16,17,18,19,20,21,12], P. E. Karchin[52], A. Kasmi[31], Y. Kato[69], W. Ketchum[34], J. Keung[62], B. Kilminster[6], D. H. Kim[15,16,17,18,19,20,21], H. S. Kim[6], J. E. Kim[15,16,17,18,19,20,21], M. J. Kim[7], S. H. Kim[60], S. B. Kim[15,16,17,18,19,20,21], Y. J. Kim[15,16,17,18,19,20,21], Y. K. Kim[34], N. Kimura[10], M. Kirby[6], K. Kondo[10]§, D. J. Kong[15,16,17,18,19,20,21], J. Konigsberg[43], A. V. Kotwal[29]*, M. Kreps[68], J. Kroll[62], M. Kruse[29], T. Kuhr[68], M. Kurata[60], A. T. Laasanen[23], S. Lammel[6], M. Lancaster[41], K. Lannon[63], G. Latino[25,26], H. S. Lee[15,16,17,18,19,20,21], J. S. Lee[15,16,17,18,19,20,21], S. Leo[42], S. Leone[25], J. D. Lewis[6], A. Limosani[29], E. Lipeles[62], A. Lister[51], Q. Liu[23], T. Liu[6], S. Lockwitz[64], A. Loginov[64]§, D. Lucchesi[3,4], A. Lucà[7,6], J. Lueck[68], P. Lujan[22], P. Lukens[6], G. Lungu[30], J. Lys[22]§, R. Lysak[8,9], R. Madrak[6], P. Maestro[25,26], S. Malik[30], G. Manca[53], A. Manousakis-Katsikakis[57], L. Marchese[35], F. Margaroli[56], P. Marino[25,70], K. Matera[42], M. E. Mattson[52], A. Mazzacane[6], P. Mazzanti[35], R. McNulty[53], A. Mehta[53], P. Mehtala[1,2], A. Menzione[25]§, C. Mesropian[30], T. Miao[6], E. Michielin[3,4], D. Mietlicki[5], A. Mitra[49], H. Miyake[60], S. Moed[6], N. Moggi[35], C. S. Moon[15,16,17,18,19,20,21], R. Moore[6], M. J. Morello[25,70], A. Mukherjee[6], Th. Muller[68], P. Murat[6], M. Mussini[35,36], J. Nachtman[6], Y. Nagai[60], J. Naganoma[10], I. Nakano[71], A. Napier[61], J. Nett[12], T. Nigmanov[33], L. Nodulman[13], S. Y. Noh[15,16,17,18,19,20,21], O. Norniella[42], L. Oakes[14], S. H. Oh[29], Y. D. Oh[15,16,17,18,19,20,21], T. Okusawa[69], R. Orava[1,2], L. Ortolan[40], C. Pagliarone[45], E. Palencia[44], P. Palni[58], V. Papadimitriou[6], W. Parker[28], G. Pauletta[45,47,48], M. Paulini[39], C. Paus[59], T. J. Phillips[29], G. Piacentino[6], E. Pianori[62], J. Pilot[50], K. Pitts[42], C. Plager[72], L. Pondrom[28], S. Poprocki[6], K. Potamianos[22], A. Pranko[22], F. Prokoshin[11], F. Ptohos[7], G. Punzi[25,27], I. Redondo Fernández[55], P. Renton[14], M. Rescigno[56], F. Rimondi[35]§, L. Ristori[25,6], A. Robson[38], T. Rodriguez[62], S. Rolli[61], M. Ronzani[25,27], R. Roser[6], J. L. Rosner[34], F. Ruffini[25,26], A. Ruiz[44], J. Russ[39], V. Rusu[6], W. K. Sakumoto[32], Y. Sakurai[10], L. Santi[45,47,48], K. Sato[60], V. Saveliev[6], A. Savoy-Navarro[6], P. Schlabach[6], E. E. Schmidt[6], T. Schwarz[5], L. Scodellaro[44], F. Scuri[25], S. Seidel[58], Y. Seiya[69], A. Semenov[11], F. Sforza[25,27], S. Z. Shalhout[50], T. Shears[53], P. F. Shepard[33], M. Shimojima[60], M. Shochet[34], I. Shreyber-Tecker[73], A. Simonenko[11], K. Sliwa[61], J. R. Smith[50], F. D. Snider[6], H. Song[33], V. Sorin[40], R. St. Denis[38]§, M. Stancari[6], D. Stentz[6], J. Strologas[58], Y. Sudo[60], A. Sukhanov[6], I. Suslov[11], K. Takemasa[60], Y. Takeuchi[60], J. Tang[34], M. Tecchio[5], P. K. Teng[49], J. Thom[6], E. Thomson[62], V. Thukral[12], D. Toback[12], S. Tokar[8,9], K. Tollefson[37], T. Tomura[60], S. Torre[7], D. Torretta[6], P. Totaro[3], M. Trovato[25,70], F. Ukegawa[60], S. Uozumi[15,16,17,18,19,20,21], F. Vázquez[43], G. Velev[6], K. Vellidis[57], C. Vernieri[25,70], M. Vidal[23], R. Vilar[44], J. Vizán[44], M. Vogel[58], G. Volpi[7], P. Wagner[62], R. Wallny[6], S. M. Wang[49], D. Waters[41], W. C. Wester III[6], D. Whiteson[62], A. B. Wicklund[13], S. Wilbur[50], H. H. Williams[62], J. S. Wilson[5], P. Wilson[6], B. L. Winer[63], P. Wittich[6], S. Wolbers[6], H. Wolfmeister[63], T. Wright[5], X. Wu[51], Z. Wu[31], K. Yamamoto[69], D. Yamato[69], T. Yang[6], U. K. Yang[15,16,17,18,19,20,21], Y. C. Yang[15,16,17,18,19,20,21], W.-M. Yao[22], G. P. Yeh[6], K. Yi[6], J. Yoh[6], K. Yorita[10], T. Yoshida[69], G. B. Yu[15,16,17,18,19,20,21], I. Yu[15,16,17,18,19,20,21], A. M. Zanetti[45], Y. Zeng[29], C. Zhou[29], S. Zucchelli[35,36]

The mass of the *W* boson, a mediator of the weak force between elementary particles, is tightly constrained by the symmetries of the standard model of particle physics. The Higgs boson was the last missing component of the model. After observation of the Higgs boson, a measurement of the *W* boson mass provides a stringent test of the model. We measure the *W* boson mass, M_W, using data corresponding to 8.8 inverse femtobarns of integrated luminosity collected in proton-antiproton collisions at a 1.96 tera–electron volt center-of-mass energy with the CDF II detector at the Fermilab Tevatron collider. A sample of approximately 4 million *W* boson candidates is used to obtain $M_W = 80{,}433.5 \pm 6.4_{\text{stat}} \pm 6.9_{\text{syst}} = 80{,}433.5 \pm 9.4\ \text{MeV}/c^2$, the precision of which exceeds that of all previous measurements combined (stat, statistical uncertainty; syst, systematic uncertainty; MeV, mega–electron volts; *c*, speed of light in a vacuum). This measurement is in significant tension with the standard model expectation.

The observation of the Higgs boson (*1*–*4*) at the Large Hadron Collider (LHC) (*5*, *6*) has validated the last missing piece of the standard model (SM) (*7*–*9*) of elementary particle physics. This model, which incorporates quantum mechanics, special relativity, gauge symmetry, and group theory, currently describes most particle physics measurements with high accuracy. It postulates a number of experimentally established symmetries among particle properties, which tightly constrain the parameters of the model from experimental data (*10*). Given the current experimental precision and the predictive power of the SM, global fits of the model to the data render precise estimates of fundamental parameters, such as the mass of the *W* boson. As one of the mediators of the weak nuclear force, this particle is a key component of the SM framework. Its mass, one of the most important parameters in particle physics, is presently constrained by SM global fits to a relative precision of 0.01%, providing a strong motivation to test the SM by measuring the *W* boson mass to the same level of precision.

All fundamental particle masses, including that of the *W* boson, are generated in the SM through interactions with the condensate of the Higgs field in the vacuum. The formation of the condensate and the quantum excitation of this field, the Higgs boson (*2*–*4*), are parametrized but not explained by the SM. A number of hypotheses have been promulgated to provide a deeper explanation of the Higgs field, its potential, and the Higgs boson. These include supersymmetry—a spacetime symmetry relating fermions and bosons [(*11*) and references therein]—and compositeness, in which additional strong confining interactions produce the Higgs boson as a bound state [(*12*) and

references therein]. Many of these hypotheses include a source of dark matter, which is currently believed to comprise ~84% of the matter in the universe (*10*) but cannot be accounted for in the SM. Evidence for dark matter is provided by the abnormally high speeds of revolution of stars at large radii in galaxies, the velocities of galaxies in galaxy clusters, x-ray emissions sensing the temperature of hot gas in galaxy clusters, and the weak gravitational lensing of background galaxies by clusters [(*13*, *14*) and references therein]. The additional symmetries and fields in these extensions to the SM would modify (*15*–*24*) the estimated mass of the W boson (Fig. 1) relative to the SM expectation (*10*) of $M_W = 80{,}357 \pm 4_{\text{inputs}} \pm 4_{\text{theory}}$ MeV (*25*). The SM expectation is derived from a combination of analytical relations from perturbative expansions on the basis of the internal symmetries of the theory and a set of high-precision measurements of observables, including the Z and Higgs boson masses, the top-quark mass, the electromagnetic (EM) coupling, and the muon lifetime, which are used as inputs to the analytical relations. The uncertainties in the SM expectation arise from uncertainties in the data-constrained input parameters (*10*) and from missing higher-order terms in the perturbative SM calculation (*26*, *27*). An example of a nonsupersymmetric SM extension is a modified Higgs sector that includes an additional scalar field with no SM gauge interactions, which predicts an M_W shift of up to ~100 MeV (*17*), depending on the mass of the additional scalar particle and its interaction with the SM Higgs boson. A light (heavy) additional scalar particle would induce a positive (negative) M_W shift. Similar but smaller shifts of 20 to 40 MeV have been calculated in an extension that contains a second Higgs-like field with the same gauge charges as the SM Higgs field (*18*). Implications of very weakly interacting new particles such as "dark photons" (*19*), restoration of parity conservation in the weak interaction (*20*), the possible composite nature of the Higgs boson (*21*), and model-independent modifications of the Higgs boson's interactions (*22*–*24*) have also been evaluated.

Previous analyses (*28*–*44*) yield a value of $M_W = 80{,}385 \pm 15$ MeV (*45*) from the combination of Large Electron-Positron (LEP) collider and Fermilab Tevatron collider measurements. The ATLAS Collaboration has recently reported a measurement, $M_W = 80{,}370 \pm 19$ MeV (*46*, *47*), that is comparable in precision to the Tevatron results. The LEP, Tevatron, and ATLAS measurements have not yet been combined, pending evaluation of uncertainty correlations.

CDF experiment at Tevatron

The Fermilab Tevatron produced high yields of W bosons from 2002 to 2011 through quark-antiquark annihilation in collisions of protons (p) and antiprotons ($\bar{p}$) at a center-of-mass

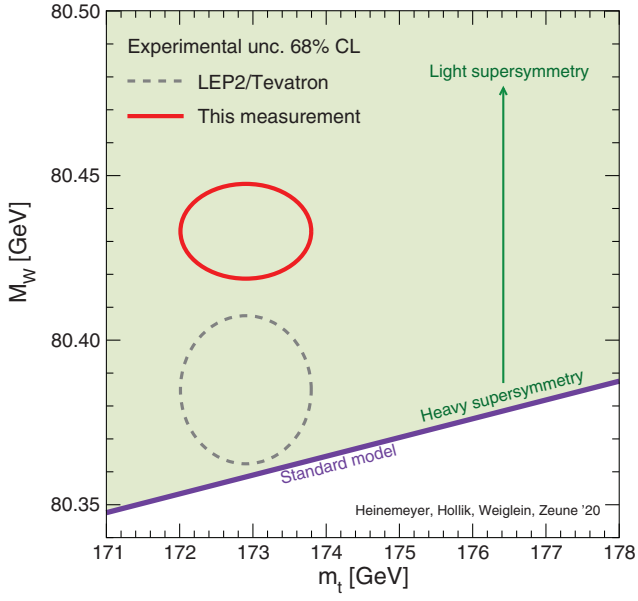


Fig. 1. Experimental measurements and theoretical predictions for the *W* boson mass. The red continuous ellipse shows the M_W measurement reported in this paper and the global combination of top-quark mass measurements, $m_t = 172.89 \pm 0.59$ GeV (*10*). The correlation between the M_W and m_t measurements is negligible. The gray dashed ellipse, updated (*16*) from (*15*), shows the 68% confidence level (CL) region allowed by the previous LEP-Tevatron combination $M_W = 80{,}385 \pm 15$ MeV (*45*) and m_t (*10*). That combination includes the M_W measurement published by CDF in 2012 (*41*, *43*), which this paper both updates (increasing M_W by 13.5 MeV) and subsumes. As an illustration, the green shaded region (*15*) shows the predicted mass of the W boson as a function of the top-quark mass m_t in the minimal supersymmetric extension (one of many possible extensions) of the standard model (SM), for a range of supersymmetry model parameters as described in (*15*). The thick purple line at the lower edge of the green region corresponds to the SM prediction with the Higgs boson mass measured at the LHC (*10*) used as input. The arrow indicates the variation of the predicted W boson mass as the mass scale of supersymmetric particles is lowered. The supersymmetry model parameter scan is for illustrative purposes and does not incorporate all exclusions from direct searches at the LHC. unc., uncertainty.

[1]Division of High Energy Physics, Department of Physics, University of Helsinki, FIN-00014, Helsinki, Finland. [2]Helsinki Institute of Physics, FIN-00014, Helsinki, Finland. [3]Istituto Nazionale di Fisica Nucleare, Sezione di Padova, I-35131 Padova, Italy. [4]University of Padova, I-35131 Padova, Italy. [5]University of Michigan, Ann Arbor, MI 48109, USA. [6]Fermi National Accelerator Laboratory, Batavia, IL 60510, USA. [7]Laboratori Nazionali di Frascati, Istituto Nazionale di Fisica Nucleare, I-00044 Frascati, Italy. [8]Comenius University, 842 48 Bratislava, Slovakia. [9]Institute of Experimental Physics, 040 01 Kosice, Slovakia. [10]Waseda University, Tokyo 169, Japan. [11]Joint Institute for Nuclear Research, Dubna RU-141980, Russia. [12]Mitchell Institute for Fundamental Physics and Astronomy, Texas A&M University, College Station, TX 77843, USA. [13]Argonne National Laboratory, Argonne, IL 60439, USA. [14]University of Oxford, Oxford OX1 3RH, UK. [15]Center for High Energy Physics, Kyungpook National University, Daegu 702-701, Korea. [16]Seoul National University, Seoul 151-742, Korea. [17]Sungkyunkwan University, Suwon 440-746, Korea. [18]Korea Institute of Science and Technology Information, Daejeon 305-806, Korea. [19]Chonnam National University, Gwangju 500-757, Korea. [20]Chonbuk National University, Jeonju 561-756, Korea. [21]Ewha Womans University, Seoul 120-750, Korea. [22]Ernest Orlando Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA. [23]Purdue University, West Lafayette, IN 47907, USA. [24]The Johns Hopkins University, Baltimore, MD 21218, USA. [25]Istituto Nazionale di Fisica Nucleare Pisa, I-56127 Pisa, Italy. [26]University of Siena, I-53100 Siena, Italy. [27]University of Pisa, I-56126 Pisa, Italy. [28]University of Wisconsin-Madison, Madison, WI 53706, USA. [29]Duke University, Durham, NC 27708, USA. [30]The Rockefeller University, New York, NY 10065, USA. [31]Baylor University, Waco, TX 76798, USA. [32]University of Rochester, Rochester, NY 14627, USA. [33]University of Pittsburgh, Pittsburgh, PA 15260, USA. [34]Enrico Fermi Institute, University of Chicago, Chicago, IL 60637, USA. [35]Istituto Nazionale di Fisica Nucleare Bologna, I-40127 Bologna, Italy. [36]University of Bologna, I-40127 Bologna, Italy. [37]Michigan State University, East Lansing, MI 48824, USA. [38]Glasgow University, Glasgow G12 8QQ, UK. [39]Carnegie Mellon University, Pittsburgh, PA 15213, USA. [40]Institut de Fisica d'Altes Energies, ICREA, Universitat Autonoma de Barcelona, E-08193 Bellaterra (Barcelona), Spain. [41]University College London, London WC1E 6BT, UK. [42]University of Illinois, Urbana, IL 61801, USA. [43]University of Florida, Gainesville, FL 32611, USA. [44]Instituto de Fisica de Cantabria, CSIC-University of Cantabria, 39005 Santander, Spain. [45]Istituto Nazionale di Fisica Nucleare Trieste, I-34127 Trieste, Italy. [46]Harvard University, Cambridge, MA 02138, USA. [47]Gruppo Collegato di Udine, I-33100 Udine, Italy. [48]University of Udine, I-33100 Udine, Italy. [49]Institute of Physics, Academia Sinica, Taipei, Taiwan 11529, Republic of China. [50]University of California, Davis, Davis, CA 95616, USA. [51]University of Geneva, CH-1211 Geneva 4, Switzerland. [52]Wayne State University, Detroit, MI 48201, USA. [53]University of Liverpool, Liverpool L69 7ZE, UK. [54]University of Trieste, I-34127 Trieste, Italy. [55]Centro de Investigaciones Energeticas Medioambientales y Tecnologicas, E-28040 Madrid, Spain. [56]Istituto Nazionale di Fisica Nucleare, Sezione di Roma 1, I-00185 Roma, Italy. [57]National and Kapodistrian University of Athens, 157 71 Athens, Greece. [58]University of New Mexico, Albuquerque, NM 87131, USA. [59]Massachusetts Institute of Technology, Cambridge, MA 02139, USA. [60]University of Tsukuba, Tsukuba, Ibaraki 305, Japan. [61]Tufts University, Medford, MA 02155, USA. [62]University of Pennsylvania, Philadelphia, PA 19104, USA. [63]The Ohio State University, Columbus, OH 43210, USA. [64]Yale University, New Haven, CT 06520, USA. [65]Istituto Nazionale di Fisica Nucleare Pavia, I-27100 Pavia, Italy. [66]University of Pavia, I-27100 Pavia, Italy. [67]Sapienza Università di Roma, I-00185 Roma, Italy. [68]Institut für Experimentelle Kernphysik, Karlsruhe Institute of Technology, D-76131 Karlsruhe, Germany. [69]Osaka City University, Osaka 558-8585, Japan. [70]Scuola Normale Superiore, I-56126 Pisa, Italy. [71]Okayama University, Okayama 700-8530, Japan. [72]University of California, Los Angeles, Los Angeles, CA 90024, USA. [73]Institution for Theoretical and Experimental Physics, ITEP, Moscow 117259, Russia.

***Corresponding author. Email: ashutosh.kotwal@duke.edu**

†All listed authors are members of the collaboration. ‡Visitors' institutions are listed in the supplementary materials. §Deceased.

energy of 1.96 TeV. The (anti)quark momentum distributions in the (anti)proton are the best-measured among all constituent partons of the colliding particles. The use of proton-antiproton collisions reduces uncertainties on the momenta of the partons and the corresponding M_W uncertainty relative to the LHC, where W bosons are produced from quarks or antiquarks and gluons, the latter of which have less precisely known momentum distributions. The moderate collision energy at the Tevatron further restricts the parton momenta to a range in which their distributions are known more precisely, compared with the relevant range at the LHC. The LHC detectors partially compensate with larger lepton rapidity coverage. The improved lepton resolution at the LHC detectors has a minor impact on the M_W uncertainty. Although the LHC dataset is much larger, the lower instantaneous luminosity at the Tevatron and in dedicated low-luminosity LHC runs helps to improve the resolution on certain kinematic quantities, compared with the typical LHC runs.

The data sample corresponds to an integrated luminosity of 8.8 inverse femtobarns (fb^{-1}) of $p\bar{p}$ collisions collected by the CDF II detector (*43*) between 2002 and 2011 and supersedes the earlier result obtained from a quarter of these data (*41*, *43*). In this cylindrical detector [figure 3 of (*43*)], trajectories of charged particles (tracks) produced in the collisions are measured by means of a wire drift chamber (a central outer tracking drift chamber, or COT) (*48*) immersed in a 1.4-T axial magnetic field. Energy and position measurements of particles are also provided by EM and hadronic calorimeters surrounding the COT. The calorimeter elements have a projective tower geometry, with each tower pointing back to the average beam collision point at the center of the detector. Additional drift chambers (*49*) surrounding the calorimeters identify muon candidates as penetrating particles. The momentum perpendicular to the beam axis (cylindrical z axis) is denoted as p_T (if measured in the COT) or E_T (if measured in the calorimeters). The measurement uses high-purity samples of electron and muon (together referred to as lepton) decays of the $W^{\pm}$ bosons, $W \to e\nu$ and $W \to \mu\nu$, respectively (e, electron; ν, neutrino; μ, muon).

W and Z boson event selection

Events with a candidate muon with $p_T > 18$ GeV or electron with $E_T > 18$ GeV (*50*) are selected online by the trigger system for offline analysis. The following offline criteria select fairly pure samples of $W \to \mu\nu$ and $W \to e\nu$ decays. Muon candidates must have $p_T >$ 30 GeV, with requirements on COT-track quality, calorimeter-energy deposition, and muon-chamber signals. Cosmic-ray muons are rejected with a targeted tracking algorithm (*51*). Electron candidates must have a COT track with $p_T > 18$ GeV and an EM calorimeter-energy deposition with $E_T > 30$ GeV and must meet requirements for COT track quality, matching of position and energy measured in the COT and in the calorimeter ($E_T/p_T < 1.6$), and spatial distributions of energy depositions in the calorimeters (*43*). Leptons are required to be central in pseudorapidity ($|\eta| < 1$) (*50*) and within the fiducial region where the relevant detector systems have high efficiency and uniform response. When selecting the W boson candidate sample, we suppress the Z boson background by rejecting events with a second lepton of the same flavor. Events that contain two oppositely charged leptons of the same flavor with invariant mass in the range of 66 to 116 GeV and with dilepton $p_T < 30$ GeV provide Z boson control samples ($Z \to ee$ and $Z \to \mu\mu$) to measure the detector response, resolution, and efficiency, as well as the boson p_T distributions. Details of the event selection criteria are described in (*43*).

The W boson mass is inferred from the kinematic distributions of the decay leptons (ℓ). Because the neutrino from the W boson decay is not directly detectable, its transverse momentum p_T^{ν} is deduced by imposing transverse momentum conservation. Longitudinal momentum balance cannot be imposed because most of the beam momenta are carried away by collision products that remain close to the beam axis, outside the instrumented regions of the detector. By design of the detector, such products have small transverse momentum. The transverse momentum vector sum of all detectable collision products accompanying the W or Z boson is defined as the hadronic recoil $\vec{u} = \Sigma_i E_i \sin(\theta_i)\hat{n}_i$, where the sum is performed over calorimeter towers (*52*) with energy E_i, polar angle θ_i, and transverse directions specified by unit vectors $\hat{n}_i$. Calorimeter towers containing energy deposition from the charged lepton(s) are excluded from this sum. The transverse momentum vector of the neutrino $\vec{p}_T^{\,\nu}$ is inferred as $\vec{p}_T^{\,\nu} \equiv -\vec{p}_T^{\,\ell} - \vec{u}$ from $\vec{p}_T$ conservation, where $\vec{p}_T^{\,\ell}$ is the vector p_T (E_T) of the muon (electron). In analogy with a two-body mass, the W boson transverse mass is defined using only the transverse momentum vectors as $m_T = \sqrt{2\left(p_T^{\ell} p_T^{\nu} - \vec{p}_T^{\,\ell} \cdot \vec{p}_T^{\,\nu}\right)}$ (*53*). High-purity samples of W bosons are obtained with the requirements $30 < p_T^{\ell} < 55$ GeV, $30 < p_T^{\nu} < 55$ GeV, $|\vec{u}| < 15$ GeV, and $60 < m_T < 100$ GeV. This selection retains samples containing precise M_W information and low backgrounds. The final samples of W and Z bosons consist of 1,811,700 (66,180) $W \to e\nu$ ($Z \to ee$) candidates and 2,424,486 (238,534) $W \to \mu\nu$ ($Z \to \mu\mu$) candidates.

Simulation of physical processes

The data distributions of m_T, p_T^{ℓ}, and p_T^{ν} are compared with corresponding simulated line shapes ("templates") as functions of M_W from a custom Monte Carlo simulation that has been designed and written for this analysis. A binned likelihood is maximized to obtain the mass and its statistical uncertainty. The kinematic properties of W and Z boson production and decay are simulated using the RESBOS program (*54*–*56*), which calculates the differential cross section with respect to boson mass, transverse momentum, and rapidity for boson production and decay. The calculation is performed at next-to-leading order in perturbative quantum chromodynamics (QCD), along with next-to-next-to-leading logarithm resummation of higher-order radiative quantum amplitudes. RESBOS offers one of the most accurate theoretical calculations available for these processes. The nonperturbative model parameters in RESBOS and the QCD interaction coupling strength α_s are external inputs needed to complete the description of the boson p_T spectrum and are constrained from the high-resolution dilepton $p_T^{\ell\ell}$ spectrum of the Z boson data and the p_T^W data spectrum. EM radiation from the leptons is modeled with the PHOTOS program (*57*), which is calibrated to the more accurate HORACE program (*58*, *59*). We use the NNPDF3.1 (*60*) parton distribution functions (PDFs) of the (anti)proton, as they incorporate the most complete relevant datasets of the available next-to-next-to-leading order (NNLO) PDFs. Using 25 symmetric eigenvectors of the NNPDF3.1 set, we estimate a PDF uncertainty of 3.9 MeV. We find that the CT18 (*61*), MMHT2014 (*62*), and NNPDF3.1 NNLO PDF sets produce consistent results for the W boson mass, within ±2.1 MeV of the midpoint of the interval spanning the range of

Table 1. Individual fit results and uncertainties for the M_W measurements. The fit ranges are 65 to 90 GeV for the m_T fit and 32 to 48 GeV for the p_T^{ℓ} and p_T^{ν} fits. The χ^2 of the fit is computed from the expected statistical uncertainties on the data points. The bottom row shows the combination of the six fit results by means of the best linear unbiased estimator (*66*).

Distribution	W boson mass (MeV)	χ^2/dof
$m_T(e, \nu)$	$80{,}429.1 \pm 10.3_{stat} \pm 8.5_{syst}$	39/48
$p_T^{\ell}(e)$	$80{,}411.4 \pm 10.7_{stat} \pm 11.8_{syst}$	83/62
$p_T^{\nu}(e)$	$80{,}426.3 \pm 14.5_{stat} \pm 11.7_{syst}$	69/62
$m_T(\mu, \nu)$	$80{,}446.1 \pm 9.2_{stat} \pm 7.3_{syst}$	50/48
$p_T^{\ell}(\mu)$	$80{,}428.2 \pm 9.6_{stat} \pm 10.3_{syst}$	82/62
$p_T^{\nu}(\mu)$	$80{,}428.9 \pm 13.1_{stat} \pm 10.9_{syst}$	63/62
Combination	$80{,}433.5 \pm 6.4_{stat} \pm 6.9_{syst}$	7.4/5

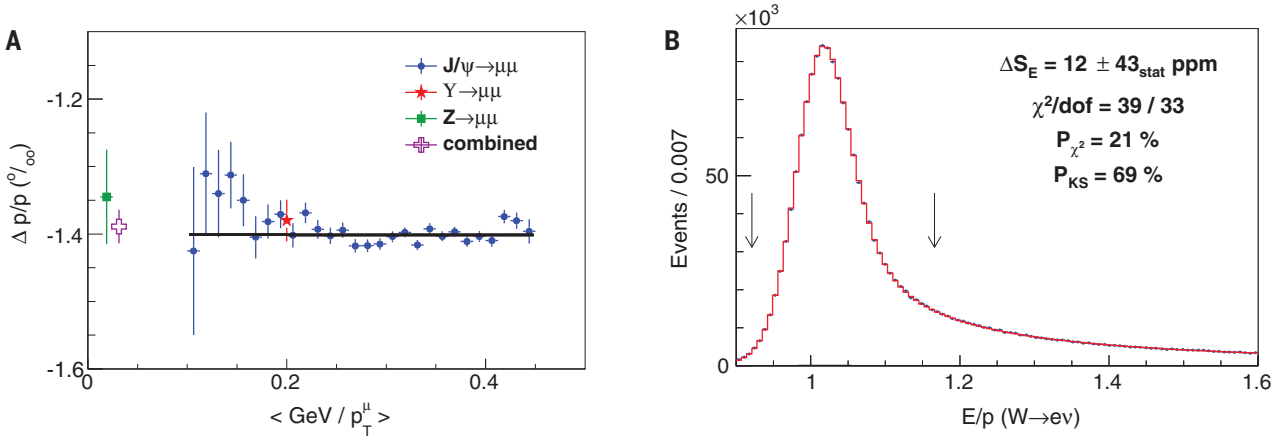


Fig. 2. Calibration of track momentum and electron's calorimeter energy. (**A**) Fractional deviation of momentum $\Delta p/p$ (per mille) extracted from fits to the $J/\psi \rightarrow \mu\mu$ resonance peak as a function of the mean muon unsigned curvature $\langle 1/p_T^{\mu} \rangle$ (blue circles). A linear fit to the points, shown in black, has a slope consistent with zero (17 ± 34 keV). The corresponding values of $\Delta p/p$ extracted from fits to the $\Upsilon \rightarrow \mu\mu$ and $Z \rightarrow \mu\mu$ resonance peaks are also shown. The combination of all of these $\Delta p/p$ measurements yields the momentum correction labeled "combined," which is applied to the lepton tracks in *W* boson data. Error bars indicate the uncorrelated uncertainties (total uncertainty) for the individual boson measurements (combined correction). (**B**) Distribution of E/p for the $W \rightarrow e\nu$ data (points) and the best-fit simulation (histogram) including the small background from hadrons misreconstructed as electrons. The arrows indicate the fitting range used for the electron energy calibration. The relative energy correction ΔS_E, averaged over the calibrated *W* and *Z* boson data [see fig. S13 in (*63*)], is compatible with zero. In this and other figures, P_{KS} refers to the Kolmogorov-Smirnov probability of agreement between the shapes of the data and simulated distributions.

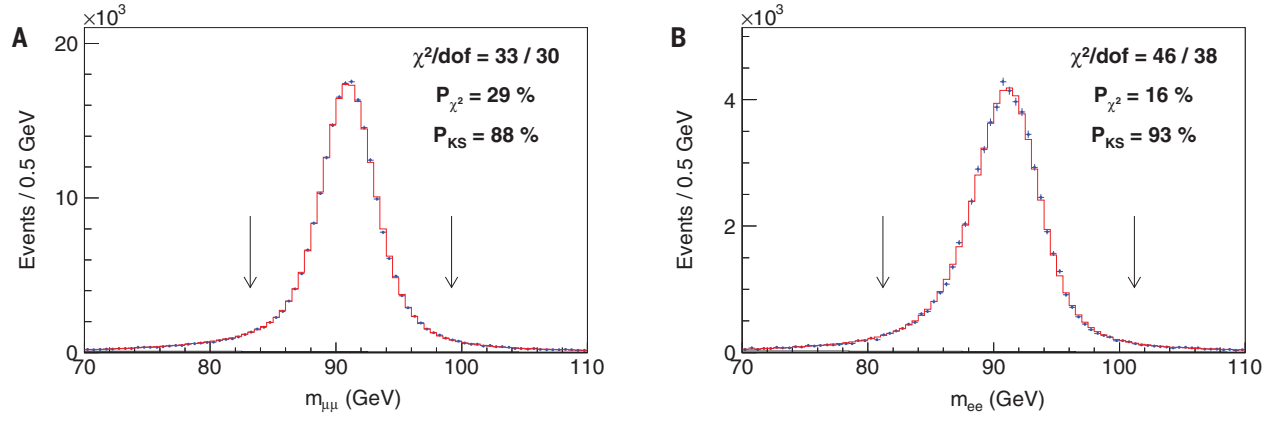


Fig. 3. Decay of the *Z* boson. (**A** and **B**) Distribution of (A) dimuon and (B) dielectron mass for candidate $Z \rightarrow \mu\mu$ and $Z \rightarrow ee$ decays, respectively. The data (points) are overlaid with the best-fit simulation template including the photon-mediated contribution (histogram). The arrows indicate the fitting range.

values. The model-dependent nature of the analysis implies that future improvements or corrections in any relevant theoretical modeling can be used to update our measurement quantifiably [see section IV of (*63*)].

The custom simulation includes a detailed calculation of the lepton and photon interactions in the detector (*39*, *43*, *64*), as well as models describing their individual position measurements within the COT. The COT position resolution as a function of radius is determined using muon tracks from Υ meson, *W* boson, and *Z* boson decays. All wire positions in the COT are measured with 1-μm precision using an in situ sample of cosmic ray muons (*65*), in addition to the electron tracks from *W* boson decays. The difference between electron and positron track momenta relative to their measured energy in the calorimeter (which is independent of charge) strongly constrains certain modes of internal misalignment in the COT.

Momentum and energy calibration

The track momentum measurement in the COT is calibrated by measuring the masses of the J/ψ and $\Upsilon(1S)$ mesons reconstructed in their dimuon decays and comparing them with the known values (*10*). These meson mass measurements are performed with maximum-likelihood fits to the dimuon mass distributions from data, using templates obtained from the custom simulation. Measurements of these masses as functions of muon momenta are used to correct for small inaccuracies in the magnetic field map, the COT position measurements, and the modeling of the energy loss by particles traversing the detector. A mismodeling of the energy loss would lead to a bias linear in the mean inverse p_T of the two muons. No such bias is observed after applying the magnetic field nonuniformity, COT, and energy-loss corrections (Fig. 2A). The curvature q/p_T measured by the COT, where q is the particle charge, is an analytic function of the true curvature. The curvature response function analytically yields a linear dependence of the measured invariant mass on p_T^{-1}, and higher-order terms in p_T^{-1} are negligible. The correction for the fractional deviation of the measured momentum from its correct value, $\Delta p/p \equiv p_{\text{measured}}/p_{\text{true}} - 1$, is inferred from the comparison of the measured meson masses to their more-precise world-average masses. The $\Delta p/p$ corrections extracted from the individual J/ψ and $\Upsilon(1S)$ invariant mass fits are consistent with each other, and the results

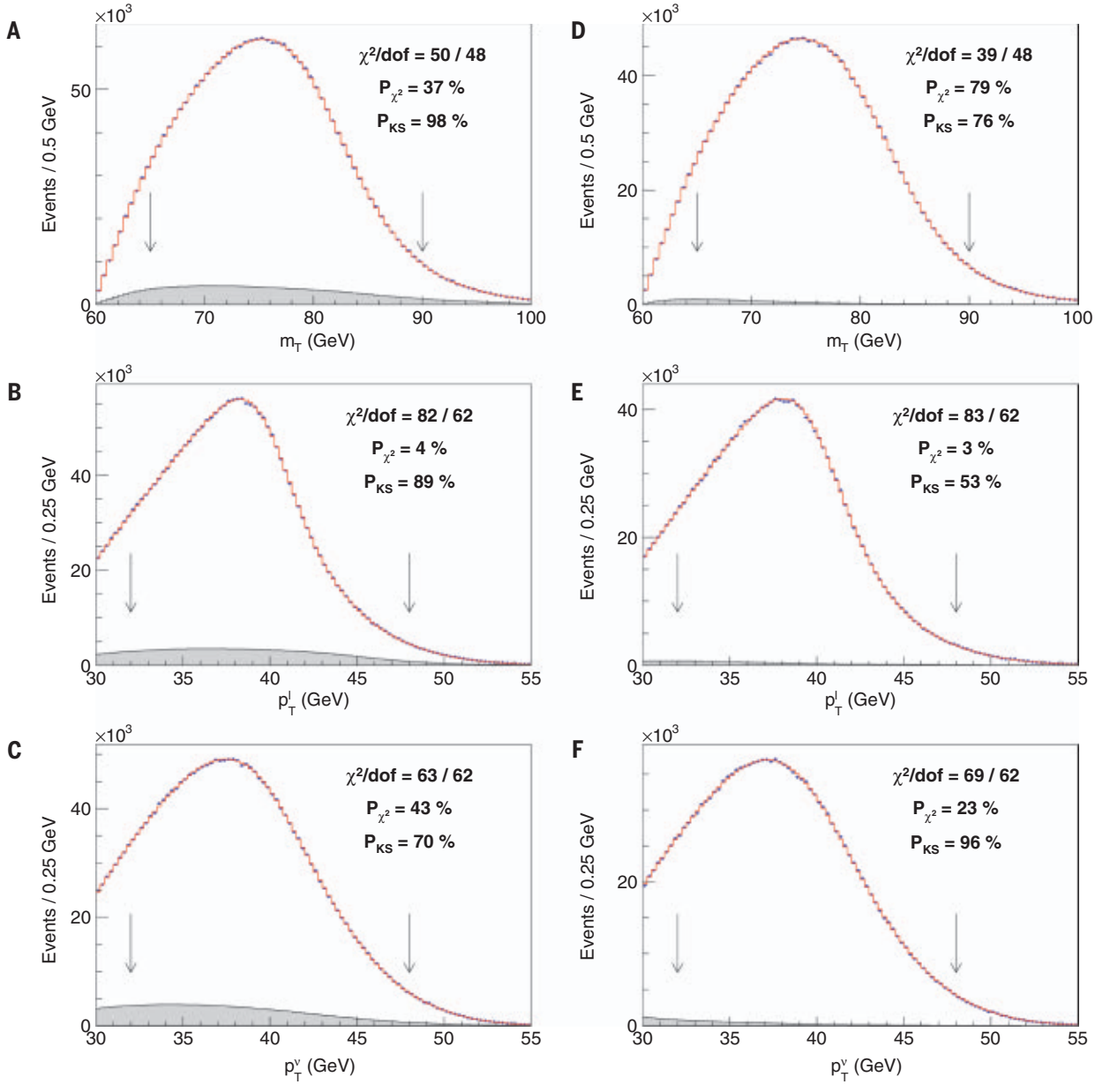


Fig. 4. Decay of the *W* boson. (**A** to **C**) Distributions for m_T (A), p_T^ℓ (B), and p_T^ν (C) for the muon channel. (**D** to **F**) Same as in (A) to (C) but for the electron channel. The data (points) and the best-fit simulation template (histogram) including backgrounds (shaded regions) are shown. The arrows indicate the fitting range.

are combined to obtain $\Delta p/p = (-1393 \pm 26)$ parts per million (ppm).

The combined momentum calibration is used to measure the Z boson mass in the dimuon channel (Fig. 3A), which is blinded with a random offset in the range of −50 to 50 MeV until all analysis procedures are established. The unblinded measurement is $M_Z = 91{,}192.0 \pm 6.4_{\text{stat}} \pm 4.0_{\text{syst}}$ MeV (stat, statistical uncertainty; syst, systematic uncertainty), which is consistent with the world average of 91,187.6 ± 2.1 MeV (*10*, *44*) and therefore provides a precise consistency check. Systematic uncertainties on M_Z result from uncertainties on the longitudinal coordinate measurements in the COT (1.0 MeV), the momentum calibration (2.3 MeV), and the QED radiative corrections (3.1 MeV). The latter two sources are correlated with the M_W measurement. The $Z \rightarrow \mu\mu$ mass measurement is then included in the final momentum calibration. The systematic uncertainties stemming from the magnetic field nonuniformity dominate the total uncertainty of 25 ppm in the combined momentum calibration.

After track momentum (p) calibration, the electron's calorimeter energy (E) is calibrated using the peak of the E/p distribution in $W \rightarrow e\nu$ (Fig. 2B) and $Z \rightarrow ee$ [fig. S13 in (*63*)] data. Fits to this peak in bins of electron E_T determine the electron energy calibration and its dependence on E_T. The radiative region of the E/p distribution ($E/p > 1.12$) is fitted to measure a small correction (≈5%) to the amount of radiative material traversed in the tracking volume. The EM calorimeter resolution is measured using the widths of the E/p peak in the $W \rightarrow e\nu$ sample and of the mass peak of the $Z \rightarrow ee$ sample.

We use the calibrated electron energies to measure the Z boson mass in the dielectron channel (Fig. 3B), which is also blinded with the same offset as used for the dimuon channel. The unblinded result, $M_Z = 91{,}194.3 \pm 13.8_{\text{stat}} \pm 7.6_{\text{syst}}$ MeV, is consistent with the world average, providing a stringent consistency check of the electron energy calibration. Systematic uncertainties on M_Z are caused by uncertainties on the calorimeter energy

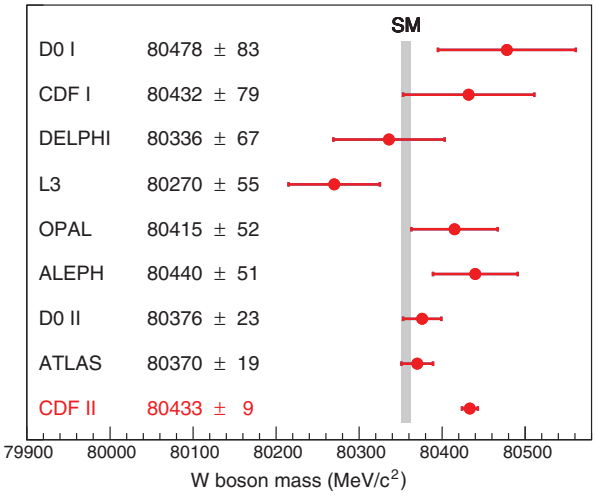


Fig. 5. Comparison of this CDF II measurement and past M_W measurements with the SM expectation. The latter includes the published estimates of the uncertainty (4 MeV) due to missing higher-order quantum corrections, as well as the uncertainty (4 MeV) from other global measurements used as input to the calculation, such as m_t. c, speed of light in a vacuum.

(6.5 MeV) and track momentum (2.3 MeV), on the z coordinate measured in the COT (0.8 MeV), and on QED radiative corrections (3.1 MeV). Measurements of the Z boson mass using the dielectron track momenta, and comparisons of mass measurements using radiative and nonradiative electrons, provide consistent results. The final calibration of the electron energy is obtained by combining the E/p-based calibration with the $Z(\rightarrow ee)$ mass-based calibration, taking into account the correlated uncertainty on the radiative corrections.

The spectator partons in the proton and antiproton, as well as the additional (≈3) $p\bar{p}$ interactions in the same collider bunch crossing, contribute visible energy that degrades the resolution of $\vec{u}$. These contributions are measured from events triggered on inelastic $p\bar{p}$ interactions and random bunch crossings, reproducing the collision environment of the W and Z boson data. Because there are no high-p_T neutrinos in the Z boson data, the $\vec{p}_T$ imbalance between the $\vec{p}_T^{\,\ell\ell}$ and $\vec{u}$ in $Z \rightarrow \ell\ell$ events is used to measure the calorimeter response to, and resolution of, the initial-state QCD radiation accompanying boson production. The simulation of the recoil vector $\vec{u}$ also requires knowledge of the distribution of the energy flow into the calorimeter towers impacted by the leptons, because these towers are excluded from the computation of $\vec{u}$. This energy flow is measured from the W boson data using the event-averaged response of towers separated in azimuth from the lepton direction.

Extracting the W boson mass

Kinematic distributions of background events passing the event selection are included in the template fits with their estimated normalizations. The W boson samples contain a small contamination of background events arising from QCD jet production with a hadron misidentified as a lepton, $Z \rightarrow \ell\ell$ decays with only one reconstructed lepton, $W \rightarrow \tau\nu \rightarrow \ell\nu\bar{\nu}\nu$, pion and kaon decays in flight to muons (DIF), and cosmic-ray muons (τ, tau lepton; $\bar{\nu}$, antineutrino). The jet, DIF, and cosmic-ray backgrounds are estimated from control samples of data, whereas the $Z \rightarrow \ell\ell$ and $W \rightarrow \tau\nu$ backgrounds are estimated from simulation. Background fractions for the muon (electron) datasets are evaluated to be 7.37% (0.14%) from $Z \rightarrow \ell\ell$ decays, 0.88% (0.94%) from $W \rightarrow \tau\nu$ decays, 0.01% (0.34%) from jets, 0.20% from DIF, and 0.01% from cosmic rays.

The fit results (Fig. 4) are summarized in Table 1. The M_W fit values are blinded during analysis with an unknown additive offset in the range of −50 to 50 MeV, in the same manner as, but independent of, the value used for blinding the Z boson mass fits. As the fits to the different kinematic variables have different sensitivities to systematic uncertainties, their consistency confirms that the sources of systematic uncertainties are well understood. Systematic uncertainties, propagated by varying the simulation parameters within their uncertainties and repeating the fits to these simulated data, are shown in Table 1. The correlated uncertainty in the m_T (p_T^ℓ, p_T^ν) fit between the muon and electron channels is 5.8 (7.9, 7.4) MeV. The mass fits are stable with respect to variations of the fitting ranges.

Table 2. Uncertainties on the combined M_W result.

Source	Uncertainty (MeV)
Lepton energy scale	3.0
Lepton energy resolution	1.2
Recoil energy scale	1.2
Recoil energy resolution	1.8
Lepton efficiency	0.4
Lepton removal	1.2
Backgrounds	3.3
p_T^Z model	1.8
p_T^W/p_T^Z model	1.3
Parton distributions	3.9
QED radiation	2.7
W boson statistics	6.4
Total	9.4

Simulated experiments are used to evaluate the statistical correlations between fits, which are found to be 69% (68%) between m_T and p_T^ℓ (p_T^ν) fit results and 28% between p_T^ℓ and p_T^ν fit results (*43*). The six individual M_W results are combined (including correlations) by means of the best linear unbiased estimator (*66*) to obtain $M_W = 80{,}433.5 \pm 9.4$ MeV, with χ^2/dof = 7.4/5 corresponding to a probability of 20%. The m_T, p_T^ℓ, and p_T^ν fits in the electron (muon) channel contribute weights of 30.0% (34.2%), 6.7% (18.7%), and 0.9% (9.5%), respectively. The combined result is shown in Fig. 1, and its associated systematic uncertainties are shown in Table 2.

Discussion

The dataset used in this analysis is about four times as large as the one used in the previous analysis (*41*, *43*). Although the resolution of the hadronic recoil is somewhat degraded in the new data because of the higher instantaneous luminosity, the statistical precision of the measurement from the larger sample is still improved by almost a factor of 2. To achieve a commensurate reduction in systematic uncertainties, a number of analysis improvements have been incorporated, as described in table S1. These improvements are based on using cosmic-ray and collider data in ways not employed previously to improve (i) the COT alignment and drift model and the uniformity of the EM calorimeter response, and (ii) the accuracy and robustness of the detector response and resolution model in the simulation. Additionally, theoretical inputs to the analysis have been updated. Upon incorporating the improved understanding of PDFs and track reconstruction, our previous measurement is increased by 13.5 MeV to 80,400.5 MeV; the consistency of the latter with the new measurement is at the percent probability level.

In conclusion, we report a new measurement of the W boson mass with the complete dataset collected by the CDF II detector at the Fermilab Tevatron, corresponding to 8.8 fb^{-1} of integrated luminosity. This measurement, $M_W = 80{,}433.5 \pm 9.4$ MeV, is more precise than all previous measurements of M_W combined and subsumes all previous CDF measurements from 1.96-TeV data (*38*, *39*, *41*, *43*). A comparison with the SM expectation of $M_W = 80{,}357 \pm 6$ MeV (*10*), treating the quoted uncertainties as independent, yields a difference with a significance of 7.0σ and suggests the possibility of improvements to the SM calculation or of extensions to the SM. This comparison, along with past measurements, is shown in Fig. 5. Using the method described in (*45*), we obtain a combined Tevatron (CDF and D0) result of $M_W = 80{,}427.4 \pm 8.9$ MeV. Assuming no correlation between the Tevatron

and LEP measurements, their average becomes $M_W = 80{,}424.2 \pm 8.7$ MeV.

REFERENCES AND NOTES

1. P. W. Anderson, *Phys. Rev.* **130**, 439–442 (1963).
2. F. Englert, R. Brout, *Phys. Rev. Lett.* **13**, 321–323 (1964).
3. P. W. Higgs, *Phys. Rev. Lett.* **13**, 508–509 (1964).
4. G. S. Guralnik, C. R. Hagen, T. W. B. Kibble, *Phys. Rev. Lett.* **13**, 585–587 (1964).
5. G. Aad *et al.*; ATLAS Collaboration, *Phys. Lett. B* **716**, 1–29 (2012).
6. S. Chatrchyan *et al.*; CMS Collaboration, *Phys. Lett. B* **716**, 30–61 (2012).
7. S. Glashow, *Nucl. Phys.* **22**, 579–588 (1961).
8. A. Salam, J. C. Ward, *Phys. Lett.* **13**, 168–171 (1964).
9. S. Weinberg, *Phys. Rev. Lett.* **19**, 1264–1266 (1967).
10. P. A. Zyla *et al.*, *Prog. Theor. Exp. Phys.* **2020**, 083C01 (2020).
11. J. Feng, *Annu. Rev. Nucl. Part. Sci.* **63**, 351–382 (2013).
12. R. Contino, T. Krämer, M. Son, R. Sundrum, *J. High Energy Phys.* **2007**, 074 (2007).
13. G. Bertone, D. Hooper, J. Silk, *Phys. Rep.* **405**, 279–390 (2005).
14. J. L. Feng, *Annu. Rev. Astron. Astrophys.* **48**, 495–545 (2010).
15. S. Heinemeyer, W. Hollik, G. Weiglein, L. Zeune, *J. High Energy Phys.* **2013**, 84 (2013).
16. S. Heinemeyer, "Electroweak precision observables and BSM physics," presented at the Snowmass EF04 meeting, 17 July 2020; https://indico.fnal.gov/event/43577/contributions/191539/attachments/131503/161060/sven.pdf.
17. D. López-Val, T. Robens, *Phys. Rev. D* **90**, 114018 (2014).
18. D. López-Val, J. Sola, *Eur. Phys. J. C* **73**, 2393 (2013).
19. D. Curtin, R. Essig, S. Gori, J. Shelton, *J. High Energy Phys.* **2015**, 157 (2015).
20. J. Chakrabortty, J. Gluza, R. Sevillano, R. Szafron, *J. High Energy Phys.* **2012**, 38 (2012).
21. B. Bellazzini, C. Csáki, J. Serra, *Eur. Phys. J. C* **74**, 2766 (2014).
22. A. Pomarol, F. Riva, *J. High Energy Phys.* **2014**, 151 (2014).
23. G. F. Giudice, C. Grojean, A. Pomarol, R. Rattazzi, *J. High Energy Phys.* **2007**, 045 (2007).
24. S. F. Ge, H. J. He, R. Q. Xiao, *J. High Energy Phys.* **2016**, 7 (2016).
25. We use the convention $\hbar = c \equiv 1$ throughout this paper.
26. M. Awramik, M. Czakon, A. Freitas, G. Weiglein, *Phys. Rev. D* **69**, 053006 (2004).
27. J. Erler, M. Schott, *Prog. Part. Nucl. Phys.* **106**, 68–119 (2019).
28. T. Affolder *et al.*; CDF Collaboration, *Phys. Rev. D* **64**, 052001 (2001).
29. B. Abbott *et al.*; DO Collaboration, *Phys. Rev. D* **58**, 092003 (1998).
30. B. Abbott *et al.*; DO Collaboration, *Phys. Rev. Lett.* **84**, 222–227 (2000).
31. B. Abbott *et al.*; DO Collaboration, *Phys. Rev. D* **62**, 092006 (2000).
32. V. M. Abazov *et al.*; DO Collaboration, *Phys. Rev. D* **66**, 012001 (2002).
33. V. M. Abazov *et al.*; CDF Collaboration, DO Collaboration, *Phys. Rev. D* **70**, 092008 (2004).
34. S. Schael *et al.*; ALEPH Collaboration, *Eur. Phys. J. C* **47**, 309–335 (2006).
35. J. Abdallah *et al.*; DELPHI Collaboration, *Eur. Phys. J. C* **55**, 1 (2008).
36. P. Achard *et al.*; L3 Collaboration, *Eur. Phys. J. C* **45**, 569–587 (2006).
37. G. Abbiendi *et al.*; OPAL Collaboration, *Eur. Phys. J. C* **45**, 307–335 (2006).
38. T. Aaltonen *et al.*; CDF Collaboration, *Phys. Rev. Lett.* **99**, 151801 (2007).
39. T. Aaltonen *et al.*; CDF Collaboration, *Phys. Rev. D* **77**, 112001 (2008).
40. V. M. Abazov *et al.*; DO Collaboration, *Phys. Rev. Lett.* **103**, 141801 (2009).
41. T. Aaltonen *et al.*; CDF Collaboration, *Phys. Rev. Lett.* **108**, 151803 (2012).
42. V. M. Abazov *et al.*; DO Collaboration, *Phys. Rev. Lett.* **108**, 151804 (2012).
43. T. Aaltonen *et al.*; CDF Collaboration, *Phys. Rev. D* **89**, 072003 (2014).
44. ALEPH Collaboration, CDF Collaboration, DO Collaboration, DELPHI Collaboration, L3 Collaboration, OPAL Collaboration, SLD Collaboration, LEP Electroweak Working Group, Tevatron Electroweak Working Group, SLD electroweak heavy flavour groups, arXiv:1012.2367 [hep-ex] (2010) and references therein.
45. T. Aaltonen *et al.*; CDF Collaboration, DO Collaboration, *Phys. Rev. D* **88**, 052018 (2013).
46. M. Aaboud *et al.*; ATLAS Collaboration, *Eur. Phys. J. C* **78**, 110 (2018).
47. M. Aaboud *et al.*; ATLAS Collaboration, *Eur. Phys. J. C* **78**, 898 (2018).
48. T. Affolder *et al.*, *Nucl. Instrum. Methods Phys. Res. A* **526**, 249–299 (2004).
49. G. Ascoli *et al.*, *Nucl. Instrum. Methods Phys. Res. A* **268**, 33–40 (1988).
50. The CDF II detector is centered on the beam (z) axis, which points in the proton direction. The $+x$ axis points outward and the $+y$ axis points upward, respectively, from the Tevatron ring. Corresponding cylindrical coordinates are defined with $r \equiv \sqrt{x^2 + y^2}$ and azimuthal angle $\phi \equiv \tan^{-1}(y/x)$. Pseudorapidity is defined as $\eta = -\ln[\tan(\theta/2)]$, where θ is the polar angle from the z axis. Energy (momentum) transverse to the beam is denoted as E_T (p_T).
51. A. V. Kotwal, H. K. Gerberich, C. Hays, *Nucl. Instrum. Methods Phys. Res. A* **506**, 110–118 (2003).
52. F. Abe *et al.*; CDF Collaboration, . *Nucl. Instrum. Methods Phys. Res. A* **271**, 387 (1988).
53. J. Smith, W. L. van Neerven, J. A. M. Vermaseren, *Phys. Rev. Lett.* **50**, 1738–1740 (1983).
54. C. Balázs, C.-P. Yuan, *Phys. Rev. D* **56**, 5558–5583 (1997).
55. G. A. Ladinsky, C. Yuan, *Phys. Rev. D* **50**, R4239–R4243 (1994).
56. F. Landry, R. Brock, P. M. Nadolsky, C.-P. Yuan, *Phys. Rev. D* **67**, 073016 (2003).
57. P. Golonka, Z. Was, *Eur. Phys. J. C* **45**, 97–107 (2006).
58. C. M. Carloni Calame, G. Montagna, O. Nicrosini, A. Vicini, *J. High Energy Phys.* **2007**, 109 (2007).
59. A. V. Kotwal, B. Jayatilaka, *Adv. High Energy Phys.* **2016**, 1615081 (2016).
60. R. D. Ball *et al.*, *Eur. Phys. J. C* **77**, 663 (2017).
61. T. J. Hou *et al.*, *Phys. Rev. D* **103**, 014013 (2021).
62. L. A. Harland-Lang, A. D. Martin, P. Motylinski, R. S. Thorne, *Eur. Phys. J. C* **75**, 204 (2015).
63. Supplementary materials.
64. A. V. Kotwal, C. Hays, *Nucl. Instrum. Methods Phys. Res. A* **729**, 25–35 (2013).
65. A. V. Kotwal, C. Hays, *Nucl. Instrum. Methods Phys. Res. A* **762**, 85–99 (2014).
66. L. Lyons, D. Gibaut, P. Clifford, *Nucl. Instrum. Methods Phys. Res. A* **270**, 110–117 (1988).
67. T. Aaltonen *et al.* (CDF Collaboration), High precision measurement of the W-boson mass with the CDF II detector, Zenodo (2022); https://doi.org/10.5281/zenodo.6245867.

ACKNOWLEDGMENTS

We thank A. Accardi, C. Carloni Calame, S. Carrazza, G. Ferrera, S. Forte, Y. Fu, L. Harland-Lang, J. Isaacson, P. Nadolsky, J. Rojo, N. Sato, S. Sen, R. Thorne, A. Vicini, Z. Was, G. Watt, and C.-P. Yuan for helpful discussions. This document was prepared by the CDF Collaboration using the resources of the Fermi National Accelerator Laboratory (Fermilab), a US Department of Energy, Office of Science, HEP User Facility. Fermilab is managed by Fermi Research Alliance, LLC (FRA), acting under contract no. DE-AC02-07CH11359. We thank the Fermilab staff and the technical staffs of the participating institutions for their vital contributions. **Funding:** This work was supported by the US Department of Energy and National Science Foundation; the Italian Istituto Nazionale di Fisica Nucleare; the Ministry of Education, Culture, Sports, Science and Technology of Japan; the Natural Sciences and Engineering Research Council of Canada; the National Science Council of the Republic of China; the Swiss National Science Foundation; the Alfred P. Sloan Foundation; the Bundesministerium für Bildung und Forschung, Germany; the National Research Foundation of Korea; the Science and Technology Facilities Council and the Royal Society, UK; the Russian Foundation for Basic Research; the Ministerio de Ciencia e Innovación, and Programa Consolider-Ingenio 2010, Spain; the Slovak R&D Agency; the Academy of Finland; and the Australian Research Council (ARC). **Author contributions:** All authors contributed to various aspects of the experiment's construction and operation, data acquisition and reconstruction, review of the analysis, and approval of the manuscript. A.V.K. led the analysis and wrote the paper. **Competing interests:** All authors declare that they have no competing interests. **Data and materials availability:** No materials are involved in the results presented. Data and code have been deposited in the Zenodo repository (*67*) and are based on the functionality of the CERN ROOT analysis package version 5.34/12.

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abk1781
Authors and Affiliations
Supplementary Text
Figs. S1 to S41
Tables S1 to S10
References (*68–110*)

27 June 2021; accepted 11 March 2022
10.1126/science.abk1781

REPORTS

DEVELOPMENTAL BIOLOGY

Functional primordial germ cell–like cells from pluripotent stem cells in rats

Mami Oikawa[1,2]†, Hisato Kobayashi[3], Makoto Sanbo[2], Naoaki Mizuno[4], Kenyu Iwatsuki[1,5], Tomoya Takashima[3,6], Keiko Yamauchi[2], Fumika Yoshida[2], Takuya Yamamoto[7,8,9], Takashi Shinohara[10], Hiromitsu Nakauchi[4,11], Kazuki Kurimoto[3], Masumi Hirabayashi[2,12]*, Toshihiro Kobayashi[1,2]*†

The in vitro generation of germ cells from pluripotent stem cells (PSCs) can have a substantial effect on future reproductive medicine and animal breeding. A decade ago, in vitro gametogenesis was established in the mouse. However, induction of primordial germ cell–like cells (PGCLCs) to produce gametes has not been achieved in any other species. Here, we demonstrate the induction of functional PGCLCs from rat PSCs. We show that epiblast-like cells in floating aggregates form rat PGCLCs. The gonadal somatic cells support maturation and epigenetic reprogramming of the PGCLCs. When rat PGCLCs are transplanted into the seminiferous tubules of germline-less rats, functional spermatids—that is, those capable of siring viable offspring—are generated. Insights from our rat model will elucidate conserved and divergent mechanisms essential for the broad applicability of in vitro gametogenesis.

In mammals, primordial germ cells (PGCs), which are the precursors of sperm and eggs, emerge from the pregastrulating epiblast. Studies using genetically modified mice have uncovered key inductive signals and transcriptional regulators that are essential for PGC fate (*1*). However, low numbers (~40 in mice) of PGCs in early embryos offer a limited amount of material to access the specific time window when germ cells are specified. A pioneering study from 2011 reconstituted mouse germ cell specification in vitro by differentiating mouse pluripotent stem cells (PSCs) into PGC-like cells (PGCLCs) capable of gametogenesis in

vivo, yielding normal offspring through assisted reproductive technology (*2*). Similar in vitro systems for other mammalian PSCs, including humans, have revealed conserved and divergent mechanisms underlying PGC specification (*3*–*6*). Although the mouse PGCLC (mPGCLC) study was conducted a decade ago, fully functional in vitro–derived PGCLCs capable of producing gametes have not been reported for any other species. In this study, we demonstrate the successful generation of functional PGCLCs from PSCs in rats (*Rattus norvegicus*).

Rats and mice share important features; however, they are distinct species with substantial differences in physiology, pharmacology, cognition, and behavior (*7*). Although mouse embryonic stem cells (ESCs) were derived more than 40 years ago, isolating rat germline-competent ESCs has proven to be much more challenging because of stringent culture requirements (*8*, *9*). Hence, mice represent the preeminent rodent model system. Recently, we have made considerable progress in understanding germline development in rats using mutant strains and xenogenic models (*10*, *11*). These advances enable us to explore rat in vitro gametogenesis.

After implantation, the rat blastocyst, similar to the mouse blastocyst, forms an egg-cylinder structure that contains a pluripotent epiblast from which germ cells arise (Fig. 1A). We tested whether we could use the culture conditions established for the mouse [N2B27 medium with 1% knockout serum replacement (KSR), activin, and basic fibroblast growth factor (bFGF)] to direct rat PSCs (rPSCs) toward the epiblast-like cell (EpiLC) fate for the specification of PGCs. To monitor the transition out of the pluripotent state, we used *Prdm14-H2BVenus* rat ESCs (rESCs) because *Prdm14-H2BVenus* specifically marks the naïve pluripotent epiblast and ESCs, but not the postimplantation formative or primed epiblast (*10*). We found that rESCs do not grow like mouse ESCs (mESCs), which grow as an adherent monolayer in EpiLC medium (fig. S1A). Instead, undifferentiated rESCs attach loosely to the feeder cells (fig. S1, B and C) (*9*). We reasoned that a floating aggregate culture might support survival and exit from the naïve pluripotent state in rESCs. Therefore, we seeded trypsinized rESCs into low-attachment U-bottom plates and cultured them for 72 hours in EpiLC medium. rESCs readily form aggregate-like embryoid bodies without extensive cell death (fig. S1D). By 72 hours of culture, the aggregates show reduced levels of Prdm14-H2BVenus, increased levels of OTX2 (a postimplantation epiblast marker) and CD47 (a plasma membrane marker up-regulated in mouse epiblast stem cells) (*12*), and steady levels of OCT3/4 (a core pluripotency factor) (Fig. 1, A and B, and fig. S1D). Thus, our culture conditions induced key features of EpiLC fate in the rat.

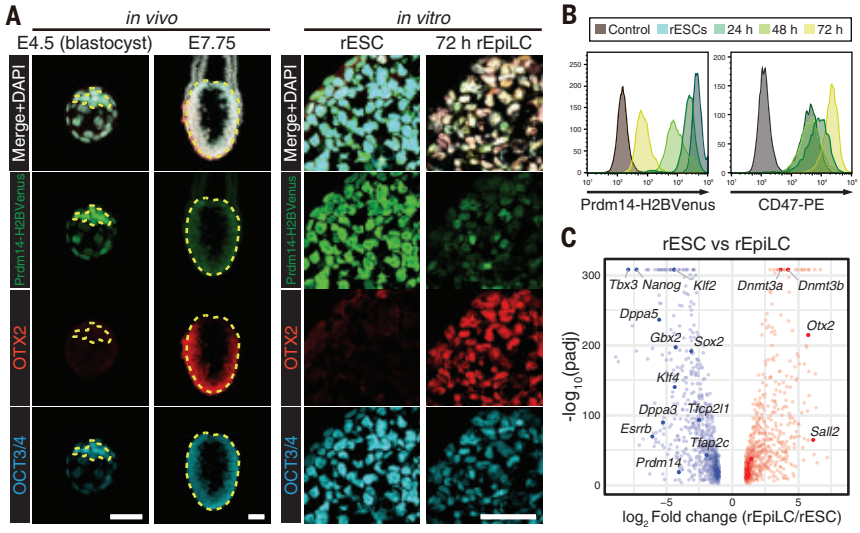


Fig. 1. Induction of rEpiLCs from rESCs. (**A**) IF images of rat blastocysts at E4.5 and a postimplantation embryo at E7.75 (whole mount), rESCs, and rEpiLCs (cryosection). The yellow dashed lines indicate the inner cell mass at E4.5 and the epiblast at E7.75. DAPI, 4′,6-diamidino-2-phenylindole. (**B**) FACS patterns for Prdm14-H2BVenus and CD47, with nonreporter and nonstaining rESCs used as controls, respectively. (**C**) Volcano plot showing DEGs between rESCs and rEpiLCs. Scale bars are 50 μm. padj, adjusted *p* value.

To examine the global gene expression in rat EpiLCs (rEpiLCs), we performed RNA sequencing (RNA-seq) on rEpiLCs and compared them with rESCs. We identified differentially expressed genes (DEGs) among rESCs and rEpiLCs (Fig. 1C). Each group contained naïve or formative and primed associated genes (highlighted in Fig. 1C). Taken together, we conclude that rEpiLCs in spherical aggregates induced from naïve rESCs recapitulate features of the in vivo postimplantation epiblast. It is not clear as to why rPSCs do not form an adherent two-dimensional (2D) culture; however, the floating aggregates seem to physiologically resemble in vivo 3D epiblasts. Indeed, the same 3D system can also be applied to mESCs (fig. S1E)

Next, we tested whether the rEpiLCs induced from rESCs are competent for PGC fate. We isolated ex vivo epiblast from rat embryos at embryonic day 7.75 (E7.75), which is before rat PGCs (rPGCs) are specified (*10*), and optimized culture conditions to maintain cell viability and induce PGC fate from the epiblast (rEpiPGCs). We determined the optimal PGCLC medium composition to be that containing N2B27 medium with 5% KSR, bone morphogenetic protein-4 (BMP4), stem cell factor (SCF), leukemia inhibitory factor (LIF), and epidermal growth factor (EGF) (methods and fig. S1, F to I). To exclude potential contamination with pluripotent rESCs, which also highly express *Prdm14* (figs. S1D and S2D), we generated *Nanos3-T2A-tdTomato* reporter rats to monitor the expression of *Nanos3*, a highly conserved germ cell marker. *Nanos3-T2A-tdTomato* is specifically expressed in E9.5 to E15.5 rPGCs, rEpiPGCs, and spermatogonia in the adult testes, but not in pre- and postimplantation epiblasts (fig. S2, A to I). rESCs derived from *Nanos3-T2A-tdTomato* reporter rats (N3T-rESCs) did not show expression of tdTomato in an undifferentiated state (fig. S2D). We also confirmed that N3T-rESCs efficiently contribute to the germline in vivo after injection into blastocysts (fig. S2, J and K). Therefore, we used N3T-rESCs for the induction of rEpiLCs and subsequent rat PGCLCs (rPGCLCs).

Dissociated rESCs were cultured for 48 to 72 hours in EpiLC medium to form aggregates, which were transferred into PGCLC medium containing BMP4, a cytokine that is critical for PGC fate (*13*) (Fig. 2A). Within 2 days of culture in the PGCLC medium, a proportion of cells in the aggregates started to show expression of Nanos3-T2A-tdTomato in response to BMP4 (Fig. 2B and fig. S3, A to E). The expression peaked at days 2 and 3 and then gradually

[1]Division of Mammalian Embryology, Center for Stem Cell Biology and Regenerative Medicine, The Institute of Medical Science, The University of Tokyo, Minato-ku, Tokyo 108-8639, Japan. [2]Center for Genetic Analysis of Behavior, National Institute for Physiological Sciences, Okazaki, Aichi 444-8787, Japan. [3]Department of Embryology, Nara Medical University, Kashihara, Nara 634-0813, Japan. [4]Division of Stem Cell Therapy, Distinguished Professor Unit, The Institute of Medical Science, The University of Tokyo, Minato-ku, Tokyo 108-8639, Japan. [5]Graduate School of Medicine, Science and Technology, Shinshu University, Ueda, Nagano 386-8567, Japan. [6]Department of Bioscience, Tokyo University of Agriculture, Setagaya-ku, Tokyo 156-8502, Japan. [7]Center for iPS Cell Research and Application, Kyoto University, Sakyo-ku, Kyoto 606-8507, Japan. [8]Institute for the Advanced Study of Human Biology, Kyoto University, Sakyo-ku, Kyoto 606-8501, Japan. [9]Medical-risk Avoidance Based on iPS Cells Team, RIKEN Center for Advanced Intelligence Project, Sakyo-ku, Kyoto 606-8507, Japan. [10]Department of Molecular Genetics, Graduate School of Medicine, Kyoto University, Sakyo-ku, Kyoto 606-8501, Japan. [11]Institute for Stem Cell Biology and Regenerative Medicine, Department of Genetics, Stanford University School of Medicine, Stanford, CA 94305, USA. [12]The Graduate University of Advanced Studies, Okazaki, Aichi 444-8787, Japan.
***Corresponding author. Email: tkoba@g.ecc.u-tokyo.ac.jp (T.K.); mhirarin@nips.ac.jp (M.H.)**
†These authors contributed equally to this work.

declined by day 5, likely owing to low proliferative activity of nascent rPGCLCs, as in mice (fig. S3C) (*2*). We used Immunofluorescence (IF) staining to confirm that N3T$^+$ cells coexpress the PGC and pluripotency markers *Tfap2c*, *Oct3/4*, and *Sox2*, indicating that they resemble in vivo rPGCs (fig. S3, F and G). rEpiLCs cultured for 48 to 60 hours showed the highest numbers of rPGCLCs (fig. S3H). This time is longer than that for PGCLC induction in mice, which peak around 36 to 48 hours (*2*, *14*). The time lag may be attributed to the 1.5- to 2-day difference in gestation period for the mouse versus rat (*10*, *15*).

We next analyzed the transcriptome of day 3 (d3) rPGCLCs by RNA-seq and compared it with the transcriptomes of rESCs, rEpiLCs, in vivo rat epiblast, and rPGCs (*10*). Hierarchical clustering and correlation coefficient evaluation of the samples showed that d3 rPGCLCs closely correlated with E9.5 to E11.5 early rPGCs (fig. S4, A and B). In the principal components analysis (PCA), the PC2-PC3 plot reflects the progression of epiblast toward germline fate both in vivo and in vitro (Fig. 2C). The d3 rPGCLCs expressed all the PGC specifiers and pluripotency genes, whereas late PGC marker expression is lower than that in E15.5 gonadal rPGCs (fig. S4, C and D). Taken together, we conclude that the induced rPGCLCs might be equivalent to the migratory stage of in vivo rPGCs.

To investigate the potential of rPGCLCs to mature into late PGCs, we reconstituted a gonadal environment using rPGCLCs and gonadal somatic cells, as described for mice (*14*) (Fig. 2A). We used E15.5 rat gonads because their sex can be clearly distinguished morphologically. To eliminate endogenous rPGCs in gonads, we explored rPGC-specific cell surface markers. Notably, stage-specific embryonic antigen 1 (SSEA1), a widely used surface marker for mouse PGCs (mPGCs), is not expressed in rPGCs (fig. S5A). From our transcriptome dataset, we found that *c-Kit* is highly up-regulated in both in vitro and in vivo rPGCs (fig. S4D). The expression of c-KIT overlaps with *Prdm14* and *Nanos3* reporters in d3 rPGCLCs and E15.5 gonadal rPGCs (fig. S5, B to E). Day 3 male N3T$^+$ rPGCLCs were aggregated with c-KIT$^+$ rPGC-depleted male or female gonadal somatic cells from wild-type rats and cultured for 3 to 6 days (ag3 to ag6; fig. S6A). Male rPGCLCs that aggregated with male gonadal somatic cells lost Nanos3-T2A-tdTomato expression by day 3 (fig. S6, A and B), indicating the need for further optimization, as has recently been demonstrated for organ culture of neonatal rat testes (*16*). By contrast, female gonadal somatic cells could support the survival of male N3T$^+$ rPGCLCs (fig. S6, A and B). d3ag3 rPGCLCs show an up-regulation of the markers for late PGCs and some meiosis-related genes, unlike d3 rPGCLCs (Fig. 2D and figs. S4, C and D, and S6C). Notably, the transcriptome of d3ag3 rPGCLCs is similar to that of E12.5 to E15.5 late rPGCs, and the PCA

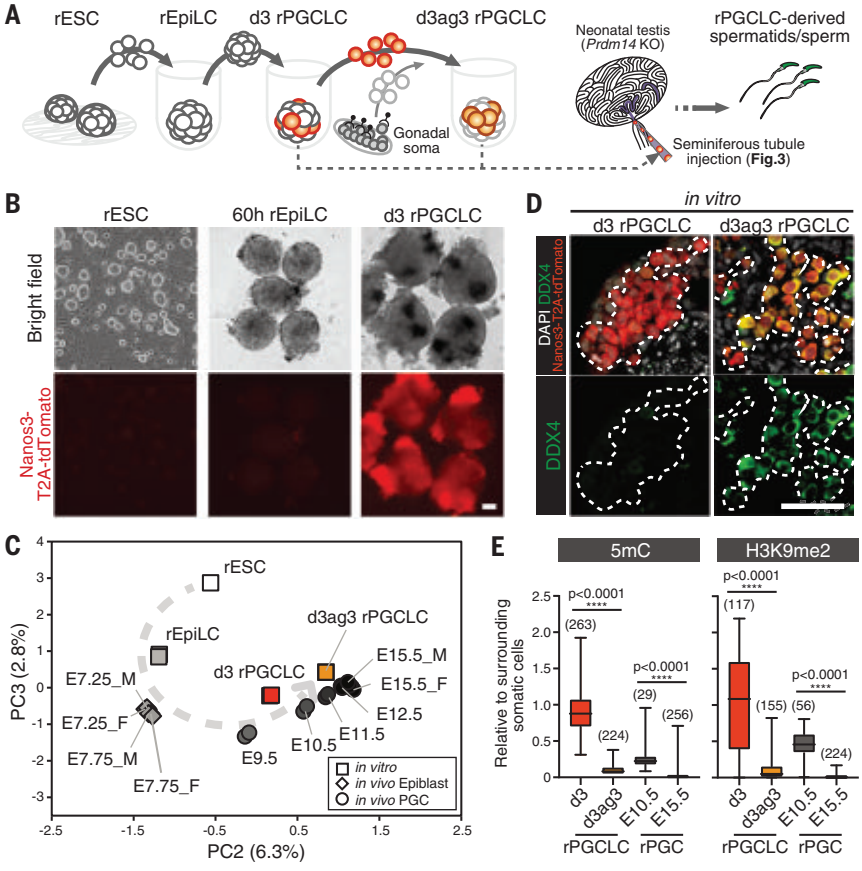


Fig. 2. Induction and maturation of rPGCLCs from rEpiLCs. (**A**) Experimental design for rPGCLC induction. (**B**) Morphology of aggregates during rPGCLC induction from N3T-rESCs visualized by bright-field (top) and fluorescence imaging (bottom). (**C**) PCA to compare in vitro and in vivo samples. The gray dashed line represents a trajectory of germline development. (**D**) IF images of a cryosection showing DDX4 expression during rPGCLC maturation in vitro. The white dashed lines indicate N3T$^+$ rPGCLCs. (**E**) Quantification of the indicated epigenetic marks. The averages and SD are shown. Numbers in parentheses indicate the number of rPGCs or rPGCLCs counted from IF images. Significance was determined using the Mann-Whitney test. 5mC, 5-methylcytosine; H3K9me2, dimethylated histone 3 lysine 9. Scale bars are 100 μm in (B) and 50 μm in (D).

showed a comparable trajectory to germline development in vivo (Fig. 2C and fig. S4, A and B). Because PGCs undergo extensive epigenetic reprogramming during development, we next examined DNA methylation and histone methylation dynamics in culture. The dynamics of the epigenetic changes in culture closely correlate with in vivo rPGC development (Fig. 2E and fig. S6, D and E), suggesting that d3 rPGCLCs mature in vitro toward the gonadal stage with stepwise progression of epigenetic reprogramming.

Finally, we investigated whether the male rPGCLCs undergo spermatogenesis in vivo after transplantation into testes (Fig. 2A). To monitor germ cell progression in the recipient testes, we generated *Acr3-EGFP* (AG) transgenic rats that show expression of enhanced green fluorescent protein (EGFP) specifically in spermatocytes, round spermatids, and mature sperm in the testis under the control of the *Acrosin* promoter (fig. S7, A to E). We derived rESCs from blastocysts obtained by crossing a *Nanos3-T2A-tdTomato* rat with an *Acr3-EGFP* rat (hereafter, N3T/AG-rESCs). N3T$^+$ day 3 to 4 (d3-4) rPGCLCs or d3ag3 rPGCLCs sorted by fluorescence-activated cell sorting (FACS) were transplanted into the seminiferous tubules of *Prdm14* knockout (*Prdm14* KO) neonatal rats that completely lacked endogenous germ cells (*11*) (fig. S8A). Eight to 11 weeks after transplantation, we detected Acr3-EGFP expression in the seminiferous tubules in both d3-4 and d3ag3 rPGCLC transplanted testes (Table 1, Fig. 3A, and fig. S8, B, C, and E). The testicular spermatozoon showed EGFP in the nucleus (fig. S8D). In the sections, we observed peanut agglutinin (PNA)–positive round spermatids and mature sperm (Fig. 3B), demonstrating that rPGCLCs can complete spermatogenesis in vivo.

The spermatogenic capacity of rPGCs and rPGCLCs is comparable to that of mPGCs and mPGCLCs (*2*) but lower than mouse spermatogonial and germline stem cells, likely because of their developmental differences (*17*, *18*). Obtaining rPSC-derived offspring through natural mating may require further maturation

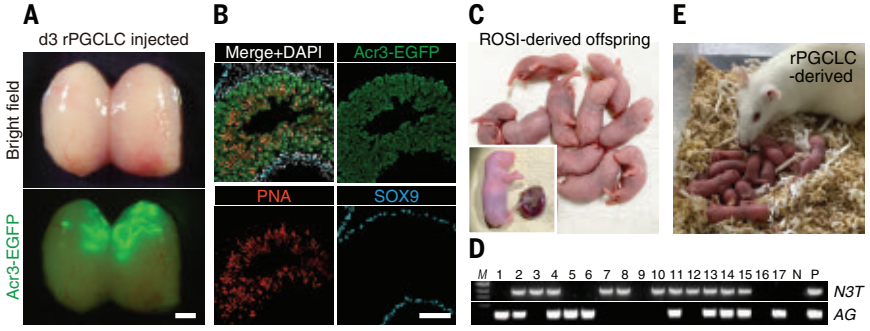


Fig. 3. Functional validation of rPGCLCs. (**A**) *Prdm14* KO rat testis at 10 weeks after transplantation of day 3 male N3T/AG-rPGCLCs, visualized by bright-field (top) and fluorescence imaging (bottom). (**B**) IF of a cryosection showing testis 10 weeks after transplantation of N3T/AG-rPGCLCs. (**C**) Offspring from rPGCLC-derived spermatids generated by ROSI. The inset shows an offspring with placenta. (**D**) Representative genotyping result of rPGCLC-derived offspring. M, molecular marker; 1 to 17, samples obtained from individual rPGCLC-derived offspring; N, negative control (water); P, positive control (N3T/AG-rESCs). (**E**) Female rat derived from N3T/AG-rPGCLCs and its offspring. Scale bars are 2 mm in (A) and 100 μm in (B).

Table 1. Spermatogenesis efficiency after rPGCLC transplantation.

rPGC or rPGCLC stage	Parental cells	Number of testes transplanted	Number of testes with successful transfer	Number of testes with EGFP-positive tubules	Number of EGFP-positive tubules in each testis
d3 rPGCLC	N3T/AG-rESCs no. 3	13	9 of 13 (69%)	6 of 9 (67%)	>5, >5, 4, 4, 1, 1
d3 rPGCLC	N3T/AG-rESCs no. 2	12	7 of 12 (58%)	6 of 7 (86%)	>5, >5, >5, >5, 4, 3
d4 rPGCLC	N3T/AG-rESCs no. 11	4	4 of 4 (100%)	2 of 4 (50%)	>5, 2
d3ag3 rPGCLC	N3T/AG-rESCs no. 3	2	1 of 2 (50%)	1 of 1 (100%)	>5
d3ag3 rPGCLC	N3T/AG-rESCs no. 2	2	1 of 2 (50%)	1 of 1 (100%)	>5
In vivo rPGC	N3T/AG E15.5 male gonad	4	2 of 4 (50%)	2 of 2 (100%)	>5, >5

of rPGCLCs into these stem cells; this merits additional investigation for future animal-breeding applications. Instead, we confirmed the developmental potential of rPGCLC-derived testicular germ cells by injecting round spermatid and testicular sperm into the oocytes obtained from wild-type rats using round spermatid injection (ROSI) and testicular sperm extraction with intracytoplasmic sperm injection (TESE-ICSI), respectively. At full term after embryo transfer, 18 (ROSI) and 6 (TESE-ICSI) live offspring were born and appeared healthy (Fig. 3C, fig. S8F, and table S3). Both *N3T* and *AG* transgenes originating from rESCs were successfully transmitted to the offspring (Fig. 3D). Whereas the body weights of the offspring were in the normal range, the placenta that was derived from ROSI offspring was significantly larger than that from control rats (fig. S8, G to I). Nevertheless, the offspring developed into fertile and normal adults (Fig. 3E and fig. S8J), suggesting that the induced rPGCLCs in vitro are functional and capable of producing mature gametes.

In vitro systems that differentiate rPSCs to rPGCLCs could become a useful platform to examine the function of key transcriptional regulators during the transition of naïve-to-formative pluripotency and during PGC specification. As exemplified by PSC research (*19*), insights from rats, a distinctive alternative model to the mouse, will help to define conserved or divergent principles in germ cell development within rodents and across mammals. In primates, PGCLCs can mature to the gonadal stage in vitro or in vivo (*20*, *21*) but do not progress to the gamete stage, perhaps owing to limitations in culture conditions or the lack of suitable models to test their function in vivo. However, rodents provide an excellent system for readily testing the fertility and developmental potential of in vitro germ cells. Because rats are physiologically more similar to humans than mice (*7*), our in vitro gametogenesis system offers the opportunity to screen causative factors in inter- or transgenerationally inherited disorders. Advances in the rat model should take us a step closer to achieving applicable systems for other species in domestic animal breeding and reproductive medicine.

REFERENCES AND NOTES

1. M. Saitou, M. Yamaji, *Cold Spring Harb. Perspect. Biol.* **4**, a008375 (2012).
2. K. Hayashi, H. Ohta, K. Kurimoto, S. Aramaki, M. Saitou, *Cell* **146**, 519–532 (2011).
3. N. Irie *et al.*, *Cell* **160**, 253–268 (2015).
4. T. Kobayashi *et al.*, *Nature* **546**, 416–420 (2017).
5. T. Kobayashi *et al.*, *Cell Rep.* **37**, 109812 (2021).
6. K. Sasaki *et al.*, *Cell Stem Cell* **17**, 178–194 (2015).
7. T. J. Aitman *et al.*, *Nat. Genet.* **40**, 516–522 (2008).
8. M. Buehr *et al.*, *Cell* **135**, 1287–1298 (2008).
9. P. Li *et al.*, *Cell* **135**, 1299–1310 (2008).
10. T. Kobayashi *et al.*, *Development* **147**, dev.183798 (2020).
11. T. Kobayashi *et al.*, *Nat. Commun.* **12**, 1328 (2021).
12. P. J. Rugg-Gunn *et al.*, *Dev. Cell* **22**, 887–901 (2012).
13. Y. Ohinata *et al.*, *Cell* **137**, 571–584 (2009).
14. K. Hayashi *et al.*, *Science* **338**, 971–975 (2012).
15. M. A. Hill, Embryology–Main page; https://embryology.med.unsw.edu.au/embryology/index.php/Main_Page.
16. T. Matsumura *et al.*, *Sci. Rep.* **11**, 3458 (2021).
17. H. Ohta, T. Wakayama, Y. Nishimune, *Biol. Reprod.* **70**, 1286–1291 (2004).
18. Y. Ishikura *et al.*, *Cell Rep.* **17**, 2789–2804 (2016).
19. Q. L. Ying, A. Smith, *Stem Cell Reports* **8**, 1457–1464 (2017).
20. E. Sosa *et al.*, *Nat. Commun.* **9**, 5339 (2018).
21. C. Yamashiro *et al.*, *Science* **362**, 356–360 (2018).

ACKNOWLEDGMENTS

We thank members of the Hirabayashi lab, in particular, M. Hashimoto and N. Niizeki for help with animals and M. Ohnishi for secretarial support. We also thank R. Sengupta for editing and providing critical input to the manuscript. We thank T. Hayama for his advice on the reaggregation of rat gonads. We thank S. Matoba for his advice on the culture of reconstructed gonads and histological analysis of placentas. We thank the Spectrography and Bioimaging Facility, National Institute for Basic Biology (NIBB) Core Research Facilities, for technical support. Flow cytometry was performed in the National Institute for Physiological Sciences (NIPS), Sciences–Exploratory Research Center on Life and Living Systems (ExCELLS); and in The Institute of Medical Science, The University of Tokyo (IMSUT) FACS Core laboratory. We also thank the Pathology Core Laboratory in IMSUT for technical support and the Single-Cell Genome Information Analysis Core (SignAC) in the Institute for the Advanced Study of Human Biology (ASHBi) for RNA sequence analysis. **Funding:** This work was supported by Grants-in-Aid for Scientific Research (KAKENHI) from the Japan Society for the Promotion of Science grants 18H02367 to M.H. and T.K., 18H05548 to T.K., 18H05544 to T.K. and K.K., 19K23711 to M.O., and 21H02382 to H.K.; Japan Agency for Medical Research and Development (AMED) grants JP18gm0010002 to H.N. and M.H. and JP18bm0704022 to T.K.; The Sumitomo Foundation grant 210348 to T.K.; and a NIPS research grant for young scientists to M.O. This work was also supported by grants from the Cooperative Study Program (21-147) of NIPS and the Cooperative Research Grant of the Genome Research for BioResource, NODAI Genome Research Center, Tokyo University of Agriculture. **Author contributions:** M.O. and T.K. designed the experiments. M.O., H.K., M.S., K.I., K.Y., F.Y., M.H., and T.K. performed the experiments. H.K., T.T., T.Y., and K.K. contributed to the RNA-seq analyses. N.M., T.S., and H.N. contributed resources. M.O., H.K., K.K., and T.K. analyzed and interpreted the data. M.O. and T.K. wrote the manuscript. **Competing interests**: The authors declare no competing interests. **Data and materials availability:** RNA-seq data have been deposited in the Gene Expression Omnibus (GEO) under accession number GSE178701. Rat strains *Nanos3-T2A-tdTomato* (knock-in reporter rat) and *Acr3-EGFP* (transgenic reporter rat) and cell lines N3T-rESC and N3T/AG-rESC are available from T. Kobayashi and M. Hirabayashi under a material transfer agreement with The University of Tokyo or NIPS.

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abl4412
Methods
Figs. S1 to S8
Tables S1 to S3
References (*22*–*35*)
MDAR Reproducibility Checklist

13 July 2021; accepted 4 March 2022
10.1126/science.abl4412

PLANT SCIENCE

TaCol-B5 modifies spike architecture and enhances grain yield in wheat

Xiaoyu Zhang[1]†, Haiyan Jia[1,2]†, Tian Li[1,3]†, Jizhong Wu[1,4], Ragupathi Nagarajan[1], Lei Lei[1], Carol Powers[1], Chia-Cheng Kan[1], Wei Hua[5], Zhiyong Liu[6], Charles Chen[7], Brett F. Carver[1]*, Liuling Yan[1]*

Spike architecture influences grain yield in wheat. We report the map-based cloning of a gene determining the number of spikelet nodes per spike in common wheat. The cloned gene is named *TaCOL-B5* and encodes a CONSTANS-like protein that is orthologous to *COL5* in plant species. Constitutive overexpression of the dominant *TaCol-B5* allele but without the region encoding B-boxes in a common wheat cultivar increases the number of spikelet nodes per spike and produces more tillers and spikes, thereby enhancing grain yield in transgenic plants under field conditions. Allelic variation in *TaCOL-B5* results in amino acid substitutions leading to differential protein phosphorylation by the protein kinase *Ta*K4. The *TaCol-B5* allele is present in emmer wheat but is rare in a global collection of modern wheat cultivars.

Common wheat (*Triticum aestivum*, 2*n* = 6*x* = 42, AABBDD genome) grain yields are influenced by three major components: spikes per unit land area, grains per spike, and grain weight (*1*). An increase in any one of these components can improve grain yield. The number of spikes can be increased through promotion of tillering, as fertile tillers eventually form spikes (*2*, *3*). The number of grains per spike can be physically and genetically dissected into two subcomponents: spikelets per spike and grains per spikelet (*2*, *3*). A normal spike can generate between 16 and 25 spikelet nodes, and within a spikelet, grains at the first and second positions are larger than those at the third, fourth, or higher positions (*4*–*6*). Therefore, understanding spikelet developmental patterns and generating more spikelet nodes per spike (hereafter referred to as SNS) increases grain number without decreasing average grain weight (*7*). The SNS trait is genetically controlled in any given wheat cultivar (*4*, *5*); however, the genetic basis of spikelet development is largely unknown. In this study, we mapped a quantitative trait locus (QTL) for SNS and then cloned the gene responsible for the QTL. We found that the cloned gene increased both SNS and spike number and further increased field-based grain yield in transgenic wheat.

We initially performed a single cross between two common wheat cultivars, CItr 17600 and Yangmai18, which have different spike morphologies (fig. S1, A and B). A population of 186 F_2 plants was genotyped using the genotyping-by-sequencing (GBS) approach (table S1), and the population of F_2-derived F_3 ($F_{2:3}$) lines was phenotyped under field conditions. On the basis of single-year phenotypic data, a potential major QTL associated with SNS was mapped to chromosome 7B (hereafter called *QSns.osu-7B*), having a log of the odds (LOD) value of 15.3 and accounting for 43% of the total phenotypic variation in the field-tested population (Fig. 1A).

To clone *QSns.osu-7B*, we screened 1857 individual F_5 plants derived from a single F_3 plant, WF112 (fig. S1, C and D), and identified 21 F_5 recombinant plants using two flanking markers (fig. S2). We also developed simple sequence repeat (SSR) and single-nucleotide polymorphism (SNP) markers (fig. S3) for fine mapping of the recombinant plants. We determined the genotypes and phenotypes of four F_6 populations derived from the recombinant F_5 plants (Fig. 1, B to F, and fig. S2). The gene responsible for *QSns.osu-7B* was flanked by two markers, SNS-M1 and SNS-G2M3, which spanned a genomic region of 318,786 base pairs (bp) encompassing two genes, *TraesCS7B02G400600* and *TraesCS7B02G400700*, according to International Wheat Genome Sequencing Consortium (IWGSC) RefSeq v2.1 sequences (Fig. 1B).

Next, we focused on allelic variation in the targeted region sequences (figs. S4 and S5) and concluded that *TraesCS7B02G400600* is most likely the gene responsible for *QSns.osu-7B*. *TraesCS7B02G400600* encodes a CONSTANS-like (COL) protein and is orthologous to *COL5* in plant species. We therefore named this wheat gene *TaCOL-B5*. We observed dominant effects of *TaCol-B5*, representing the CItr 17600 allele, over *Tacol-B5*, representing the Yangmai18 allele, on SNS and spike length (fig. S2). We also observed 10 SNPs along the sequenced 2014-bp region between the two alleles (fig. S4). We validated the functions of *TaCol-B5* using a transgenic approach in wheat.

We transformed Yangmai18 with the cloned cDNA of *TaCol-B5* from CItr 17600 and obtained four independent transgenic events (T_0 plants) that showed changed phenotypes in the T_1 generation (Fig. 2A and fig. S6, A to D). We confirmed the overexpression of transgenic *TaCol-B5* in the four independent transgenic events using quantitative real-time polymerase chain reaction (qRT-PCR) (fig. S6E). Additionally, we observed the expression of both transgenic *TaCol-B5* and native *Tacol-B5* in the same spike sample of the transgenic plants (fig. S6F). In the greenhouse, the transgenic plants, averaged across the four transgenic T_1 families expressing *TaCol-B5*, produced 3.5 more spikelet nodes per spike (Fig. 2B) and 3.4 more grains per spike (Fig. 2C) than did nontransgenic plants. Further, overexpression of *TaCol-B5* promoted tillering, resulting in an additional 1.3 spikes per plant and higher single-plant productivity (fig. S7). This observed increase in single-plant productivity (fig. S7) led us to test the transgenic wheat plants under field conditions.

First, we tested the effects of *TaCol-B5* in T_2 transgenic plants at a reduced seeding rate (40 plants/m^2), owing to limited grain availability (Fig. 2D). In comparison with nontransgenic plants, the transgenic plants produced larger and longer spikes, showed similar effects in all spikes (fig. S8, A to H), and generated longer but narrower grains (fig. S8, I to K). Averaged across the four T_2 families, the transgenic plants set an additional 2.4 spikelet nodes per spike (Fig. 2E) and increased spike length by 7.3 cm over the nontransgenic plants (Fig. 2F). Spikelet density was therefore lower in transgenic plants, 1.3 spikelets/cm versus 1.9 spikelets/cm in nontransgenic plants, which could contribute to greater per-plant productivity (*8*). The transgenic plants produced an additional 8.1 grains per spike (Fig. 2G) and 2.3 spikes per plant (Fig. 2H), with no compensatory loss in thousand grain weight (fig. S9A). The significant increases in single-plant productivity (fig. S9B) and single-row-plot grain yield (fig. S9C) in the presence of *TaCol-B5* (table S2) led us to further investigate its effect on grain yield following standard wheat yield trial procedures.

We analyzed the genetic effects of *TaCol-B5* in four T_3 transgenic lines at a higher seeding rate (130 plants/m^2) in the field, each in a 6-m^2 plot with three replicates. The phenotypes of the *TaCol-B5* transgenic plants were stable at the population level (Fig. 2, I and J). Compared

[1]Department of Plant and Soil Sciences, Oklahoma State University, Stillwater, OK 74078, USA. [2]The Applied Plant Genomics Laboratory, National Key Laboratory of Crop Genetics and Germplasm Enhancement, Nanjing Agricultural University, Nanjing 210095, Jiangsu, China. [3]Key Laboratory of Crop Gene Resources and Germplasm Enhancement, Institute of Crop Sciences, Chinese Academy of Agricultural Sciences, Beijing 100081, China. [4]Institute of Germplasm Resources and Biotechnology, Jiangsu Academy of Agricultural Sciences, Nanjing 210014, Jiangsu, China. [5]Zhejiang Academy of Agricultural Sciences, Hangzhou 310021, China. [6]Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, Beijing 100101, China. [7]Department of Biochemistry and Molecular Biology, Oklahoma State University, Stillwater, OK 74078, USA.
***Corresponding author. Email: liuling.yan@okstate.edu (L.Y.); brett.carver@okstate.edu (B.F.C.)**
†These authors contributed equally to this work.

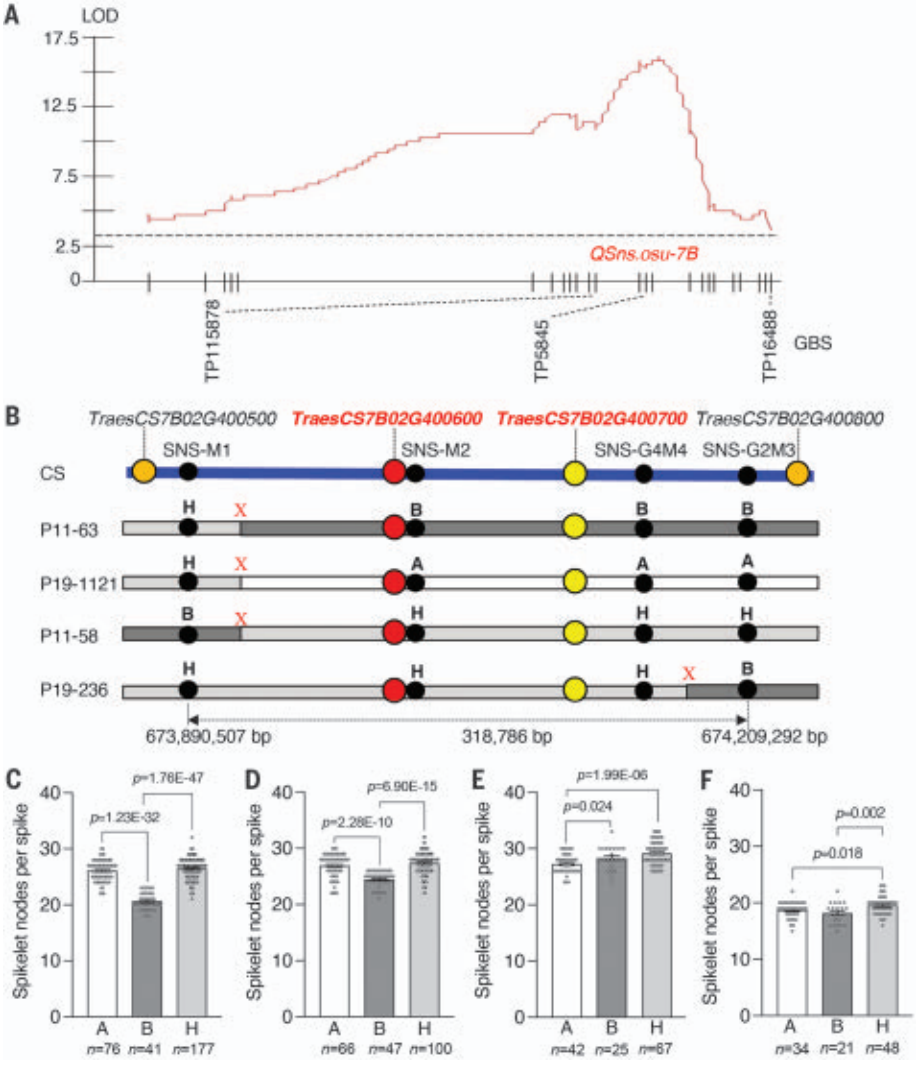


Fig. 1. Mapping and positional cloning of *QSns.osu-7B*. (**A**) Mapping of *QSns.osu-7B*. Physical locations of the GBS markers are provided in table S1. The horizontal dashed line represents a threshold LOD value of 3.0. (**B**) Physical map of crossovers detected in four critical recombinant plants. The populations derived from these plants were mapped with black dots representing PCR markers (fig. S3), red dots representing *TraesCS7B02G400600*, and yellow dots representing *TraesCS7B02G400700*. The red "X" indicates a crossover between markers. "A" represents the CItr 17600 allele; "B," the Yangmai18 allele; and "H," heterozygotes. CS, Chinese Spring. (**C** to **F**) Average SNS in the four F_6 populations. Populations P11-58 (C) and P19-236 (D) each show significant segregation for SNS. Populations P19-1121 (E) and P11-63 (F) show no association between SNS and two candidate genes, *TraesCS7B02G400600* and *TraesCS7B02G400700*. More detailed phenotypic analyses of these populations are provided in fig. S2.

with nontransgenic plants, the T_3 transgenic plants showed a greater number of spikelets per spike (0.9 spikelets; fig. S10A), longer spike length (4.4 cm; fig. S10B) and more spikes per plant (0.29 spikes; fig. S10C), again with no compensatory loss in thousand grain weight (fig. S10E). No significant difference was observed, however, in grain number per spike (fig. S10D). The net effect of these spike and grain traits was that grain yield increased between 7.8% and 19.8% among the four transgenic Yangmai18 lines (fig. S10F) in standard yield plots, averaging an 11.9% increase over nontransgenic Yangmai18 (table S3). We concluded that the constitutive overexpression of *TaCol-B5* in T_1, T_2, and T_3 transgenic plants modified spike architecture and increased the numbers of spikelets and spikes, but its positive effect on the number of grains in the T_3 population was suppressed by plant density or environmental effects. *TaCol-B5* as a single gene increased grain yield in the indigenous cultivar Yangmai18.

TaCOL-B5 was primarily expressed in the shoot apex and tiller bud, consistent with its potential role in promoting tillering and thus more spikes, but it was also expressed in leaves and roots of juvenile plants at the five- to six-leaf stage (fig. S11). However, there was no significant difference in the spatial or temporal expression of *TaCOL-B5* between the two alleles (fig. S11), excluding the possibility that the traits described above were regulated by *TaCOL-B5* at the transcript level and leading to the alternative hypothesis that phenotypic differences were probably determined by differences in *Ta*COL-B5 at the protein level.

Three amino acid substitutions, Phe^{243}/Leu^{243}, Ser^{269}/Gly^{269}, and Ala^{338}/Thr^{338}, were found between *Ta*Col-B5 and *Ta*col-B5 proteins (fig. S4). We next investigated whether any of these amino acid substitutions affected the interaction of *Ta*Col-B5 or *Ta*col-B5 with other proteins. From a wheat yeast two-hybrid (Y2H) library, we identified a clone encoding *Traes*CS4D02G196100, or *Ta*K4, which is an ortholog of rice *OsK4* encoding a serine/threonine protein kinase (GenBank Q852Q1) (*9*). Indeed, *Ta*Col-B5 and *Ta*col-B5 showed differential interactions with *Ta*K4 in the Y2H system (Fig. 3A and fig. S12) and in a transient expression system in tobacco leaves (fig. S13). Protein sequence analysis suggests phosphorylation sites in the amino acid substitutions (Fig. 3B). Furthermore, comparative in vitro phosphorylation interaction studies showed that the Ser^{269}/Gly^{269} substitution in *Ta*Col-B5 and *Ta*col-B5 resulted in potential differential protein phosphorylation by *Ta*K4 (Fig. 3C). This study provides an example that protein phosphorylation may be involved in spike architecture and grain yield in plants.

Constitutive overexpression of *TaCol-B5* also was found to regulate heading date (earlier) in the greenhouse and plant height (taller) in the greenhouse and field (fig. S14). We tested whether the CCT (CONSTANS, CO-like, and TOC1) domain of *Ta*COL-B5 is manifested

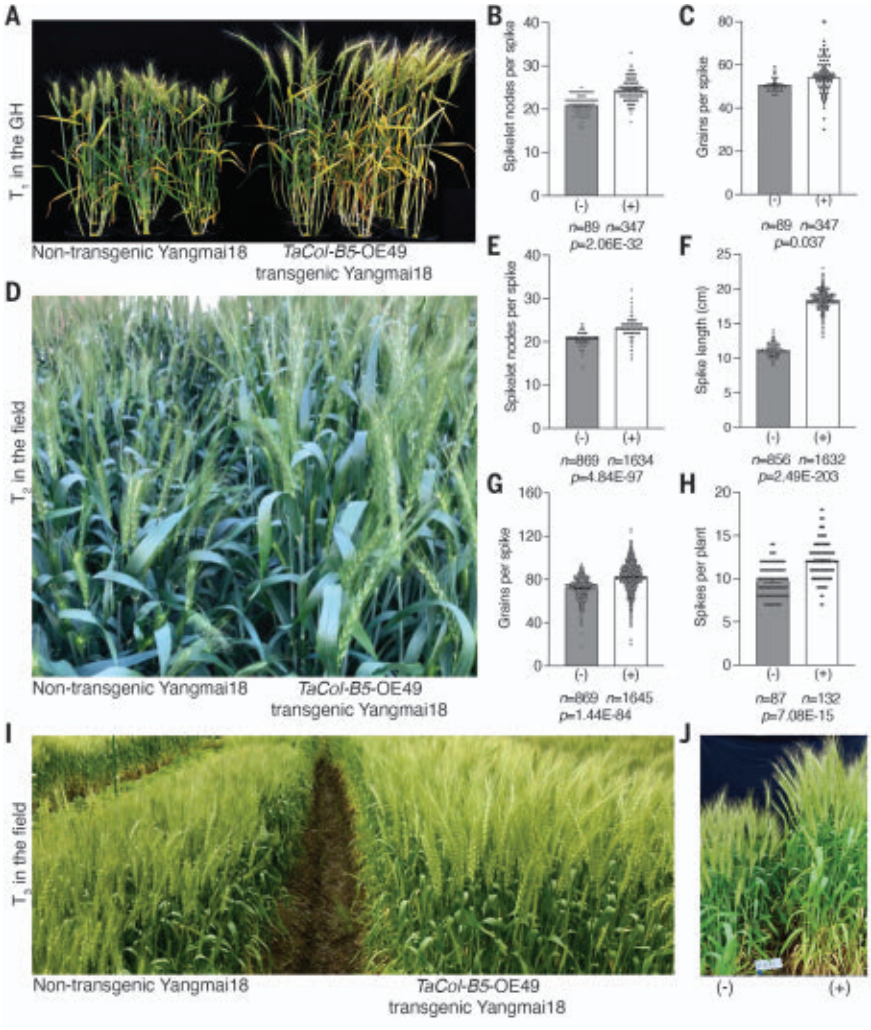


Fig. 2. Effects of *TaCol-B5* in T_1, T_2, and T_3 transgenic wheat plants. (**A** to **C**) Assessment of *TaCol-B5* in T_1 transgenic wheat in the greenhouse (GH). In addition to visual differences in adult plant phenotype (transgenic event *TaCol-B5*-OE49) (A), significant differences were observed between transgenic plants (+) and nontransgenic plants (−) for SNS (B) and grain number (C), averaged across the four *TaCol-B5* overexpression families. (**D** to **H**) Assessment of *TaCol-B5* in T_2 transgenic wheat in the field. From single-row field plots in Jiangsu, China (D), significant differences were observed between T_2 transgenic plants (+) and nontransgenic plants (−) in the number of spikelet nodes per spike (E), spike length (F), the number of grains per spike (G), and the number of spikes per plant (H). A two-tailed Student's *t* test was used to determine the significance level, and actual *P* values are shown in the figures. The *n* values indicate the number of spikes [(B), (C), and (E) to (G)] or plants (H) as characterized. (**I** and **J**) Assessment of *TaCol-B5* in T_3 transgenic wheat in the field. Adult transgenic plants (+) and nontransgenic plants (−) are shown in standard field plots, with a drainage ditch arranged between plots (I). Same comparison as in (I), but with a black cloth used as a background (J). Additional phenotypic information is provided in table S3.

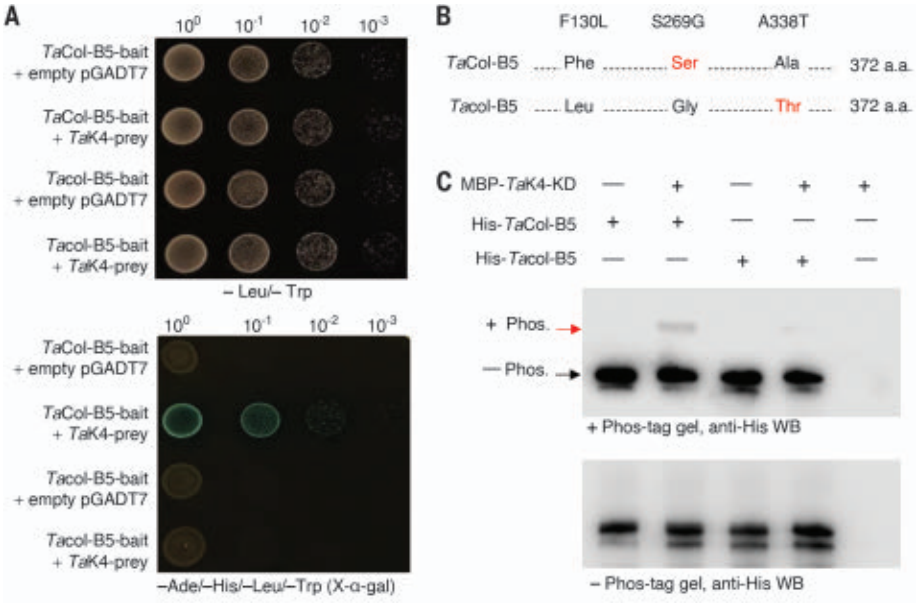


Fig. 3. Interaction and phosphorylation of *Ta*Col-B5 by *Ta*K4. (**A**) Cotransformed cells were grown on plates lacking two amino acids (–Leu/–Trp) and on plates lacking four amino acids (–Ade/–His/–Leu/–Trp). Four colony solutions diluted to different levels (10^0 to 10^{-3}) were inoculated on the same plate for each protein pair. Only yeast cells harboring the *Ta*Col-B5 bait and *Ta*K4 prey combination grew. (**B**) Three amino acid (a.a.) substitutions between *Ta*Col-B5 and *Ta*col-B5, and two of them that could be differentially phosphorylated, are indicated in red. (**C**) Phosphorylation of *Ta*Col-B5 by *Ta*K4. An in vitro kinase assay was performed with purified His-tagged *Ta*COL-B5 proteins. The phosphorylated (+Phos) and nonphosphorylated (–Phos) proteins were separated on a Phos-tag (50 μM) 10% polyacrylamide gel (top image), and the proteins from the same reaction were run on a non-Phos-tag gel used as a control (bottom image). The +Phos and –Phos proteins were analyzed by Western blotting (WB) with an anti-His tag antibody. The +Phos *Ta*Col5 protein showed lower electrophoretic mobility (indicated by a red arrow) than its –Phos counterparts (indicated by a black arrow).

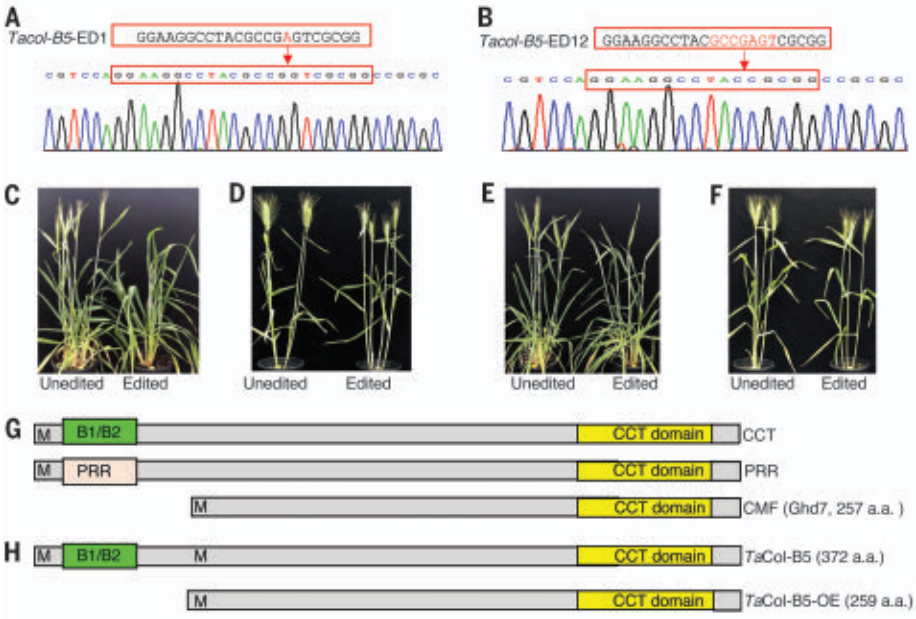


Fig. 4. Functional domains and pleiotropic effects of *Ta*COL-B5 proteins. (**A** to **F**) Effects of edited *Tacol-B5*. Edited sequences (in red) include a 1-bp deletion (*Tacol-B5*-ED1) (A) and a 7-bp deletion (*Tacol-B5*-ED12) (B). Images were taken to show effects of *Tacol-B5*-ED1 on heading date (C) and plant height (D) and effects of *Tacol-B5*-ED12 on heading date (E) and plant height (F). (**G**) Diagram of the structures of three plant CCT proteins. M, methionine as the first amino acid at the N terminus. (**H**) Structure of *Ta*Col-B5 protein with the predicted B1/B2-box. *Ta*Col-B5-OE without the predicted B-boxes was tested in transgenic plants.

visually in wheat by editing the sequence specific to *Tacol-B5* in Yangmai18 (not *TaCol-B5* in CItr 17600, which is not yet transformable). Damage to the *Ta*col-B5 CCT domain in two independent editing events, *Ta*col-B5-ED1 and *Ta*col-B5-ED12 (Fig. 4, A and B), delayed heading and reduced plant height (Fig. 4, C to F). The edited CCT domain in *Ta*col-B5 showed effects on plant productivity in the greenhouse but not in the field (fig. S15 and table S2), suggesting that *Ta*COL-B5 might regulate multiple agronomic traits through its different domains (*10*).

We found that 10 sequenced wheat genomes and five tetraploid *T. durum* wheat accessions have *Ta*col-B5, whereas the tetraploid wild emmer wheat (*T. turgidum*) cultivar Zavitan has *Ta*Col-B5 (fig. S16). We developed a diagnostic marker for the SNP involving the Ser^{269}/Gly^{269} substitution between *Ta*Col-B5 and *Ta*col-B5 (fig. S17). *TaCol-B5* was found in only 33 of 1657 accessions in a global collection of modern wheat cultivars and germplasm (table S4). It remains plausible that the rare allele *TaCol-B5*, accessible from modern wheat cultivars in different continents, could be used to enhance grain yield in a diverse array of genetic backgrounds and environments.

Numerous CCT proteins have been found in different plant species, and they are classified into three families according to their domains: COL, PRR (Pseudo-response regulator), and CMF (CCT motif family) (Fig. 4G) (*11*, *12*). The CCT proteins mostly regulate flowering through pathways of photoperiod (*13*–*15*), circadian rhythms (*16*), or vernalization (*17*), but a few proteins, including ZmCCT in maize (*18*) and Hd1 in rice (*19*), are reported to be involved in spike development. *Ghd7* in rice is a gene that affects heading date, grain number, and plant height, and Ghd7 protein does not have any B-box or PRR (Fig. 4G) (*20*). *TaCol-B5* is not orthologous to *Ghd7* in sequence, and the predicted full length of the *Ta*COL-B5 protein has a conserved CCT domain (fig. S18) and B1/B2-boxes (Fig. 4H and fig. S19, A and B). We attempted to transform the same host plant with *TaCol-B5* with the region encoding the predicted B1/B2-boxes and without this region to identify the functions of the B-boxes (fig. S19C), but only the latter construct was successful in transformation. The results demonstrated that the expressed *Ta*Col-B5 protein without the predicted B-boxes was able to maintain its functions in transgenic wheat. However, the dominant *TaCol-B5* allele without the region encoding B-boxes was driven by the maize ubiquitin promoter, which could produce pleiotropic effects in wheat (*21*, *22*). Future studies should investigate functions of the predicted B-boxes in *Ta*Col-B5, oligomeric states and structure of *Ta*Col-B5 proteins with or without the B-boxes, and their downstream genes in wheat plants. The cloned *TaCol-B5* has a general role in promoting cell proliferation and differentiation, leading to an overall increase in spikelet number and spike length, as well as tiller and spike number and plant size, and it is thus a growth regulator in plant species.

REFERENCES AND NOTES

1. G. M. Wolde, M. Mascher, T. Schnurbusch, *Mol. Genet. Genomics* **294**, 457–468 (2019).
2. Y. Wang, F. Miao, L. Yan, *PLOS ONE* **11**, e0151656 (2016).
3. S. Sakuma, T. Schnurbusch, *New Phytol.* **225**, 1873–1882 (2020).
4. S. Kuzay *et al*., *Theor. Appl. Genet.* **132**, 2689–2705 (2019).
5. K. P. Voss-Fels *et al*., *Theor. Appl. Genet.* **132**, 2707–2719 (2019).
6. Z. Guo, G. A. Slafer, T. Schnurbusch, *J. Exp. Bot.* **67**, 4221–4230 (2016).
7. S. Sakuma *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **116**, 5182–5187 (2019).
8. Z. Guo *et al*., *Sci. Rep.* **8**, 14435 (2018).
9. X. Sun *et al*., *PLOS Genet.* **12**, e1005927 (2016).
10. J. Putterill, F. Robson, K. Lee, R. Simon, G. Coupland, *Cell* **80**, 847–857 (1995).
11. S. Griffiths, R. P. Dunford, G. Coupland, D. A. Laurie, *Plant Physiol.* **131**, 1855–1867 (2003).
12. J. Cockram *et al*., *PLOS ONE* **7**, e45307 (2012).
13. S. Wenkel *et al*., *Plant Cell* **18**, 2971–2984 (2006).
14. M. Yano *et al*., *Plant Cell* **12**, 2473–2483 (2000).
15. A. Turner, J. Beales, S. Faure, R. P. Dunford, D. A. Laurie, *Science* **310**, 1031–1034 (2005).
16. C. Strayer *et al*., *Science* **289**, 768–771 (2000).
17. L. Yan *et al*., *Science* **303**, 1640–1644 (2004).
18. G. Xu *et al*., *New Phytol.* **214**, 852–864 (2017).
19. N. Endo-Higashi, T. Izawa, *Plant Cell Physiol.* **52**, 1083–1094 (2011).
20. W. Xue *et al*., *Nat. Genet.* **40**, 761–767 (2008).
21. M. J. Cornejo, D. Luth, K. M. Blankenship, O. D. Anderson, A. E. Blechl, *Plant Mol. Biol.* **23**, 567–581 (1993).
22. A. H. Christensen, P. H. Quail, *Transgenic Res.* **5**, 213–218 (1996).

ACKNOWLEDGMENTS

Funding: This project was supported by Agriculture and Food Research Initiative Competitive Grants 2017-67007-25939, 2017-67007-25932, and 2022-68013-36439 from the USDA National Institute of Food and Agriculture (NIFA). This project was also supported by grants from the Oklahoma Center for Advanced Science and Technology (OCAST, AR17-020-03), the Oklahoma Wheat Research Foundation, the Oklahoma Agricultural Experiment Station, and the Dillon and Lois Hodges Professorship. H.J. received grants from the "111" Project and the Collaborative Innovation Center for Modern Crop Production (CIC-MCP) cosponsored by Province and Ministry, China. We thank E. Akhunov for providing the pBUN421 plasmid and M. Tadege for valuable discussion. **Author contributions:** X.Z. and H.J. performed experimental procedures and analyzed results. T.L. contributed to protein interaction and phosphorylation analyses. R.N., L.L., and C.-C. K. contributed to gene expression, protein identification, and greenhouse experiments. H.J., J.W., W.H., Z.L., C.P., and B.F.C. contributed to field trials and phenotypic analyses. C.P. and C.C. contributed to initial GBS marker development and QTL mapping. L.Y. conceived of the idea, designed experiments, analyzed genotypic and phenotypic data, and interpreted results. X.Z., H.J., and L.Y. wrote the manuscript. B.F.C. edited the manuscript. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** All data are available in the main text or the supplementary materials.

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abm0717
Materials and Methods
Figs. S1 to S19
Tables S1 to S5
References (*23*–*49*)
MDAR Reproducibility Checklist
Data S1

8 September 2021; accepted 22 February 2022
10.1126/science.abm0717

OPTICS

Topological engineering of terahertz light using electrically tunable exceptional point singularities

M. Said Ergoktas[1,2], Sina Soleymani[3], Nurbek Kakenov[4]†, Kaiyuan Wang[1,2], Thomas B. Smith[5]‡, Gokhan Bakan[1,2], Sinan Balci[6], Alessandro Principi[5], Kostya S. Novoselov[5], Sahin K. Ozdemir[3,7]*, Coskun Kocabas[1,2,8]*

The topological structure associated with the branch point singularity around an exceptional point (EP) can provide tools for controlling the propagation of light. Through use of graphene-based devices, we demonstrate the emergence of EPs in an electrically controlled interaction between light and a collection of organic molecules in the terahertz regime at room temperature. We show that the intensity and phase of terahertz pulses can be controlled by a gate voltage, which drives the device across the EP. Our electrically tunable system allows reconstruction of the Riemann surface associated with the complex energy landscape and provides topological control of light by tuning the loss imbalance and frequency detuning of interacting modes. Our approach provides a platform for developing topological optoelectronics and studying the manifestations of EP physics in light–matter interactions.

The ability to understand and control light-matter interactions is fundamental to a wide range of applications in the classical and quantum domains, including but not limited to sensing, imaging, light generation, information processing, and computation. The light component in these interactions is usually in the form of electromagnetic modes confined in a resonator, whereas the matter component involves a single or a mesoscopic number of oscillators. Changing the number of oscillators coupled to a resonator is one route for achieving strong or weak light-matter coupling (*1*); however, this is not desirable in many practical settings as it does not lend itself to tunable and finely controllable platforms that can enable study of both weak and strong coupling regimes as well as transitions between them. The alternative is to keep the number of oscillators fixed while tuning the coupling strength and loss imbalance between the oscillators and the resonator such that the coupled oscillator-resonator system is steered between the weak and strong coupling regimes. Such non-Hermitian engineering of the system inevitably gives rise to non-Hermitian degeneracies known as exceptional points (EPs), which coincide with the crossover point between the weak and strong coupling regimes (*2*–*4*). EPs are substantially different from the degeneracies of Hermitian systems, known as diabolic points (DPs) (*5*). At a DP, only the eigenvalues coalesce but the corresponding eigenstates remain orthogonal. By contrast, at an EP both the eigenvalues and the associated eigenvectors coalesce, considerably modifying the energy landscape of the system and thus resulting in reduced dimensionality and skewed topology. This, in turn, enhances the system's response to perturbations (*6*–*9*), modifies the local density of states leading to the enhancement of spontaneous emission rates (*10*, *11*), and leads to a plethora of counterintuitive phenomena such as loss-induced lasing (*12*), topological energy transfer (*13*), enhanced chiral absorption (*14*), linewidth enhancement in lasers (*15*), unidirectional emission in ring lasers (*16*), and asymmetric mode switching (*17*).

We demonstrate the emergence of EPs in an electrically tunable platform that enables non-Hermitian engineering of the interaction of light with a collection of organic molecules in the terahertz (THz) regime. In contrast to previous demonstrations in optical (*18*–*20*), optomechanical (*13*, *15*, *21*), electronic (*22*), acoustic (*23*), and thermal systems (*24*)—where EPs emerge in a parameter space constructed from measurements of samples with different geometrical parameters—we observe EPs in a single fully electrically tunable device. This electrical control allows us to finely tune the losses as well as detune the system to construct voltage-controlled parameter space.

Our platform is a graphene-based tunable terahertz resonator (*25*), with the gate electrode forming a bottom reflective mirror and the graphene layer placed a distance away from it forming a tunable top mirror (Fig. 1A). A nonvolatile ionic liquid electrolyte layer is placed between the mirrors to achieve reversible gating of graphene by an applied voltage V_1 (i.e., effective gate voltage from the Dirac point), enabling an electrically tunable reflectivity and hence resonator loss. The gate electrode (a 100-nm gold film evaporated on a 50-μm-thick Kapton film) is placed on a piezo stage driven by an applied voltage V_2, forming a moveable mirror that can be used to vary the cavity length and hence tune the resonance frequency. Details regarding device fabrication are provided in (*26*). α-lactose crystals that support collective intermolecular vibrations at $\omega_{vib} = 0.53$ THzwith a very narrow linewidth of $\gamma_{vib} = 0.023$ THz are embedded in the resonator to allow for study of the emergence of EPs in light-matter interactions (i.e., coupling between the resonator field and the α-lactose crystals) in the THz regime. α-lactose was chosen over other materials, as its smaller damping rate makes it possible to achieve strong coupling at room temperature with our graphene THz resonator.

The dynamics of this coupled system, in which an ensemble of N identical molecular vibrations of frequency ω_{vib} are coupled to a resonator mode of frequency ω_c with the same coupling strength g, are given by the complex eigenfrequencies $\omega_\pm = (\Delta + 2\omega_{vib})/2 - i(\Gamma + 2\gamma_{vib})/4 \pm \Omega/4$. The nonorthogonal eigenmodes are $|\psi_\pm\rangle \propto \begin{pmatrix} \omega_\pm \\ \sqrt{N}g \end{pmatrix}$. Here, $\Delta = \omega_c - \omega_{vib}$ is the frequency detuning and $\Gamma = \gamma_c - \gamma_{vib}$ represents the loss imbalance between the molecular oscillators and the resonator, whereas γ_c and γ_{vib} are the decay rates of the resonator and molecular vibrations, respectively. Finally, $\Omega = \sqrt{16Ng^2 + (2\Delta + i\Gamma)^2}$ denotes the effective coupling strength between two systems. Analysis of this expression reveals that for $\Delta = 0$ (i.e., when the field is resonant with molecular vibrations) and $\sqrt{N}g > \Gamma/4$ (i.e., strong coupling regime), the complex eigenfrequencies exhibit splitting in their real parts whereas their imaginary parts remain coalesced. On the other hand, for $\sqrt{N}g < \Gamma/4$ (i.e., weak coupling regime) they exhibit splitting in their imaginary parts whereas the real parts coalesce, implying the modification of the decay rates of the eigenstates. For $\sqrt{N}g = \pm\Gamma/4$, the complex eigenfrequencies coalesce both in their real and imaginary parts, i.e., $\omega_\pm = \omega_{EP} = (\omega_c + \omega_{vib})/2 - i(\gamma_c + \gamma_{vib})/4$, and in their associated eigenmodes, i.e., $|\psi_\pm\rangle = |\psi_{EP}\rangle \propto \begin{pmatrix} \omega_{EP} \\ \Gamma_{EP} \end{pmatrix}$ with $\Gamma_{EP} = \pm 4\sqrt{N}g$, implying the emergence of two EPs.

In our system (Fig. 1A), the knobs V_1 and V_2 are used to finely tune Γ and Δ, respectively, and allow us to observe the transition between the strong and weak coupling regimes through the EP. Plotting the complex energy landscape (i.e., real and imaginary parts of the complex eigenfrequencies $\omega_\pm$) as V_1 and V_2 are varied yields two intersecting Riemann sheets wrapped

[1]Department of Materials, University of Manchester, Manchester, M13 9PL, UK. [2]National Graphene Institute, University of Manchester, Manchester, M13 9PL, UK. [3]Department of Engineering Science and Mechanics, Pennsylvania State University, University Park, PA 16802 USA. [4]Department of Physics, Bilkent University, Ankara, Turkey. [5]Department of Physics and Astronomy, University of Manchester, Manchester, M13 9PL, UK. [6]Department of Photonics, Izmir Institute of Technology, Izmir, Turkey. [7]Materials Research Institute, Pennsylvania State University, University Park, PA 16802, USA. [8]Henry Royce Institute for Advanced Materials, University of Manchester, Manchester M13 9PL, UK.
*Corresponding author. Email: sko9@psu.edu (S.K.O.); coskun.kocabas@manchester.ac.uk (C.K.)
†Present address: Department of Physics, Technical University of Denmark, DK-2800 Kongens Lyngby, Denmark.
‡Present address: Brainpool AI Dudley House 169 Piccadilly, St. James's, London, W1J 9EH, UK.

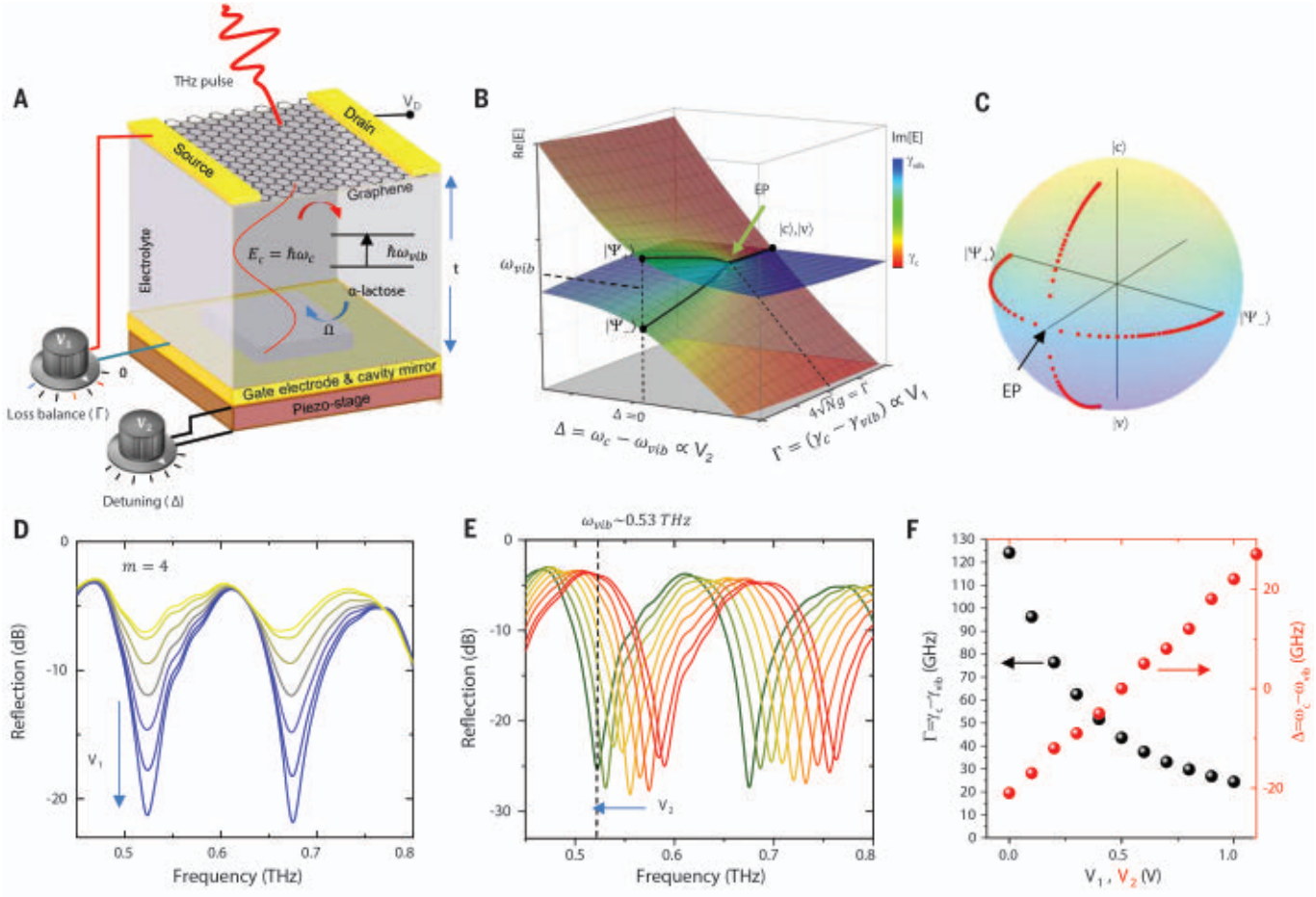


Fig. 1. Electrically tunable EP device. (**A**) Schematic of the electrolyte-gated graphene transistor embedded with lactose microcrystals. The tunable coupling between the resonator mode $E_c = \omega_c + i\gamma_c$ and the intermolecular vibrations of lactose crystals $E_{vib} = \omega_{vib} + i\gamma_{vib}$ forms an electrically tunable two-parameter framework to realize EP devices. The gate voltage V_1 controls the loss imbalance Γ between the cavity and intermolecular vibrations by tuning the charge density on graphene, and V_2 controls the detuning frequency Δ by changing the cavity size. (**B**) Riemann surface obtained through numerical simulations shows the complex energy eigenvalues of the device plotted on the two-parameter voltage space defined by V_1 and V_2. EP emerges when the coupling strengths compensates the loss imbalance $\sqrt{N}g = \pm\Gamma/4$, when the cavity field and the intermolecular vibrations are on resonant $\Delta = \omega_c - \omega_{vib} = 0$. (**C**) Visualization of the evolution of the supermodes of the coupled system on a Bloch sphere as the gate voltage V_1 is varied (loss imbalance Γ is tuned). The azimuthal angle on the sphere indicates the relative phase, the polar angle represents the relative intensity of the uncoupled cavity (photon mode), and the collective molecular vibrations (matter mode) are represented by the eigenmodes $|c\rangle$ and $|v\rangle$, respectively. (**D** and **E**) THz reflection spectrum of the graphene cavity without lactose molecules but with the electrolyte showing the dependence of the cavity mode $|c\rangle$ on V_1 and V_2, respectively. (**F**) Voltage dependence of the loss imbalance Γ and detuning Δ of the system.

around a second-order EP right in the center where the two complex eigenfrequencies of the system coalesce (Fig. 1B). Representing the eigenstates of the system on the Bloch sphere (Fig. 1C) allows us to monitor the evolution of the state of the system during the transition from weak to strong coupling through the EP. In largely detuned or large loss imbalance cases (i.e., $\Delta \to \infty$ or $\Gamma \gg \sqrt{N}g$, that is, the limit of the uncoupled modes), the two supermodes of the system approach to the individual uncoupled electromagnetic mode (cavity photonic mode) and the matter mode (vibrational mode), which are located at the north and the south poles of the Bloch sphere, respectively. For $\Delta = 0$, varying V_1 and hence Γ gradually shifts the supermodes from the poles distributing them across the cavity and the matter (α-lactose crystals). The supermode close to the north pole mostly resides in the cavity (cavity-like mode) whereas the supermode close to the south pole mostly resides in the matter (matter-like mode). With further tuning of Γ, the cavity-like mode $|c\rangle$ moves downward from the north pole, whereas the matter-like mode $|v\rangle$ moves upward from the south pole toward the equator. These modes then coalesce to the single mode $|\psi_{EP}\rangle$ on the equator at the critical value $\Gamma_{EP} = \pm 4\sqrt{N}g$, where dual EPs emerge.

We first confirm the effects of tuning knobs V_1 and V_2 (Fig. 1A) on the reflectivity of the empty THz resonator. As the voltage V_1—which controls the cavity loss (and hence the loss imbalance Γ of the couple)—is increased, the resonance frequency ω_c of the resonator remains intact, but the linewidth (proportional to the decay rate γ_c) of the cavity resonance becomes narrower and the resonance depth increases, approaching critical coupling (Fig. 1D). The second knob V_2 (cavity voltage) controls the length of the resonator and its resonance frequency ω_c by moving a piezo stage (hence the gate electrode) with respect to the graphene transistor with a resolution of <6 nm. This helps finely adjust the frequency detuning Δ. It is clearly seen that as V_2 is varied, the resonance frequency ω_c of the THz resonator shifts with no considerable variation in the resonance linewidth (Fig. 1E). Because these processes do not have any effect on the vibrational frequency and decay rate of the molecules, knobs V_1 and V_2 effectively control the two-dimensional parameter space of Δ and Γ. We observed a tunability of ~±25 GHz in Δ and

100 GHz in Γ when V_1 and V_2 were increased from 0 to 1 V (Fig. 1F). As a result, the knobs enable non-Hermitian engineering of the light-matter interaction between the THz resonator field and the collective intermolecular vibrations and allow us to map the complex energy landscape of the hybrid system.

Next, time-domain THz spectroscopy demonstrates the tunable transition between the weak and strong coupling regimes through an EP. We first tuned V_2 to have Δ = 0 and then varied the gate voltage V_1, which controls the loss imbalance of the couples. As V_1 is increased, the formation of the characteristic polariton branching around ω_{vib} is clearly observed in the reflectivity map of the device (Fig. 2A). This branching takes place at two symmetric EPs $V_{EP} = \pm 0.2V$ as a result of the ambipolar electrical conduction of graphene. A cross section of this reflectivity map around one of these EPs reveals the transition from a split

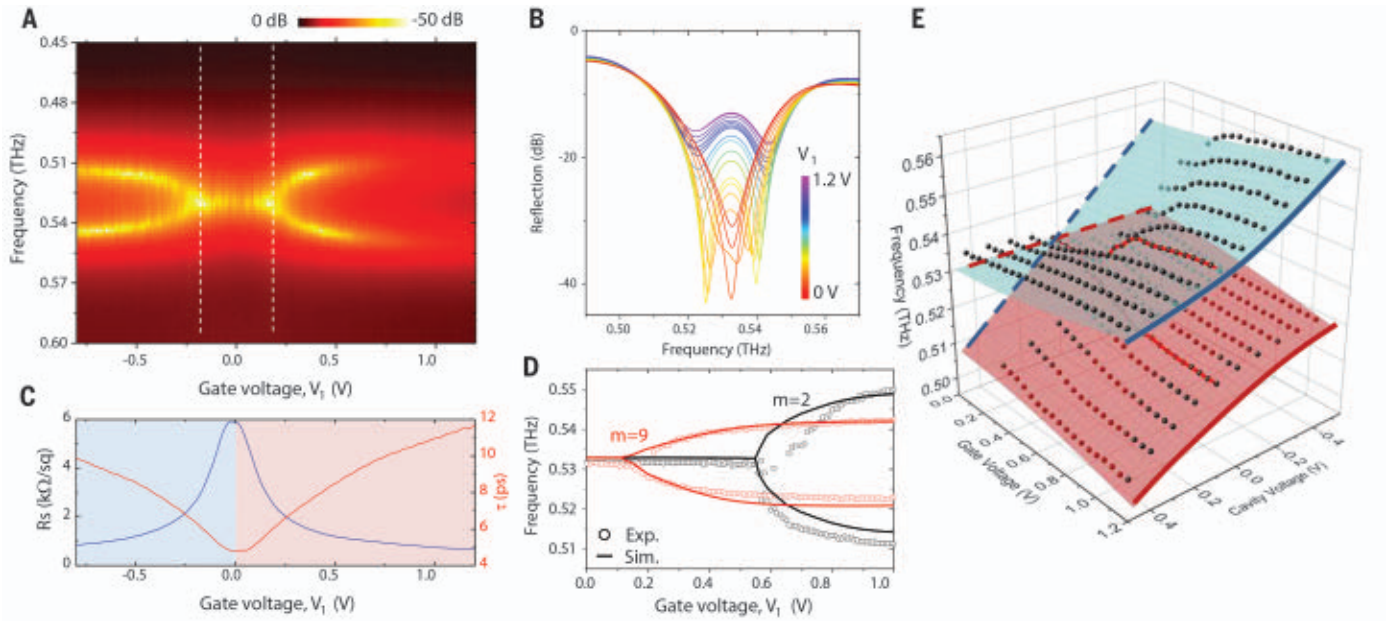


Fig. 2. Spectroscopic characterization of the EP device. (**A** and **B**) Reflectivity map and spectra of the device showing the transition from the weak (coalesced modes) to the strong coupling (split modes) regimes through an EP as V_1 is varied (Γ is tuned) at constant V_2, satisfying Δ = 0. Because of the ambipolar conduction of graphene, the device goes through two EPs at $V_{EP1} = -0.2V$ (electron doping) and $V_{EP2} = 0.2V$ (hole doping). (**C**) Sheet resistance of graphene and cavity decay time plotted against the gate voltage. Increasing the gate voltage enhances the THz reflectivity of the graphene mirror, leading to a longer cavity decay time. (**D**) Position of the EP and the amount of splitting vary with the mode number *m*. EPs emerge at smaller gate voltages for higher *m*. (**E**) Riemann surfaces obtained experimentally (black dotted) and through calculations (blue and red sheets) showing the real part of complex eigenvalues of the device in the voltage-controlled parameter space.

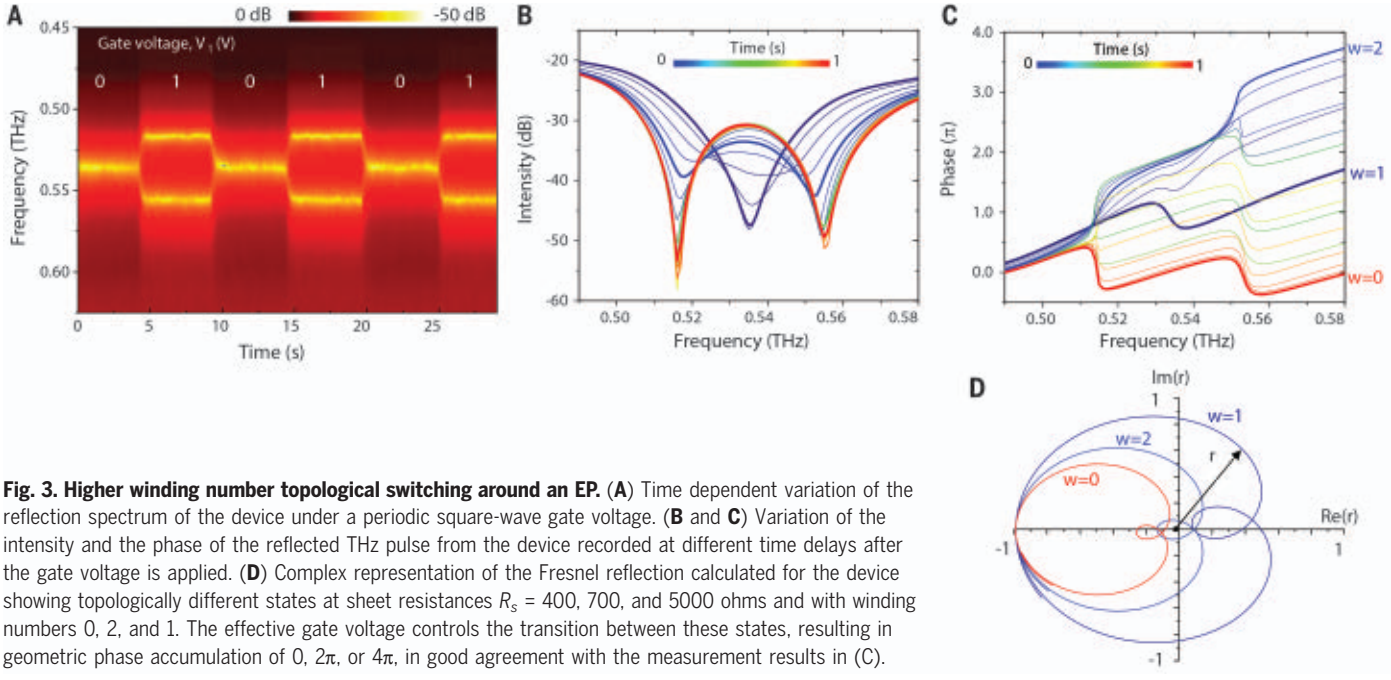


Fig. 3. Higher winding number topological switching around an EP. (**A**) Time dependent variation of the reflection spectrum of the device under a periodic square-wave gate voltage. (**B** and **C**) Variation of the intensity and the phase of the reflected THz pulse from the device recorded at different time delays after the gate voltage is applied. (**D**) Complex representation of the Fresnel reflection calculated for the device showing topologically different states at sheet resistances R_s = 400, 700, and 5000 ohms and with winding numbers 0, 2, and 1. The effective gate voltage controls the transition between these states, resulting in geometric phase accumulation of 0, 2π, or 4π, in good agreement with the measurement results in (C).

mode spectrum (i.e., strong coupling regime) to a coalesced mode spectrum (i.e., weak coupling regime) through the EP (Fig. 2B). The transition between these two regimes as V_1 is varied can be attributed to the variation of the optical conductivity of graphene and the corresponding cavity decay time (Fig. 2C). This dependence on V_1 clarifies our ability to control loss imbalance between the couples through the control of the resonator losses.

Experiments with different cavity modes (from m = 2 to 9, adjusted by tuning the cavity size) satisfying $\Delta = 0$ reveal that the transition from the split modes to coalesced modes occurs at different V_1 voltages for different cavity modes (Fig. 2D): The higher the mode number m, the smaller the required gate voltage V_1 to arrive at the EP. This behavior may be attributed to (i) the larger mode volume (and hence lower field strength) and thus the reduced effective coupling strength at higher m or (ii) the smaller γ_c of higher-order modes and thus smaller initial loss imbalance between the couples. As a result, the amount of additional loss imbalance required to satisfy the EP condition $\sqrt{N}g = \Gamma/4$ is smaller for higher-order cavity modes, implying that modes with higher m require smaller gate voltage V_1 to reach EP. Because the EP is a singularity point in the two-parameter space, we have finely tuned Γ and Δ through the knobs V_1 and V_2 for a fixed mode m and reconstructed the Riemann surface associated with the complex energy landscape of the system (Fig. 2E). The topology of two intersecting Riemann sheets centered around an EP is clearly seen (Figs. 1B and 2E). From the experimentally determined maximum frequency splitting values, we estimate the number of molecules contributing to the process as ~10^{18} for all cavity modes (*26*).

Next, we investigate the electrical control of EP and its effect on the intensity and the phase of the reflected THz light. For this purpose, we prepare the system at $\Delta = 0$ and dynamically modulate the loss imbalance Γ by applying a periodic square-wave gate voltage V_1. The time-dependent reflection spectra clearly show periodic splitting and coalescence of the modes (Fig. 3A). The system gradually transits from the coalesced modes ~0.535 THz to split modes with a splitting of ~40 GHz in 0.2 s after the gate voltage is set to the "ON" state. We recorded the intensity (Fig. 3B) and the phase (Fig. 3C) of the reflected THz pulse from the device at different time delays after the ON signal is applied. We must point out that the measured phase depends on the reference plane; however, the phase difference is uniquely defined. We observe a phase accumulation of 0, 2π, and 4π across the free spectral range of the resonator during the transition through the EP. This geometrical (i.e., Berry) phase is the result of the topology of the Fresnel reflectivity $r(\omega)$. Here the topological invariant is the winding number $n = \frac{1}{2\pi i}\oint \frac{dr}{r}$ of the complex Fresnel reflectivity around the perfect absorption singularity (r = 0; critical coupling) in which the reflection phase is undefined. Calculated reflection (Fig. 3D) for our device at three different sheet resistances reveals three topologically different reflectivities identified by winding numbers n = 0, 1, and 2 and the associated Berry phases of 0, 2π, or 4π, respectively, agreeing with the phases measured in the experiments (Fig. 3C). These results provide the first direct evidence for the electrically switchable reflection topology.

One of the most notable features of an EP is the exchange of the eigenstate when it is adiabatically encircled. This contrasts with encircling a DP in Hermitian systems where the eigenstate acquires a geometric phase and no state flip takes place. Although one loop around the EP flips the eigenstate, only the second loop returns the system to its initial state apart from a Berry phase π. State flip when encircling EPs has been experimentally demonstrated with static measurements from a series of samples including microwave cavities (*27*), optical resonators (*28*), exciton-polariton systems (*19*, *29*), and acoustic systems (*23*). Here, we probe our system when it is steered on cyclic paths encircling an EP by tuning Γ and Δ with the knobs V_1 and V_2. This is possible in our system because the two finely controlled knobs are independent. By varying V_1 and V_2 in steps of 25 mV such that an EP is encircled in the clockwise or counterclockwise directions, we monitor how the final state of the system is affected by the encircling process. In order to do this, we defined a loop by the points $\{\Delta_{max}, \Gamma_{min}\}$, $\{\Delta_{max}, \Gamma_{max}\}$, $\{\Delta_{min}, \Gamma_{max}\}$, $\{\Delta_{min}, \Gamma_{min}\}$ returning back to $\{\Delta_{max}, \Gamma_{min}\}$ after ~20 s. Similarly, in the parameter space of V_1 and V_2, the loop is defined by the corresponding voltage points as $\{V_{2max}, V_{1min}\}$, $\{V_{2max}, V_{1max}\}$, $\{V_{2min}, V_{1max}\}$, $\{V_{2min}, V_{1min}\}$ returning back to $\{V_{2max}, V_{1min}\}$. When we choose a control loop that does not enclose the EP, the system returns to the same state at the end of the loop (Fig. 4A), regardless of whether the loop is clockwise or counterclockwise. By contrast, when the loop encircles the EP, we observe that a trajectory starting on one of the Riemann sheets ends on the other sheet (Fig. 4B), resulting in eigenstate exchange (state flip): $|\psi_+\rangle \rightarrow |\psi_-\rangle$ and $|\psi_-\rangle \rightarrow |\psi_+\rangle$. To gain more insight on these dynamics, we illustrate the evolution of the eigenstates of the system on Bloch spheres for closed loops that do (Fig. 4D) and do not

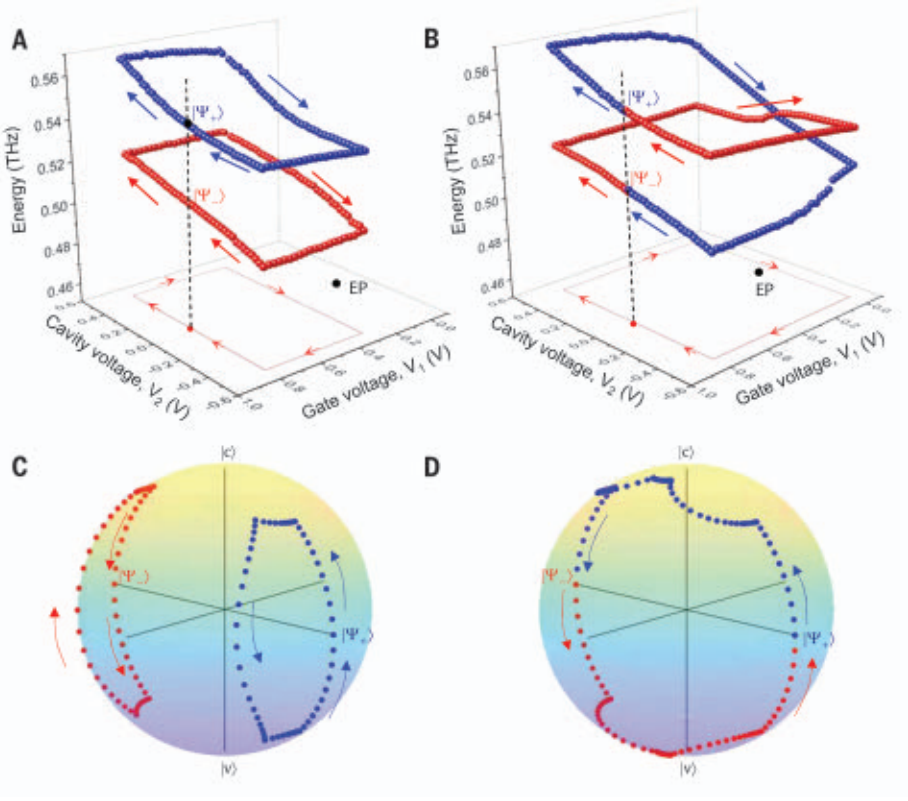


Fig. 4. Voltage-controlled encircling of EP. (**A** and **B**) Evolution of the energy of the coupled system along the trajectories traced by varying the voltages V_1 and V_2 in small steps. (A) A trajectory starting on one of the Riemann sheets stays on the same sheet if it does not encircle the EP. (B) A trajectory starting on one of the Riemann sheets ends on the other sheet (state exchange) if it encircles the EP. (**C** and **D**) Evolution of eigenstates of the system on the Bloch sphere for the trajectories shown in (A) and (B), respectively.

(Fig. 4C) encircle the EP. When the system is initially in the state $|\psi_+\rangle = (|c\rangle + |v\rangle)/\sqrt{2}$, which is the equal to the superposition of the cavity $|c\rangle$ and vibrational $|v\rangle$ modes, the final state after a closed loop encircling the EP becomes $|\psi_-\rangle = (|c\rangle - |v\rangle)/\sqrt{2}$, which is orthogonal to the initial state $|\psi_+\rangle$. A second loop around the EP brings the system back to its initial state $|\psi_+\rangle$ apart from a geometrical phase. As seen in the Bloch sphere (Fig. 4D), these two loops around the EP cut the Bloch sphere directly in half and correspond to a solid angle of 2π, which in turn implies that the acquired geometrical phase is π (i.e., the geometrical phase is the half of the solid angle enclosed by the curve connecting the initial and final states).

We have demonstrated a non-Hermitian optical device to study EP in the collective interaction of vibrational modes of organic molecules with a THz field. Through use of fully electrically tunable independent knobs, we can steer the system through an EP that enables electrical control on reflection topology. Our results provide a platform for the topological control of light-matter interactions around an EP, with potential applications ranging from topological optoelectronic devices to topological control of physical and chemical processes.

REFERENCES AND NOTES

1. P. A. Thomas, W. J. Tan, H. A. Fernandez, W. L. Barnes, *Nano Lett.* **20**, 6412–6419 (2020).
2. M. A. Miri, A. Alù, *Science* **363**, eaar7709 (2019).
3. S. K. Özdemir, S. Rotter, F. Nori, L. Yang, *Nat. Mater.* **18**, 783–798 (2019).
4. R. El-Ganainy *et al.*, *Nat. Phys.* **14**, 11–19 (2018).
5. C. M. Bender, S. Boettcher, *Phys. Rev. Lett.* **80**, 5243–5246 (1998).
6. W. Chen, Ş. K. Özdemir, G. Zhao, J. Wiersig, L. Yang, *Nature* **548**, 192–196 (2017).
7. H. Hodaei *et al.*, *Nature* **548**, 187–191 (2017).
8. Y. H. Lai, Y. K. Lu, M. G. Suh, Z. Yuan, K. Vahala, *Nature* **576**, 65–69 (2019).
9. M. P. Hokmabadi, A. Schumer, D. N. Christodoulides, M. Khajavikhan, *Nature* **576**, 70–74 (2019).
10. A. Pick *et al.*, *Opt. Express* **25**, 12325–12348 (2017).
11. Q. Zhong, A. Hashemi, Ş. K. Özdemir, R. El-Ganainy, *Phys. Rev. Res.* **3**, 013220 (2021).
12. B. Peng *et al.*, *Science* **346**, 328–332 (2014).
13. H. Xu, D. Mason, L. Jiang, J. G. E. Harris, *Nature* **537**, 80–83 (2016).
14. S. Soleymani *et al.*, *Nat. Commun.* **13**, 599 (2022).
15. J. Zhang *et al.*, *Nat. Photonics* **12**, 479–484 (2018).
16. B. Peng *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **113**, 6845–6850 (2016).
17. J. Doppler *et al.*, *Nature* **537**, 76–79 (2016).
18. J. B. Khurgin, *Optica* **7**, 1015 (2020).
19. W. Gao, X. Li, M. Bamba, J. Kono, *Nat. Photonics* **12**, 362–367 (2018).
20. B. Peng *et al.*, *Nat. Phys.* **10**, 394–398 (2014).
21. H. Jing *et al.*, *Phys. Rev. Lett.* **113**, 053604 (2014).
22. M. Geiser, G. Scalari, F. Castellano, M. Beck, J. Faist, *Appl. Phys. Lett.* **101**, 141118 (2012).
23. K. Ding, G. Ma, M. Xiao, Z. Q. Zhang, C. T. Chan, *Phys. Rev. X* **6**, 021007 (2016).
24. Y. Li *et al.*, *Science* **364**, 170–173 (2019).
25. N. Kakenov, M. S. Ergoktas, O. Balci, C. Kocabas, *2D Materials* **5**, 035018 (2018).
26. Materials and methods are available as supplementary materials.
27. C. Dembowski *et al.*, *Phys. Rev. Lett.* **86**, 787–790 (2001).
28. S. B. Lee *et al.*, *Phys. Rev. Lett.* **103**, 134101 (2009).
29. T. Gao *et al.*, *Nature* **526**, 554–558 (2015).
30. M. S. Ergoktas *et al.*, Topological engineering of terahertz light using electrically tunable exceptional point singularities, version 1, Zenodo (2022); https://doi.org/10.5281/zenodo.6105427.

ACKNOWLEDGMENTS

Funding: This work was funded through the European Research Council through ERC-Consolidator grant 682723, SmartGraphene (C.K.), the Air Force Office of Scientific Research (AFOSR) Multidisciplinary University Research Initiative (MURI) Award on Programmable systems with non-Hermitian quantum dynamics (Award FA9550-21-1-0202) (S.K.O.), and the Air Force Office of Scientific Research (AFOSR) Award FA9550-18-1-0235 (S.K.O.) A.P. acknowledges support from the European Commission under the EU Horizon 2020 MSCA-RISE-2019 programme (project 873028 HYDROTRONICS) and from the Leverhulme Trust under the grant RPG-2019-363. **Author contributions:** M.S.E., C.K., and S.K.O. conceived the idea. M.S.E. synthesized the graphene samples and fabricated the devices. M.S.E. and C.K. performed the experiments. S.S. and S.K.O. performed the simulation and developed the theory. N.K., G.B., and K.W. helped with the measurements. A.P. and T.B.S. provided theoretical support. M.S.E., S.K.O., S.B., and C.K. analyzed the data and wrote the manuscript with input from all the authors. All authors discussed the results and contributed to the scientific interpretation as well as to the writing of the manuscript. **Competing interests:** The authors declare no competing financial interests. **Data and materials availability:** All data discussed in the main text and code used in the simulations are available at Zenodo (*30*).

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abn6528
Materials and Methods
Supplementary Text
Figs. S1 to S13
References (*31*–*35*)

10 December 2021; accepted 10 March 2022
10.1126/science.abn6528

BIOMATERIALS

Mineralization generates megapascal contractile stresses in collagen fibrils

Hang Ping[1,2], Wolfgang Wagermaier[2]*, Nils Horbelt[2], Ernesto Scoppola[2], Chenghao Li[2], Peter Werner[2], Zhengyi Fu[1]*, Peter Fratzl[2]*

During bone formation, collagen fibrils mineralize with carbonated hydroxyapatite, leading to a hybrid material with excellent properties. Other minerals are also known to nucleate within collagen in vitro. For a series of strontium- and calcium-based minerals, we observed that their precipitation leads to a contraction of collagen fibrils, reaching stresses as large as several megapascals. The magnitude of the stress depends on the type and amount of mineral. Using in-operando synchrotron x-ray scattering, we analyzed the kinetics of mineral deposition. Whereas no contraction occurs when the mineral deposits outside fibrils only, intrafibrillar mineralization generates fibril contraction. This chemomechanical effect occurs with collagen fully immersed in water and generates a mineral-collagen composite with tensile fibers, reminiscent of the principle of reinforced concrete.

Biological hybrid materials such as bone elegantly combine hard inorganic nanometer-sized minerals and soft organic matrices into hierarchical architectures to achieve specific properties and functions (*1*, *2*). Such complex structures, ranging from nanoscale to macroscale, result in superior mechanical properties of biomineralized materials compared to their artificial counterparts (*3*, *4*). Collagen is the main constituent of extracellular tissues in our bodies, from tendon and bone to skin and arterial walls. In bone, collagen is reinforced by nanometer-sized particles of carbonated hydroxyapatite (*5*, *6*). Collagen fibrils can also be infiltrated in vitro with hydroxyapatite (*7*–*9*) and with other minerals such as calcium carbonate (*10*), silica (*11*), or iron hydroxides (*12*).

Effective prestressing strategies at the nanoscale are known to strengthen many materials and biominerals in particular (*13*, *14*). As an example, local compressive or tensile stresses can interact with cracks in minerals and deflect them and consequently enhance the materials' toughness (*13*). Prestresses in natural collagen-based tissues contribute substantially to their overall mechanical properties (*15*, *16*). Collagen molecules in bone contract in length when dehydrated or under osmotic stress (*16*), but how this is associated with mineral deposition can only be speculated. The present work shows that this is likely not a specific interaction with hydroxyapatite, because similar effects were observed for a broad range of mineral types.

Intrafibrillar collagen mineralization can be achieved in vitro by applying negatively charged macromolecules that help the penetration of fibrils by forming mineral-protein complexes. These disordered mineral precursors, sometimes called polymer-induced liquid precursor (*17*), are known to penetrate collagen fibrils and form mineral particles that resemble those in vivo (*7*–*9*). We adopted this principle for in vitro mineralization of a collagen substrate with various minerals—$SrCO_3$, $SrWO_4$, $SrSO_4$, CaF_2, and $CaCO_3$—and measured the

[1]State Key Laboratory of Advanced Technology for Materials Synthesis and Processing, Wuhan University of Technology, Luoshi Road No. 122, Wuhan 430070, China. [2]Department of Biomaterials, Max Planck Institute of Colloids and Interfaces, Am Mühlenberg 1, 14476 Potsdam, Germany.
***Corresponding author. Email: wolfgang.wagermaier@mpikg.mpg.de (W.W.); zyfu@whut.edu.cn (Z.F.); peter.fratzl@mpikg.mpg.de (P.F.)**

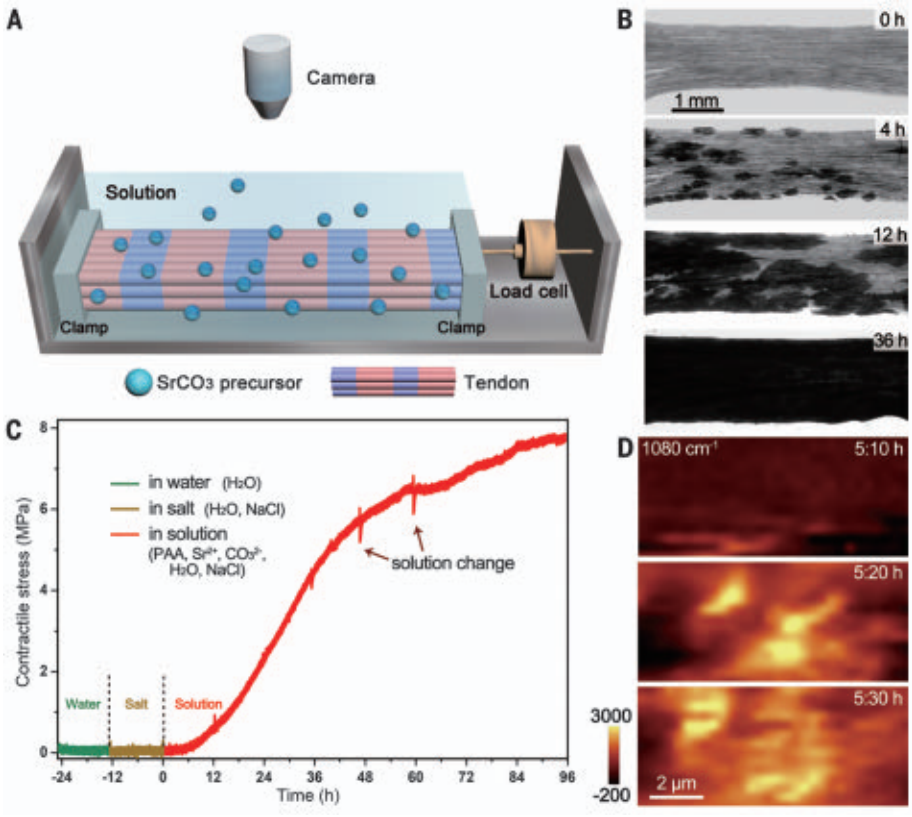


Fig. 1. Stress generation in tendons during mineralization with $SrCO_3$. (**A**) Schematic of the in-operando mechanical testing setup. The two ends of a tendon slice are fixed by clamps. One of the clamps is connected to a load cell to monitor the change of stress. The tendon slices are immersed in a reaction chamber containing mineralizing solution; the shape evolution of tendon slices is recorded by optical microscopy from the top. (**B**) Time series of optical images from the top view of a tendon slice at different reaction times of mineralization. The black spots are mineralized regions. They gradually grow to realize the full mineralization of the tendon. (**C**) Contractile stress curves of a tendon slice in various media (water, salt, solution). Small peaks in the curve are caused by the exchange of solution. (**D**) In-operando Raman mapping of a mineralized region on the surface of a tendon sample. The plotted signal (1080 cm^{-1}) indicates the progression of crystalline $SrCO_3$.

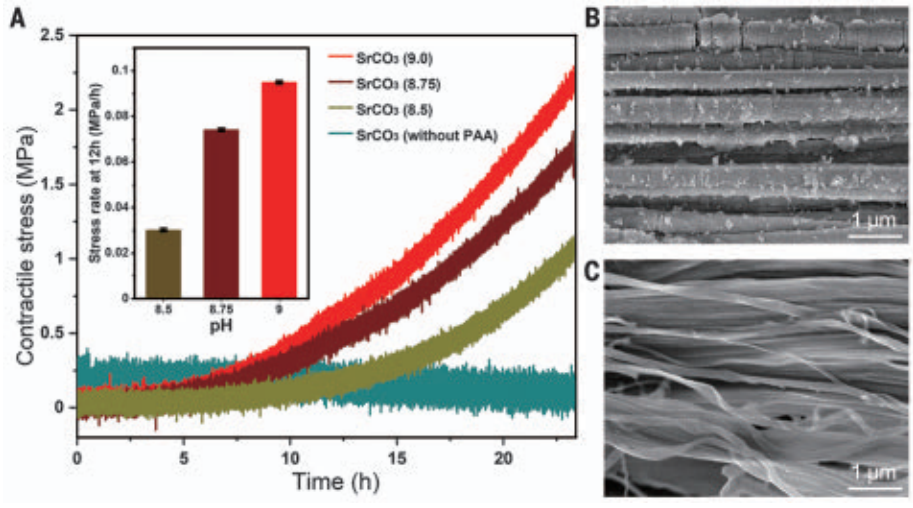


Fig. 2. Stress generation in tendons in $SrCO_3$ solution with different pH values. (**A**) Contractile stress of tendon slices as a function of time in mineralizing solution (pH = 9.0, red curve; pH = 8.75, dark red curve; pH = 8.5, olive curve) and without PAA (teal curve). The inset shows the slope (stress rate) of three curves with different pH values at 12 hours after the start of mineralization. (**B**) Scanning electron microscopy (SEM) image of mineralized tendon. Fibrils show some extrafibrillar mineral but did not shrink by dehydration. (**C**) SEM image of tendon treated in salt solution without PAA. Collagen fibrils are laterally shrunk by dehydration and look almost like ribbons, because there is no mineral precipitated inside.

resulting contractile stresses. The corresponding mineralization processes of the collagen matrix and the formation of internal stresses were monitored by in-operando x-ray scattering in the case of $SrCO_3$. A custom-made mechanical testing setup equipped with a reaction chamber and an optical microscopy system was used to investigate the stress generation during mineralization of unmineralized turkey leg tendons with $SrCO_3$ (Fig. 1A and fig. S1). Additionally, the local formation and morphology of the mineral were imaged by in-operando Raman spectroscopy and electron microscopy, respectively.

As a biological source of parallel collagen fibrils known to be able to mineralize in vivo, we used slices of unmineralized turkey tendon and immersed them into a $SrCO_3$ precursor phase. In our experiments, we used a solution of 200 μg/ml polyacrylic acid (PAA), 10 mM Sr^{2+}, and CO_3^{2-}, whereby the PAA molecules stabilize the ions in solution through the polymer-induced liquid precursor process (*7*). These precursors first infiltrated the collagen fibrils of the tendon and subsequently acted as nucleation sites, and thereby led to gradual mineral deposition. Initial mineral agglomerates could be observed after 4 hours (Fig. 1B) with a roughly ovoid morphology. This might be caused by the faster diffusion of the precursor along the longitudinal direction of tendon fibers than along the transverse direction. Fully impregnated tendon slices

were observed after 36 hours (Fig. 1B and movies S1 and S2).

The progression of mineralization was monitored by in-operando Raman scanning microscopy of regions close to the mineralizing front (fig. S2). The spatiotemporal variation of the line intensity from the symmetric stretching mode v1 of carbonate groups (1080 cm^{-1}) (*18*) is shown (fig. S2D). Signal intensity started to increase at 5 hours, and 20 min later a strong increase of the signal was observed. This indicated that the mineralization front crossed the observation window at this moment. Moreover, the gradual appearance of a crystalline phase was indicated by an increasing intensity of the characteristic band of $SrCO_3$ at 1080 cm^{-1} (*19*) (Fig. 1D).

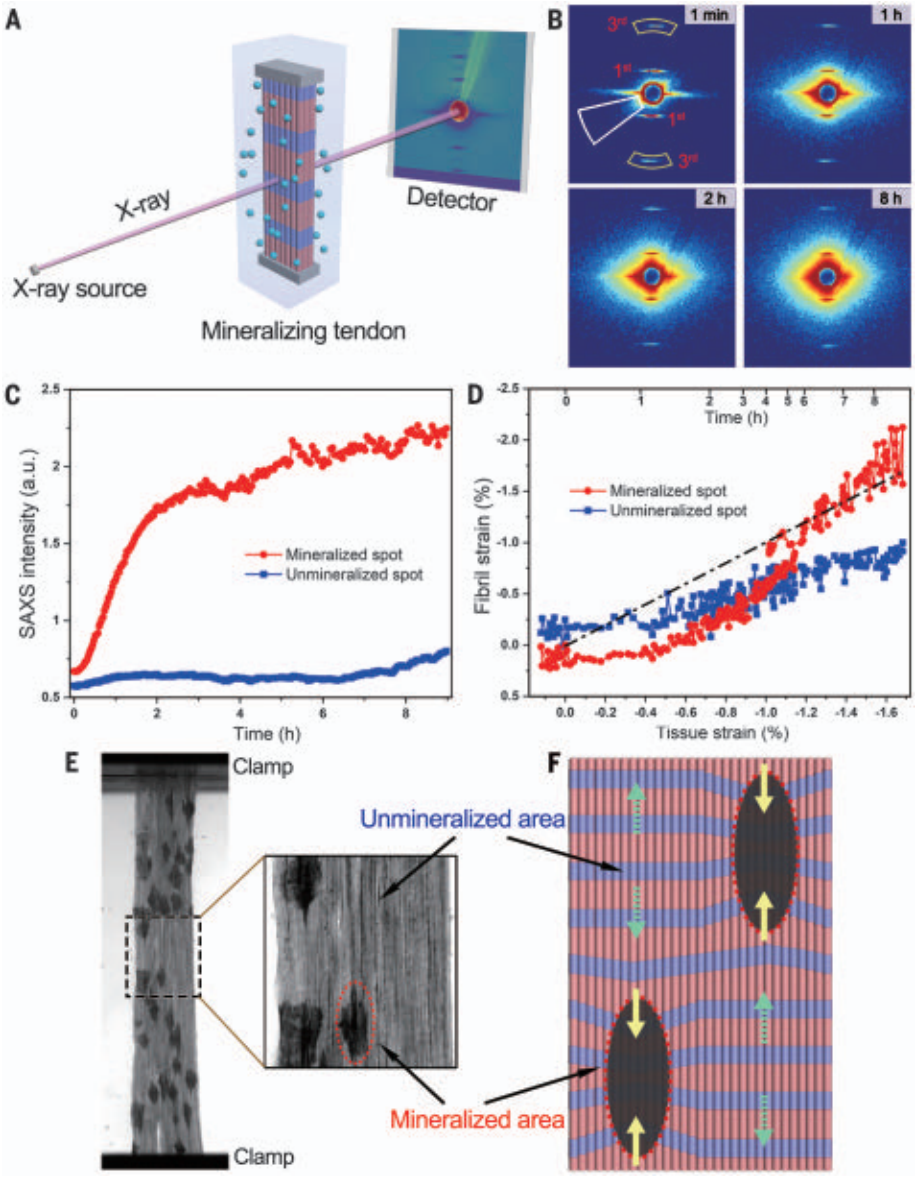


Fig. 3. In-operando synchrotron SAXS during tendon mineralization of $SrCO_3$ (force constant mode, zero stress). (**A**) Schematic of in-operando setup for synchrotron SAXS. (**B**) Two-dimensional SAXS patterns of mineralized tendon at different reaction times. First- and third-order reflections from collagen axial staggering are marked in the upper left pattern. Two peak positions of the third order (marked by yellow boxes) are acquired by Gauss fitting of scattering peaks after radial integration. (**C**) Integrated SAXS intensity of two regions of the tendon, one that was mineralized and another that remained unmineralized after 8 hours, as a function of time. Intensities were obtained by integrating within the sector area marked by a white line in (B). (**D**) Correlation between tissue and fibril strain in the mineralized and unmineralized spots (negative strains indicate contraction). For easier comparison with (C), time is also indicated on the upper horizontal axis. (**E**) Optical snapshot of a tendon after 4 hours of mineralization. (**F**) Schematic of the evolution of strains during tendon mineralization. In the high-magnification image and the schematic, the mineralized regions are marked by red dashed circles. In mineralized regions, a stress (yellow arrows) toward the central areas is generated by the contraction of the collagen matrix (decreasing *D* spacing). Owing to the inhomogeneity of mineralization, strains in the tendon are also inhomogeneous.

When a tendon slice was placed between force gauges, no stress developed in an aqueous solution or a salt solution (Fig. 1C). When transferred to the mineralizing solution, the development of a contractile stress appeared, gradually increasing to a maximum of 7.8 MPa within ~96 hours. Figure 2A shows the measured contractile stress as the result of intrafibrillar mineralization and the influence of pH value on the stress rate. The stress generation is correlated with the formation of (nanometer-sized) crystals inside the collagen fibrils.

In the presence of PAA, the $SrCO_3$ precursor infiltrated the collagen fibrils, as evidenced by a comparison of Fig. 2, B and C. In Fig. 2C, where mineral did not nucleate within fibrils, dehydration flattened them, whereas in Fig. 2B, the fibrils stayed nicely cylindrical even after dehydration because they were filled with $SrCO_3$ mineral (fig. S3F). The surface of collagen fibrils was also covered by $SrCO_3$ nanoparticles (Fig. 2B and fig. S3F), corresponding to an extrafibrillar mineral coating.

We found that the content of $SrCO_3$ crystals in mineralized tendon amounts to about 90 wt % after mineralization at final conditions (fig. S4). Microcomputed tomography (μCT) demonstrated that minerals were deposited throughout the whole body of the tendons (fig. S5A). Wide-angle x-ray scattering (WAXS) on mineralized tendons and the integration of the resulting two-dimensional (2D) patterns revealed a strong scattering peak at 17.9 nm^{-1}, corresponding to the (111) lattice planes of $SrCO_3$ crystals (fig. S5B). The orientation of minerals in tendons was evaluated by analyzing azimuthal profiles of the integrated small-angle x-ray scattering (SAXS) patterns (fig. S5C) (*20*) to determine the ρ parameter, which generally is used to characterize the degree of alignment of platelet-like minerals in mineralized tissues. The average ρ parameter of mineralized tendon is 0.45 ± 0.03 (n = 30), which is comparable to that of lamellar bone (*21*), indicating a relatively high degree of organization of $SrCO_3$ mineral particles along the long axis of collagen fibrils.

We describe three aspects that have a strong influence on the mineralization process. First, when using a precursor solution without PAA, no stress generation was observed. In this case no mineral phase was formed inside the fibrils (Fig. 2C and fig. S3A), which indicated that molecular interactions between collagen and mineral inside fibrils were a prerequisite for contraction. Small $SrCO_3$ particles were nucleated only at the surface of the tendon, as

expected for a mineralization process in supersaturated solution (*22*) (fig. S6).

Second, tendon samples were immersed into different solutions with the same total concentration of ions, to test whether the interaction with ions or the mineralization process was the origin of the contraction (fig. S7). No stress generation occurred if only either Sr^{2+} or CO_3^{2-} ions were present in the solution. Stress increased solely if both mineralizing ions (Sr^{2+} and CO_3^{2-}) were present and caused intrafibrillar mineralization. Therefore, the deposition of minerals inside collagen fibrils (intrafibrillar mineralization) plays a dominant role for stress generation. Third, by modifying the pH value of the mineralizing solution, the degree of mineralization could be controlled. A higher pH value of the solution led to a fast increase of stress in the initial stage of tendon mineralization, ranging from 0.03 MPa/hour at a pH of 8.5 to 0.095 MPa/hour at a pH of 9 (inset of Fig. 2A). That pH influences the rate of stress generation could also indicate a role of collagen charges in the process.

The generation of contractile stress during mineralization corresponds to a contraction of tendons along their longitudinal direction. To reveal structural changes of tendon tissue during the mineralization, we performed in-operando synchrotron SAXS measurements (Fig. 3A and fig. S8). Throughout these measurements, a constant force of 0.06 N was applied to the tendon during mineralization (fig. S9A), and the motor position was recorded to evaluate the tissue strain (fig. S9B). After water was replaced by a $SrCO_3$ solution, the free ions quickly diffused into the interfibrillar matrix, causing a slight expansion of the tendon. Subsequently, a fast drop of tissue strain to −0.44% could be observed at 1.1 hours, and a slow contraction to −1.7% after 9 hours of mineralization (fig. S9B).

The axial staggering of tropomolecules in collagen fibrils results in an alternation of stripes with high and low molecular density, which not only can be visualized by transmission electron microscopy (TEM) but also can be measured by SAXS (*23*, *24*). The SAXS patterns of original tendons in water exhibited a series of Bragg peaks (fig. S10A). The q-positions of the nth order (n = 1, 3, 5…) corresponds to $q_n = 2\pi n/D$, where D is the periodic spacing (~67 nm) within collagen fibrils according to the gap and overlap zones. The nanoscopic variation of the D spacing (inverse relation to q) is an ideal measure to evaluate the microscopic stress generation in tendons during mineralization. Because in comparison to the first-order scattering peak, the relatively strong third-order peak is more sensitive to changes of D spacing, it was employed to determine the peak positions by fitting with a Gaussian (fig. S10).

In-operando SAXS measurements were performed to evaluate both (i) the degree of mineralization at different regions across the tendon sample and (ii) changes in D spacing during the process. The latter measurement is related to the internal, "microscopic" strain of the collagen fibrils (fibril strain), whereas the mechanical setup (fig. S8) yields the "macroscopic" strain of the tissue (tissue strain), when the clamps are moved so as to keep the total stress in the tendon close to zero. If clamps are kept fixed instead, the setup can measure the macroscopic contractile stress that develops during mineralization.

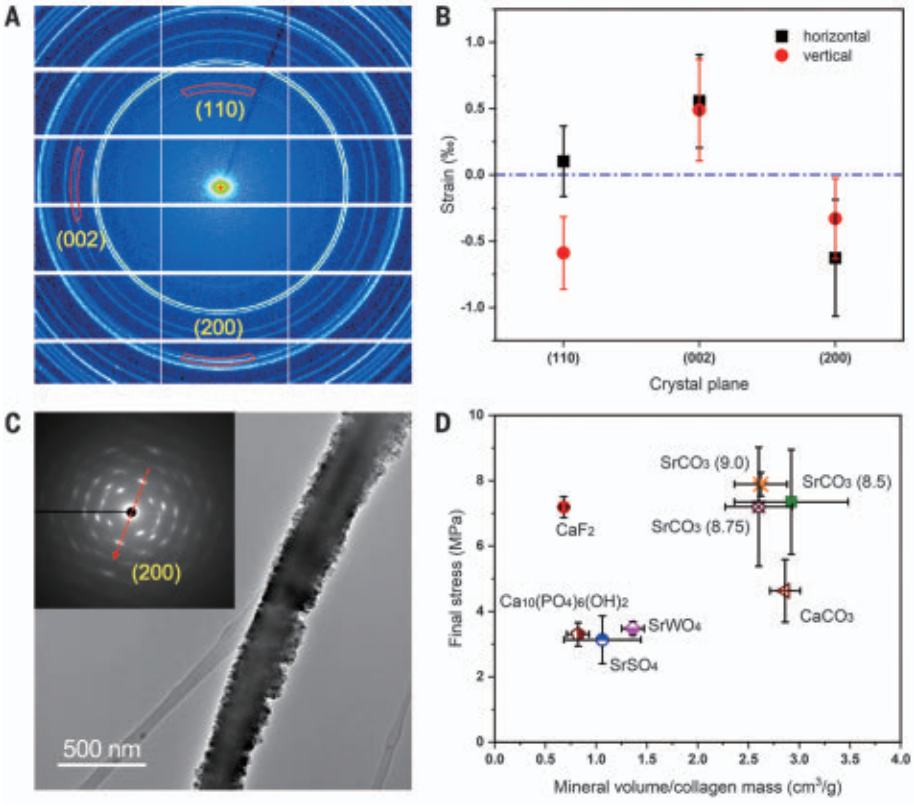


Fig. 4. Lattice strains of nanocrystals in collagen tissues. (**A**) Two-dimensional WAXS pattern of $SrCO_3$-mineralized tendon. The direction of the mineralized tendon is vertical. To detect peak shifts along and perpendicular to the collagen fibril direction, the integrated areas of (110), (002), and (200) rings are marked by red boxes. (**B**) Lattice strains as calculated from scattering patterns shown in (A) for the three indicated lattice directions. (**C**) TEM image of an isolated $SrCO_3$-mineralized collagen fibril. Inset shows the corresponding SAED pattern. The orientation of nanocrystals along the (200) direction is marked by a red arrow. (**D**) Generality of stress generation in collagen tissue. The final stress and mineral volume per collagen mass are shown for the following minerals: $SrCO_3$, $SrWO_4$, $SrSO_4$, CaF_2, $CaCO_3$, and $Ca_{10}(PO_4)_6(OH)_2$.

Figure 3B shows SAXS patterns where the intensity of the reflection in the trapezoidal box was monitored, starting from 1 min to 8 hours (see also fig. S11A). The integrated SAXS intensity increased with the mineral content in some areas but does not distinguish between crystalline or amorphous. In other regions no mineralization occurred, as evidenced by the SAXS intensity that remained low (Fig. 3C). Clear differences between mineralized and unmineralized regions could be observed (fig. S11B). The time evolution of the D spacing in collagen fibrils and the calculated fibril strain during mineralization is summarized in fig. S11, C to D. Both mineralized and unmineralized regions exhibited a decrease of the D spacing. However, mineralized regions showed a faster decrease of fibril strain than the unmineralized ones (see also fig. S12).

This fibril strain can be correlated with the macroscopic tissue strain measured through the movement of the clamps (Fig. 3D), which is a weighted average of all local strains. The black dashed line indicates a 1:1 correlation of fibril and tissue strain. Deviations from this value are due to the local inhomogeneities in mineralization and in fibril strain. In the region that mineralized (red symbols in Fig. 3, C and D), the SAXS intensity (which monitors the amount of mineral) increased sharply in the first 2 hours, with only a slight increase in local contractile fibril strain. After this, mineral content continued to increase more slowly, with an associated strong contraction of the fibril (up to nearly 2% contraction). In the region that did not mineralize (blue symbols),

there was only minimal contraction. Although there is a correlation between local strains and amount of intrafibrillar mineral, it cannot be perfect because strains induced by mineralization obviously extend through elastic interactions over larger areas. This is illustrated schematically in Fig. 3E. Finally, the macroscopic contraction does not increase linearly with time. It is faster in the first hour and then changes approximately linearly (fig. S9B).

The origin of this chemomechanical effect is not obvious, but it is most likely due to the replacement of water by mineral during the mineralization process. This is probably similar to bone mineralization (i.e., when the mineral is hydroxyapatite), where it has been proposed that water is pushed from intrafibrillar to extrafibrillar compartments and replaced by precursors under the effect of the Gibbs-Donnan equilibrium (*25*) or capillary transport (*7*). Dehydration experiments have shown that the removal of water shortens the collagen molecules by changing their conformation (*16*). This suggests that the contraction due to dehydration and mineralization has the same origin. Indeed, mineralization causes dehydration and water replacement, thereby changing the osmotic equilibrium that leads to a shrinkage of the triple-helical pitch in certain regions of the collagen molecules (*16*).

To find out whether the contraction of collagen also leads to a compression of the $SrCO_3$ mineral particles embedded in the fibrils, we used synchrotron wide-angle x-ray scattering to extract 2D WAXS patterns (Fig. 4A). The crystallography of the orthorhombic unit cell is presented in fig. S13. To determine a potential compression or dilatation of the lattice, we compared the position of (110), (002), and (200) rings with the peak positions from a reference sample (table S1), which was obtained by a heat treatment of a $SrCO_3$-mineralized tendon to induce a relaxation of any strains in the crystal lattice by thermal degradation of the organic matrix. The morphology and the crystal orientation in respect to the orientation of the collagen fibrils were also analyzed at the end of the in-operando measurement using TEM. A typical example of a mineralized collagen fibril is shown in Fig. 4C. The TEM images revealed that the mineral matrix consisted of small $SrCO_3$ nanocrystals, which were well coaligned, forming a nearly single-crystalline matrix. This was demonstrated especially by selected area electron diffraction (SAED). As the dominant crystallographic orientation of nanocrystals parallel to the collagen fibrils, a <100> direction was determined (Fig. 4C, inset).

In mineralized tendons, WAXS measurements showed a pronounced compression of crystals along the <200> direction, but an elongation in the perpendicular <002> direction (Fig. 4B). A compressive strain in the <200> lattice direction was measured to be −0.033% and −0.063% along vertical and horizontal directions, respectively. In the perpendicular <002> direction, an expansion was measured to be 0.049 and 0.056% along vertical and horizontal directions, respectively. The orientation of the collagen molecules was along the vertical direction. Taking 62 GPa as a bulk modulus of $SrCO_3$ (*26*), the prestress on nanocrystals in collagen fibrils was estimated to be between 20 and 40 MPa. This amounts to the same magnitude as compressive load on hydroxyapatite minerals in bone (*27*).

We also studied the mineralization with $Ca_{10}(PO_4)_6(OH)_2$, CaF_2, $CaCO_3$, $SrWO_4$, and $SrSO_4$. In all these cases, mineralization occurred (fig. S14A and S15), and a corresponding contractile stress was measured in the collagen fibers (Fig. 4D and fig. S14B). The stress values at the end of the mineralization process, as shown in fig. S14B, have been plotted as "final stress" in Fig. 4D. In some cases ($SrSO_4$, $SrWO_4$), the final stress corresponds to a plateau value; in other cases, it continued to increase up to 120 hours, when the experiments were terminated. The kinetics of stress generation varied for the different minerals (fig. S14). For the $SrCO_3$ system, the final stress was determined by the final mineral content, independently of the pH of the solution, which regulated the stress rate in the initial stage. There seems to be a linear relation between final stress and mineral volume fraction in strontium-based inorganic species (Fig. 4D). Although $Ca_{10}(PO_4)_6(OH)_2$ is close to this hypothetical line, CaF_2 and $CaCO_3$ are far away from it. This indicates that the precipitation of different minerals leads to contraction of different degree. Finally, it is interesting to note that the apatite content in our artificially mineralized tendon is about 72 wt % (fig. S4), which is of the same order of magnitude as compact bone, which has ~65% mineral and 25% collagen by weight, with the rest being water (*28*). With other minerals, the inorganic content can even be somewhat higher, up to 88 wt % (fig. S4). However, we cannot exclude some residual mineral on the surface of all specimens in the TGA measurements, so that the mineral content may generally be overestimated.

This work demonstrates that chemomechanical coupling between precipitation and collagen contraction that was previously observed for hydroxyapatite in bone occurs for a wide range of minerals. Furthermore, our investigations also reveal that the stress of the collagen fibrils is transferred to the embedded mineral. As a result, its crystal lattice is strongly compressed parallel to the fibrils in the range of 20 to 40 MPa. This phenomenon not only reveals an intriguing property of collagen; it also provides an exciting concept for enhancing the mechanical properties of hybrid materials through internal stresses similar to concrete that is prestressed by steel fibers.

REFERENCES AND NOTES

1. M. Eder, S. Amini, P. Fratzl, *Science* **362**, 543–547 (2018).
2. J. J. Xie *et al*., *Prog. Mater. Sci.* **105**, 100571 (2019).
3. L. B. Mao *et al*., *Science* **354**, 107–110 (2016).
4. E. Munch *et al*., *Science* **322**, 1516–1520 (2008).
5. S. Weiner, H. D. Wagner, *Annu. Rev. Mater. Sci.* **28**, 271–298 (1998).
6. P. Fratzl, H. S. Gupta, E. P. Paschalis, P. Roschger, *J. Mater. Chem.* **14**, 2115–2123 (2004).
7. M. J. Olszta *et al*., *Mater. Sci. Eng. Rep.* **58**, 77–116 (2007).
8. F. Nudelman *et al*., *Nat. Mater.* **9**, 1004–1009 (2010).
9. Y. Wang *et al*., *Nat. Mater.* **11**, 724–733 (2012).
10. M. J. Olszta, E. P. Douglas, L. B. Gower, *Calcif. Tissue Int.* **72**, 583–591 (2003).
11. L.-N. Niu *et al*., *Angew. Chem. Int. Ed.* **50**, 11688–11691 (2011).
12. B. M. Oosterlaken *et al*., *ACS Biomater. Sci. Eng.* **7**, 3123–3131 (2021).
13. E. Seknazi, B. Pokroy, *Adv. Mater.* **30**, e1707263 (2018).
14. I. Polishchuk *et al*., *Science* **358**, 1294–1298 (2017).
15. J.-B. Forien *et al*., *Nano Lett.* **15**, 3729–3734 (2015).
16. A. Masic *et al*., *Nat. Commun.* **6**, 5942 (2015).
17. L. B. Gower, D. J. Odom, *J. Cryst. Growth* **210**, 719–734 (2000).
18. J. J. M. Lenders *et al*., *J. Am. Chem. Soc.* **134**, 1367–1373 (2012).
19. C. C. Lin, L. G. Liu, *J. Phys. Chem. Solids* **58**, 977–987 (1997).
20. G. Benecke *et al*., *J. Appl. Crystallogr.* **47**, 1797–1803 (2014).
21. T. Tang *et al*., *Acta Biomater.* **89**, 330–342 (2019).
22. H. Colfen, M. Antonietti, *Mesocrystals and Nonclassical Crystallization* (Wiley, 2008).
23. H. S. Gupta *et al*., *Phys. Rev. Lett.* **93**, 158101–158104 (2004).
24. H. S. Gupta *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **103**, 17741–17746 (2006).
25. L.-N. Niu *et al*., *Nat. Mater.* **16**, 370–378 (2017).
26. M. Wang *et al*., *Phys. Chem. Miner.* **42**, 517–527 (2015).
27. L. Bertinetti *et al*., *J. Mech. Behav. Biomed. Mater.* **52**, 14–21 (2015).
28. J. D. Currey, *Bones: Structure and Mechanics* (Princeton Univ. Press, 2006).
29. H. Ping *et al*., Contractile stresses in collagen fibrils, Edmond, Version 1 (2022); https://doi.org/10.17617/3.94.

ACKNOWLEDGMENTS

We thank S. Amini for help with the heat treatment of samples, P. Leibner for support with mechanical testing, D. Werner for assistance with micro-CT measurements, C. Schmitt for help during Raman mapping, and W. Fang for help with intrafibrillar mineralization of various inorganic materials. **Funding:** This work was financially supported by the National key Research and Development Program of China (2021YFA0715700), by the Max Planck Society, and by the National Natural Science Foundation of China (51832003, 51902236). P.F. also acknowledges support by the German Research Foundation within SFB1444 and through the Cluster of Excellence "Matters of Activity. Image Space Material" EXC 2025. **Author contributions:** H.P., W.W., and P.F. conceived the main idea of the project; Z.F. codirected the project. H.P., Z.F., and W.W. designed and performed the experiments; N.H. participated in the construction of the mechanical testing setup; and E.S. and C.L. participated in the collection and analysis of data. H.P., W.W., and P.F. analyzed the experimental results, prepared the figures, and wrote the manuscript. All authors discussed the results and commented on the manuscript. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All data are available in the main text or the supplementary materials, as well as at Edmond–Open Research Data Repository of the Max Planck Society (*29*).

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abm2664
Materials and Methods
Supplementary Text
Figs. S1 to S15
Table S1
References (*30*–*32*)
Movies S1 and S2

5 September 2021; accepted 17 February 2022
10.1126/science.abm2664

GRAPHENE

Orderly disorder in magic-angle twisted trilayer graphene

Simon Turkel[1], Joshua Swann[1], Ziyan Zhu[2], Maine Christos[2], K. Watanabe[3], T. Taniguchi[4], Subir Sachdev[2,5], Mathias S. Scheurer[6], Efthimios Kaxiras[2,7], Cory R. Dean[1], Abhay N. Pasupathy[1,8]*

Magic-angle twisted trilayer graphene (TTG) has recently emerged as a platform to engineer strongly correlated flat bands. We reveal the normal-state structural and electronic properties of TTG using low-temperature scanning tunneling microscopy at twist angles for which superconductivity has been observed. Real trilayer samples undergo a strong reconstruction of the moiré lattice, which locks layers into near–magic-angle, mirror symmetric domains comparable in size with the superconducting coherence length. This relaxation introduces an array of localized twist-angle faults, termed twistons and moiré solitons, whose electronic structure deviates strongly from the background regions, leading to a doping-dependent, spatially granular electronic landscape. The Fermi-level density of states is maximally uniform at dopings for which superconductivity has been observed in transport measurements.

The prediction of a magic angle in twisted trilayer graphene (TTG) (*1*, *2*) was soon followed by the observations of superconductivity and field-dependent quantum interference (*3*–*5*). This set of properties makes TTG the only moiré heterostructure outside of magic-angle twisted bilayer graphene (MATBG) to exhibit signatures of both a superconducting transition and macroscopic quantum phase coherence. Because TTG and MATBG share the distinctive attribute of twofold rotational symmetry C_{2z}, it has been proposed that this symmetry is essential to establishing superconductivity in twisted graphenes (*4*, *6*). Superconductivity in TTG appears to be even more robust than in MATBG, with critical temperature (T_c) reaching up to 2.9 K in the first generation of devices. This has led to speculation that magic-angle TTG is structurally more stable than MATBG, locking experimental devices into a mirror symmetric configuration that possesses the crucial C_{2z} symmetry. Theoretical works have proposed several exotic orders for the mirror symmetric configuration, including spontaneous flavor-symmetry breaking, nematic superconductivity, and spin triplet pairing (*7*–*9*). To date, however, there is little experimental information about the atomic or electronic structure of this material; there remains no direct experimental confirmation of even the most basic hypothesis that superconducting devices possess the mirror symmetric stacking on which theoretical predictions are based.

TTG is formed by consecutively stacking three layers of graphene so that the bottom layer (B) is rotated at an angle θ_{BM} relative to the middle layer (M) and the top layer (T) is rotated at an angle θ_{TM} relative to the middle layer; both outer layers are rotated in the same direction relative to the middle layer (Fig. 1A, inset). Each rotation θ_{ij} gives rise to a periodic density modulation, or moiré pattern, at wavelength $\lambda_{ij} \sim a/\theta_{ij}$, where a = 0.246 nm is the graphene lattice constant (*10*–*12*). For the special case of mirror symmetric stacking, $\theta_{BM} = \theta_{TM} = \theta$ (T and B are aligned, and M is twisted relative to these by an angle θ), TTG is predicted to host two sets of flat bands whose band velocity vanishes at a magic angle of $\theta \sim 1.56°$ (*1*, *2*). As in MATBG, the quenched kinetic energy of charge carriers in these bands is expected to favor the formation of strongly correlated states of matter. Recent transport measurements have confirmed the importance of electronic correlations in TTG with the observation of superconductivity by two groups with similar phenomenology (*3*, *4*).

Several obstacles can stand in the way of achieving perfect mirror symmetry. Despite state-of-the-art fabrication techniques, the highest-quality TTG heterostructures will inevitably have a small mismatch between θ_{TM} and θ_{BM} over macroscopic length scales, as was the case in at least one superconducting device (*4*). In the limit of perfectly rigid graphene layers (neglecting lattice relaxation), such a misalignment will produce a beating pattern between the top-middle (TM) and bottom-middle (BM) moirés at a "moiré of moiré" wavelength $\Lambda \sim a/\delta_\theta$, where $\delta_\theta = |\theta_{TM} - \theta_{BM}|$ (Fig. 1A). In regions where the two moirés are in phase, TM AA sites sit atop BM AA sites, resulting in a locally mirror symmetric AtA ("A-twist-A") trilayer configuration composed of AAA, ABA, and BAB stacking sites. Where the two moirés are out of phase, by contrast, the AA sites of one bilayer align with the AB sites of the other, generating a local AtB configuration (*13*, *14*), comprising ABB, AAB, and BAC stacking sites; AtB is related to the AtA configuration through translation of the top layer (Fig. 1, B and C). The emergent structures of the trilayer moirés in these two regions are distinguished by their different symmetry classes, as visualized by their predicted topographic profiles in Fig. 1B. We estimate the out-of-plane corrugations for AtA and AtB domains as a superposition of sinusoidal functions of local bilayer stackings, with maxima on AA and minima on AB sites (*15*); we found that whereas the AtB regions host a honeycomb moiré lattice, the moiré pattern in the AtA domains is expected to be hexagonal.

In this work, we used the atomic-scale imaging capabilities of ultrahigh-vacuum scanning tunneling microscopy and spectroscopy (STM/S) at temperatures from 4.8 to 7.2 K to directly characterize the electronic structure of magic-angle TTG. Our devices were fabricated by using the "cut and stack" technique, and electrical contact was made with a preplaced graphite finger to which Field's metal μ-solder was subsequently affixed (fig. S1). STM topography of a TTG sample is shown in Fig. 1D, in which two distinct moiré wavelengths, $\lambda \sim 9$ nm and $\Lambda \sim 70$ nm, are clearly visible, corresponding to the bilayer moiré and moiré of moiré length scales, respectively. The corresponding angle mismatch $\delta_\theta \sim a/\Lambda$ for this region is ~0.2°, which is nearly identical to the mismatch of ~0.3° measured in a superconducting TTG device (*4*). At such small δ_θ, the moiré of moiré is not expected to give rise to strong direct signatures in transport, rendering microscopic probes such as STM one of the few ways of detecting it.

The presence of two moiré patterns is a generic feature over large areas of our sample (figs. S2 and S3) and represents a deviation from the three moirés (λ_{TM}, λ_{BM}, and Λ) that are expected on the basis of a simple rigid model (fig. S6). The STM signal in constant current mode was dominated by structural height variations across the sample surface (figs. S4 and S5), so that we could identify the global stacking configuration as AtA by the smaller moiré lattice in Fig. 1D being hexagonal rather than honeycomb at each point in space. The bright spots in topography therefore correspond to regions of local AAA stacking and were surrounded by alternating ABA and BAB domains, which we confirmed with line-cut spectroscopy (fig. S7).

The absence of AtB domains in a sample with nonzero angle mismatch δ_θ implies that TTG undergoes a reconstruction on the scale

[1]Department of Physics, Columbia University, New York, NY 10027, USA. [2]Department of Physics, Harvard University, Cambridge, MA 02138, USA. [3]Research Center for Functional Materials, National Institute for Materials Science, 1-1 Namiki, Tsukuba 305-0044, Japan. [4]International Center for Materials Nanoarchitectonics, National Institute for Materials Science, 1-1 Namiki, Tsukuba 305-0044, Japan. [5]School of Natural Sciences, Institute for Advanced Study, Princeton, NJ 08540, USA. [6]Institute for Theoretical Physics, University of Innsbruck, A-6020 Innsbruck, Austria. [7]John A. Paulson School of Engineering and Applied Sciences, Harvard University, Cambridge, MA 02138, USA. [8]Condensed Matter Physics and Materials Science Department, Brookhaven National Laboratory, Upton, NY 11973, USA.
*Corresponding author. Email: apn2108@columbia.edu

of the moiré lattice that favors the lower-energy (*16*) AtA configuration. Close examination of Fig. 1D reveals that this moiré lattice reconstruction (MLR) produces a periodic warping of the AAA site positions to enforce AtA stacking over the entire sample area. The observed warping of the moiré lattice can be understood at the atomic scale as arising from variations in the local twist angle (θ_x) and strain (ε_x) of the individual graphene layers. $\theta_x \sim a/\sqrt{A_x}$ is plotted in Fig. 1, E and F, for two nearby sample regions, where A_x is the area of the moiré unit cell centered on position x. For small-angle mismatch δ_θ, the system segregates into highly uniform triangular domains (Fig. 1, E and F, blue areas) bounded by sharp point-like irregularities in the local twist angle (Fig. 1, E and F, red areas).

For $\Lambda \gtrsim 30$nm, the average twist angle internal to each triangular domain (θ_I) saturates to a common value of ~1.5° that is independent of the moiré of moiré wavelength (Fig. 1G). This implies that the MLR not only enforces AtA stacking but also tends to lock the lattice to a constant local twist angle, even as θ_{TM} or θ_{BM} is varied. To shed light on this behavior, we performed structural relaxation calculations for TTG at a range of interlayer twists. The results (fig. S10) indicate that for $\delta_\theta \lesssim 0.5°$, θ_I locks to the smaller of θ_{TM} and θ_{BM} because a stronger interlayer coupling exists at a lower twist angle interface. The additional twist angle degree of freedom therefore enables TTG to locally conform to the mirror symmetric magic angle structure while "absorbing" twist angle inhomogeneity at the larger moiré of moiré length scale. The effect of the MLR on the local electronic structure is profound, as evidenced by the Fermi-level local density of states (LDOS) map (Fig. 1D, inset), which shows large modulations in the tunneling conductivity across regions of the MLR. It was therefore necessary, in considering the potentialities of TTG as a platform for correlated phases, to analyze the electronic structure on both the sub- and supra-Λ length scales.

In Fig. 2A, we present STM topography of a 250-nm^2 area, which is part of an even larger region with only a single-moiré wavelength corresponding to a twist angle of $\theta = 1.55°$. The extreme degree of homogeneity in this area is conveyed by the local twist angle histogram (Fig. 2A, inset), showing a standard deviation of 0.03° over the entire field of view. This indicates a twist angle mismatch of $\delta_\theta < 0.05°$, which provided us with the opportunity to study a single domain of the MLR as well as to investigate the spectroscopic properties of a large patch of magic-angle TTG that approaches the size of a transport device.

The high energy resolution of STS permitted us to directly probe the structure of the flat bands. A series of STS measurements acquired at 7.2 K on a single AAA site is shown in Fig. 2B for a range of voltages (V_g) applied to the graphite back gate (additional twist angles are

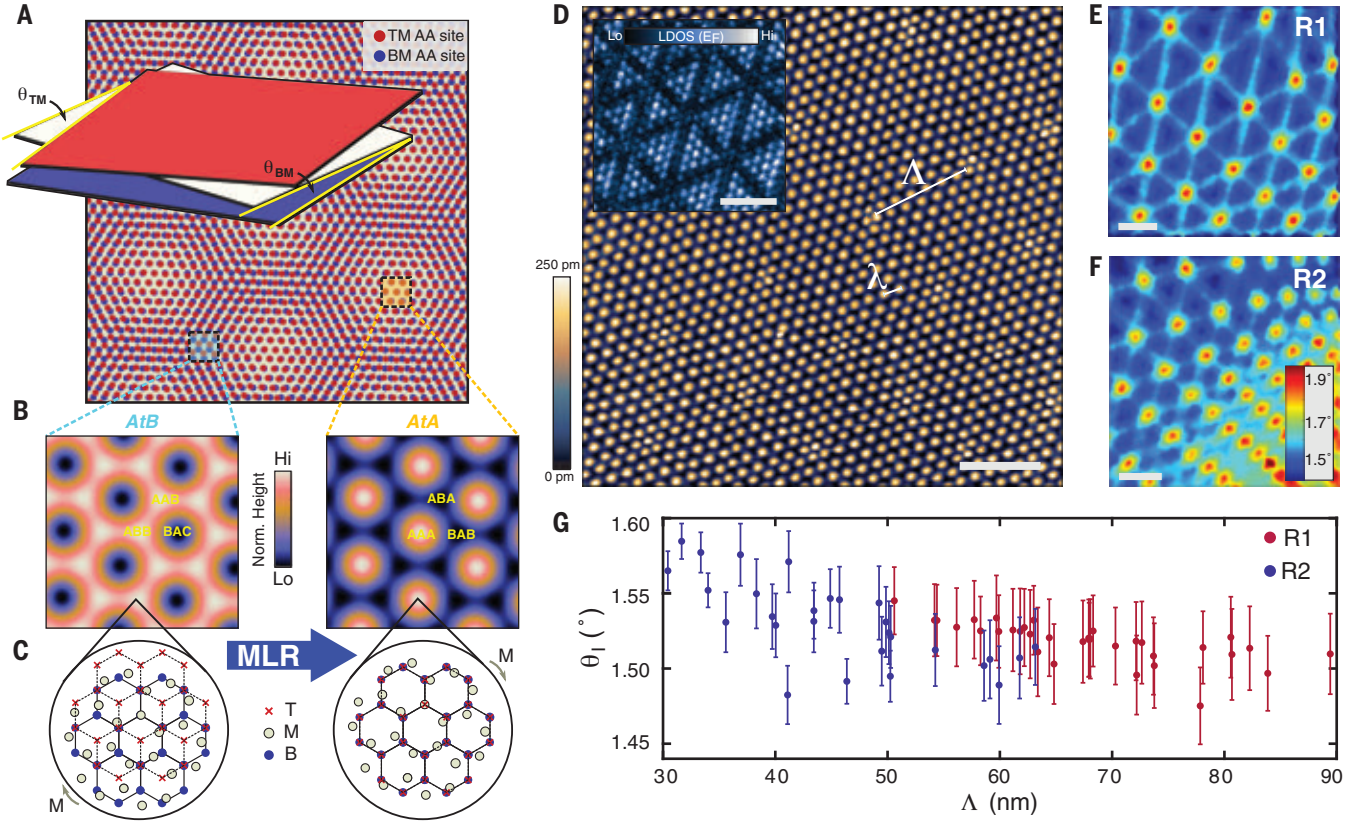


Fig. 1. STM on three twisted graphene layers. (**A**) Illustration of the moiré of moiré pattern in TTG for $\theta_{TM} \neq \theta_{BM}$ in the absence of lattice relaxation. Local AtA and AtB domains are formed, creating two characteristic length scales. (Inset) Illustration of the two independent twist angles expected in a general three-layer stack. (**B**) Normalized out-of-plane corrugation calculated (*18*) for AtB and AtA stacking configurations, showing the local domain structure of each configuration. (**C**) Schematic of the atomic stacking structure of AtA and AtB TTG. The two configurations are related by translation of the top layer. In real devices, a MLR makes it energetically favorable for AtB domains to warp into AtA. (**D**) STM topography of TTG at an average twist of 1.56°. (Inset) Charge-neutral local density of states map acquired at the Fermi level, showing electronic inhomogeneity caused by the MLR. Set voltage (V_{set}) = 300 mV, set current (I_{set}) = 120 pA, and modulation voltage (V_{mod}) = 2 mV. (**E** and **F**) Local twist-angle maps over the region shown in (D) (R1) and a nearby sample area (R2). The local twist angle is extracted from the cell areas of the Voronoi tessellation generated by the AAA site positions. (**G**) Plot of the internal twist angle (θ_I) within a MLR domain as a function of domain size (Λ) for regions R1 (E) and R2 (F). Error bars represent 1 SD of the local twist angle within a given domain. Scale bars, 50 nm.

shown in fig. S12). The measured spectrum did not change appreciably upon cooling to 4.8 K (fig. S13). At the charge neutrality point (CNP), the spectrum was dominated by a pair of overlapping resonances that arose from the partially overlapped conduction (CB) and valence (VB) flat bands. Additional soft humps at higher energy (Fig. 2B, black arrows) correspond to the edges of the next available (remote) bands. Each flat band was expected to host a saddle point in its momentum space structure, giving rise to a sharp peak, or van Hove singularity (VHS), in the density of states. We extracted the energy positions and widths of these VHSs by fitting our spectra with the sum of two Lorentzian curves and found that at CNP, the CB and VB VHSs are separated by ~18 meV and have an average width [full width at half maximum (FWHM)] of ~23 meV.

Varying V_g systematically alters the shape of the quasiparticle spectrum, changing the intensities, separations, and widths of the flat-band VHSs. In particular, we found a transfer of spectral weight between the two VHSs upon reversing the sign of V_g (Fig. 2B). Moreover, the width of each flat band was reduced when doped to the Fermi level, saturating to a minimum width of ~15 meV at $\nu \sim \pm 2$ (Fig. 2C). Last, the VHS separation is an increasing function of doping away from CNP with a distinct asymmetry between filling of electrons and holes (Fig. 2D). In general, such gate-dependent spectral shifts can be attributed either to the single-particle effect of the displacement field ($D = V_g/2d$) on the material's band structure or to variations in the quasiparticle interaction strength as a function of band filling ($\nu = 4n/n_s$), where d is the dielectric thickness, n is the induced carrier density, and n_s is the carrier density at full filling of a fourfold degenerate moiré band.

We examined the role of interactions in determining the band structure of TTG by comparing the experimental spectrum with continuum model (*1*, *2*, *17*) calculations for a uniform mirror symmetric AtA stacking configuration at $\theta = 1.55°$. In Fig. 2E and fig. S14A, we compare the measured VHS separation and widths at CNP with those predicted with three separate calculations. Using inter- and intralayer tunneling parameters (*18*) derived from ab initio computations (*13*) severely underestimates both the separation and widths of the VHSs (SP1). Enhancing the monolayer graphene Fermi velocity by ~30% (SP2) enables us to reproduce the VHS separation but predicts widths that are still a factor of ~6 smaller than those found in experiment. Because a doping-dependent calculation that includes electron interactions (*7*) is beyond the scope of this primarily experimental work, we restricted the theoretical analysis of interactions to the CNP. Apart from spontaneous symmetry breaking, interactions can have two effects on the quasiparticle spectrum. Coulomb repulsion between electrons can change the energy landscape for a quasiparticle moving through the heterostructure, leading to a renormalization of the band structure. In addition, inelastic scattering events can lead to a finite lifetime for quasiparticle excitations, which, because of the quantum uncertainty between energy and time, causes the excitation spectrum to be broadened. Using the self-consistent Hartree-Fock procedure of (*7*, *18*), we found that similar to the situation in MATBG (*19*, *20*), the interaction-induced band renormalization without additional symmetry breaking accurately accounts for the separation between the peaks and roughly 70% of their widths. An additional lifetime

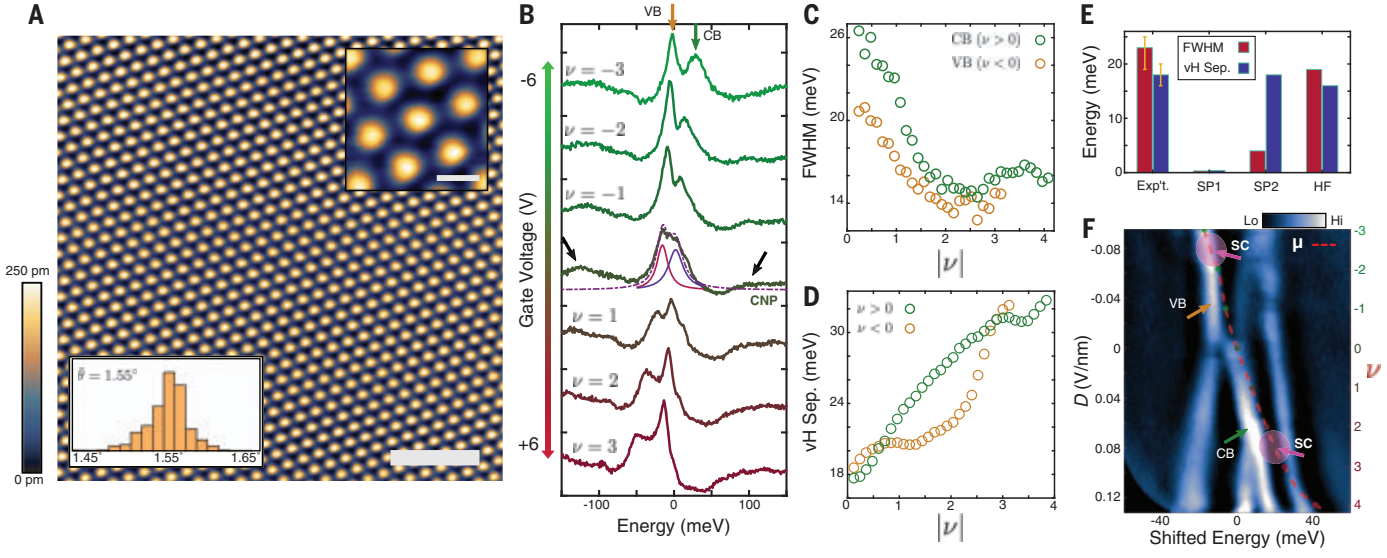


Fig. 2. Spectroscopy on a uniform 1.55° region. (**A**) STM topography of a uniform area presenting a single moiré wavelength corresponding to a twist angle of 1.55°. Scale bar, 50 nm. (Top right inset) Zoomed-in topography of a single moiré unit cell showing bright AAA sites surrounded by alternating ABA and BAB domains. Scale bar, 8 nm. (Bottom left inset) Histogram of local twist angle values extracted for each moiré unit cell. Local twist angle values are as in Fig. 1, E and F. (**B**) AAA site STS spectra showing the evolution of the flat band structure at 1.55° twist as a function of applied gate voltage. Each curve represents the average of 10 measurements performed on a single AAA site from the region in (A). Gold and green arrows indicate the valence and conduction flat bands, respectively. Black arrows indicate the edges of the remote bands. Charge-neutral spectrum shows Lorentzian fits to the valence and conduction bands. Curves are offset vertically for clarity and are plotted on the same vertical scale. V_{set} = 300 mV, I_{set} = 150 pA, and V_{mod} = 1 mV. (**C**) FWHM of the conduction and valence band VHSs from (B) as a function of respective band filling. Each band grows flatter as it is doped to the Fermi level. (**D**) Separation between conduction and valence band peaks from (B) as a function of doping. (**E**) Comparison of VHS separation and widths at charge neutrality between experiment and three continuum model calculations. SP1 and SP2 are single-particle calculations with different inter- and intralayer hopping parameters, and HF includes electronic interactions through Hartree-Fock corrections to the continuum model, resulting in a band renormalization and lifetime broadening (*18*). Only the interacting calculation reproduces the experimental spectrum. (**F**) High-resolution AAA site spectra shifted as described in the text and with a smooth background subtracted to emphasize the evolution of the flat bands with doping. Red dashed line indicates the position of the chemical potential for which a given spectrum was acquired. Pink arrows indicate optimal doping for superconductivity. V_{set} = 200 mV, I_{set} = 200 pA, and V_{mod} = 0.5 mV.

broadening of 4 meV is sufficient to reproduce the widths quantitatively (Fig. 2E and fig. S14A).

Tuning V_g away from zero produces systematic changes in the intensities, separation, and widths of the VHSs. We examined the effect of D on the single-particle band structure and found that although it does account for the shift in relative intensity of the VHSs, values of D within the experimentally accessible range fail to produce notable changes in either the predicted separation or widths of the VHSs (fig. S14B). The inability of single-particle calculations to reproduce the measured quasiparticle spectrum, combined with the strong doping dependence of the latter, provides clear evidence for a pronounced band renormalization in TTG near the magic angle caused by strong quasiparticle interactions.

To isolate and better visualize this reconfiguration of the band structure, we plotted in Fig. 2F gate-dependent STS, with each curve shifted so that the flat bands remain centered on zero energy (*18*). The position of the chemical potential at each doping is indicated in Fig. 2F with the red dashed line. Unlike MATBG, in which superconductivity occurs in the vicinity of multiple integer filling factors of the moiré bands (*21*), observations of superconductivity in TTG have been strictly limited to within the vicinity of $|\nu| = 2$, with optimal doping occurring in a roughly particle-hole symmetric fashion for $2 < |\nu| < 3$ (*3*, *4*). In MATBG, superconductivity occurs when the chemical potential is embedded in the moiré flat bands, leading to a large density of states at the Fermi level. The enhancement of the density of states relative to pristine graphene or graphite has been hypothesized to support conventional electron-phonon–mediated superconductivity (*22*, *23*). In TTG, however, our measurements (Fig. 2F) show that large densities of states occur at multiple fillings between $\nu = -4$ and $\nu = 4$, even though superconductivity has not been observed in all of these regions in transport measurements. It is thus clear that it is not the density of states alone that controls the superconducting dome observed in transport.

One aspect of the LDOS spectrum studied in Fig. 2 is that it breaks particle-hole symmetry in a way that is not expected on the basis of noninteracting calculations. This is apparent in that the CB remains considerably broader than the VB for all measured dopings (Fig. 2C). One implication of this particle-hole asymmetry is that the chemical potential crosses the VHSs at different filling factors for electron as compared with hole doping, which is reproduced by our Hartree-Fock calculations (fig. S16). For hole doping, the chemical potential crosses the VHS in the vicinity of the parent state at $\nu \sim -2.5$, whereas for electron doping, the chemical potential has already crossed the VHS by $\nu \sim 1$. The enhancement of the Fermi level density of states for the hole-doped parent state may contribute to the comparative robustness of the hole-doped superconducting dome (*3*, *4*). However, the particle-hole asymmetry of the tunneling spectrum stands in contrast to the approximate particle-hole symmetry of the superconducting phase diagram measured in transport. This apparent discrepancy suggests that additional factors may be relevant in determining the boundaries of the superconducting phase.

Having analyzed the electronic structure at the sub-Λ length scale, we next turned to a detailed study of the MLR at twist angles near those for which robust superconductivity has

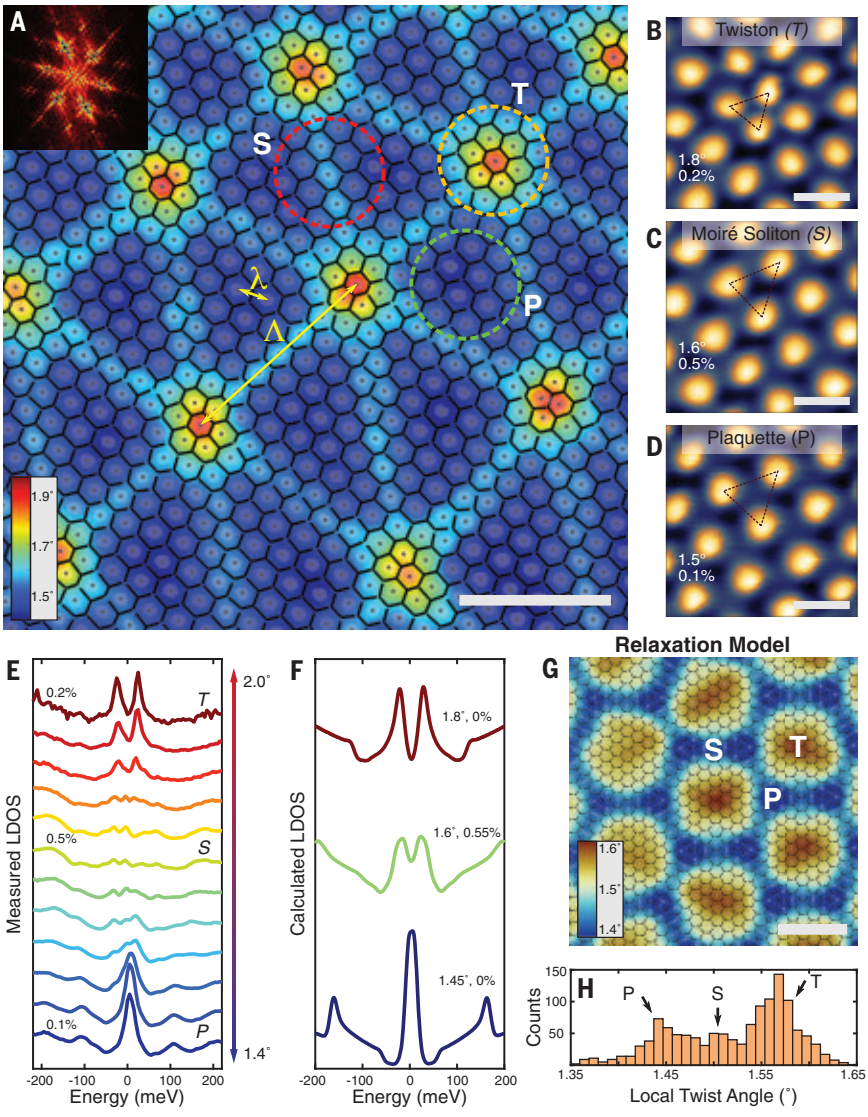


Fig. 3. Moiré lattice reconstruction. (**A**) STM topography colored in proportion to the local twist angle. Scale bar, 50 nm. (Inset) FFT of 320-nm^2 topograph centered on this field of view, showing two sets of moiré wave vectors. (**B** to **D**) Zoomed-in topography of the circled regions in (A), illustrating the local structure of the MLR. Numbers indicate local twist-angle and heterostrain values extracted from dashed moiré lattice vectors. Scale bars, 10 nm. (**E**) Experimental AAA site LDOS spectra extracted from conductance maps taken over the field of view of (A), displaying the change in electronic structure over different regions of the MLR. Curves are offset vertically for clarity and are plotted on the same vertical scale. Percentages denote characteristic heterostrain values for each MLR region. (**F**) Continuum model (SP2) TTG densities of states for three sets of structural parameters (θ, ε). Calculations at finite heterostrain preserve mirror symmetry by applying a uniaxial strain to the middle layer only. (**G**) Local twist angle as determined with nearest-neighbor AAA site distance for structural relaxation calculation with $\theta_{TM} = 1.5°$ and $\theta_{BM} = 1.69°$. Scale bar, 50 nm. (**H**) Histogram of the twist angles present in (G) showing three populations corresponding to plaquette, soliton, and twiston sites.

been observed (*3*, *4*). Large area topography of a sample region with angle mismatch $\delta_\theta \sim 0.25°$ is shown in Fig. 3A and fig. S8A. Overlaid on the topography of Fig. 3A is a map of the local twist angle, giving a spatially averaged value of $\bar{\theta} = 1.55°$. The MLR segregates the system into domains of uniform moiré plaquettes arranged in a honeycomb lattice with $\theta_x \sim 1.5°$, separated by quasi–one-dimensional moiré solitons. Populating the nodes of this soliton network is a hexagonal lattice of point-like faults in the local twist angle corresponding to topological moiré defects that we term "twistons." Zoomed-in topographs of these three types of region of the MLR are shown in Fig. 3, B to D. Of the three sites, only the moiré solitons showed considerable breaking of C_3 rotational symmetry, which is consistent with a distinctly large value of local heterostrain $\varepsilon_x \gtrsim 0.5\%$ on these structures (*24*).

The large variation in θ_x and ε_x on the Λ scale has a dramatic effect on the local electronic structure of TTG. θ_x serves as a convenient parameter to quantitatively classify

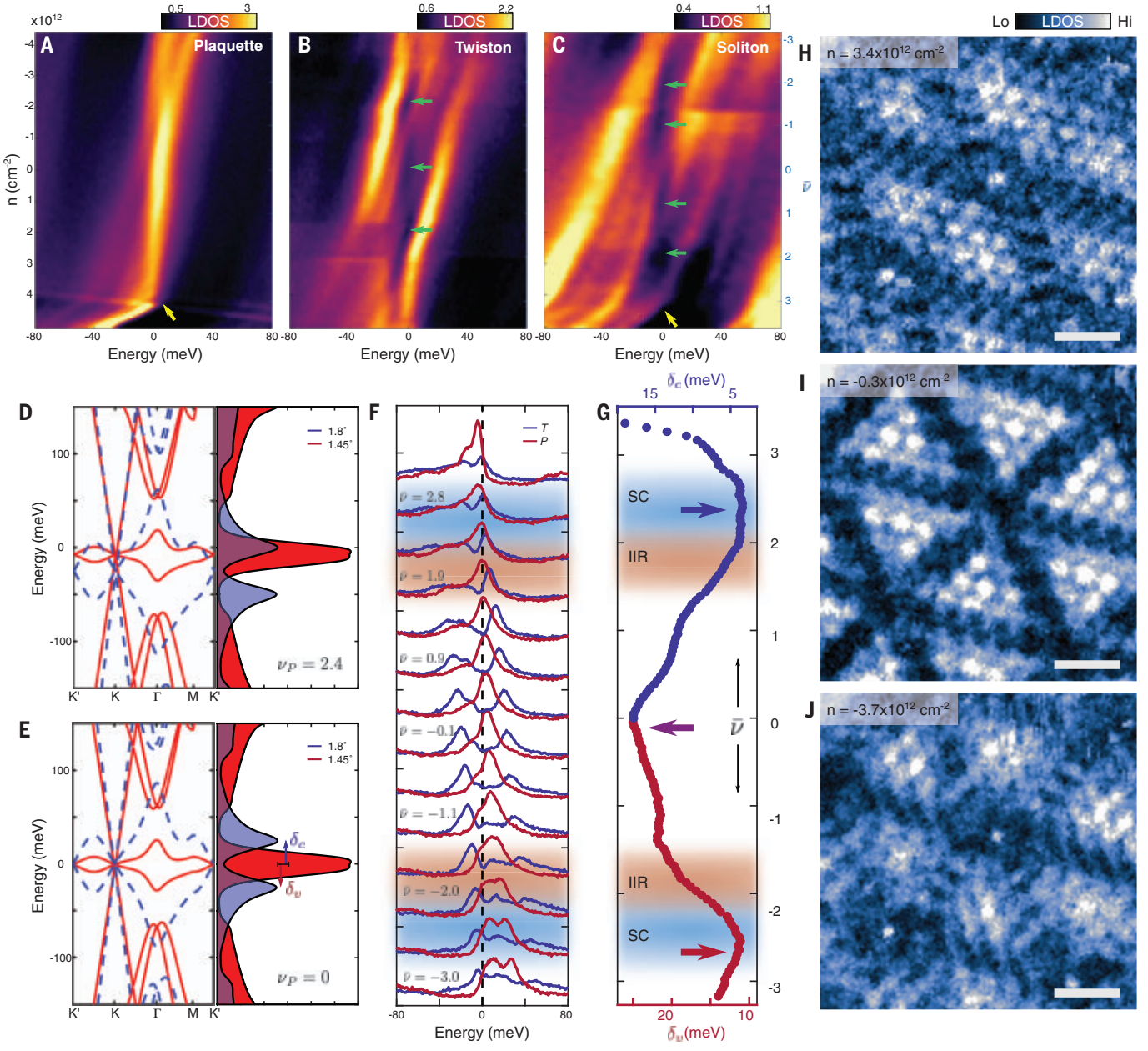


Fig. 4. Correlated gaps and flat-band resonance. (**A** to **C**) Gate-dependent LDOS spectroscopy on the plaquette, twiston, and moiré soliton regions. Yellow arrows in (A) and (C) indicate full filling of the moiré superlattice. Green arrows in (B) and (C) indicate correlated gaps that are confined to the twiston and soliton regions. (**D** and **E**) Continuum model calculations showing (left) the band structure and (right) density of states for twist angles of 1.45° (red) and 1.8° (blue) at two different fillings. The zero of energy corresponds to the position of the chemical potential. Flat band resonance (D) for these angles occurs when the moiré plaquette (1.45°) superlattice is filled to $\nu_P = 2.4$. (**F**) Doping-dependent LDOS spectroscopy on the twiston (blue) and plaquette (red) regions showing flat-band resonance at $|\bar{\nu}| \sim 2.5$. Curves are offset vertically for clarity and are plotted on the same vertical scale. The vertical dashed line indicates the Fermi level. (**G**) Extracted values of flat-band energy splitting between twiston and plaquette sites, $\delta_{c/v}$ [see (E)], as a function of doping. Minima correspond to flat-band resonances and the resulting reduction in real-space electronic disorder. (**H** to **J**) Fermi-level LDOS maps at (I) charge neutrality and at the (H) electron- and (J) hole-doped flat-band resonances. Scale bars, 25 nm. V_{set} = 300 mV, I_{set} = 120 pA, and V_{mod} = 2 mV.

different regions of the larger moiré because each region roughly corresponds to a specific value. LDOS spectra acquired on AAA sites are shown in Fig. 3E as a function of increasing θ_x. Alongside these, in Fig. 3F we plot continuum model (SP2) densities of states for a series of structural parameters (θ, ε) that approximate those found in the respective experimental topography. For low relative twists corresponding to the plaquettes, the spectrum approximates that expected for TTG, with a uniform $\theta \sim 1.45°$. As we increased θ_x, moving onto the moiré solitons, the spectral intensity of the flat bands was progressively attenuated to the point of being practically indistinct, as expected for a highly strained TTG system. Our calculation assumes a uniaxial strain applied only to the middle layer (*25*, *26*), which likely underestimates the experimental effect of ε, for which strain is distributed in a nonuniform way throughout all three layers. Increasing θ_x still further (hence decreasing ε_x), we found that the flat bands regained their intensity but were now split apart in energy by ~40 meV, which is consistent with our calculation for unstrained TTG at $\theta \sim 1.8°$. These observations indicate that the local electronic structure of TTG at small but finite δ_θ is primarily determined by the local values of heterostrain and twist angle given by the MLR. This form of twist angle disorder in TTG does not, therefore, result in a smooth and random fluctuation of the electronic structure, as it does in MATBG (*27*), but rather leads to the formation of electronic grains whose size depends directly on the experimental parameter δ_θ, yielding an inherently controllable type of moiré disorder.

To confirm the nature of the reconstructed moiré lattice, we have performed structural relaxation calculations (*28*) for TTG. The system is characterized by two twist angles that, in the absence of relaxation, are situated at the interfaces of adjacent graphene layers (θ_{TM} and θ_{BM}). As the relaxation strength was turned on, we found that the top and bottom layers locally aligned to enforce universal AtA stacking (fig. S9E). In the process, the system spontaneously organizes into patches of distinct local twist angles, as illustrated by Fig. 3G and the trimodal distribution in Fig. 3H, which is in agreement with the experimental topography (compare Fig. 3A and fig. S8B). Thus, the original twist-angle mismatch between θ_{TM} and θ_{BM} is rotated by the lattice relaxation into the plane of the sample to create a twist-angle texture between adjacent regions of the reconstructed moiré lattice. Although this calculation provides a correct qualitative description of the experimental observation of plaquettes, solitons, and twistons, it fails to accurately predict the relative sizes of these three regions, possibly because of neglecting the effects of out-of-plane corrugations and the interaction between top and bottom layers.

We explored the implications of this structural and electronic inhomogeneity for the correlated states at partial fillings by performing STS measurements as a function of V_g. Characteristic filling-dependent spectroscopy measured on AAA sites of the plaquette, twiston, and soliton is shown in Fig. 4, A to C, respectively. Full filling of the moiré superlattice can be identified as the carrier density at which the derivative of the chemical potential, $d\mu/dn$, undergoes a rapid step-like increase (Fig. 4, A and C, yellow arrows). In TTG, each moiré band is fourfold degenerate, so that full filling corresponds to a density $n_s = 4/A$, where A is the moiré unit cell area (*3*, *4*, *29*). In our case, the size of the moiré unit cell is a function of position in the MLR [$A \rightarrow A(x)$], so that we must refer to a local filling factor $\nu_x = nA(x)$. To facilitate comparisons between our local measurements and the phase diagram gleaned from bulk probe assays, we provide in fig. S8C a chart of the statistical prevalence of local filling factors as a function of induced carrier density. The area-weighted average value $\bar{\nu}$ is an approximation of the quantity probed in transport.

In spectroscopic measurements, correlation-induced insulating states typically appear as spectral gaps centered on the Fermi level that emerge and disappear as a function of induced carrier density (*19*, *20*, *30*–*32*). Unlike MATBG, in which strong correlated insulating states emerge, TTG displays only weakly resistive behavior near integer fillings. It is possible that these interaction-induced resistive states (IIRs) in TTG remain relatively undeveloped because of the coexistence of an ungapped Dirac band that serves as an alternate conducting pathway (*3*, *4*). In this scenario, we would still expect to see a suppression of the Fermi-level density of states in our spectroscopic measurements caused by the opening of an energy gap within the flat bands. In our measurements, however, we did not observe spectral gaps in uniform regions near the magic angle (Figs. 2B and 4A and fig. S13). Instead, we found that spectral gaps emerge at certain dopings near integer fillings on the twiston and soliton sites (Fig. 4, B and C, green arrows). These features of the spectrum are not expected on the basis of single-particle calculations and therefore present clear signatures of electronic correlations that are confined to particular regions of the MLR. The modulation of correlation effects by the reconstructed moiré landscape indicates the importance of the lattice reconstruction in determining the correlated phases and suggests that the microscopic structure of the MLR may have unanticipated effects on bulk properties.

We next reexamined the parent state out of which superconductivity emerges, in the context of the observed MLR. The differential rates of band filling on the regions of the Λ-modulation (fig. S8C) mean that as we add charge to the system we are simultaneously tuning the twiston and plaquette flat bands relative both to the chemical potential and to one another. This is illustrated in Fig. 4, D and E, which shows calculated band fillings at two values of n for twist angles of 1.45° and 1.8°. In Fig. 4F, we overlay the flat band spectra on twiston and plaquette sites for the full range of measured fillings. There exists a small range of n for which the two sets of flat bands are maximally overlapped and in approximate resonance with one another, giving rise to an enhanced Fermi-level density of states, which favors electronic correlations. In Fig. 4G, we quantify this flat band resonance by plotting the energy difference between spatially separated flat bands as a function of doping. The resonance condition is satisfied for $2 \lesssim |\bar{\nu}| \lesssim 3$, which is roughly aligned with the region of optimal doping for superconductivity (*3*, *4*). We expect the range of resonant dopings to be largely independent of the particular value of the twist-angle mismatch δ_θ in a given sample, given the observed relaxation phenomenon described above (Fig. 1, E to G, and fig. S10), so that the regime of optimal doping would be roughly constant across samples with $\delta_\theta \lesssim 0.5°$. Moreover, the flat band resonance occurs at dopings in between the plaquette and twiston VHSs, which is consistent with transport measurements that found superconductivity to be bounded by VHSs in doping space.

We gained further insight into the nature of the parent state by examining the effect of the flat band resonance on the real-space electronic structure through doping-dependent LDOS mapping. LDOS maps acquired at the Fermi level are shown in Fig. 4, H to J, for the three carrier densities indicated with arrows in Fig. 4G (energy dependence is provided in fig. S17). The sample displays considerable disorder at charge neutrality (Fig. 4I). The angle mismatch δ_θ in this region, as in superconducting devices (*4*), is ~0.3°, leading to magic-angle plaquettes of lateral dimension ~50 nm, which is similar in magnitude to the superconducting coherence length (*3*, *4*). As we tuned the carrier density toward the flat band resonance, however, the LDOS maps became increasingly homogeneous (Figs. 4, H and J), indicating a reduction in the strength of the disorder potential. TTG is therefore distinct among moiré-engineered materials in that varying V_g provides a means to systematically tune electronic disorder. The co-occurrence of the flat-band resonance condition, with its resulting minimization of electronic disorder and optimal doping for superconductivity, raises the possibility that the superconducting phase boundary along the doping axis is disorder

driven. If true, this would have certain implications for the symmetry of the superconducting order parameter (*33*, *34*). Recent transport measurements that indicate reentrant superconductivity at high magnetic field are compatible with a spin-triplet order parameter that would be sensitive to disorder of the type we observed (*5*).

Confirmation of this hypothesis requires direct measurements of the effect of disorder on superconductivity. Future work that systematically explores this expanded phase space by controllably tuning moiré defect density through the angle mismatch δ_θ has the potential to further shed light on the pairing mechanism in TTG by determining its sensitivity to nonmagnetic impurity scattering, as has been done in a range of other unconventional systems (*35*–*37*).

REFERENCES AND NOTES

1. E. Khalaf, A. J. Kruchkov, G. Tarnopolsky, A. Vishwanath, *Phys. Rev. B* **100**, 085109 (2019).
2. X. Li, F. Wu, A. H. MacDonald, arXiv:1907.12338 [cond-mat.mtrl-sci] (2019).
3. J. M. Park, Y. Cao, K. Watanabe, T. Taniguchi, P. Jarillo-Herrero, *Nature* **590**, 249–255 (2021).
4. Z. Hao *et al*., *Science* **371**, 1133–1138 (2021).
5. Y. Cao, J. M. Park, K. Watanabe, T. Taniguchi, P. Jarillo-Herrero, *Nature* **595**, 526–531 (2021).
6. A. Yazdani, *Science* **371**, 1098–1099 (2021).
7. M. Christos, S. Sachdev, M. S. Scheurer, arXiv:2106.02063 [cond-mat.str-el] (2021).
8. A. Fischer *et al*., *npj Quantum Mater.* **7**, 5 (2022).
9. W. Qin, A. H. MacDonald, *Phys. Rev. Lett.* **127**, 097001 (2021).
10. M. Yankowitz *et al*., *Nat. Phys.* **8**, 382–386 (2012).
11. G. Li *et al*., *Nat. Phys.* **6**, 109–113 (2010).
12. D. Wong *et al*., *Phys. Rev. B Condens. Matter Mater. Phys.* **92**, 155409 (2015).
13. C. Lei, L. Linhart, W. Qin, F. Libisch, A. H. MacDonald, *Phys. Rev. B* **104**, 035139 (2021).
14. Z. Wu, Z. Zhan, S. Yuan, *Sci. China Phys. Mech. Astron.* **64**, 267811 (2021).
15. M. Koshino *et al*., *Phys. Rev. X* **8**, 031087 (2018).
16. S. Carr *et al*., *Nano Lett.* **20**, 3030–3038 (2020).
17. R. Bistritzer, A. H. MacDonald, *Proc. Natl. Acad. Sci. U.S.A.* **108**, 12233–12237 (2011).
18. Materials and methods are available as supplementary materials.
19. A. Kerelsky *et al*., *Nature* **572**, 95–100 (2019).
20. Y. Xie *et al*., *Nature* **572**, 101–105 (2019).
21. X. Lu *et al*., *Nature* **574**, 653–657 (2019).
22. B. Lian, Z. Wang, B. A. Bernevig, *Phys. Rev. Lett.* **122**, 257002 (2019).
23. F. Wu, E. Hwang, S. Das Sarma, *Phys. Rev. B* **99**, 165112 (2019).
24. N. P. Kazmierczak *et al*., *Nat. Mater.* **20**, 956–963 (2021).
25. Z. Bi, N. F. Q. Yuan, L. Fu, *Phys. Rev. B* **100**, 035448 (2019).
26. C. Rubio-Verdú *et al*., *Nat. Phys.* **18**, 196–202 (2022).
27. U. Zondiner *et al*., *Nature* **582**, 203–208 (2020).
28. Z. Zhu, P. Cazeaux, M. Luskin, E. Kaxiras, *Phys. Rev. B* **101**, 224107 (2020).
29. Y. Cao *et al*., *Nature* **556**, 43–50 (2018).
30. Y. Jiang *et al*., *Nature* **573**, 91–95 (2019).
31. Y. Choi *et al*., *Nat. Phys.* **15**, 1174–1180 (2019).
32. C. Zhang *et al*., *Nat. Commun.* **12**, 2516 (2021).
33. M. S. Scheurer, R. Samajdar, *Phys. Rev. Res.* **2**, 033062 (2020).
34. P. W. Anderson, *J. Phys. Chem. Solids* **11**, 26–30 (1959).
35. A. P. Mackenzie *et al*., *Phys. Rev. Lett.* **80**, 161–164 (1998).
36. J. Li *et al*., *Phys. Rev. B Condens. Matter Mater. Phys.* **85**, 214509 (2012).
37. T. Hotta, *J. Phys. Soc. Jpn.* **62**, 274–280 (1993).
38. S. Turkel *et al*., Replication data for: Orderly disorder in magic angle twisted trilayer graphene, *Harvard Dataverse* (2022); doi:10.7910/DVN/QWIFR1.

ACKNOWLEDGMENTS

We thank D. Halbertal, H. Ochoa, and A. Tsvelik for fruitful discussions. **Funding:** Studies of the electronic structure of twisted trilayer graphene were supported as part of Programmable Quantum Materials, an Energy Frontier Research Center funded by the US Department of Energy (DOE), Office of Science, Basic Energy Sciences (BES), under award DE-SC0019443. The synthesis of the trilayer samples was supported by the NSF MRSEC program through Columbia in the Center for Precision-Assembled Quantum Materials (PAQM) DMR-2011738 and by DMR-2004691. Support for cryogenic STM measurements was provided by the Air Force Office of Scientific Research through grant FA9550-21-1-0378. Z.Z. and E.K. are supported by STC Center for Integrated Quantum Materials, NSF grant DMR-1231319, ARO MURI grant W911NF14-0247, and NSF DMREF grant 1922165. Relaxation calculations were performed on the Odyssey cluster supported by the FAS Division of Science, Research Computing Group at Harvard University. M.C. and S.S. are supported by NSF grant DMR-2002850. K.W. and T.T. acknowledge support from the Elemental Strategy Initiative conducted by the MEXT, Japan (grant JPMXP0112101001) and JSPS KAKENHI (grants 19H05790 and JP20H00354). **Author contributions:** S.T. and J.S. conceived the experiment. J.S. fabricated samples and performed AFM and PFM measurements. S.T. performed STM measurements, analyzed the experimental data, and performed single-particle band structure calculations. Z.Z. performed structural relaxation calculations. M.C. and M.S.S. performed interacting band structure calculations. K.W. and T.T. provided the BN crystals. S.S., E.K., C.R.D., and A.N.P. advised. S.T. wrote the manuscript, with contributions from all authors. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** Data and calculations that support the conclusions of this work are available in (*38*).

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abk1895
Materials and Methods
Supplementary Text
Figs. S1 to S17
References (*39*–*53*)

27 June 2021; accepted 9 March 2022
10.1126/science.abk1895

PHYSICS

Measurement of a helium tune-out frequency: an independent test of quantum electrodynamics

B. M. Henson[1]†, J. A. Ross[1]†, K. F. Thomas[1], C. N. Kuhn[2], D. K. Shin[1], S. S. Hodgman[1], Yong-Hui Zhang[3], Li-Yan Tang[3]*, G. W. F. Drake[4]*, A. T. Bondy[4], A. G. Truscott[1], K. G. H. Baldwin[1]*

Despite quantum electrodynamics (QED) being one of the most stringently tested theories underpinning modern physics, recent precision atomic spectroscopy measurements have uncovered several small discrepancies between experiment and theory. One particularly powerful experimental observable that tests QED independently of traditional energy level measurements is the "tune-out" frequency, where the dynamic polarizability vanishes and the atom does not interact with applied laser light. In this work, we measure the tune-out frequency for the 2^3S_1 state of helium between transitions to the 2^3P and 3^3P manifolds and compare it with new theoretical QED calculations. The experimentally determined value of 725,736,700(260) megahertz differs from theory [725,736,252(9) megahertz] by 1.7 times the measurement uncertainty and resolves both the QED contributions and retardation corrections.

Quantum electrodynamics (QED) describes the interaction between matter and light. It is so ubiquitous that the theory is considered a cornerstone of modern physics. QED has been remarkably predictive in describing fundamental processes, such as spontaneous emission rates of photons from atoms and the anomalous electron magnetic moment (*1*). However, as the precision of atomic spectroscopy approaches the part-per-trillion level, discrepancies between such predictions and experiments have come to light, such as the "proton radius puzzle" (*2*). Spectroscopic measurements [of muonic hydrogen (*3*), hydrogen (*4*, *5*), and muonic deuterium (*6*)] yield determinations of the proton radius that disagree with other approaches [electron-proton scattering (*7*) and hydrogen spectroscopy (*8*)] by up to five standard deviations.

Helium is an ideal testing ground for QED because its simple two-electron structure makes high-precision predictions tractable and testable. Notably, helium also presents a nuclear "puzzle," with precision measurement of isotope shifts of the $2^3S_1 \rightarrow 2^3P_{(0,1,2)}$ (*9*) and $2^3S_1 \rightarrow 2^1S_0$ (*10*) transitions disagreeing by two standard deviations in the derived nuclear charge radius. Further, recent measurements of the ionization energy for the helium 2^1S_0 state (*11*) confirm similar discrepancies in the Lamb shift to those recently revealed theoretically (*12*). These puzzles raise the possibility that the issue lies with QED itself (*13*). Thus, we look to challenge QED directly by precision spectroscopy in helium beyond the usual energy interval measurements.

An atom in an optical field experiences an energy shift in proportion to the real part of the frequency-dependent polarizability, a fundamental atomic property dictated by the

[1]Department of Quantum Science and Technology, Research School of Physics, The Australian National University, Canberra, ACT 2601, Australia. [2]Centre for Quantum and Optical Science, Swinburne University of Technology, Melbourne, VIC 3122, Australia. [3]State Key Laboratory of Magnetic Resonance and Atomic and Molecular Physics, Innovation Academy for Precision Measurement Science and Technology, Chinese Academy of Sciences, Wuhan 430071, People's Republic of China. [4]Department of Physics, University of Windsor, Windsor, Ontario N9B 3P4, Canada.
***Corresponding author. Email: kenneth.baldwin@anu.edu.au (K.G.H.B.); gdrake@uwindsor.ca (G.W.F.D.); lytang@wipm.ac.cn (L.-Y.T.)**
†These authors contributed equally to this work.

position of energy levels and the strengths of the transitions between them (Fig. 1). A "tune-out" frequency (f_{TO}) occurs between transition frequencies at the point where the contributions to the dynamic polarizability [$\alpha(f)$] by all transitions below that frequency are balanced by all those above it [$\alpha(f) = 0$] (*14*). This balance point is therefore fixed by the strength and frequency of every transition in the atomic spectrum and provides a precise constraint on the ratio of transition dipole matrix elements (DMEs). Similarly, "magic" wavelengths [wherein the light shift of a transition cancels (*15*), rather than the light shift of a level, as is the case for a tune-out wavelength] have yielded absolute and relative determinations of DMEs (*16*, *17*).

As a test of QED, a tune-out frequency is advantageous because it is a null measurement, which does not require calibration of the light intensity or a measurement of excitation probability. These factors have previously limited the precision of direct transition strength measurements (*18*–*20*). In comparison, previous tune-out measurements (*16*, *17*, *21*–*23*) have indicated the potential for measuring QED effects.

In this work, we measured the tune-out of the metastable 2^3S_1 state of helium (denoted He*) that lies between transitions to the 2^3P and 3^3P manifolds (denoted $2^3S_1 - 2^3P/3^3P$) at ~726 THz (413 nm). We chose this particular tune-out frequency because the two neighboring transitions are more than an octave apart in frequency, causing the gradient of atomic polarizability with optical frequency to be small at the tune-out. Thus, this tune-out frequency is especially sensitive to higher-order QED effects. We achieved a 20-fold improvement in precision compared with the sole previous measurement (*23*).

For an unambiguous comparison, we also present a new theoretical estimate of the $2^3S_1 - 2^3P/3^3P$ tune-out in helium. In the wake of the first prediction (*24*) and measurement (*23*) of the tune-out, a vigorous campaign of theoretical studies (*25*–*29*) has reduced the uncertainty in the predicted frequency, which limited comparison with experiment. Our work represents a 10-fold improvement in precision over previous calculations, and its uncertainty now surpasses the experimental state-of-the-art.

Measuring a tune-out frequency involves measuring the potential energy of a light field interacting with an atom, known as an optical dipole potential (*30*), and precisely identifying the frequency at which it vanishes (Fig. 1). The experimental approach taken here measures the optical dipole potential via changes in the spatial oscillation frequency (also called the trap frequency) of Bose-Einstein condensates (BECs) in a harmonic magnetic trap when overlapped with a laser probe beam (Fig. 2). The net potential energy is the sum of a harmonic magnetic potential and a Gaussian optical potential, which is approximately harmonic for the small oscillation amplitudes we considered. In this approximation, the oscillation frequency is given by $\Omega_{net}^2 = \Omega_{mag}^2 + \Omega_{probe}^2$, where Ω_{mag}, Ω_{probe}, and Ω_{net} denote the trap frequency of the magnetic, probe, and combined potentials, respectively. For a Gaussian beam profile, as used here, the probe perturbation scales as $\Omega_{probe}^2 \propto \alpha(f)I$, where I is the intensity of the probe beam. With the probe beam power stabilized, the difference of squared trapping frequencies $\Omega_{net}^2 - \Omega_{mag}^2 \propto \alpha(f)$ produces a response that is linearly proportional to the dynamic polarizability. Having measured the transverse and longitudinal profiles of the probe beam, we find that the shift in trapping frequency completely specifies the optical dipole potential.

We determined the trap frequency of our BECs with a novel method (*31*) that repeatedly samples the momentum of an oscillating BEC with a pulsed atom laser (*32*) (Fig. 2A). Each measurement was started by generating a new He* BEC, which was set in motion by applying a field gradient, and was then depleted over the duration of the trap frequency measurement (1.2 s) (Fig. 2B). The starting sample of atoms was cooled to ~80 nK, well below the critical temperature, to reduce the damping that ultimately limits the interrogation time and, in turn, uncertainty in the trapping frequency. We alternated between measurements of trapping frequency with and without the optical potential to calibrate for any long-term drift in Ω_{mag}. We then measured the change in (squared) trap frequency due to the probe beam, Ω_{probe}^2, as a function of the probe beam (optical) frequency f near the tune-out frequency at ~726 THz (413 nm). The small laser frequency scan range used in our experiment allowed us to determine the tune-out frequency, f_{TO}, through linear interpolation from the measured response of Ω_{probe}^2 (Fig. 2C).

The dynamic atomic polarizability consisted of the frequency-dependent scalar, vector, and tensor components [$\alpha^S(f)$, $\alpha^V(f)$, $\alpha^T(f)$, respectively]. The total polarizability (and therefore the tune-out) also depends on the degree of linear and circular polarization in the atom's reference frame, given by the second and fourth Stokes parameters, $\mathcal{Q}_\mathcal{A}$ and $\mathcal{V}$, respectively, and on the angle θ_k between the laser propagation direction and the magnetic field vector (*33*). The tune-out frequency for the 2^3S_1 state and arbitrary polarization is

$$f_{TO}(\mathcal{Q}_\mathcal{A}, \mathcal{V}) = f_{TO}^S + \frac{1}{2}\beta^V \cos(\theta_k)\mathcal{V} - \frac{1}{2}\beta^T\left[3\sin^2(\theta_k)\left(\frac{1}{2} + \frac{\mathcal{Q}_\mathcal{A}(\mathcal{Q}_\mathcal{L}, \theta_\mathcal{L})}{2}\right) - 1\right] \quad (1)$$

where f_{TO}^S is the tune-out frequency for the scalar polarizability $\alpha^S(f)$, and $\mathcal{Q}_\mathcal{A}(\mathcal{Q}_\mathcal{L}, \theta_\mathcal{L})$ is the second Stokes parameter in terms of the laboratory measurement of the second Stokes parameter, $\mathcal{Q}_\mathcal{L}$, and the angle between the lab and atomic frames, $\theta_\mathcal{L}$. Here, β^V and β^T

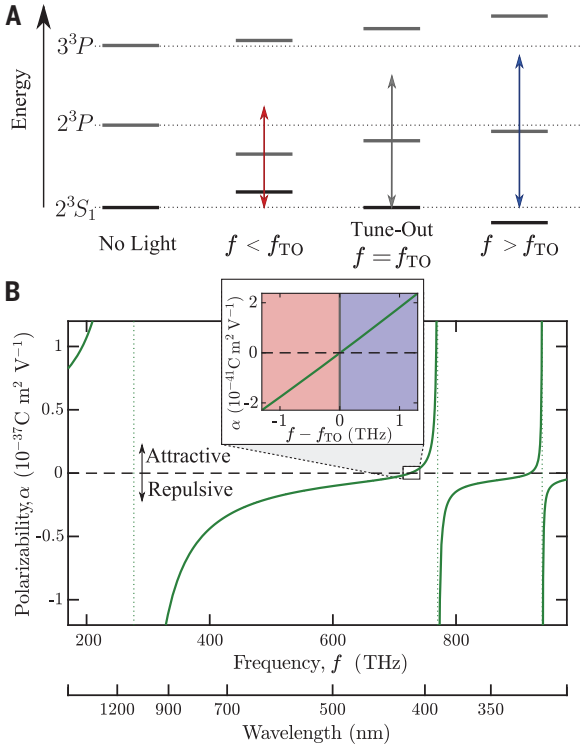


Fig. 1. Tune-out in atomic helium. (**A**) Atomic energy level shift of the dominant state (manifolds) around the tune-out. When an optical field of frequency f (arrows) is applied to the atom, the individual levels shift depending on the difference between f and the transition frequency. At the tune-out frequency, f_{TO} (middle right), the shifts to the 2^3S_1 state energy cancel. Energy spacing and shifts are not to scale. (**B**) Theoretical frequency–dependent polarizability of 2^3S_1 helium, for a constant light polarization, indicating that the polarizability vanishes near 726 THz, the tune-out frequency measured in this paper. Vertical dotted lines show, from left to right, the transitions to the 2^3P, 3^3P, and 4^3P manifolds. Inset shows the approximately linear polarizability with frequency around the tune-out.

are the vector and tensor polarizabilities divided by the gradient of the scalar polarizability (with respect to frequency) at the tune-out [see supplementary materials (SM) section 2.3].

We measure the tune-out $f_{TO}(-1,0)$, corresponding to a linearly polarized light field whose polarization axis is perpendicular to both the laser propagation and the magnetic field. For this configuration, the sensitivity to θ_k and $\theta_{\mathcal{L}}$ is minimized, and the atomic polarizability simplifies to

$$\alpha(f) = \alpha^{S}(f) - \frac{1}{2}\alpha^{T}(f) \qquad (2)$$

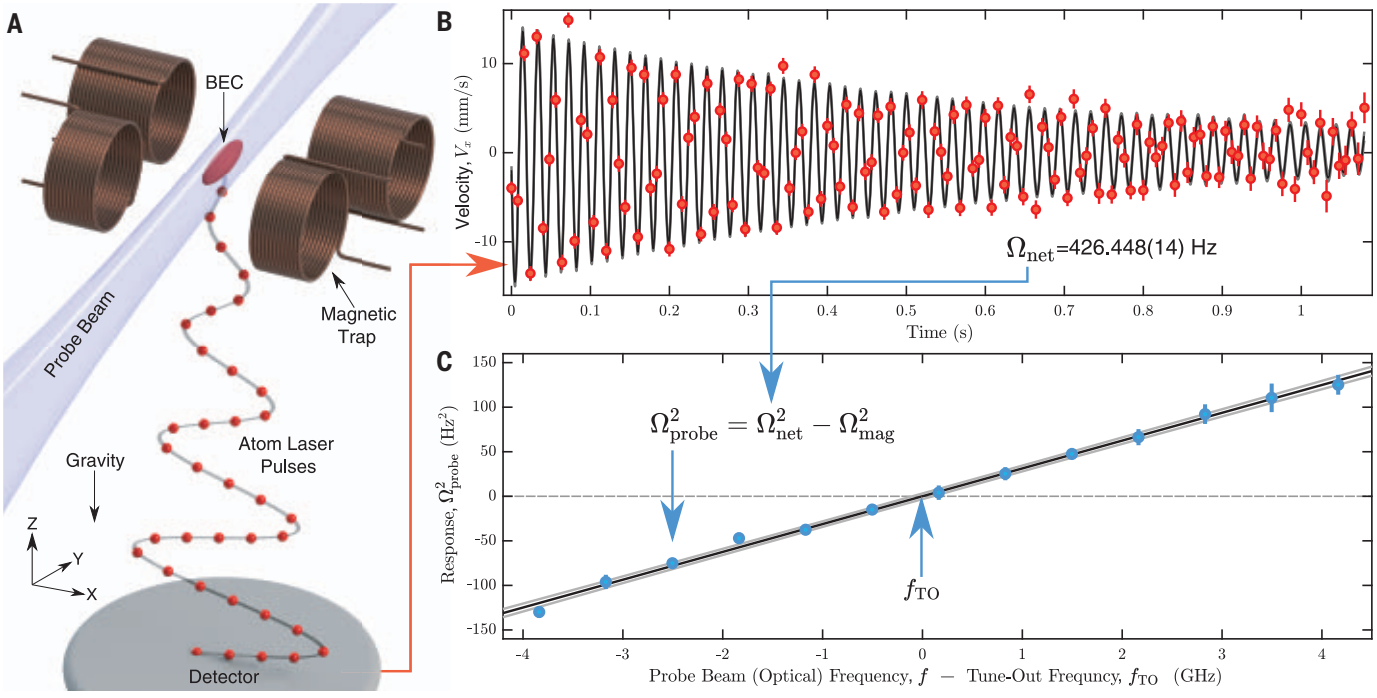


Fig. 2. Experimental procedure. Method to determine the tune-out for a fixed probe beam polarization. (**A**) A magnetically trapped BEC of metastable helium atoms was illuminated with a probe laser beam with an adjustable (optical) frequency. A sequence of atom laser pulses was outcoupled from the BEC to sample the oscillation. (**B**) The mean velocity of each pulse in the *x* direction (v_x) was used to trace out the oscillation over time (red points) and extract the oscillation frequency with a dampened sine wave fit (solid line). A single experimental realization is shown. (**C**) The squared probe beam trap frequency (response) was found using a separate measurement of the magnetic trap frequency. This measurement was repeated over a small range of optical frequencies. The tune-out was extracted by finding the *x* intercept of the response as a function of probe beam frequency using a linear fit (solid black line). Light-gray lines show the model 1σ confidence intervals. All error bars represent the standard error in the mean.

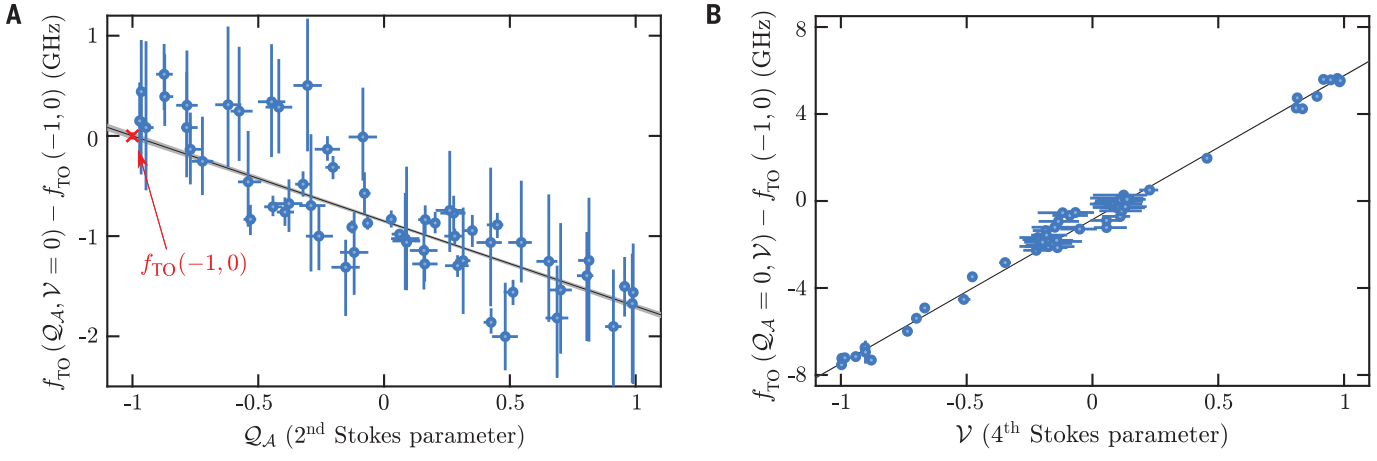


Fig. 3. Tune-out dependence on probe beam polarization. (**A**) Dependence of the measured tune-out on $\mathcal{Q}_{\mathcal{A}}$ when interpolated to $\mathcal{V} = 0$. (**B**) Dependence of the measured tune-out on $\mathcal{V}$ when interpolated to $\mathcal{Q}_{\mathcal{A}} = 0$. The linear fit to all scans is in the form of Eq. 1, with fit parameters $f_{TO}(-1,0)$ = 725,736,700(40) MHz, $\beta^{V}\cos(\theta_k)$ = 13,240(70) MHz, $\beta^{T}\sin^2(\theta_k)$ = 1140(20) MHz, and χ^2/degree of freedom = 0.9968. Horizontal error bars show polarization state uncertainty, and vertical error bars show the standard error of the measurement combined with the propagated polarization state uncertainty from the interpolated axis. For a visualization of the combined dependence, see fig. S4. The shaded regions in (A) show the model 1σ confidence interval, which is too small to be visible in (B). The point marked with a red cross in (A) shows the reference value $f_{TO}(-1,0)$ (error bar not visible at this scale).

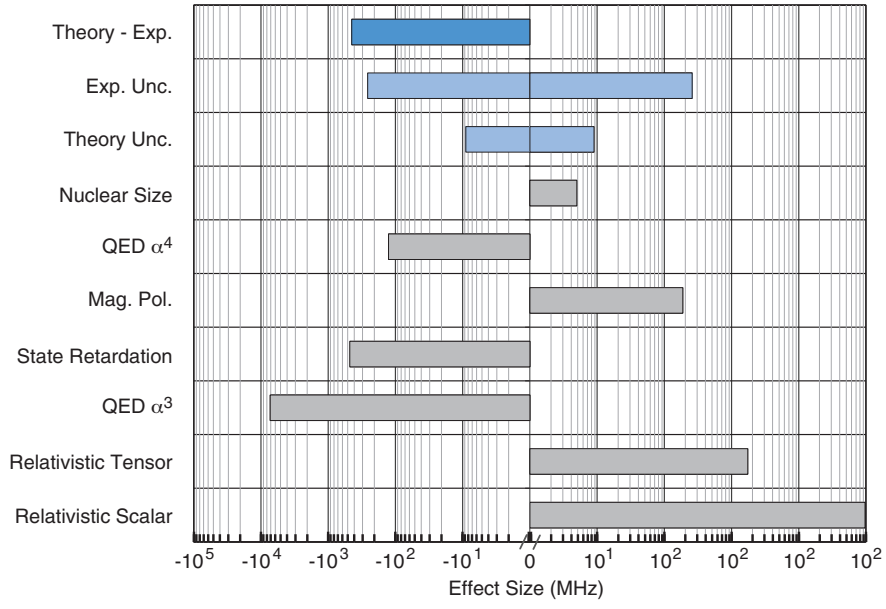


Fig. 4. Experimental and theoretical sensitivity. Comparison of uncertainties in the theoretical and experimental determinations of the $2^3S_1 - 2^3P/3^3P$ tune-out frequency and the various theoretical contributions to the tune-out value. Exp. Unc., experimental uncertainty; Theory Unc., theoretical uncertainty; Mag. Pol., magnetic polarizability.

We measured $f_{\mathrm{TO}}(\mathcal{Q}_{\mathcal{A}}, \mathcal{V})$ as a function of the probe beam polarization parameters $\mathcal{Q}_{\mathcal{A}}$ and $\mathcal{V}$ and interpolated using Eq. 1 to determine $f_{\mathrm{TO}}(-1,0)$ (Fig. 3). We took the sign of β^{T} from theory but used no other predictions in our calculation. Thus, we determined a value of 725,736,700 MHz for the $f_{\mathrm{TO}}(-1,0)$ tune-out with a statistical uncertainty of 40 MHz and a systematic uncertainty of 260 MHz (SM section 4).

The dominant systematic effect in our measurement was the uncertainty in the light polarization. The probe beam passed through a vacuum window before it interacted with the atoms, which may have subtly altered the laser polarization relative to measurements made outside the vacuum chamber. We constrained this error to be <200 MHz by measuring the probe beam polarization before entering, and after exiting, the vacuum system (SM section 4.1).

Separately, we improved on the state-of-the-art calculation (*28*) of the tune-out frequency by accounting for finite nuclear mass, relativistic, QED, finite nuclear size, and finite wavelength retardation effects (*27*, *29*). We achieved a 10-fold improvement in precision and found a theoretical value of 725,736,252 (9) MHz for $f_{\mathrm{TO}}(-1,0)$. The major contribution to the theoretical uncertainty stems from the nonradiative QED corrections (±6 MHz) of order α^4 Ry, which was an order of magnitude less than the systematic experimental uncertainty. We show a comparison of our experimental and theoretical uncertainties to the main contributions of interest to the theoretical value in Fig. 4, to demonstrate the contributions to which our measurement was sensitive.

Our experimental determination is a 20-fold improvement over the previous experimental determination and is larger than the theoretical prediction by 1.7 times the measurement uncertainty (herein, σ). Our measurement corresponds to a relative precision in oscillator strength ratio of 6 parts per million (SM section 6), which is a factor of two improvement over the previous record (*17*). The combined theoretical and experimental uncertainties (~260 MHz) were able to discern the contribution of QED effects (~30σ) and are similar to the retardation corrections to the dipole interaction (~2σ) but much greater than the contribution of finite nuclear size effects (5 MHz). Furthermore, our method for measuring the dipole potential was able to discern a peak potential energy of as little as 10^{-35} J. This is, to our knowledge, the most sensitive measurement of potential energy reported to date.

Our measurement was sensitive to the retardation corrections not normally included in the theory of the frequency-dependent polarizability (*27*, *29*). The result was an ~1.7σ difference between experiment and theory, which took into account the estimated uncertainty from terms not currently included in the theoretical calculation. It is notable that by ignoring the retardation correction term—proposed in (*29*) and included here in tune-out frequency calculations—the difference between theory and experiment fell to ~0.1σ. If the experimental precision is increased by an order of magnitude, then the effect of the retardation contribution could be more stringently tested.

Future experimental improvements could include more precise laser polarization calibrations, likely using in-vacuum optics, and a finer measurement of the angle between the laser propagation and the magnetic field. These would allow an independent comparison of the predicted and measured scalar, vector, and tensor polarizabilities, providing further information on the structure of the helium atom and QED theory itself.

Our method could be easily applied to other tune-out frequencies in helium and used as an investigative tool for other problems in QED theory. If the precision of future measurements reaches the megahertz level, the tune-out frequency could determine the nuclear charge radius of helium. Further improvements and use of our method may thus continue to challenge and elucidate QED theory.

REFERENCES AND NOTES

1. T. Aoyama, M. Hayakawa, T. Kinoshita, M. Nio, *Phys. Rev. D* **91**, 033006 (2015).
2. H. Gao, M. Vanderhaeghen, *Rev. Mod. Phys.* **94**, 015002 (2022).
3. R. Pohl *et al.*, *Nature* **466**, 213–216 (2010).
4. N. Bezginov *et al.*, *Science* **365**, 1007–1012 (2019).
5. A. Beyer *et al.*, *Science* **358**, 79–85 (2017).
6. R. Pohl *et al.*, *Science* **353**, 669–673 (2016).
7. X. Zhan *et al.*, *Phys. Lett. B* **705**, 59–64 (2011).
8. H. Fleurbaey *et al.*, *Phys. Rev. Lett.* **120**, 183001 (2018).
9. X. Zheng *et al.*, *Phys. Rev. Lett.* **119**, 263002 (2017).
10. R. J. Rengelink *et al.*, *Nat. Phys.* **14**, 1132–1137 (2018).
11. G. Clausen *et al.*, *Phys. Rev. Lett.* **127**, 093001 (2021).
12. V. Patkóš, V. A. Yerokhin, K. Pachucki, *Phys. Rev. A* **103**, 042809 (2021).
13. R. J. Hill, *EPJ Web Conf.* **137**, 01023 (2017).
14. L. J. LeBlanc, J. H. Thywissen, *Phys. Rev. A* **75**, 053612 (2007).
15. Y.-H. Zhang, L.-Y. Tang, J.-Y. Zhang, T.-Y. Shi, *Phys. Rev. A* **103**, 032810 (2021).
16. C. D. Herold *et al.*, *Phys. Rev. Lett.* **109**, 243003 (2012).
17. R. H. Leonard, A. J. Fallon, C. A. Sackett, M. S. Safronova, *Phys. Rev. A* **92**, 052501 (2015).
18. N. Bouloufa, A. Crubellier, O. Dulieu, *Phys. Scr.* **134**, 014014 (2009).
19. F. Vogt *et al.*, *Eur. Phys. J. D* **44**, 73–79 (2007).
20. K. F. Thomas *et al.*, *Phys. Rev. Lett.* **125**, 013002 (2020).
21. W. F. Holmgren, R. Trubko, I. Hromada, A. D. Cronin, *Phys. Rev. Lett.* **109**, 243004 (2012).
22. F. Schmidt *et al.*, *Phys. Rev. A* **93**, 022507 (2016).
23. B. M. Henson *et al.*, *Phys. Rev. Lett.* **115**, 043004 (2015).
24. J. Mitroy, L.-Y. Tang, *Phys. Rev. A* **88**, 052515 (2013).
25. Y.-H. Zhang, L.-Y. Tang, X.-Z. Zhang, T.-Y. Shi, *Phys. Rev. A* **93**, 052516 (2016).
26. J. Manalo, thesis, University of Windsor, Ontario, Canada (2017).
27. G. W. F. Drake, J. G. Manalo, P.-P. Zhang, K. G. H. Baldwin, *Hyperfine Interact.* **240**, 31 (2019).
28. Y.-H. Zhang *et al.*, *Phys. Rev. A* **99**, 040502 (2019).
29. K. Pachucki, M. Puchalski, *Phys. Rev. A* **99**, 041803 (2019).
30. R. Grimm, M. Weidemüller, Y. B. Ovchinnikov, *Adv. At. Mol. Opt. Phys.* **42**, 95–170 (2000).
31. B. M. Henson *et al.*, arXiv:2201.10021 [cond-mat.quant-gas] (2022).
32. M.-O. Mewes *et al.*, *Phys. Rev. Lett.* **78**, 582–585 (1997).
33. F. Le Kien, P. Schneeweiss, A. Rauschenbeutel, *Eur. Phys. J. D* **67**, 92 (2013).
34. B. M. Henson, K. F. Thomas, Replication data for: Testing quantum electrodynamics by measuring a tune-out frequency

in atomic helium, version 1, Harvard Dataverse (2022); https://doi.org/10.7910/DVN/KQEIOW.
35. B. M. Henson, J. A. Ross, K. F. Thomas, Tune out v2 code, GitHub (2022); https://github.com/HeBECANU/Tune_out_v2.

ACKNOWLEDGMENTS

We thank M. Bromley for instructive discussion regarding the hyperpolarizability, D. Cocks for careful reading of the manuscript, C. J. Vale and S. Hoinka for the loan of the laser, T.-Y. Shi for helpful discussions regarding the theoretical calculations, and K. Pachucki for helpful correspondence concerning the relativistic and retardation corrections to the tune-out frequency. **Funding:** This work was supported through Australian Research Council (ARC) Discovery Project grants DP160102337 and DP180101093, as well as Linkage Project LE180100142. K.F.T. and D.K.S. were supported by Australian Government Research Training Program (RTP) scholarships. S.S.H. was supported by ARC Discovery Early Career Researcher Award DE150100315. L.-Y.T. was supported by the National Key Research and Development Program of China under grant 2017YFA0304402, the Strategic Priority Research Program of the Chinese Academy of Sciences under grant XDB21030300, and the National Natural Science Foundation of China under grants 12174402 and 12121004. G.W.F.D. acknowledges support by the Natural Sciences and Engineering Research Council of Canada (NSERC) and by SHARCNET. **Author contributions:** B.M.H., J.A.R., K.F.T., L.-Y.T., G.W.F.D., A.G.T., and K.G.H.B. conceived of the work. B.M.H., J.A.R., K.F.T., A.G.T., and K.G.H.B. designed the experiments. B.M.H., J.A.R., K.F.T., C.N.K., and D.K.S. ran the experiments. B.M.H., J.A.R., and K.F.T. analyzed and visualized the data. Y.-H.Z., L.-Y.T., G.W.F.D., and A.T.B. developed the theoretical methods. B.M.H., J.A.R., K.F.T., S.S.H., L.-Y.T., G.W.F.D., A.G.T., and K.G.H.B. wrote the manuscript. All authors reviewed the results and commented on the manuscript. S.S.H., A.G.T., and K.G.H.B. secured funding for and supervised the project. **Competing interests:** None declared. **Data and materials availability:** All experimental data along with the associated processing code are available online (*34*, *35*). All other data needed to evaluate the conclusions in the paper are present in the paper or the supplementary materials.

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abk2502
Materials and Methods
Supplementary Text
Figs. S1 to S7
Tables S1 to S3
References (*36*–*66*)

5 July 2021; resubmitted 24 January 2022
Accepted 8 March 2022
10.1126/science.abk2502

NANOMATERIALS

Three-dimensional visualization of nanoparticle lattices and multimaterial frameworks

Aaron Michelson[1], Brian Minevich[2], Hamed Emamy[2], Xiaojing Huang[3], Yong S. Chu[3], Hanfei Yan[3]*, Oleg Gang[1,2,4]*

Advances in nanoscale self-assembly have enabled the formation of complex nanoscale architectures. However, the development of self-assembly strategies toward bottom-up nanofabrication is impeded by challenges in revealing these structures volumetrically at the single-component level and with elemental sensitivity. Leveraging advances in nano-focused hard x-rays, DNA-programmable nanoparticle assembly, and nanoscale inorganic templating, we demonstrate nondestructive three-dimensional imaging of complexly organized nanoparticles and multimaterial frameworks. In a three-dimensional lattice with a size of 2 micrometers, we determined the positions of about 10,000 individual nanoparticles with 7-nanometer resolution, and identified arrangements of assembly motifs and a resulting multimaterial framework with elemental sensitivity. The real-space reconstruction permits direct three-dimensional imaging of lattices, which reveals their imperfections and interfaces and also clarifies the relationship between lattices and assembly motifs.

The self-assembly of nanomaterials is an attractive means of creating three-dimensional (3D) nanostructures for novel applications in photonics, catalysis, and biomaterials (*1*, *2*) without the limitations of conventional nanofabrication methods. Recent advances in nanoparticle assemblies were achieved through tailoring interparticle interactions (*3*) and nanoparticle shapes (*4*, *5*) or by constructing directional interparticle bonds (*6*, *7*). Although nanoparticle superlattices can be formed, the complexity of binding modes (*5*, *6*) and crystallization pathways (*8*) can lead to metastable states that typically obfuscate the assembly. These result in disordered regions and different imperfections. An ability to reveal formed structures volumetrically on a single-particle level is critical for advancing self-assembly approaches toward creating fully engineered nanomaterials. For example, understanding the relationship between assembly motifs and assembled organization, or between an assembly process and defect types, requires imaging that can uncover global and local structure in three dimensions. As the capabilities of forming continuous (framework) and discrete (particle) organizations (*9*–*14*) and templating them with inorganic materials (*13*, *15*, *16*) increase, there is a concomitant need for 3D nanoscale visualization.

Recent advances in electron microscopy allowed for direct 3D nano-imaging of polymers (*17*) and nanoparticles (*18*, *19*). However, its application for large-scale assemblies is challenging because of the high absorption of electrons. In contrast, hard x-rays offer excellent penetration, but x-ray imaging suffers from limited resolution. Tomography based on coherent x-ray diffractive imaging (CXDI) was applied to visualize colloidal crystals with 80-nm resolution (*20*) and an integrated circuit with 15-nm resolution using ptychography (*21*). These phase retrieval–based methods, however, lack elemental sensitivity. In contrast, raster-scan imaging with a nanobeam performed in scanning hard x-ray microscopy (SHXM) can provide simultaneous elemental and morphological visualization through direct fluorescence imaging and ptychography reconstruction, and has the potential to exceed optical limitations. Previously, correlative 3D x-ray microscopy with a resolution in the range of 100 nm has been demonstrated (*22*).

For particle-by-particle analysis of superlattices, we first assembled a face-centered cubic (fcc) lattice using DNA origami tetrahedra frames whose vertices possess DNA complementarity to single-stranded DNAs grafted to 20-nm gold nanoparticles (AuNPs) (*18*). For visualization of the DNA assembly motif, we used a pair of tetrahedra with complementary DNA–encoded vertices to form a diamond lattice (*14*), where each 15-nm AuNP is located at the tetrahedron center (fig. S13). Surveyed assembled structures displayed a mixture of ordered and disordered aggregates (Fig. 1C and figs. S10, S12, and S23). Samples (~2 μm in diameter) were mounted on a tungsten needle tip, and a focused ion beam (FIB) was used to trim the sample while preserving surface features (Fig. 1C and fig. S12). This sample geometry allows for the collection of images from a full range of angles for a complete 3D tomogram.

We used a monochromatic x-ray beam at 12 keV, focused by a set of crossed multilayer Laue lenses (*23*), to produce a 13-nm nanobeam (*24*) for SHXM studies and a specially designed microscope with high stiffness and thermal stability (*25*). A schematic of the experimental setup is shown in Fig. 1A; details are described in the supplementary materials. At each projection, both fluorescent and far-field diffraction images were obtained. The latter were analyzed with a ptychography reconstruction algorithm to retrieve both the complex-valued probe and object functions (*26*). For acquired fluorescence spectra, we performed fluorescence peak fitting to remove background and separate overlapped peaks (fig. S18), and further refined the data using a probe function retrieved from ptychography analysis (fig. S2). Consequently, elemental maps

[1]Department of Applied Physics and Applied Mathematics, Columbia University, New York, NY 10027, USA. [2]Department of Chemical Engineering, Columbia University, New York, NY 10027, USA. [3]National Light Source II, Brookhaven National Laboratory, Upton, NY 11973, USA. [4]Center for Functional Nanomaterials, Brookhaven National Laboratory, Upton, NY 11973, USA.
***Corresponding author. Email: hyan@bnl.gov (H.Y.); og2226@columbia.edu (O.G.)**

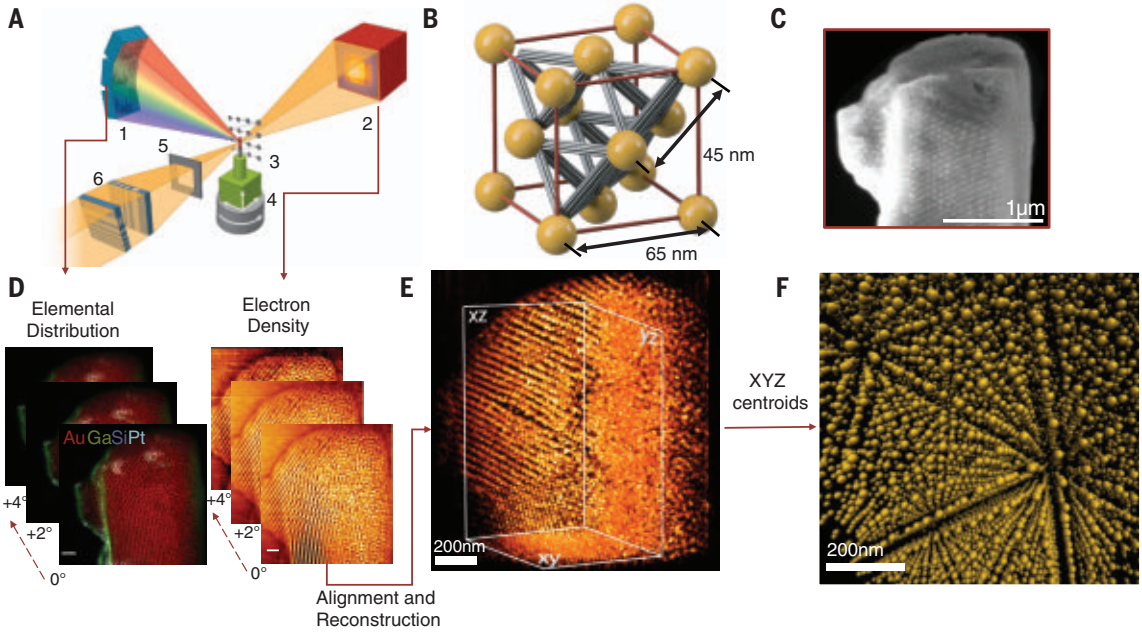


Fig. 1. Hard x-ray nanoprobe tomography and revealed 3D organization of a nanoparticle lattice. (**A**) Schematic of hard x-ray nanoprobe beamline, showing (1) fluorescence detector, (2) pixel-array detector, (3) translation stage, (4) rotation stage, (5) order-sorting aperture, and (6) multilayer Laue (MLL) optics. The MLL focuses hard x-rays into a 13-nm x-ray beam used for collecting the fluorescence and transmission ptychography data simultaneously. (**B**) Unit cell of lattice with tetrahedron origami and AuNP at vertices. (**C**) Scanning electron micrograph of sample. (**D**) Left: Elemental distribution 2D imaging. Right: Electron density maps. Scale bar, 200 nm. (**E**) 3D reconstruction of lattice with 3D region of interest removed to view interior grain structure, showing two grains and a disordered region toward the center. (**F**) 3D perspective view of AuNP superlattice with ~10^4 particles. The image is generated with centroid coordinates from the 3D reconstruction and shown with idealized spheres representing 20 nm (up to scale) AuNP. See movies S4 and S5 for different rotation views of the nanoparticle superlattice.

and electron density maps (the latter from the phase of the complex object function) were obtained at each projection angle (Fig. 1D). Tomography reconstruction (Fig. 1E), performed with assistance from ImageJ-Fiji-1.5 and Tomviz1.9 (*27*) software packages, yielded ~10^4 individual nanoparticle coordinates. To enhance a 3D lattice visualization, we performed segmentation and determined nanoparticle centroids, and replaced NPs with identical spheres (see Fig. 1F, figs. S1 to S10, and movies S1 to S3). A visualized structure allows us to identify both lattice order and imperfections.

In Fig. 2, we show representative ptychography reconstructions and results of fluorescence imaging at one projection angle of the fcc assembly. As a consequence of the weak x-ray absorption of AuNPs, the amplitude of the object function contains many background fluctuations, but its phase is clean (Fig. 2A) and was used for the tomographic reconstruction of the electron density map. Elemental distribution at the same projection (Fig. 2B) shows superlattice planes with Ga ions on the periphery, deposited during FIB processing. Volumetric views of the tomography reconstruction produced from the phase, which reflects the electron density variation and Au fluorescence signals, are presented in Fig. 2C. Both images are consistent in exhibiting ordered and disordered domains. The phase variation is attributed to AuNPs, silica bonds, Pt, and Ga, whereas the Au fluorescence pinpoints the location of individual NPs and removes ambiguity in the phase image (Fig. 2D). Because of the weak Si fluorescence signal, silica struts were not reconstructed.

The 3D reconstruction depicts two crystalline grains and one amorphous grain, as well as various defects. To quantify the achieved resolution, we performed an analysis in reciprocal space to determine the cutoff frequency at which the signal is dropped to the noise level. Figure 2, E and F, shows power spectral density variations projected on three orthogonal planes along with spherical shells in reciprocal space for phase and Au fluorescence reconstructions, respectively. We determined a half-pitch resolution of 7 nm × 7 nm × 9 nm for the phase and 9 nm × 9 nm × 15 nm for the fluorescence (figs. S4 to S6). These estimations agree well with our expectations because the optics used a numerical aperture of 5 mrad, equivalent to 10-nm resolution (Rayleigh criterion) at 12 keV. We stress that ptychography helps break the resolution barrier imposed by the numerical aperture, and the fluorescence image is also greatly improved by accurate knowledge of the point-spread function. This aspect is also reflected in the sectioned Fourier shell correlation of the data (fig. S6), which shows the resolution of the reconstructed data to be between 5 nm (single pixel) and 14 nm for the sectioned regions, and globally 9 nm at a conservative one-bit resolution threshold for ptychography (*28*).

The achieved tomographic reconstruction at a single-particle level allowed us to inspect and analyze volumetrically the occurring defects in superlattices (Fig. 3). We found that point (0D), line (1D), planar (2D), and bulk (3D) defects at the nanoscale resemble their atomic analog in atomic crystals, although self-assembly and atomic crystal growth have different mechanisms and occur at different length scales. We stress an important distinction between the assembled lattice here and lattices of isotropic nanoparticles: The geometric constraints and directionality of interactions provided by frames (Fig. 1B) can result in specific defect types.

A commonly observed imperfection is a vacancy, or 0D defect, which might be energetically favorable as a result of entropic forces. Vacancies similarly appear in our assembled superlattice (Fig. 3, A and B, blue spheres), yet their origin is likely different. Atomically, vacancies nucleate from the diffusion of atoms in the lattice at temperatures even well below the melting point; by contrast, the nanoparticles in a superlattice are held in place by four tetrahedra with each vertex connected to an AuNP by up to six DNA bonds, which are stable at room temperature (Fig. 3B). This makes it unlikely for an AuNP, once fully bonded, to have sufficient energy for diffusion at room temperature; in turn, this suggests that defects form during lattice annealing. Another source for 0D defects in our structure comes from deviant packing of tetrahedra around the AuNP. If there are more than four tetrahedra, then the unit cell will distort. This effect was observed in our experiment (Fig. 3C), as viewed in the (111) plane. Perfect (green) and imperfect (yellow) packing of nanoparticles is reminiscent of a particle

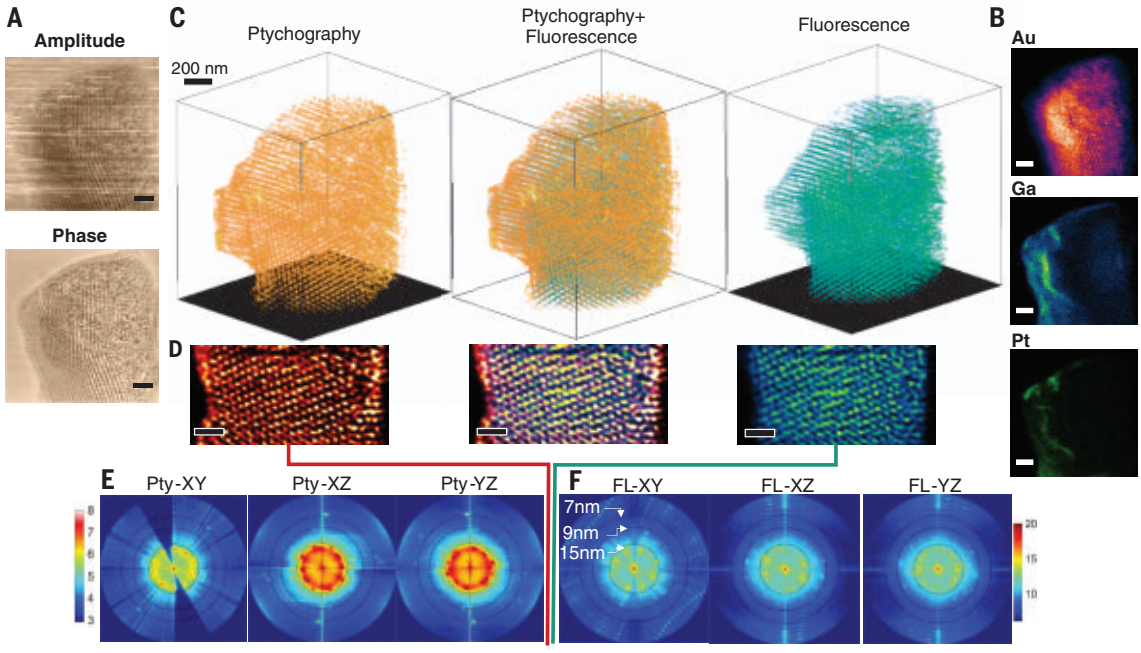


Fig. 2. 3D renderings of 20-nm nanoparticle superlattice. (**A**) X-ray ptychography–reconstructed amplitude and phase. (**B**) Fluorescence signal from Au, Pt, and Ga (introduced by FIB milling) at the same angle. (**C**) 3D ptychography reconstruction and fluorescence-Au signal channel reconstruction and central multimodal 3D model with both reconstructions simultaneously displayed. (**D**) Internal slice from each reconstruction displaying the central capability to resolve elemental and morphological features at high resolution. (**E**) Power spectral density of ptychography reconstruction with resolution rings at 15 nm (inner ring), 9 nm (middle ring), and 7 nm (outer ring); density is clear in the *xz* and *yz* projections. (**F**) Power spectral density of fluorescence with resolution at 15 nm. In the *xy* plane, the observed streak-like features are artifacts resulting from the limited number of projections (see supplementary materials for more detail). All scale bars, 200 nm.

with a correct and an aberrant number of nearest-neighbor frames, respectively. The key parameter controlling the number of tetrahedra per particle is the particle's diameter (*18*). For a 20-nm AuNP, more than four frames can occasionally be coordinated (Fig. 3D), resulting in this defect, consistent with previous computational studies (*29*).

We further investigated line and screw dislocations; these are 1D defects, which on the atomic scale are sources of stress and strain within a lattice and can lead to long-range imperfections. We show an example of observed screw dislocation in Fig. 3E. In the 2D projections, the screw terminates on the lattice surface and extends to the internal interface with the second grain. On the basis of the orientation of the screw dislocation, the Burgers vector is a $1/3a$ [111] displacement, typical of a Frank partial dislocation in atomic systems. The figure overlay for the (112) plane shows the particle positions (red lines) for the overall planes (blue lines) crossing over to different planes of (111) with a right-handed screw. Similar to the vacancy defects, the screw dislocations are energetically unable to diffuse through the superlattice, which means they likely originated during lattice annealing. In another example, the presence of an inclusion particle (Fig. 3G) disrupts the lattice by two to three lattice spacings. The spatially skewed extent of the void space created around the inclusion suggests that a misaligned tetrahedron fills the missing wedge, but the silicate struts cannot be visualized. The superlattice surface is partly preserved, thus allowing for the observation of steps, ridges, and adatoms (Fig. 3F). The coordination of particles shows that growth of the surface proceeds in the [111] direction. The preference for the growth in this direction is in line with the expectation that the (111) face is the most energetically favorable for the fcc superlattice.

For atomic crystals, thermal excitations result in the oscillations of atoms near their equilibrium position. In our system, AuNPs are attached to the vertices of tetrahedra through flexible single-stranded DNA motifs required for crystallization (*3*, *14*, *18*); thus, an AuNP can oscillate in a native solution. Upon lattice mineralization, particles become immobilized. The frozen state (Fig. 3G) is a snapshot of particle fluctuations and stress fields in lattices during the mineralization. The captured 3D distribution of NPs might represent phonon modes available to nanoparticles in the lattice. The analysis of nanoparticle positions using a pair distribution function (fig. S16) indicated oscillation of ~10 nm from their mean placement.

We applied tomographic imaging to explore a 3D structure of grain boundaries. To determine what domain a particle belonged to, we used the Fourier transform of the raw data to inspect the ordered domain peaks related to the fcc structure in reciprocal space (*k*-space). This showed two sets of lattice reflections at a slight angle to each other. Each set of *k*-space points corresponding to one of the lattices was then masked and inverse-transformed, resulting in each crystal domain being specified, and thus allowing us to assign a particle to domains A or B (figs. S7 to S10). The corresponding crystalline planes, (111) and (100), show the relative orientation of these grains as they meet at the grain boundary (Fig. 3H).

Contrary to atomic systems, the two ordered domains did not have identical orthonormality as identified from centroids in *k*-space (Fig. 3I). Such tolerance for angular distortion indicates the enhanced flexibility of assembly. The interface between the ordered grains is faceted along [111] and [100] directions, and the angle between them is ~13° (Fig. 3J). Recent simulation on the expected Wulff shape of nanoparticles assembled by tetrahedra suggests that the (111) and (100) faces are energetically favorable, following the broken bond theory (*30*) wherein the minimal surface energy facet of the superlattice mimics fcc with $\gamma_{(111)}$ and $\gamma_{(100)}$ preference. Electron microscopy probing shows similar faceting (fig. S12), which results in a pyramidal shape. The intergrain interface region (green) presents common particles between the lattice grains. We have typically observed only a single-particle layer between the two grains, which suggests that lattice points are shared. This scenario generally corresponds to a semicoherent grain boundary; however, the angle between the grains does not correspond to a common low-index normal vector. Thus, we hypothesize that flexibility of frame-particle bonds allows for a greater tolerance between the grains.

Whereas the particle positions reveal 3D organization on a particle-by-particle level, the assembly motif (DNA frame) remains invisible. We further explored a different, more

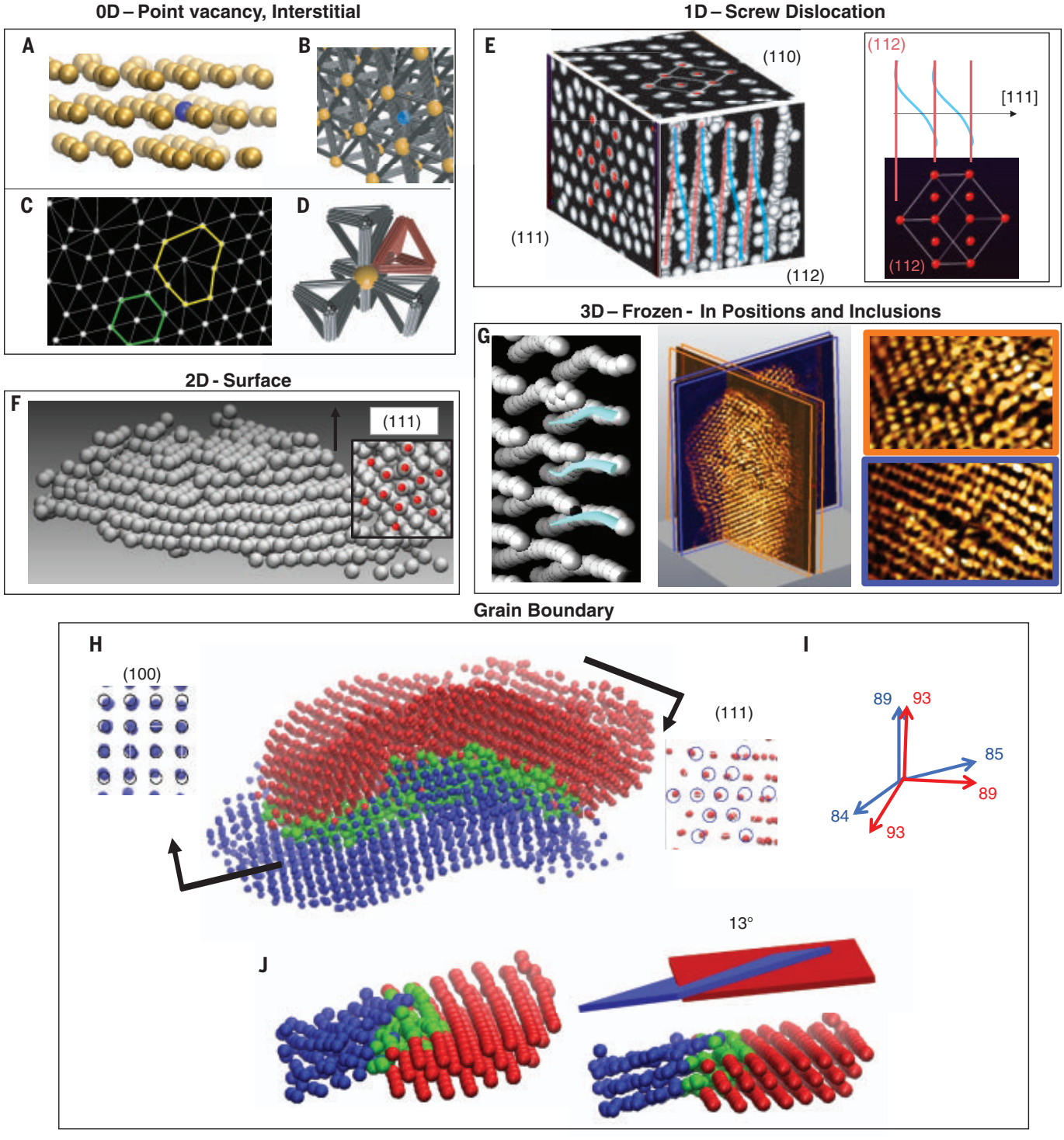


Fig. 3. Observation of defects, dislocations, and distortions in the 3D nanoparticle superlattice. (**A**) Observed and reconstructed particle superlattice and vacancy defect (blue). (**B**) Lattice model with tetrahedron cages, with vacancy (blue) visible. (**C**) Unit cell distortion with nominal neighbors (green) and distorted cell (yellow). (**D**) Model packing of sphere with tetrahedron, with extra tetrahedron cage surrounding nanoparticle. (**E**) Observed screw dislocation, viewed from the [112] perspective with red/blue overlay. The *a*/3 [111] Frankel-type defect runs from left to right in a right-handed curl; overlaid are model (111), (110), and (112) planes. (**F**) View of reconstructed 2D surface defects with [111] direction indicated. (**G**) Left: Imaging of frozen-in particle positions. Center and right: Inclusion viewed from 90° perspectives. Spheres represent 20-nm Au nanoparticles. (**H**) Cropped portion of reconstructed superlattice grain boundary showing faceting along [111] and [100] directions. Perspective views of the (100) and (111) planes are overlaid with open circles corresponding to idealized lattice positions, with blue and red dots representing the experimentally observed particle positions. (**I**) Orthonormal vectors for red grain and blue grain. (**J**) Low-angle mismatch between lattices from two perspectives alongside a model with the calculated angle between the grains. Spheres represent 20-nm Au nanoparticles.

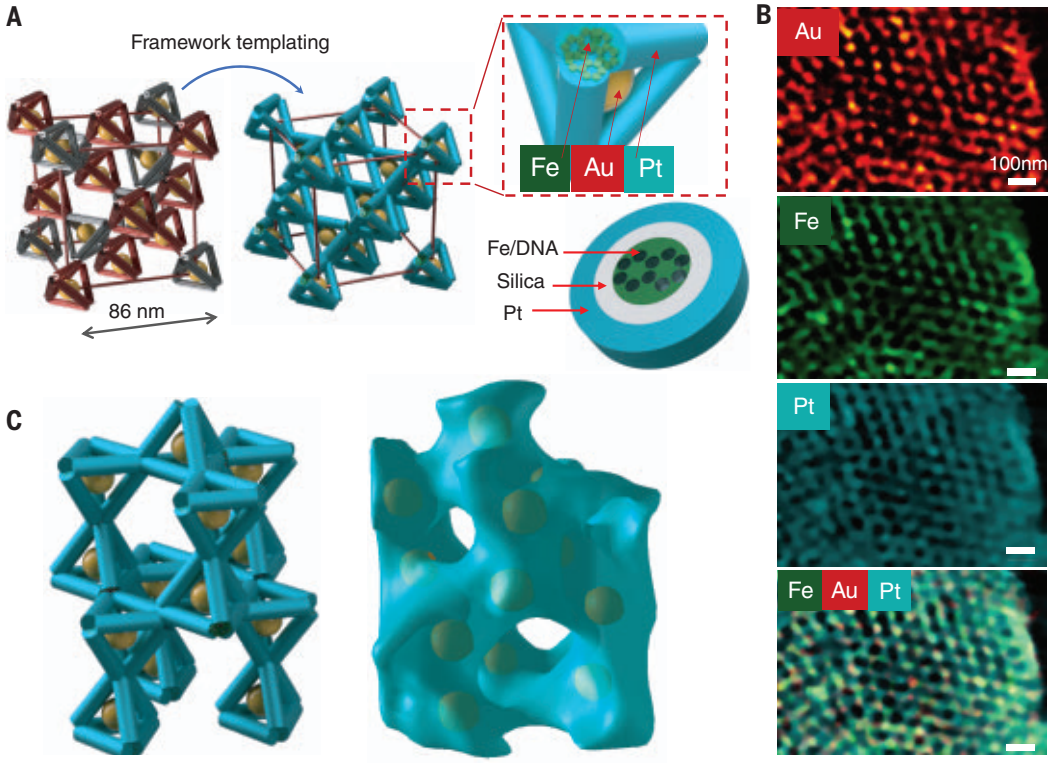


Fig. 4. Multielement continuous framework with embedded nanoparticles based on a diamond superlattice assembled from tetrahedron motifs. (**A**) Model of a diamond unit cell with complementary DNA–encoded tetrahedra in red and gray with nanoparticles inset, multielement templated (teal) framework of tetrahedra, and zoom-in of the templated tetrahedra with iron internal along the DNA structure, silica and platinum outer coating. (**B**) 3D reconstructed slice with Au, Fe, and Pt channels along with a composite image demonstrating the particle/matrix visibility, directly showing particle-to-particle bonds. Scale bars, 100 nm. (**C**) 3D model alongside 3D reconstructed data with segmented representation of the continuous framework formed by Fe/silica/Pt tetrahedra and Au nanoparticles.

complex organization where 15-nm AuNPs were encapsulated in tetrahedra that were assembled in a diamond lattice (Fig. 4A) through inter-vertex bonds (*14*). We used this superlattice to create a multimaterial (iron/silica/platinum) continuous framework based on the tetrahedron motif while preserving nanoparticle placements. Such complex inorganic structures are desirable for diverse material applications and present an important challenge for visualization. After assembling the diamond lattice, we templated the DNA framework by absorption of iron ions into the charged DNA backbone, followed by silicification and platinum coating. The elemental sensitivity and spatial acuity of our imaging allowed us to produce gold, iron, and platinum maps (Fig. 4B) of the continuous framework. We then reconstructed these 3D multimaterial frameworks and AuNP lattice (Fig. 4C), where a close correspondence to a model structure was observed. This approach clearly shows the relation between the nanoparticles and the tetrahedra motif, for both global and local arrangements. The platinum and iron coatings offer a complete view into the framework and complement each other where either platinum or iron did not coat fully (movie S6 and figs. S19 and S20).

The developed methods make it possible to create DNA-prescribed discrete and continuous inorganic lattices that can find use in catalytical, optical, and energy material applications. The demonstrated characterization approach will provide unprecedented opportunities to understand and perfect a broad range of self-assembled nanomaterials.

REFERENCES AND NOTES

1. E. Yablonovitch, *J. Opt. Soc. Am. B* **10**, 283–295 (1993).
2. E. Auyeung *et al.*, *J. Am. Chem. Soc.* **137**, 1658–1662 (2015).
3. D. Nykypanchuk, M. M. Maye, D. van der Lelie, O. Gang, *Nature* **451**, 549–552 (2008).
4. M. R. Jones *et al.*, *Nat. Mater.* **9**, 913–917 (2010).
5. G. van Anders, N. K. Ahmed, R. Smith, M. Engel, S. C. Glotzer, *ACS Nano* **8**, 931–940 (2014).
6. F. Romano, F. Sciortino, *Nat. Commun.* **3**, 975 (2012).
7. M. B. Zanjani, I. C. Jenkins, J. C. Crocker, T. Sinno, *ACS Nano* **10**, 11280–11289 (2016).
8. J. J. De Yoreo, *MRS Bull.* **42**, 525–536 (2017).
9. Y. Tian *et al.*, *Nat. Mater.* **15**, 654–661 (2016).
10. R. J. Macfarlane *et al.*, *Science* **334**, 204–208 (2011).
11. X. Liu *et al.*, *Nature* **559**, 593–598 (2018).
12. I. Coropceanu, M. A. Boles, D. V. Talapin, *J. Am. Chem. Soc.* **141**, 5728–5740 (2019).
13. E. Auyeung, R. J. Macfarlane, C. H. Choi, J. I. Cutler, C. A. Mirkin, *Adv. Mater.* **24**, 5181–5186 (2012).
14. Y. Tian *et al.*, *Nat. Mater.* **19**, 789–796 (2020).
15. L. Nguyen, M. Döblinger, T. Liedl, A. Heuer-Jungemann, *Angew. Chem. Int. Ed.* **58**, 912–916 (2019).
16. P. W. Majewski *et al.*, *Sci. Adv.* **7**, eabf0617 (2021).
17. K. W. Gotrik *et al.*, *Adv. Funct. Mater.* **24**, 7689–7697 (2014).
18. W. Liu *et al.*, *Science* **351**, 582–586 (2016).
19. A. Mayence, D. Wang, G. Salazar-Alvarez, P. Oleynikov, L. Bergström, *Nanoscale* **6**, 13803–13808 (2014).
20. A. G. Shabalin *et al.*, *Phys. Rev. Lett.* **117**, 138002 (2016).
21. M. Holler *et al.*, *Nature* **543**, 402–406 (2017).
22. J. Deng *et al.*, *Sci. Adv.* **4**, eaau4548 (2018).
23. H. Yan, R. Conley, N. Bouet, Y. S. Chu, *J. Phys. D* **47**, 263001 (2014).
24. H. Yan *et al.*, *Nano Futures* **2**, 011001 (2018).
25. E. Nazaretski *et al.*, *J. Synchrotron Radiat.* **24**, 1113–1119 (2017).
26. H. Yan, *New J. Phys.* **22**, 023035 (2020).
27. Y. Jiang *et al.*, *Microsc. Microanal.* **23** (suppl. 1), 222–223 (2017).
28. S. Koho *et al.*, *Nat. Commun.* **10**, 3103 (2019).
29. H. Emamy, O. Gang, F. W. Starr, *Nanomaterials* **9**, 661 (2019).
30. E. Auyeung *et al.*, *Nature* **505**, 73–77 (2014).
31. L. Li *et al.*, *Proc. SPIE* **10389**, 103890U (2017).
32. Y. Romano, M. Elad, P. Milanfar, *SIAM J. Imaging Sci.* **10**, 1804–1844 (2017).
33. K. Dabov, A. Foi, V. Katkovnik, K. Egiazarian, *IEEE Trans. Image Process.* **16**, 2080–2095 (2007).
34. Avizo 9 [computer software], ThermoFisher Scientific Inc.
35. Dragonfly 2020.2 [computer software], Object Research Systems (ORS) Inc.

ACKNOWLEDGMENTS

We thank the Imaging Facility of CUNY Advanced Science Research Center for instrument use and technical assistance. This research used resources of the Center for Functional Nanomaterials and the Hard X-ray Nanoprobe Beamline (HXN) at 3-ID. Small-angle scattering was collected at the Complex Matter Scattering (CMS) instrument at 11-BM and Soft Matter Interfaces (SMI) at 12-ID of the National Synchrotron Light Source II, which are part of the US Department of Energy, Office of Science facilities at Brookhaven National Laboratory under contract DE-SC0012704. **Funding:** The DNA-assembly work and defect analysis were supported by the US Department of Energy, Office of Basic Energy Sciences, grant DE-SC0008772. The inorganic templating and structural analysis work were supported by the US Department of Defense, Army Research Office, W911NF-19-1-0395. **Author contributions:** A.M., H.Y., and O.G. conceived and designed the experiments. A.M. performed the assembly and mineralization experiments. A.M., B.M., H.Y., and X.H. prepared and contributed to sample preparation for the beamline. A.M., H.Y., X.H., and Y.S.C. performed SHXM experiments and assisted in data preparation. A.M., H.Y., and H.E. performed data processing and analysis. A.M., H.Y., and O.G. wrote the paper. O.G. and H.Y. supervised the project. All authors discussed the results, assisted with assembling figures, and reviewed or edited the manuscript. **Competing interests:** The authors claim no competing Interests **Data and materials availability:** Synchrotron data files will be made available upon reasonable request to the corresponding authors. Ptychography (*26*) and PyXRF (*31*) software packages are available via GitHub https://github.com/NSLS-II/ptycho_gui, https://github.com/NSLS-II/PyXRF. All data needed to evaluate the conclusions are present in the paper and the supplementary materials.

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abk0463
Materials and Methods
Supplementary Text
Figs. S1 to S25
Movies S1 to S6
References (*32*–*35*)

28 June 2021; accepted 8 March 2022
10.1126/science.abk0463

Cube Rack Reader
The Ziath DataPaq Cube Rack Reader is a 2D data-matrix tube camera reader with a scan-and-decode time of 1–2 seconds on a normal computer. The Cube can read all racks on the market, including Cryoboxes and SBS (Society for Biomolecular Screening) racks—even 384 racks. The reader is easy to set up, arrives fully calibrated, and is ready to read all makes of SBS format and Cryobox 2D barcoded racks and tubes. Our patented DataPaq Cube is a camera-based instrument and is significantly quicker than scanner-based 2D barcode readers: Just one second and you can load the next rack! Our DataPaq software makes it simple to export data to Excel, XML, or text, and scanned images can also be saved. The DataPaq software can connect with Oracle, SQL Server, MySQL, Postgres, and other databases.
Ziath
For info: +1-858-880-6920
www.ziath.com

Deep-Well Plates for Biobanking Specimen Storage
Porvair Sciences' 96-well polypropylene microplates provide the perfect vessel for biobanks and biorepositories looking to store their valuable specimens and maintain specimen integrity over prolonged periods of time. We design and manufacture 96-well round deep-well plates with a 2-mL liquid capacity per well in an automation-compatible ANSI/SLAS footprint. Measuring just 45 mm in height, this innovative design prevents locking when stacked and enables easy heat sealing. Manufactured under class 100,000 conditions from ultrapure-grade polypropylene, our extensive range of 24-, 48-, and 96-well deep-well plates are certified as RNase/DNase free and contain no measurable contaminants that could otherwise leach out and affect biological specimens. For long-term storage at −80ºC, Porvair deep-well plates can be heat sealed with a wide choice of foils and seals, including DMSO-safe seals, using an Ultraseal range thermal sealer to provide high-integrity biological stored specimens.
Porvair Sciences
For info: +1-800-552-3696
www.microplates.com/deep-well-round

Reduced Growth Factor Basement Membrane Extract
Cultrex UltiMatrix Reduced Growth Factor Basement Membrane Extract (RGF BME) meets the cell-culture scaffolding demands of stem-cell and organoid researchers by delivering an extracellular matrix substrate that contains a high protein concentration, optimized stiffness, reduced growth factor composition, and consistent performance in 3D and 2D cell-culture applications. With rigorous quality-control testing against standard and difficult-to-grow 3D tissue culture, Cultrex UltiMatrix has proven resilience that supports dome formation for organoid culture, ultralow adhesion embedding for spheroids, and thin-layer coating for embryonic stem cell (ESC) or induced pluripotent stem cell (iPSC) expansion and maintenance. Be the first to enhance the performance and consistency of your organoid, spheroid, and pluripotent stem cell cultures by using the newest and most optimized Cultrex BME matrix from R&D Systems.
R&D Systems
For info: +1-800-343-7475
www.rndsystems.com

T Cell Metabolic Profiling Kit
The Agilent Seahorse XF T Cell Metabolic Profiling kit allows for robust, accurate measurements of both glycolytic and mitochondrial activities in T cell populations, providing a complete picture of T cell energy metabolism. These measurements can be linked to antitumor properties of T cell therapy products and are therefore valuable in designing and optimizing therapy development processes to improve T cell persistence or avoid exhaustion in the tumor microenvironment. The kit not only provides improved reagents, but also features a streamlined assay workflow, reducing assay preparation time and minimizing the need for uncoupler optimization. The XF T Cell Metabolic Profiling kit is also integrated with Wave Pro and Seahorse Analytics software to simplify data analysis, visualization, and interpretation. Each assay kit contains sufficient materials for six full-plate tests and is available in two packaging sizes. The XF kit is for use with Agilent Seahorse XF Pro and XFe/XF96 analyzers. The XFp kit is for use with the Agilent Seahorse HS Mini and XFp analyzers.
Agilent
For info: +1-800-227-9770
www.agilent.com

HLA-Typed CD34+ Cells for Humanized Mouse Models
Lonza provides human cord blood CD34+ hematopoietic stem cells (CB-CD34+ HSCs) in large batch sizes, meeting a critical and rapidly expanding market need. Lots in a range of sizes are also now available with high-resolution human leukocyte antigen (HLA) type information, removing the requirement for cumbersome HLA screening after lot purchase. CB-CD34+ HSCs are the preferred cell choice for creating humanized mouse models, which are critical for preclinical safety testing of a range of immunotherapies. Customers can create larger mouse model cohorts of the exact HLA type they need, expanding testing throughput capabilities and unlocking predictive results more quickly and at a significantly lower cost. The breadth of Lonza's inventory will also allow researchers to obtain all their CB-CD34+ HSCs from a single supplier, ensuring consistency and reliable quality in their processes. Lonza's large cell lots are guaranteed to be >90% pure, contain ≥2 million viable cells per lot, and come complete with a certificate of analysis. Cell customers will also receive Lonza's renowned global technical support, ensuring they can quickly overcome hurdles and achieve optimal outcomes in their mouse model creation.
Lonza
For info: +41-(0)-61-316-81-11
www.lonza.com/cd34cells-with-hla-information

Electronically submit your new product description or product literature information! Go to www.science.org/about/new-products-section for more information.

Newly offered instrumentation, apparatus, and laboratory materials of interest to researchers in all disciplines in academic, industrial, and governmental organizations are featured in this space. Emphasis is given to purpose, chief characteristics, and availability of products and materials. Endorsement by *Science* or AAAS of any products or materials mentioned is not implied. Additional information may be obtained from the manufacturer or supplier.

By **Alexandria Hughes**

In search of plan B, I found plan A

On the day of my first interview for a postdoctoral position, I was excited to finally discuss the detailed research ideas and questions I had feverishly prepared in the preceding weeks. When the principal investigator (PI) instead asked me, "If academia doesn't work out, what is your plan B?" I froze. This hadn't been part of my interview prep. I knew few postdocs get tenure-track positions and how important it was to have a backup plan, yet I did not. I don't remember what answer I eventually sputtered out. Over postinterview beers with my partner, I gazed into my swirling pint like a psychic's crystal ball. "Well, Alex, what is your plan B?" I asked myself. "Do you really have no idea what else you might like to do?"

I hit LinkedIn in search of inspiration. My approach was similar to a kid in an arcade with a messy fistful of tickets: What can I get with this neuroscience Ph.D.? I searched for jobs that matched my research experience—familiarity with molecular techniques, writing, some programming—but the results were overwhelmingly varied. After a couple of weeks with 50 tabs open, I didn't feel any closer to finding a plan B.

It slowly dawned on me that I didn't just need to identify a plan B job, but an entire plan B career path. It was time to close the tabs and take a more fundamental approach.

I began to think deeply about the work experiences I had most enjoyed up to that point. I remembered that in college, I loved working in the math department's drop-in tutoring center, helping students with any problems they brought through the door. In grad school, I felt the same purposeful warmth when I worked with peers to develop ways to measure biological phenomena. In my own research, I was more interested in rigorously applying methods than deeply investigating a particular research area. As I thought about it more, I realized helping other scientists learn from their data was more alluring than leading a lab of my own. In search of my backup plan, I instead found a new contender for plan A: helping scientists make their own science better by providing quantitative support as a statistician.

I wasn't ready to close the door on an academic career—perhaps I could fit statistical consulting into my work as a PI. But I knew that either way, I would benefit from additional training to broaden the analytic skills I had developed during my Ph.D. To keep the doors to both plans propped open while I deliberated, I decided to pursue a master's in statistics part-time during my postdoc.

"Helping other scientists learn from their data was more alluring than leading a lab of my own."

It was a grueling balance of research and study. But the further I advanced in my statistics degree, the more certain I became that I had found what I wanted to do. I returned to LinkedIn, this time with a steadier hand and clear goals: to meet statisticians and learn about opportunities.

I applied for jobs that I felt underqualified for, having not yet completed my statistics degree—and in some cases was rejected without an interview. But I also got some encouraging responses. Interviewers from outside academia told me they valued my Ph.D.; it showed I know how to solve problems and can drive projects forward. They were also impressed that I had taken stock of my goals and changed direction. I eventually found a perfect union of my knowledge and interests as a biostatistician in public health at a research firm, and I'm on track to finish my statistics degree in 2023.

As for that first postdoc interview, apparently I didn't blow it with my sputtering about my career plan B. The PI still offered me a position, which I accepted. And despite my angst in the moment, I'm grateful he challenged me to think about my career development. It was just the push I needed. ■

Alexandria Hughes is a senior biostatistician at Westat in Washington, D.C. Send your career story to SciCareerEditor@aaas.org.

ILLUSTRATION: ROBERT NEUBECKER

CALL FOR PAPERS

SPJ
SCIENCE PARTNER JOURNALS

spj.sciencemag.org/bmef

BMEF

OPEN ACCESS

BMEF is a **Science Partner Journal** distributed by the **American Association for the Advancement of Science** (AAAS) in collaboration with the **Suzhou Institute of Biomedical Engineering and Technology, Chinese Academy of Sciences** (SIBET CAS). *BMEF* serves the multidisciplinary community of biomedical engineering by publishing breakthrough original Research Articles, Rapid Reports, Reviews, Perspectives, and Editorials. The journal also publishes research in the fields of pathogenic mechanisms as well as disease prevention, diagnosis, treatment, and assessment.

The Science Partner Journals (SPJ) program was established by the American Association for the Advancement of Science (AAAS), the nonprofit publisher of the *Science* family of journals. The SPJ program features high-quality, online-only, open access publications produced in collaboration with international research institutions, foundations, funders and societies. Through these collaborations, AAAS expands its efforts to communicate science broadly and for the benefit of all people by providing top-tier international research organizations with the technology, visibility and publishing expertise that AAAS is uniquely positioned to **offer as the world's largest general science membership society**.

Submit your research to *BMEF* today!
Learn more at: **spj.sciencemag.org/bmef**

ARTICLE PROCESSING CHARGES WAIVED UNTIL 2023

SCIENCE FOR HUMANITY

AAAS | ANNUAL MEETING

The 2023 AAAS Annual Meeting will be held in-person in Washington, D.C. and online March 2-5, 2023. The meeting will highlight groundbreaking multi-disciplinary research that advances knowledge and responds equitably to the needs of humanity. Submit a proposal for one of these types of sessions before June 16, 2022:

Scientific Session Panels

Experts from different facets of the science and technology community assemble to compare notes during discussions about groundbreaking multi-disciplinary research that advances knowledge and responds to the needs of society.

10-Minute Lightning Talks

These live in-person short presentations offer individuals the opportunity to offer data and insights on any sci/tech topic relevant to those attending the AAAS meeting—discoveries, innovation, or policy.

Workshops

These instructional or informational sessions will highlight opportunities or resources available to enhance or augment careers paths or advocacy efforts.

aaas.org/meetings | #AAASmtg