Catching speciation in action *pp. 157 & 224*

Gene therapy's long journey to success *p. 175*

Buried water ice sheets on Mars *p. 199*

Science

$15
12 JANUARY 2018
sciencemag.org
AAAS

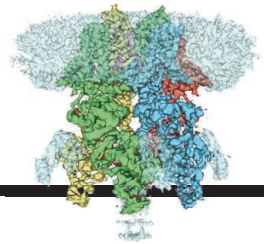
A MODERN MONSTER

The lasting legacy of Frankenstein

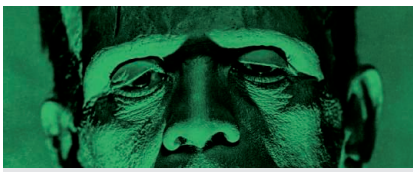
pp. 137 & 146

CONTENTS

12 JANUARY 2018 • VOLUME 359 • ISSUE 6372



160, 228, & 237
How a channel senses calcium



ON THE COVER

The cover illustration reimagines the monster first conjured in Mary Shelley's 1818 novel. Two centuries later, the story of how the fictional Dr. Frankenstein built a creature from scavenged body parts and unleashed it on the world has lost none of its power to unsettle, and Frankenstein has become a byword for scientific overreach. See pages 137 and 146. *Illustration: Craig & Karl*

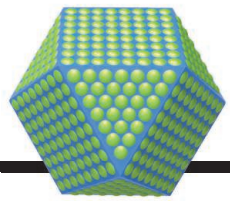
CREDITS: (LEFT TO RIGHT) UNIVERSAL PICTURES/SCREENPROD/PHOTONONSTOP/ALAMY STOCK PHOTO; NASA; H. FORD (JHU); AUTZEN *ET AL.*

156

NEWS

INSIGHTS

CONTENTS



206
Pores large and small

12 JANUARY 2018 • VOLUME 359 • ISSUE 6372

CREDITS: (TOP TO BOTTOM) YOSSI YOVEL; SHEN *ET AL.*

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. Domestic Individual membership, including subscription (12 months): $165 ($74 allocated to subscription). Domestic institutional subscription (51 issues): $1808; Foreign postage extra: Mexico, Caribbean (surface mail) $55; other countries (air assist delivery): $89 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; bulk rate on request. The identification code for *Science* is 0036-8075. *Science* is indexed in the *Reader's Guide to Periodical Literature* and in several specialized indexes.

Published by AAAS

Science
AAAS
1200 New York Avenue, NW, Washington, DC 20005
Clarendon House, Clarendon Road, Cambridge, UK CB2 8FH

SUBSCRIPTION SERVICES For change of address, missing issues, new orders and renewals, and payment questions: 866-434-AAAS (2227) or 202-326-6417, FAX 202-842-1065. Mailing addresses: AAAS, P.O. Box 96178, Washington, DC 20090-6178 or AAAS Member Services, 1200 New York Avenue, NW, Washington, DC 20005

INSTITUTIONAL SITE LICENSES 202-326-6730 **REPRINTS:** Author Inquiries 800-635-7181 **COMMERCIAL INQUIRIES** 803-359-4578 **PERMISSIONS** 202-326-6765, permissions@aaas.org **AAAS Member Central Support** 866-434-2227 www.aaas.org/membercentral.

INFORMATION FOR AUTHORS See pages 624 and 625 of the 5 February 2016 issue or access www.sciencemag.org/authors/science-information-authors

EDITORIAL

Frankenstein lives on

It was 200 years ago that Mary Shelley's *Frankenstein; or, The Modern Prometheus* was published. Over the decades, this gothic tale has captured the popular imagination through the numerous theater productions and films it inspired. The story is commonly taken to imply a dire warning about the dangers of scientific hubris. Just mention the name Frankenstein and laypersons think of scientists "playing God." In the common view, the inevitable consequence of Frankenstein's alleged transgression—bestowing life on inanimate matter—was that he created a monster that would wreak havoc on his family and friends. Frankenstein's name is repeatedly invoked in debates about emerging technologies like biotech, nanotech, synthetic biology, and artificial intelligence. However, the view of Shelley's story as a cautionary tale about scientific hubris, although dominant, is only one possible interpretation. Her novel, actually, is a multilayered story full of ambivalences and much subtler than most Hollywood versions. It naturally lends itself to diverse interpretations.

Mary Shelley

"...researchers must...assume responsibility for the vicissitudes of their technological creations..."

An alternative view holds that Frankenstein's moral transgression was not that he undertook the hubristic attempt to create a human being, but rather that he, once his work finally met with success, ran away from his creature and left him without any parental care. Left to fend for himself, the creature sought friendship and companionship from others, but again and again encountered rejection due to his hideous appearance. Only then did he become a monster intent on violently avenging the injustices done to him. On this view, then, Frankenstein's moral shortcoming was that he did not assume responsibility for his own creature and failed to give him the care he needed.

The American philosopher of technology Langdon Winner has used this interpretation as a clue for dealing more responsibly with new technologies in general. He holds that researchers must be willing to assume responsibility for the vicissitudes of their technological creations, help them to acquire a suitable role in society, and provide adequate follow-up care if necessary. With so much emphasis nowadays on responsible innovation, this message finds wide resonance. Under the heading "Love your monsters," it is also supported by the French philosopher Bruno Latour. A group of social researchers from Arizona State University, led by David Guston, concurs: "The lack of care for new creations is what ultimately destroys us, not the creations themselves." This group considers the focus on hubris unhelpful in furthering responsible innovation. Rather, the emphasis should be on openness and responsiveness to public concerns, and on the anticipation and timely modulation of possible negative effects of new technologies.

Researchers, especially in the life sciences, are understandably anxious about being tainted with the "F-word." To the charge of playing God, they usually react by professing humility, but some take a more defiant attitude. James Watson once famously declared: "If scientists don't play God, who else is going to?" Craig Venter is on record as having said about his synthesis of a microbe with a minimal genome: "Shelley would have loved this!" Such responses might be perceived as arrogant, but they are also a welcome challenge to a quasi-theological argument. Kevin Esvelt, coinventor of CRISPR-based gene drives, aligns with the proponents of responsible innovation in his pleas for openness and public engagement. For him, it is really hubris when researchers work in secret like Frankenstein and fail to seek advice from others.

The question remains whether indeed technological creations themselves can never bring us down, assuming due care is exercised. What about George Church's ambitious plan to resurrect Neandertal man? Would everything be okay if he just provided adequate fatherly care for a Neandertal baby born from a willing surrogate mother? Or would resurrection be not such a good idea after all?

–Henk van den Belt

Henk van den Belt is assistant professor of philosophy at Wageningen University, Netherlands. henk.vandenbelt@wur.nl

10.1126/science.aas9167

$2^{77,232,917} - 1$

The largest known prime number, with more than 23 million digits—1 million more than the previous record—was discovered in December 2017 by an electrical engineer, Jonathan Pace of Germantown, Tennessee, working with a consortium of volunteers.

IN BRIEF

Edited by **Jeffrey Brainard**

Barnacles reveal whales' migrations

Scientists studied barnacles like these on a humpback whale's head.

Want to know where a whale has been? Just ask the barnacles on its head. Since their earliest days, whales have hosted invertebrate hitchhikers, and certain species of barnacles favor certain species of whales. That loyalty is enabling paleobiologists to reconstruct ancient whale migrations, researchers reported 6 January at the annual meeting of the Society of Integrative and Comparative Biology in San Francisco, California. At the University of California, Berkeley, Seth Finnegan and his graduate student Larry Taylor first demonstrated that as coronulid barnacles grow on humpback whale hosts, the layers they add to their protective plates carry chemical traces of the surrounding ocean. Each barnacle offers clues about its host's migration history—and about conditions in the oceans they visited. Taylor's analysis of fossil barnacles indicates that humpbacks have traveled long distances for at least the past 4.5 million years.

PHOTO: TONY WU/MINDEN PICTURES

AROUND THE WORLD

Climate panel revived

ALBANY | New York state, in collaboration with several other states and Columbia University, is reviving an advisory committee on climate change that President Donald Trump's administration disbanded. The move is aimed at helping local governments prepare for global warming. The federal advisory committee had been slated to provide guidance based on scientific findings of the latest National Climate Assessment, which Congress requires agencies to issue every 4 years. But in August 2017, the administration allowed the panel's charter to expire. Climate scientist Richard Moss, who chaired the federal panel, will lead the revived committee, which will also include 10 of the original 15 panelists. The group expects to release a draft report by June. It will focus on presenting data that engineers can use to protect public infrastructure.

China plans AI research park

BEIJING | China's ambitions to become a world leader in artificial intelligence (AI) will bring a new incubator of innovation to its capital's western suburbs. The government plans to build a $2.1 billion R&D park for AI there within 5 years, according to state-run news agency Xinhua. The facilities will focus on high-speed data, cloud computing, biometric identification, and deep learning. The park's infrastructure will include a fifth-generation mobile network and a supercomputer. "This seems like yet another strong indication from China that they want to invest heavily in AI," said Subbarao Kambhampati, president of Association for the Advancement of Artificial Intelligence in Palo Alto, California. Chinese authors contributed 43% of all content in the top 100 AI journals in 2015.

Pfizer retreats from neuroscience

NEW YORK CITY | One of the world's largest pharmaceutical companies will abandon its pursuit of new treatments for neurological diseases. On 6 January, Pfizer announced it will halt most of its early-stage neuroscience research and development programs—most

notably its efforts to develop drugs for Alzheimer's and Parkinson's diseases. The move comes with 300 layoffs, although the company will continue a handful of neuroscience projects, including some rare disease research and clinical testing of chronic pain medications. Neuroscience has long been a risky area for pharmaceutical companies, and several promising Alzheimer's treatments have recently floundered in late-stage clinical trials. But some analysts remain optimistic about neuroscience more broadly, citing an influx of investment in 2017.

Better typhoid shot gets a boost

GENEVA, SWITZERLAND | Doses of a new-generation typhoid vaccine that is longer lasting and more effective than existing ones will soon be available for babies and others in countries hard hit by the disease. An intestinal bacterial disease spread by contaminated food and water, typhoid infects some 11 million to 20 million people a year. Children are the most susceptible. The World Health Organization (WHO) in December 2017 "prequalified" the vaccine, which fuses a typhoid bacterial protein to an immune-stimulating molecule, after concluding that it is safe and effective. That means the vaccine, developed by Bharat Biotech in India, can now be purchased by United Nations agencies and used globally. Even before WHO approval, Gavi, the Vaccine Alliance, approved $85 million to support the introduction of the vaccine in poor countries, which is expected to begin in 2019. WHO recommends the vaccine for routine use in children over 6 months of age.

Neutrino detector expands

ANTARCTICA | Physicists working at the South Pole this week completed an expansion of the Askaryan Radio Array (ARA), which aims to spot ultra–high-energy neutrinos from space. On rare occasions, neutrinos passing through polar ice create detectable cascades of electrically charged particles and telltale radio signals. Using a water drill, scientists lowered strings of receivers 200 meters into the ice, adding two new multistring stations to three previously installed. The ARA already covers six times the area of the neighboring IceCube Neutrino Observatory, which spans 1 square kilometer and detects neutrinos using light signals. The ARA is sensitive to neutrinos at least five times more energetic than those seen by IceCube. Ultimately, physicists hope to expand the ARA to 37 stations covering 200 square kilometers.

PHOTO: RADIUS IMAGES/ALAMY STOCK PHOTO

TRUMP TRACKER

Political review of grants

A senior adviser at the U.S. Department of the Interior will review certain grants and cooperative agreements of $50,000 or more to universities and nonprofit groups to ensure they "better align" with the priorities of President Donald Trump's administration. The move, which follows a similar decision last summer by the Environmental Protection Agency, was first reported by *The Washington Post*, citing a 28 December 2017 memo. The new approval process appears to be without precedent within the Interior Department, the *Post* reported. The priorities of Secretary of the Interior Ryan Zinke include securing the U.S. southern border and "utilizing our natural resources," according to an accompanying memo.

Census furor

Social scientists and civil rights advocates are urging the Census Bureau to reject a request by the Department of Justice to add a question about citizenship status to the 2020 census. They fear that such a question will decrease participation by undocumented immigrants in the national head count and that a last-minute addition to the questionnaire could undermine its quality, as well. A lower initial response rate could also require the agency to hire more field workers to track down those who haven't answered and could boost costs by hundreds of millions of dollars. A citizenship question will "destroy any chance for an accurate count," Vanita Gupta, president and CEO of The Leadership Conference on Civil and Human Rights in Washington, D.C., said in a statement.

Environmental regulations rollback

The administration continued its efforts to roll back environmental regulations finalized under former President Barack Obama. In late December 2017, it instructed federal land managers to eliminate rules that place tighter controls and reporting requirements on energy companies that use hydraulic fracturing, or fracking, on public lands. A separate directive required the government to cease fining energy companies for activities that incidentally kill small numbers of birds and other animals protected by federal law. The administration also rescinded guidance documents that encourage agencies to consider how industrial development might affect climate change. Environmental groups say they will go to court to try to overturn or delay the actions.

New plan to expand offshore drilling

The administration on 4 January said it intends to lift an existing ban on selling offshore oil leases for vast areas along the eastern, western, gulf, and arctic coasts of the United States. The unprecedented expansion is intended to create jobs and assure U.S. "energy dominance" in coming decades, the White House said. But the new plan, which will take at least a year to implement, immediately drew bipartisan condemnation from politicians in many coastal states, as well as environmentalists who fear potential spills and damage to resources important to tourism and fisheries.

Environmentalists worry about risks of a potential increase in offshore oil rigs.

STEM workers report inequities

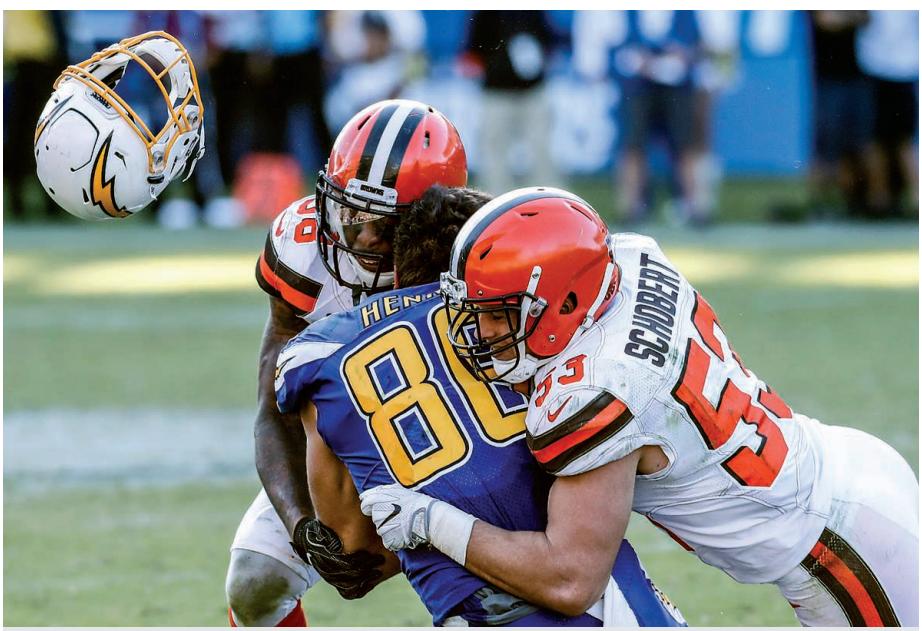
WASHINGTON, D.C. | Large percentages of women and members of minority groups employed in science jobs report workplace discrimination, a new survey found. The Pew Research Center surveyed 2344 U.S. workers in diverse fields of science, technology, engineering, and mathematics (STEM), including health care and school teaching. Fifty percent of women working in STEM jobs said they experienced gender-based discrimination at work, whereas only 19% of men in STEM jobs did. Experiences with discrimination, which included earning a lower salary and being treated as less competent, also varied among members of different races: Just over 60% of black STEM workers reported race-related discrimination in the workplace. About 40% of Asian and Hispanic STEM workers reported it, whereas only 13% of white STEM workers did, according to the survey results, released 9 January.

South Korea confronts Elsevier

SEOUL | Protesting a reported 4.5% price hike, some 100 South Korean universities have refused to renew their contracts with Dutch publishing giant Elsevier for access to its ScienceDirect database. Contracts are typically renewed in December. Negotiators "narrowed the gap" on costs, says Lee Chang Won, secretary general of the Korea University & College Library Association in Seoul, which together with the Korean Council for University Education leads a consortium representing the universities. But when talks broke down at the end of the year, the group organized the contract boycott. Talks are continuing, Lee says. Access to ScienceDirect in South Korea will continue at least until 12 January, Elsevier says.

Couple sues rare disease firms

RENO, NEVADA | The parents of 13-year-old twins with a lethal, inherited disease of cholesterol metabolism on 5 January sued several companies that developed a therapy for Niemann-Pick type C-1. In a suit filed in federal district court in Nevada, Christine and Hugh Hempel and their private firm, Solution Therapeutics, demand at least $100 million from the firms—Cydan Development of Cambridge, Massachusetts, and Vtesse, part of Sucampo Pharmaceuticals, a Rockville, Maryland, company. The Hempels allege that the firms misappropriated data, business plans, and other know-how developed using their daughters, who in 2010 were the first to receive a form of the drug cyclodextrin by spinal injection. Sucampo is conducting an advanced clinical trial of its cyclodextrin, VTS-270, which the Hempels call "virtually identical" to their version. Sucampo and Cydan said in statements they will defend their positions vigorously. Cydan added, "Their claims are without merit."



Repeated impacts to football players' heads have been associated with the degenerative brain disease chronic traumatic encephalopathy.

PHOTO: KEVIN TERRELL/AP PHOTO

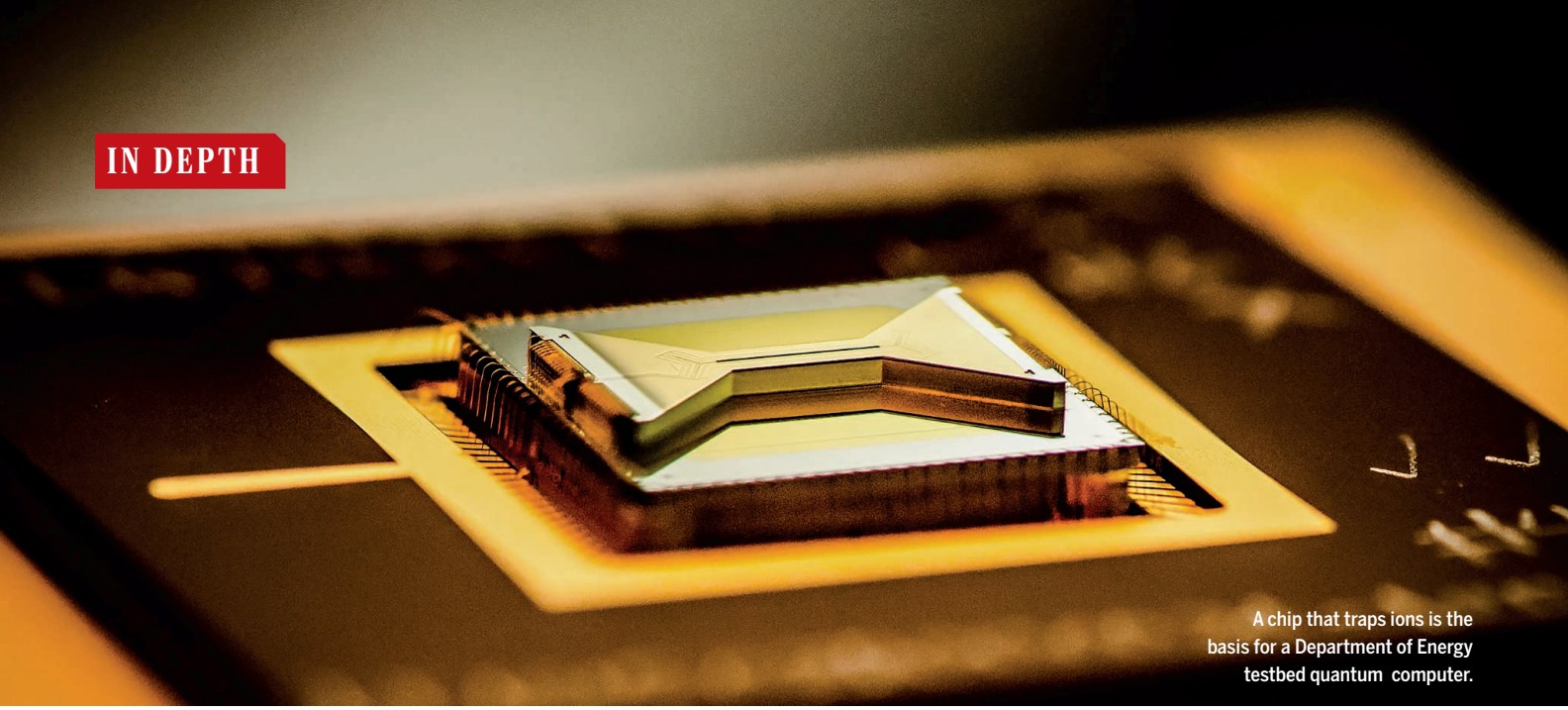
National Football League hands off research cash

The National Football League (NFL) on 5 January finally provided $16 million it promised in 2012 to study concussions and other football-related injuries. The league originally pledged the money for National Institutes of Health (NIH) research, but in 2015 refused to supply it after NIH awarded a grant to researchers led by Robert Stern at Boston University. In 2014, Stern had filed an affidavit criticizing NFL's estimated $1 billion settlement with retired players for failing to compensate men with a wide range of behavioral symptoms associated with repetitive head trauma. NIH ended up funding Stern's study itself. The $16.3 million that NFL released this month will now go to a Department of Defense study to track the effects of concussions in thousands of college athletes ($7.65 million); Transforming Research and Clinical Knowledge in Traumatic Brain Injury, or TRACK-TBI, an NIH-funded study led by the University of California, San Francisco, which aims to follow 3000 people after traumatic brain injury ($7.65 million); and NIH's National Institute on Aging for research on dementia and cognitive function ($1 million).

NEWSMAKERS

Swedish researcher disciplined

UPPSALA, SWEDEN | The Swedish Research Council on 8 January canceled a 4-year, $403,000 research grant awarded last October to ecologist **Peter Eklöv** of Uppsala University (UU), one of two authors on a paper published in *Science* in 2016. The paper, about the dangers posed by microplastics to fish, was later judged to include fabricated data. The council also barred Eklöv from applying for new grants for 2 years. In December 2017, a UU investigation concluded that Eklöv's role in the fraud didn't amount to misconduct because the data were made up by his co-author Oona Lönnstedt. But the university said both researchers engaged in misconduct by performing experiments on perch and pike without an ethical permit (*Science*, 15 December 2017, p. 1367). Eklöv declined to comment. Another Swedish funding agency, FORMAS, canceled a $367,000 grant to Lönnstedt in September 2017.



A chip that traps ions is the basis for a Department of Energy testbed quantum computer.

COMPUTATION

DOE pushes for useful quantum computing

As quantum supremacy nears, so does a desire to apply machines to science problems

By **Adrian Cho**

The U.S. Department of Energy (DOE) is joining the quest to develop quantum computers, devices that would exploit quantum mechanics to crack problems that overwhelm conventional computers. The initiative comes as Google and other companies race to build a quantum computer that can demonstrate "quantum supremacy" by beating classical computers on a test problem. But reaching that milestone will not mean practical uses are at hand, and the new $40 million DOE effort is intended to spur the development of useful quantum computing algorithms for its work in chemistry, materials science, nuclear physics, and particle physics.

"We are looking for algorithms that can advance the science," says Stephen Binkley, acting director of DOE's $5.4 billion Office of Science in Washington, D.C., who in a 29 November 2017 open letter urged researchers to submit proposals for such work.

The U.S. government already spends about $250 million per year on quantum computing, mostly through the Army Research Office, says Christopher Monroe, a physicist at the University of Maryland in College Park and co-founder of the quantum computing startup IonQ. But the DOE money will go mostly to its national laboratories. Monroe says researchers there can play a leading role in developing the machines. "Industry can't do it because they don't have the people, and academics can't do it because they don't build things."

Whereas a conventional computer manipulates bits that can be set to either 0 or 1, a quantum computer employs quantum bits or qubits that, bizarrely, can be set to 0 and 1 at the same time. A qubit can be a patch of superconducting metal that can be electrically charged to encode 1, uncharged to encode 0, or both charged and uncharged at the same time. Trapped ions, which can spin in opposite directions or both ways at once, can also serve as qubits. With their two-ways-at-once capability, just 300 qubits could simultaneously encode more numbers than there are atoms in the observable universe.

However, it is the way quantum computers solve problems that accounts for their power—and their limitations. Problems can be encoded so that potential solutions correspond to different quantum waves sloshing through the qubits. Set things up so the waves interfere the right way, and the wrong solutions will cancel one another while the right solution pops out. That's how a quantum computer could quickly factor large numbers, potentially enabling it to crack current internet encryption protocols. But the approach cannot aid every computation.

For instance, quantum computers won't help analyze the billions of records of individual particle collisions produced by atom smashers such as the Large Hadron Collider in Switzerland, says James Amundson, a computational physicist at Fermi National Accelerator Laboratory in Batavia, Illinois. Each of the records is easy to analyze, so they need only to be fed through an army of ordinary computers working in parallel, Amundson says. A quantum computer can't speed up the process.

Still, the machines hold great promise for some problems, researchers say, such as those that involve modeling or simulating inherently quantum mechanical processes. In chemistry, for example, enzymes called nitrogenases catalyze the reactions that enable nitrogen-fixing bacteria to turn nitrogen from the air into a form that plants can use. No conventional computer can calculate exactly how the process works,

A quantum computing to-do list

Researchers have several general ideas for scientific applications of quantum computers.

FIELD	TASK
Chemistry	Calculate molecules' energies and structures, model catalysis.
Materials science	Design novel materials from the atom up.
Nuclear physics	Calculate energies and structures of nuclei and particles such as protons.
Particle physics	Optimize search for subtle signals.

PHOTO: S. DEBNATH AND E. EDWARDS/JQI

but a quantum computer could, says Wibe de Jong, a computational chemist at Lawrence Berkeley National Laboratory in Berkeley, California. "There are lots of catalytic processes that are still very hard to model because of the computational complexity," he says.

Quantum computers might also aid in the design of materials from their atomic constituents on up. And they could help predict how the superdense matter in neutron stars behaves or how a proton breaks up during a particle collision. Such applications all involve the interplay of the quantum waves that describe subatomic particles. Tracking the oscillating waves swamps a conventional computer, but a quantum computer handles that aspect of a calculation automatically, explains Martin Savage, a nuclear theorist at the University of Washington in Seattle.

Researchers have only begun to figure out how to map such problems onto a quantum computer's qubits. To speed the process, DOE in September 2017 launched two testbeds to enable designers and scientists to work together on approaches to quantum computation. At the Berkeley lab, physicist Irfan Siddiqi and colleagues aim to build their own 64-qubit quantum computer using superconducting qubits. Feedback from users will influence their designs, such has how the qubits are arranged and connected with one another on a chip, Siddiqi says.

In contrast, a testbed at Oak Ridge National Laboratory in Tennessee will provide remote access to existing machines at IBM and IonQ. That approach should spark the same sort of "co-design" without requiring Oak Ridge researchers to build a machine from scratch, says Raphael Pooser, a quantum information scientist at Oak Ridge. It also more closely resembles the way DOE develops its supercomputers in partnership with industry, he says.

In the meantime, commercial machines are getting more powerful. This week, researchers at Google's laboratory in Santa Barbara, California, began testing a 50-qubit chip they think will achieve quantum supremacy, although the experiment could still take months. Yet some researchers worry that such a demonstration may mislead the public into thinking that scientists have reached the end of the road in developing a useful quantum computer. "It's not even the beginning of the road," Siddiqi says.

John Martinis, the physicist who leads Google's effort, says the company "understands that quantum supremacy is a great milestone and that it will take longer, perhaps much longer, to make something practical." DOE clearly agrees. ■

Extensive sampling shows that poliovirus lurks in many open sewers.

INFECTIOUS DISEASE

In Pakistan, surveillance for polio reveals a paradox

Cases are the lowest ever, but the virus persists in sewage

By **Leslie Roberts**

Just a year ago, poliovirus seemed on its last legs in Pakistan, one of its final strongholds. Polio cases were steadily falling, from 306 in 2014 to 54 in 2015, 20 in 2016, and, by last count, eight in 2017. Blood tests showed that, overall, immunity to the virus had never been higher, even among children aged 6 to 11 months, thanks to years of tireless vaccination campaigns. Surely, there were not enough susceptible kids to sustain transmission, and the virus would burn itself out within a year.

Unsettling new findings, however, show it is far from gone. In the most extensive effort in any country to scour the environment for traces of the virus, polio workers are finding it widely across Pakistan, in places they thought it had disappeared. They are wondering "just what the hell is going on" and how worried they should be, says epidemiologist Chris Maher of the World Health Organization (WHO) in Geneva, Switzerland, who runs polio operations in the eastern Mediterranean region. Does this mean the virus is more entrenched than anyone realized and is poised to resurge? Or is this how a virus behaves in its final days—persisting in the environment but not causing disease until it fades out?

"We have never had this level of environmental sampling anywhere else. We have nothing to compare it to," Maher says. "We don't understand the dynamic," agrees Michel Zaffran, who leads the Global Polio Eradication Initiative at WHO. "But we take it very seriously." In response to the sampling data, he and his colleagues are already changing their tactics—and their definition of success.

Along with Afghanistan and Nigeria, Pakistan is one of just three endemic countries—places where indigenous wild poliovirus has never been vanquished. With its dysfunctional government, unceasing violence, poverty, and huge numbers of people on the move, it may well be the toughest challenge for eradication. The border with Afghanistan is so porous that the two countries are considered one epidemiologic block in which the virus circulates freely. Conventional wisdom is that if Pakistan defeats polio, Afghanistan will soon follow. That could be the key to global eradication, as no virus has been detected in Nigeria for the past 15 months.

Since the eradication effort began in 1988, the gold standard for detecting poliovirus has been surveillance for acute flaccid paralysis (AFP)—finding and testing every child with a sudden weakness or floppiness in the arms or legs. The yearly case count has been the benchmark for success: After 12 months without a polio case, WHO has historically removed a country from the endemic list.

PHOTO: ©NEOC/PAK2017/ABID AHSAN

But as case numbers fell to today's low levels, AFP surveillance is no longer the only meaningful indicator. Only about one in 200 or 300 people infected with the virus becomes paralyzed; the rest show no symptoms but can still shed the virus in their stool and infect others. Environmental surveillance can detect that hidden virus.

Polio workers collect sewage samples, usually from open drainage ditches, and test them for virus. If the test is positive, that means someone in the catchment area is infected and actively excreting it. Pakistan now has 53 sampling sites, more than any other country. And at a time when cases are the lowest on record, 16% of samples from across the country are testing positive.

"It is extraordinary to have so much virus in sewage and so few cases," Zaffran says.

What makes the environmental samples so hard to interpret is that a catchment area may contain the combined feces of 50,000 or 100,000 people. "If you isolate a virus from a child, you know who is infected. When you find it in an environmental sample, you don't know if three people are infected or 3000," Maher explains.

One possible explanation for the disconnect is that AFP surveillance is missing cases. Maher doubts that the number is significant, but others suspect that too many children among the mobile populations, including the marginalized Pashtun minority, still aren't being vaccinated despite ramped up efforts to reach them. "I don't think polio is entrenched across Pakistan, but this last reservoir of 'people on the move' is sustaining the virus," says Steve Cochi, a polio expert at the U.S. Centers for Disease Control and Prevention in Atlanta.

Maher has another view. "My own suspicion is this is part of what we see at the end," he says. "The lack of cases means immunity is high, but because of the very difficult circumstance in Pakistan," the virus still has a tenuous hold. Ultimately, he says, "The virus will die out because it is not getting enough purchase."

The program is not taking any chances. The response to each positive environmental test is now as aggressive as to a case of paralysis. And the program is hammering the virus with repeated vaccination campaigns throughout the "low season," between December and May, when cold weather makes it tougher for the virus to survive. Whether the strategy works will become clear later this year when the weather turns warm. But one thing is certain: The absence of cases is no longer enough to declare victory over polio. Going forward, a country will not be considered polio-free until 12 months have passed without a case—or a positive environmental sample. ■

EARTH SCIENCE

Earth scientists list top priorities for space missions

Report urges NASA to choose medium-size Earth-observing satellites through competition to keep costs down

By **Paul Voosen**

Earth scientists hope a new priority setting effort will help them make the most of NASA's limited budget for satellite missions that watch over the planet. The so-called decadal survey, issued last week by the National Academies of Sciences, Engineering, and Medicine, lays out the community's consensus wish list, ranging from cloud monitoring to multiwavelength imaging—and recommends a strong dose of competition to keep costs down.

Thanks to an infusion of climate-focused spending from the administration of former President Barack Obama, NASA's earth science budget grew from $1.4 billion a decade ago to $1.9 billion last year, the largest of the agency's four science divisions. "We're not at the bottom of some pit like we were before," says Bill Gail, chief technology officer at the Global Weather Corporation in Boulder, Colorado, and co-chair of the committee that wrote the report.

> ***"... there's not enough money to do what we want to do."***
>
> **Waleed Abdalati**, University of Colorado

But even if the earth science budget remains flat—far from certain in the current political climate—"the simple fact is there's not enough money to do what we want to do," says committee co-chair Waleed Abdalati, director of the Cooperative Institute for Research in Environmental Sciences at the University of Colorado in Boulder. To identify the most cost-effective projects, the report recommends that NASA's earth science division set up a competition for medium-size, $350 million missions and identifies a slim list of five more expensive flagship missions that could also be open to competition.

The first decadal survey, in 2007, prioritized 15 explicitly defined missions. NASA centers had already started work on many of them, and some quickly ballooned in cost. "The assignment of missions didn't have the rigor that a competitive environment enforces," says Berrien Moore, vice president of weather and climate programs at the University of Oklahoma in Norman and co-lead of the 2007 survey. The new report sets mission goals without endorsing particular projects.

The first to fly, according to the report, should be a probe to monitor Earth's surface across a wide range of wavelengths, using a "hyperspectral" technique that detects reflective fingerprints invisible to the human eye, highlighting substances from chlorophyll to oil. The recommendation, which sets a cost cap of $650 million, is likely to give a boost to the Hyperspectral Infrared Imager, a proposed mission from NASA's Jet Propulsion Lab in Pasadena, California.

Two other major investments, each capped at $800 million, would target clouds and atmospheric particles called aerosols, the two biggest wild cards in climate change. The final two priorities are a $300 million mission to chart gravity variations—a successor to the recently ended Gravity Recovery and Climate Experiment (GRACE) and its replacement—and a $500 million satellite that would track tiny crustal movements using radar techniques.

The report also recommends that NASA solicit proposals for $350 million missions to monitor seven different parameters including greenhouse gases, ice elevations, and ocean surface winds. Three of the missions, to be called Earth System Explorers, would be selected for flight in the next 10 years. Abdalati says it's an opportunity to push scientists to develop "smaller, more agile systems."

It's far from certain that NASA will heed the recommendations, and Congress may not deliver the money to execute them. In its budget proposal last year, the administration of President Donald Trump sought to boost planetary science and cut earth science. Moore hopes that Senate opposition will forestall the cuts, and he thinks that both planetary and earth science will thrive. "I wouldn't be surprised to see both boats rise," he says. Reassuringly, Trump's nominee to lead NASA, Representative Jim Bridenstine (R–OK), has pledged to follow the decadal, but he still needs to be confirmed for the post. ■

PHOTO: YAMIL LAGE/AFP/GETTY IMAGES

CLIMATE ADAPTATION

Cuba's 100-year plan for climate change

Nation seeks assistance for project to strengthen coastal defenses and relocate villages

By **Richard Stone**, *in Havana*

On its deadly run through the Caribbean last September, Hurricane Irma lashed northern Cuba, inundating coastal settlements and scouring away vegetation. The powerful storm dealt Havana only a glancing blow; even so, 10-meter waves pummeled El Malecón, the city's seaside promenade, and ravaged stately but decrepit buildings in the capital's historic district. "There was great destruction," says Dalia Salabarría Fernández, a marine biologist here at the National Center for Protected Areas (CNAP).

As the flood waters receded, she says, "Cuba learned a very important lesson." With thousands of kilometers of low-lying coast and a location right in the path of Caribbean hurricanes, which many believe are intensifying because of climate change, the island nation must act fast to gird against future disasters.

Irma lent new urgency to a plan, called Tarea Vida, or Project Life, adopted last spring by Cuba's Council of Ministers. A decade in the making, the program bans construction of new homes in threatened coastal areas, mandates relocating people from communities doomed by rising sea levels, calls for an overhaul of the country's agricultural system to shift crop production away from saltwater-contaminated areas, and spells out the need to shore up coastal defenses, including by restoring degraded habitat. "The overarching idea," says Salabarría Fernández, "is to increase the resilience of vulnerable communities."

But the cash-strapped government had made little headway. Now, "Irma [has] indicated to everybody that we need to implement Tarea Vida in a much more rapid way," says Orlando Rey Santos, head of the environment division at Cuba's Ministry of Science, Technology and Environment (CITMA) here, which is spearheading the project. The government aims to spend at least $40 million on Project Life this year, and it has approached overseas donors for help. Italy was the first to respond, pledging $3.4 million to the initiative in November 2017. A team of Cuban experts has just finished drafting a $100 million proposal that the government plans to submit early this year to the Global Climate Fund, an international financing mechanism set up under the United Nations Framework Convention on Climate Change.

Many countries with vulnerable coastlines are contemplating similar measures, and another island nation—the Seychelles—has offered to collaborate on boosting coastal protection in Cuba. But Project Life stands out for taking a long view: It intends to prepare Cuba for climatological impacts over the next century. "It's impressive," says marine scientist David Guggenheim, president of Ocean Doctor, a nonprofit in Washington, D.C., that has projects in Cuba. "Cuba is an unusual country in that they actually respect their scientists, and their climate change policy is science driven."

Rising sea levels pose the most daunting challenge for Cuba. Over the past half-century, CITMA says, average sea levels have risen some 7 centimeters, wiping out low-lying beaches and threatening marsh vegetation, especially along Cuba's southern midsection. The coastal erosion is "already much worse than anyone expected," Salabarría Fernández says. Storms drive the rising seas farther inland, contaminating coastal aquifers and croplands.

Still worse is in store, even in conservative scenarios of sea-level rise, which forecast an 85-centimeter increase by 2100. According to the latest CITMA forecast, seawater incursion will contaminate nearly 24,000 square kilometers of land this century. About 20% of that land could become submerged. "That means several percent of Cuban land will be underwater," says Armando Rodríguez Batista, director of science, technology, and innovation at CITMA.

To shore up the coastlines, Project Life aims to restore mangroves, which constitute about a quarter of Cuba's forest cover. "They are the first line of defense for coastal communities. But so many

Habaneros wade through floodwaters near El Malecón after Hurricane Irma.

PHOTO: NASA/JPL/UNIVERSITY OF ARIZONA/USGS

mangroves are dying now," Salabarría Fernández says. Leaf loss from hurricane-force winds, erosion, spikes in salinity, and nutrient imbalances could all be driving the die-off, she says.

Coral reefs can also buffer storms. A Cuban-U.S. expedition that circumnavigated the island last spring found that many reefs are in excellent health, says Juliett González Méndez, a marine ecologist with CNAP. But at a handful of hot spots, reefs exposed to industrial effluents are ailing, she says. One Project Life target is to squelch runoff and restore those reefs.

Another pressing need is coastal engineering. Topping Cuba's wish list are jetties or other wave-disrupting structures for protecting not only the iconic Malecón, but also beaches and scores of tiny keys frequented by tourists whose spending is a lifeline for many Cubans. Cuba has appealed to the Netherlands to lend its expertise in coastal engineering.

Perhaps the thorniest element of Project Life is a plan to relocate low-lying villages. As the sea invades, "some communities will disappear," Salabarría Fernández says. The first relocations under the initiative took place in October 2017, when some 40 families in Palmarito, a fishing village in central Cuba, were moved inland.

Other communities may not need to pull up stakes for decades. But Cuban social scientists are already fanning out to those ill-fated villages to educate people on climate change and win them over on the eventual need to move. That's an easier sell in the wake of a major hurricane, Rodríguez Batista says. "Irma has helped us with public awareness," he says. "People understand that climate change is happening now." ■

PLANETARY SCIENCE

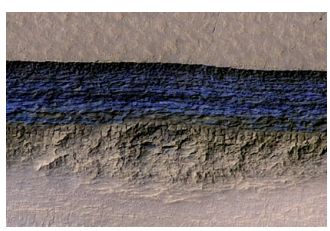
Cliffs of ice spied on Mars

Buried glaciers could be prime target for human exploration

By **Paul Voosen**

For more than a decade, Colin Dundas, a geologist at the U.S. Geological Survey in Flagstaff, Arizona, has had a daily routine: inspecting a dozen or so high-resolution images beamed back every day from the Mars Reconnaissance Orbiter (MRO). A few years ago, something surprising popped out from the planet's sea of rust: a pale sliver of blue.

What Dundas saw that day, and subsequently found at seven other sites, are steep cliffs, up to 100 meters tall, that expose what appears to be nearly pure ice. The discovery points to large stores of underground ice buried only a meter or two below the surface at surprisingly low martian latitudes. "This kind of ice is more widespread than previously thought," says Dundas, who, with his co-authors, describes the cliffs this week in *Science*. Each cliff seems to be the naked face of a glacier, tantalizing scientists with the promise of a layer-cake record of past martian climates and space enthusiasts with a potential resource for future human bases.



Thick bands of ice (blue) have been spotted in steep cliff faces.

Finding ice on Mars is nothing new. Ice covers the poles, and a radar instrument on the MRO has detected signatures of thick, buried ice across the planet's belly (*Science*, 26 February 2010, p. 1075). Some researchers suggested these deposits could be the remnants of glaciers that existed millions of years ago when the planet's spin axis and orbit were different. But the depth of the ice and whether it exists as relatively pure sheets or as granules frozen in the pore spaces of martian soil have been uncertain.

A decade ago, researchers using the MRO spotted a related clue: pools of seemingly pure ice in the floors of small craters carved out by fresh meteorite impacts. But it was unclear whether these frozen pools were connected to the buried glaciers or were merely isolated patches. At the ice cliffs, Dundas and his team could see the glaciers in cross section, and they patiently revisited the sites to see how they changed over time.

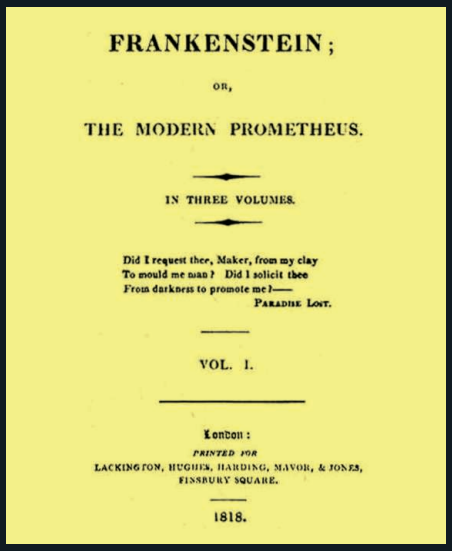
They found that the ice persisted through the martian summer, when any ephemeral frost would have vaporized. And last year, the MRO caught several boulders tumbling out of one of the cliff faces, suggesting that gradual erosion had released them from a massive ice deposit. Evidently the near-surface ice and the large subsurface deposits are one and the same, says Ali Bramson, a co-author and graduate student at the University of Arizona in Tucson. "This deep, thick, pure ice extends almost all the way up to the surface."

Banding and subtly varying shades of blue suggest that the slabs of ice are stacked. That implies that the deposits built up over many seasons as layers of snow were compressed in a previous climate cycle, says Susan Conway, a planetary geologist at the University of Nantes in France. Winds then buried the ice sheets in grit. "It's the only reasonable explanation," she says.

Drilling a core from one of these deposits and returning it to Earth would offer a treasure trove of information to geologists about the past martian climate, says G. Scott Hubbard, a space scientist at Stanford University in Palo Alto, California. "That preserved record would be of extreme importance to go back to," he says.

These sites are "very exciting" for potential human base as well, says Angel Abbud-Madrid, director of the Center for Space Resources at the Colorado School of Mines in Golden, who led a recent NASA study exploring potential landing sites for astronauts. Water is a crucial resource for astronauts, because it could be combined with carbon dioxide, the main ingredient in Mars's atmosphere, to create oxygen to breathe and methane, a rocket propellant. And although researchers suspected the subsurface glaciers existed, they would only be a useful resource if they were no more than a few meters below the surface. The ice cliffs promise abundant, accessible ice, Abbud-Madrid says.

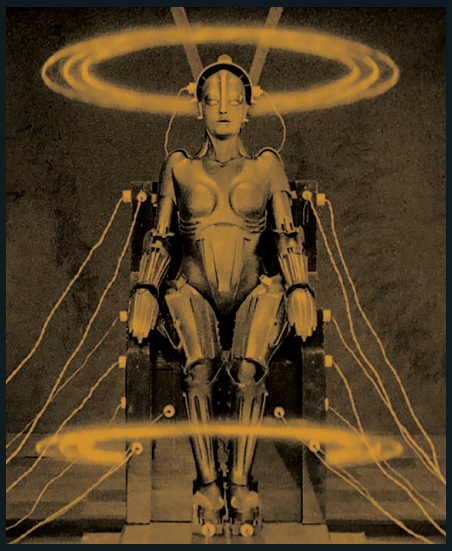
The cliffs are all found at latitudes about 55° north or south, however, which grow frigid and dark in the martian winter—unpromising latitudes for a solar-powered human base. For this reason, the NASA study was limited to sites to within 50° of the equator. Now, Hubbard wants NASA's human exploration program to look for similar cliffs closer to the equator. "What's the cutoff point?" he asks. He hopes the next surprise will be ice closer to the martian tropics. ■



1818. The cover of the first edition of *Frankenstein*, published when Mary Shelley was 20 years old.

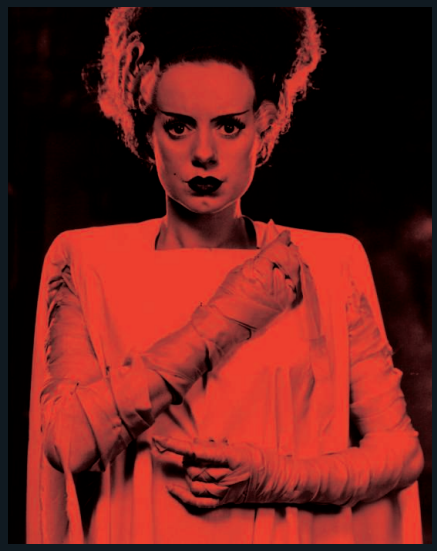
1831. Theodor von Holst's illustration for the inside cover of the third, heavily revised edition.

1910. The monster in the first movie adaptation, a 16-minute silent film directed by J. Searle Dawley.



1927. In Fritz Lang's *Metropolis*, a scientist produces a female robot; unlike Frankenstein, he loves his creation.

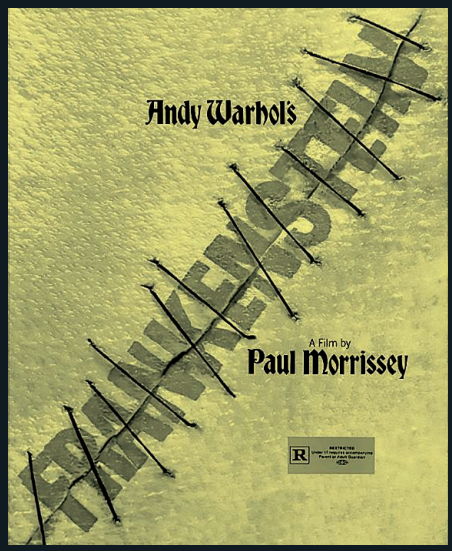
1931. Boris Karloff as the creature in the box office hit *Frankenstein*, directed by James Whale.



1935. Elsa Lanchester played both Mary Shelley and the monster's mate in the sequel *Bride of Frankenstein*.

1964. Fred Gwynne starred in *The Munsters*, a U.S. sitcom about a family of friendly monsters.

1974. The U.S. grandson of the notorious scientist creates his own monster in Mel Brooks's *Young Frankenstein*.

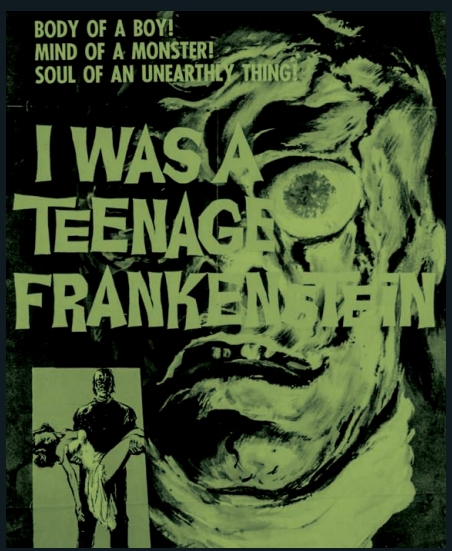


1974. Andy Warhol co-produced the movie *Flesh for Frankenstein*, also named *Andy Warhol's Frankenstein*.

PHOTOS: (TOP, LEFT TO RIGHT): PAUL FEARN/ALAMY STOCK PHOTO; IAN DAGNALL COMPUTING/ALAMY STOCK PHOTO; MOVIESTORE COLLECTION LTD/ALAMY STOCK PHOTO; (MIDDLE, LEFT TO RIGHT) FRIEDRICH-WILHELM-MURNAU STIFTUNG, WIESBADEN, GERMANY; UNIVERSAL PICTURES/SCREENPROD/PHOTONONSTOP/ALAMY STOCK PHOTO; UNIVERSAL PICTURES/AF ARCHIVE/ALAMY STOCK PHOTO; (BOTTOM, LEFT TO RIGHT) CBS/AF ARCHIVE/ALAMY STOCK PHOTO; C20TH FOX/TWENTIETH CENTURY FOX/RONALD GRANT ARCHIVE/ALAMY STOCK PHOTO; EVERETT COLLECTION, INC./ALAMY STOCK PHOTO

FEATURES

1920. *The Golem: How He Came Into The World* was part of a German film trilogy with echoes of *Frankenstein*.



1957. *I Was a Teenage Frankenstein* was a Mary Shelley–inspired follow-up to the hit *I Was a Teenage Werewolf*.

1994. *Mary Shelley's Frankenstein*, directed by Kenneth Branagh, featured Robert de Niro as the creature.

CREDITS: (PHOTOS, TOP TO BOTTOM): CHRONICLE/ALAMY STOCK PHOTO; THE ADVERTISING ARCHIVES/ALAMY STOCK PHOTO; COLUMBIA PICTURES/RONALD GRANT ARCHIVE/ALAMY STOCK PHOTO; (ILLUSTRATION) CRAIG & KARL

THE LONG SHADOW OF FRANKENSTEIN

Mary Shelley's 200-year-old tale is still essential reading for scientists

By **Kai Kupferschmidt**

In January 1818, a woman barely out of her teens unleashed a terrifying tale on the world: the story of a doctor who builds a creature from scavenged body parts, then recoils in horror, spurns it, and sees his friends and family destroyed by the monster. Two hundred years later, Mary Shelley's *Frankenstein* is still essential reading for anyone working in science. The ill-fated creator she portrays has influenced public perception of the scientific enterprise unlike any other character, forever haunting the borderland between what science can do and what it should do.

The story has mutated and it has frequently been mangled. It has spawned countless books, plays, and movies—some pictured on these pages—and even a superhero comic. It has inspired technophobes and scientists alike. "Franken-" has become a passe-partout prefix for anything deemed unnatural or monstrous.

Interpretations of the tale have also multiplied. A story of scientific hubris, a creator consumed by his creation, a male scientist trying to eliminate women's role in reproduction, an attempt by Shelley to deal with the trauma of losing a baby. To the growing group of scientists pondering the ways in which science might eventually destroy humanity, it is the earliest warning of such risks.

None of this quite captures the secret of the story's longevity. To borrow the monster's own description of indelible knowledge, Shelley's tale "clings to the mind ... like a lichen on the rock." In the preface to the 1831 edition, Shelley wrote: "Now, once again, I bid my hideous progeny go forth and prosper." It did. And it still does. ■

HOW A HORROR STORY HAUNTS SCIENCE

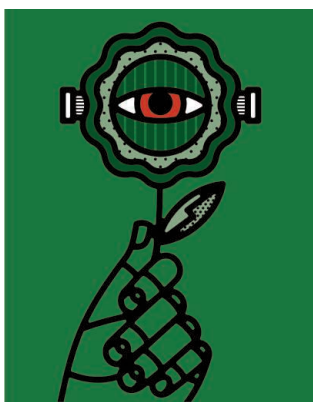
A novel offers grist for psychologists and science historians, ammunition for technophobes, and even inspiration for researchers

By **Jon Cohen**

On 1 August 1790, a precocious student named Victor Frankenstein submitted a radical proposal to an ethical panel at the University of Ingolstadt in Bavaria. Under the title "Electro-chemical Mechanisms of Animation," Frankenstein explained how he wanted to "reverse the processes of death" by collecting "a large variety of human anatomical specimens" and putting them together to try and "restore life where it has been lost."

Frankenstein assured the institutional review board (IRB) that he had the highest ethical standards. "If I do succeed in fully animating a human or human-like creature, I will provide the creature with information about the study and allow it, if it is capable, to choose whether or not to participate further in continued observation and study," noted the budding scientist. If the creature had "diminished capacity," Frankenstein promised to bring in a third party to act in its interest and treat "the being" in accordance with recognized standards.

Of course no such proposal ever went to bioethicists at the University of Ingolstadt, where the fictional Frankenstein created his monster. In 1790, even a real Frankenstein would have faced no ethical reviews. But the proposal does exist in a 2014 paper, which speculates about whether the *Frankenstein* story would have had a happier ending if 21st century safeguards had existed 2 centuries ago. It is one of many riffs on the novel to be found in biomedical literature. In conceiving her story, Mary Shelley was influenced by the nascent medical science of the day and by early experiments on electricity. In return,



Frankenstein has haunted science ever since.

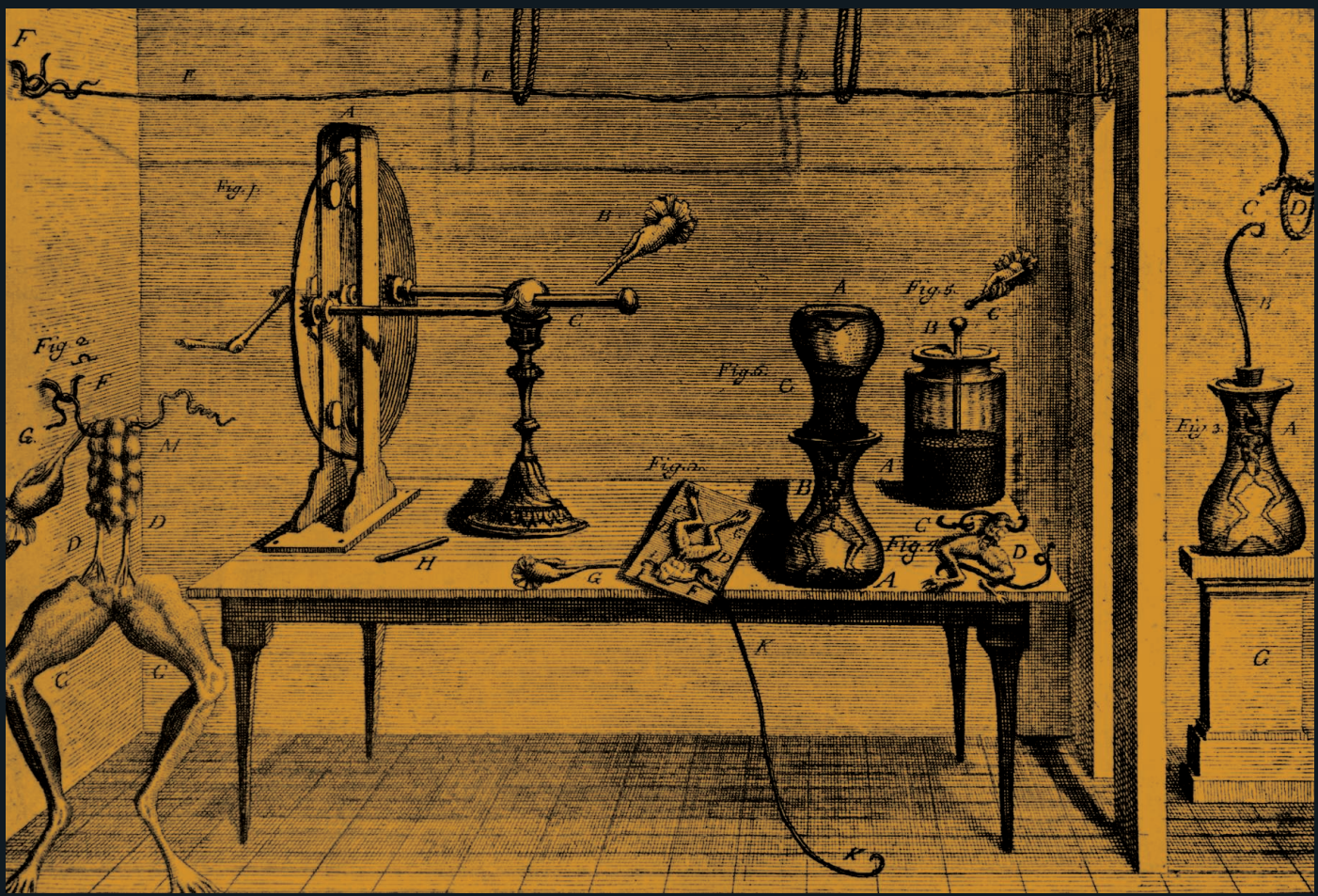
First published anonymously in 1818, the book and subsequent films and plays have become what Jon Turney, author of the book *Frankenstein's Footsteps: Science, Genetics and Popular Culture*, calls "the governing myth of modern biology": a cautionary tale of scientific hubris. And as with all long-lasting myths, it is not one myth, but many, as a search for "Frankenstein" in the PubMed database—the main catalog of life sciences papers—makes clear. Scientific literature, like the popular press, is rife with references to Frankenfood, Frankencells, Frankenlaws, Frankenswine, and Frankendrugs—most of them supposedly monstrous creations. Other papers explicitly mentioning *Frankenstein*—there are 251 of them—analyze the science behind the novel or even, in a twist that can be downright bizarre, draw inspiration from it.

Several reports in psychological journals delve into the state of mind of its author when she first imagined the tale during the summer of 1816. Then Mary Wollstonecraft Godwin, she was visiting the poet Lord Byron at Villa Diodati, a mansion he had rented on the shores of Lake Geneva in Switzerland. She was 18, accompanying her married lover, the poet Percy Bysshe Shelley. Her stepsister, Claire Clairmont, was there, as was Byron's live-in doctor, John William Polidori. It was the "year without a summer," a climatic anomaly caused by the eruption of Mount Tambora in the Dutch East Indies, and endless rain and gray skies kept the guests cooped up. Byron suggested as a party game that they each write a ghost story.

There was plenty to unsettle Mary's fertile mind. Mary and Percy had a 6-month-old baby together, but had lost another baby a year earlier. Mary's own mother had died of puerperal sepsis 11 days after giving birth to her fame-bound daughter. Percy, as a 2013 paper in *Progress in Brain Research* recounts, had been booted from the University of Oxford in the United Kingdom for "extolling the virtues of atheism" and was a believer in "free love." Another paper, in a 2015 issue of *The Journal of Analytical Psychology*, suggests that Percy, Mary, and Claire had previously formed "a ménage à trois of sorts."

The Journal of Analytical Psychology paper's author, Ronald Britton, a prominent psychoanalyst, links these tensions and griefs to the daydream in which Mary Shelley first envisioned Frankenstein's monster—"the spectre which had haunted my midnight pillow," as she later put it. The "background facts to her nightmare," Britton writes, invoking Freud, "opened a

ILLUSTRATION: CRAIG & KARL



In 1780, Italian scientist Luigi Galvani showed that a spark could make the muscles of a dead frog contract—inspiration for *Frankenstein*'s author, Mary Shelley.

PHOTO: DEAGOSTINI PICTURE LIBRARY/GETTY IMAGES

door to unconscious phantasies of a dreadful scene of childbirth." He adds that after losing her first child in 1815, Shelley wrote in her journal that she dreamed about the baby coming back to life. "I thought that if I could bestow animation upon lifeless matter, I might, in process of time, renew life where death had apparently devoted the body to corruption," she wrote the year before imagining *Frankenstein*.

More horrors were to follow for Shelley after she completed the novel. She married Percy after his first wife's suicide, only to lose him 6 years later when he drowned in a sailing accident. But she called on science, not psychology, in explaining how she "came to think of, and dilate upon, so very hideous an idea" at 18 years of age. Among the influences she cites in a preface to an 1831 edition of her novel is Luigi Galvani, who in 1780 found that an electrical charge could make a dead frog's legs twitch. It was Percy who may have acquainted her with galvanism, which Frankenstein explicitly mentions as the key to reanimation in the 1831 edition. As a boy, the poet "had dabbled with electricity (on his sister's sores and the family cat)," another study in *Progress in Brain Research* notes.

Many a paper has attempted to parse other ways in which the science of the day influenced Shelley's tale. A 2016 essay in *Nature* by a U.K. biographer noted that her novelist father was friends with electrochemist Humphry Davy and with William Nicholson, a co-discoverer of electrolysis, the technique of triggering chemical reactions using electricity. Several accounts point to the influence of Byron's physician, Polidori (who later poisoned himself with prussic acid), and his discussions of experiments on spontaneous generation by Erasmus Darwin, grandfather of Charles. A 2004 paper in the *Journal of Clinical Neurophysiology* that reviews the "electrophysiological undercurrents for Dr. Frankenstein" notes that Shelley could not have missed the widely discussed work of Giovanni Aldini, Galvani's nephew, who in 1803 zapped the heads of decapitated criminals in an attempt to reanimate them; he imagined this could be used to resuscitate people who had drowned or suffocated and possibly to help the insane.

Over time, the influence ran from the novel back to science. "From Frankenstein to the Pacemaker," in *IEEE Engineering in Medicine and Biology Magazine*, tells how 8-year-old Earl Bakken in 1932 saw the famous *Frankenstein* movie starring Boris Karloff, which "sparked Bakken's interest in combining electricity and medicine." Bakken would later found Medtronic, develop the first transistorized cardiac pacemaker, and open a museum devoted to electricity in the life sciences that's housed in a Gothic Revival style mansion in Minneapolis, Minnesota. Neighborhood kids call it Frankenstein's castle.

Indeed, many scientific studies proudly reference *Frankenstein*, mainly because they combine disparate parts to create a novel entity that the researchers present as delightfully chimeric. A milk sugar enzyme fused with a carrier protein. An atlas of the head and neck to guide radiotherapy, created by merging views from different patients. A face recognition study

that swapped the eyes, noses, and mouths of former President George W. Bush and former U.S. Secretary of State Colin Powell. A "Frankenrig" used to create 3D animations, made by mixing and matching bones from different skeletons.

In perhaps the strangest embrace of the *Frankenstein* label, a 2013 article in *Surgical Neurology International* proposes recreating Aldini's electrifying head experiments. The authors of "HEAVEN: The Frankenstein effect," note that Aldini ultimately aimed to transplant a human head, using electricity to spark it back into awareness. That's just what the authors have in mind for their project, the head anastomosis venture (HEAVEN). "On the whole, in the face of clear commitment, HEAVEN could bear fruit within a couple of years," they write. (Many scientists have called the project unfeasible and unethical, but last November, two of the co-authors announced to the media that they had performed a head transplant on a human corpse and soon planned to publish details.)

Reanimation was in fashion in 1818. Scottish doctor Andrew Ure attempted the feat on a corpse.

But by far the bulk of the scientific literature hand-wrings, ponders, and philosophizes about the most familiar form of the *Frankenstein* myth, which Shelley flicked at in her "Modern Prometheus" subtitle: the idea that mad scientists playing God the creator will cause the entire human species to suffer eternal punishment for their trespasses and hubris.

"Mary Shelley, Frankenstein, and The Dark Side of Medical Science," a 2014 essay published in the charmingly incongruous *Transactions of the American Clinical and Climatological Association*, ticks off a diverse list of recent experiments that have drawn the "Franken-" label: the cloning of Dolly the sheep, the engineering of a highly lethal H5N1 bird influenza that could more easily infect mammals, the synthesizing of an entire bacterial genome. Other triggers of *Frankenstein*-ish fears have included in vitro fertilization, proposals to transplant pig organs into humans, and tomatoes endowed with genes from fish to make them freeze-tolerant.

J. Craig Venter, a pioneer in genomics based in San Diego, California, has been called a Frankenstein for his effort to create artificial bacteria with the smallest possible genomes. Still, he's a fan of Shelley's tale. "I think she's had more influence with that one book than most authors in history," says Venter, who owns a first edition. "It affects a lot of people's thinking and fear because it represents this fundamental of 'You don't mess with Mother Nature and you don't mess with life because God will strike you down.'"

"Obviously, I don't buy into that theme," he adds.

The *Frankenstein* myth endures, he says, because "fear is easy to sell"—even when unwarranted. "Most people have a fear of what they don't understand," he says. "Synthetic cells are pretty complicated and putting a new gene into corn sounds scary." But by throwing around labels like Frankenfood and Frankencells to rally the public against potentially valuable innovations, he says, the "fear-based community will potentially do more damage to humanity than the things they fear."

Unlike the Frankenstein character, who initially didn't consider how his work might go wrong, Venter says he recognizes that editing and rewriting genomes could "contaminate the world" and cause unintended harm. "I think we need to be very smart about when we do it and how we do it," he says. He thinks Shelley "would highly appreciate" his work.

Henk van den Belt, a philosopher and ethicist at Wageningen University in the Netherlands who wrote a paper about *Frankenstein* and synthetic biology, applauds Venter for fighting back against the Frankenslur. "Very often scientists are afraid to take this position, but I think it's better to be defiant," Van den Belt says. "Rhetoricians and journalists can accuse people of playing Frankenstein, but it's a little too easy. If scientists challenge this phrase, it will have less impact."

Shelley of course couldn't have imagined any of this hubbub, and indeed her tale has been wildly distorted in the popular imagination over the past 2 centuries. Frankenstein's aim was not to rule the world à la Dr. Evil, but "to banish disease from the human frame and render man invulnerable to any but a violent death." And Britton, the psychoanalyst, notes that the creature did not begin life as a monster; he only went on a killing spree because he sought love and happiness but was abhorred by his creator, who referred to him as "devil," "fiend," "abortion," "daemon," "vile insect," and other terms that would have made an IRB contact the Office for Human Research Protections. "I was benevolent and good, misery made me a fiend," Frankenstein's creation said. "[I]mpotent envy and bitter indignation filled me with an insatiable thirst for vengeance."

A dental radiologist of all people published an insightful two-part essay in *The Journal of the Royal Society of Medicine* in 1994 that underscores what some argue is the real moral of the book: not the danger of scientists violating the natural order, but the dire fate that awaits creators who fail to care for their creations. "Read the book and weep for those we have rejected, and fear for what revenge they will exact, but shed no tears for Frankenstein," the essay advises, referring to the doctor. "Those who think, in ignorance of the book, that his is the name of the Monster are in reality more correct than not." ■

Creating a modern monster

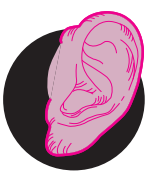
When Mary Shelley published her story of Victor Frankenstein and his misshapen monster in 1818, she provided little detail about how exactly the doctor built his creation except that "the dissecting room and the slaughter-house furnished many of [his] materials" and that he infused "a spark of being in the lifeless thing." But what if Shelley had written her book today? What technologies might give rise to her iconic creature?

Text by **David Shultz**; *Graphic by* **Adolfo Arranz**

Transplants

The kidney, first transplanted in 1950, remains the most commonly transplanted organ today, followed by the liver, heart, lung, pancreas, and intestine. A 2018 Frankenstein could also transplant tissues such as the skin, nerves, cornea, cartilage, and bones. More cutting-edge are face transplants, performed 37 times between 2005 and 2015, and penis transplants, first successfully done in 2014. The first baby to develop in a transplanted womb was born in Sweden in 2014.

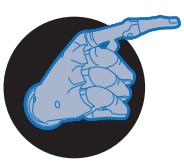
What's next: Two surgeons say they want to perform human head transplants, perhaps better called "whole body" transplants, though most scientists say reconnecting all the nerves within the spinal cord will remain science fiction for a long time.



Lab-grown organs

Skin, urethras, bladders, blood vessels, vaginas, and muscle can all be produced by taking a patient's own cells and growing them on a biodegradable scaffold in the lab. The technique works best for flat, hollow, and tubular organs.

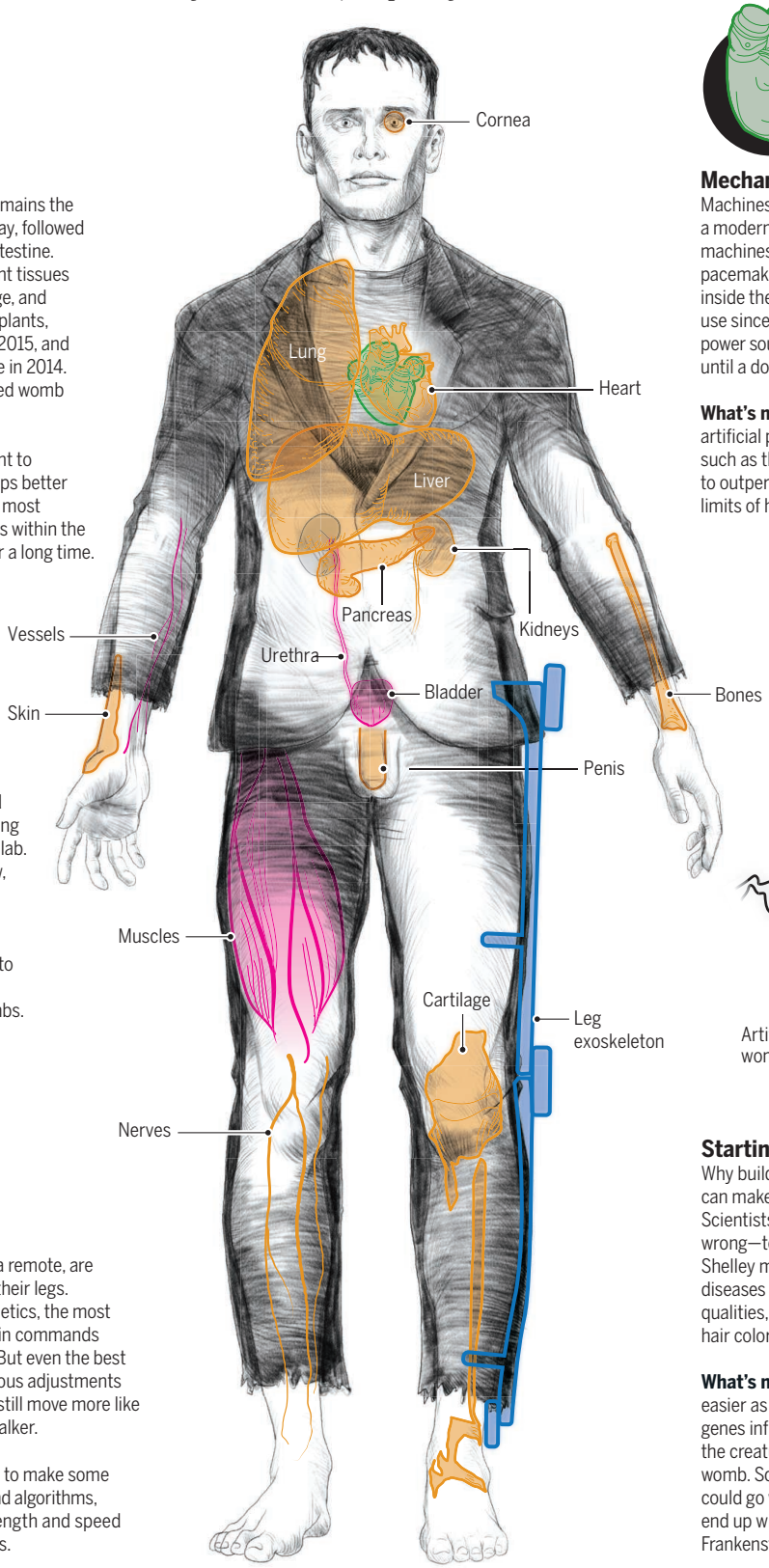
What's next: Scientists are using 3D printing and other techniques in efforts to grow more complex structures such as hearts, livers, kidneys, penises, and wombs.

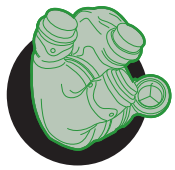


Bionics

Robotic "exoskeletons," controlled with a remote, are helping paraplegics regain control over their legs. Missing limbs can be replaced by prosthetics, the most advanced of which can directly read brain commands through electrodes placed on the skull. But even the best prosthetics can't simulate the unconscious adjustments that smooth out normal gestures. They still move more like Frankenstein's monster than Luke Skywalker.

What's next: Artificial limbs could learn to make some decisions on their own, using cameras and algorithms, to allow for smoother movements; strength and speed could be increased to superhuman levels.

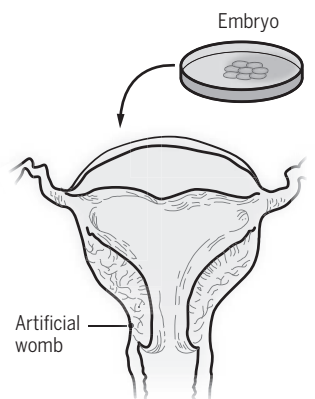




Mechanical organs

Machines could substitute for organs in a modern version of the creature. Dialysis machines function as external kidneys; pacemakers and cochlear implants work inside the body for years. Artificial hearts, in use since the 1980s, still need external power sources and are only used as a "bridge" until a donor heart becomes available.

What's next: On the horizon may be fully artificial pancreases, eyes, and lungs. Organs such as the heart and the lungs could be built to outperform natural ones, extending the limits of human performance.



Starting from scratch

Why build a human from spare parts if you can make one to order from an embryo? Scientists agree it is already feasible—albeit wrong—to clone a human. A 21st century Shelley might call on gene editing to eliminate diseases and endow the creature with specific qualities, including size, strength, and eye or hair color.

What's next: Tweaking humans will get easier as scientists further unravel how our genes influence physical traits. One day, the creature could be grown in an artificial womb. Scientists warn that countless things could go wrong along the way, and you might end up with something monstrous—just as Frankenstein did.

TAMING THE MONSTERS OF TOMORROW

A small group of researchers is studying how science could destroy the world—and how to stop that from happening

By **Kai Kupferschmidt**, *in Cambridge, U.K.*

Philosopher Nick Bostrom believes it's entirely possible that artificial intelligence (AI) could lead to the extinction of *Homo sapiens*. In his 2014 bestseller *Superintelligence: Paths, Dangers, Strategies*, Bostrom paints a dark scenario in which researchers create a machine capable of steadily improving itself. At some point, it learns to make money from online transactions and begins purchasing goods and services in the real world. Using mail-ordered DNA, it builds simple nanosystems that in turn create more complex systems, giving it ever more power to shape the world.

Now suppose the AI suspects that humans might interfere with its plans, writes Bostrom, who's at the University of Oxford in the United Kingdom. It could decide to build tiny weapons and distribute them around the world covertly. "At a pre-set time, nanofactories producing nerve gas or target-seeking mosquito-like robots might then burgeon forth simultaneously from every square meter of the globe."

For Bostrom and a number of other scientists and philosophers, such scenarios are more than science fiction. They're studying which technological advances pose "existential risks" that could wipe out humanity or at least end civilization as we know it—and what could be done to stop them. "Think of what we're trying to do as providing a scientific red team for the things that could threaten our species," says philosopher Huw Price, who heads the Centre for the Study of Existential Risk (CSER) here at the University of Cambridge.

The idea of science eliminating the human race can be traced all the way back to *Frankenstein*. In Mary Shelley's novel, the monster gets angry at his creator, Victor Frankenstein, for having spurned him. He kills Frankenstein's little brother William, but then offers the doctor a deal: Make a female companion for me and we will leave you in peace and go to South America to live out our days. Frankenstein starts working on the bride, but realizes that the couple might reproduce and outcompete humans: "A race of devils would be propagated upon the earth who might make the very existence of the species of man a condition precarious and full of terror." He destroys the half-finished female, reigniting the creature's wrath and bringing about his own demise.

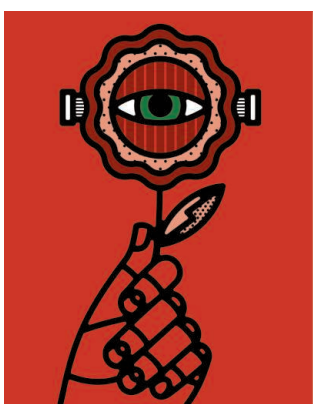


ILLUSTRATION: CRAIG & KARL

"I think *Frankenstein* illustrates the point beautifully," says physicist Max Tegmark of the Massachusetts Institute of Technology (MIT) in Cambridge, a board member of CSER and a co-founder of a similar think tank, the Future of Life Institute (FLI), near MIT. "We humans gradually develop ever-more-powerful technology, and the more powerful the tech becomes, the more careful we have to be, so we don't screw up with it."

The study of existential risks is still a tiny field, with at most a few dozen people at three centers. Not everyone is convinced it's a serious academic discipline. Most civilization-ending scenarios—which include humanmade pathogens, armies of nanobots, or even the idea that our world is a simulation that might be switched off—are wildly unlikely, says Joyce Tait, who studies regulatory issues in the life sciences at the Innogen Institute in Edinburgh. The only true existential threat, she says, is a familiar one: a global nuclear war. Otherwise, "There is nothing on the horizon."

Harvard University psychologist Steven Pinker calls existential risks a "useless category" and warns that "Frankensteinian fantasies" could distract from real, solvable threats such as climate change and nuclear war. "Sowing fear about hypothetical disasters, far from safeguarding the future of humanity, can endanger it," he writes in his upcoming book *Enlightenment Now: The Case for Reason, Science, Humanism, and Progress*.

But advocates predict the field will only get more important as scientific and technological progress accelerates. As Bostrom pointed out in one paper, much more research has been done on dung beetles or *Star Trek* than on the risks of human extinction. "There is a very good case for saying that science has basically ignored" the issue, Price says.

HUMANITY HAS ALWAYS FACED the possibility of an untimely end. Another asteroid of the size that ended the dinosaurs' reign

PHOTO: OMIKRON/SCIENCE SOURCE

A nuclear bomb is detonated at the U.S. Pacific Proving Grounds in 1958. Global nuclear warfare is one of the "existential risks" that could end civilization as we know it.

A glossary of Frankenwords

Along with fears, the *Frankenstein* story has inspired hundreds of whimsical names for products and phenomena. Here's a selection; you'll find more at http://scim.ag/Frankenwords.

By **Jon Cohen**

Frankenbrooms
Supercoated broomheads used in curling, a winter sport played on ice.

Frankencell
J. Craig Venter's attempt to create an artificial cell containing the smallest possible number of essential genes.

Frankenfears
Exaggerated concerns about transgenic food (Frankenfood) and other products of genetic engineering.

Frankenforests
Engineered trees that grow more quickly, absorbing more carbon dioxide and providing more wood and pulp without the need for toxic chemicals.

Frankengene
The *catechol-O-methyltransferase* gene. In 2008, researchers linked variants in the gene to the strong, frightful reaction some people have to horror movies.

Frankenmoth
A male diamondback moth engineered to spread a lethal gene to females, creating nonviable offspring that reduce the moth's toll on crops. Cousins include Frankenflies and Frankenmosquitoes.

Frankenmouse
A genetically or surgically altered mouse. Variations have included the "oncomouse" that's prone to cancer and the "earmouse" that had a human-shaped ear (photo, left).

Frankenmums
Mothers who freeze eggs for their infertile daughters to use.

Frankenpets
Transgenic dogs that would repel fleas; cats that would not cause allergies.

Frankenpines
Cellphone towers that resemble pine trees.

Frankenrobot
A robot, created at the University of Reading in the United Kingdom, that had a "brain" made of rat neurons placed on a multielectrode array.

Frankenshoes
Platform shoes with wedge heels or sneakers with heels.

Frankensites
University websites that have conflicting material posted by professors and their departments.

Frankenspuds
Genetically modified potatoes that are resistant to blight. Also known as Frankenfries.

Frankenstorm
A combination of storms that creates a monster event. Superstorm Sandy famously merged with another storm in 2012 to wallop the U.S. East Coast.

could hit Earth; a volcanic cataclysm could darken the skies for years and starve us all.

But existential risks arising from scientific advances were literally fiction until 16 July 1945, when the first atomic bomb was detonated. Based on some back-of-the-envelope calculations, physicist Edward Teller had concluded that the explosion might set off a global chain reaction, "igniting" the atmosphere. "Although we now know that such an outcome was physically impossible, it qualifies as an existential risk that was present at the time," Bostrom writes. Within 2 decades a real existential risk emerged, from growing stockpiles of the new weapons. Physicists had finally assembled Frankenstein's bride.

Other scientific disciplines may soon pose similar threats. "In this century we will introduce entirely new kinds of phenomena, give ourselves new kinds of powers to reshape the world," Bostrom says. Biotechnology is cheaper and easier to handle than nuclear technology has ever been. Nanotechnology is making rapid strides. And at a 2011 meeting in Copenhagen, Estonian computer programmer and Skype co-developer Jaan Tallinn told Price about his deep fears about AI during a shared taxi ride. "I'd never met anyone at that point who took that as seriously as Jaan," says Price, who was about to start working at the University of Cambridge.

Price introduced Tallinn to astronomer Martin Rees, a former president of the Royal Society, who had long warned that as science progresses, it will increasingly place the power to destroy civilization in the hands of individuals. The trio decided to launch CSER, the second such center after Bostrom's Future of Humanity Institute in Oxford, which he launched in 2005. CSER's name was "a deliberate attempt to push the idea of existential risk more towards the mainstream," Price says. "We were aware that people think of these issues as a little bit flaky."

CSER has recruited some big-name supporters: The scientific advisory board includes physicist Stephen Hawking, Harvard biologist George Church, global health leader Peter Piot, and tech entrepreneur Elon Musk. In a sign of just how small the field still is, Tallinn also co-founded FLI in 2014, and Church, Musk, Hawking, Bostrom, and Rees all serve on its scientific advisory board. (Actor Morgan Freeman, who has literally played God, is also an FLI adviser.)

Most of CSER's money comes from foundations and individuals, including Tallinn, who donated about $8 million to existential risk researchers in 2017. CSER's academic output has been "ephemeral" so far, Tallinn concedes. But the center was set up as "a sort of training ground for existential risk research," he says, with academics from elsewhere coming to visit and then "infecting" their own institutions with ideas.

THE DOZEN PEOPLE working at CSER itself—little more than a large room in an out-of-the-way building near the university's occupational health service—organize talks, convene scientists to discuss future developments, and publish on topics from regulation of synthetic biology to ecological tipping points. A lot of their time is spent pondering end-of-the-world scenarios and potential safeguards.

Church says a "crunch," in which a large part of the world population dies, is more likely than a complete wipe-out. "You don't have to turn the entire planet into atoms," he says. Disrupting electrical grids and other services on a huge scale or releasing a deadly pathogen could create chaos, topple governments, and send humanity into a downward spiral. "You end up with a medieval level of culture," Church says. "To me that is the end of humanity."

Existential risks stemming from the life sciences are perhaps easiest to imagine. Pathogens have proved capable of killing off entire species, such as the frogs that have fallen victim to the amphibian fungus *Batrachochytrium dendrobatidis*. And four influenza pandemics have swept the world in the past century, including one that killed up to 50 million people in 1918 and 1919. Researchers are already engineering pathogens that in principle could be even more dangerous. Worries about studies that made the H5N1 bird flu strain more easily transmissible between mammals led the United States to halt such research until late last year. Terrorists or rogue states could use labmade agents as a weapon, or an engineered plague could be released accidentally.

Rees has publicly wagered that by 2020, "bioterror or bioerror will lead to 1 million casualties in a single event." Harvard microbiologist Marc Lipsitch has calculated that the likelihood of a labmade flu virus leading to an accidental pandemic is between one in 1000 and one in 10,000 per year of research in one laboratory; Ron Fouchier of Erasmus MC in Rotterdam, the Nether-

The release of a dangerous pathogen might cause a "crunch" in the human population.

lands, one of the researchers involved in the H5N1 studies, has dismissed that estimate, saying the real risk is more like one in 33 billion per year and lab.

One measure against "bioerror" might be to make researchers who carry out risky experiments buy insurance; that would require an independent assessment of the risk and would force researchers to face up to it, Lipsitch says. Still, the most important countermeasure is to strengthen the world's capacity to contain an outbreak early on, he adds, for instance with vaccines. "For biological risks, short of a really massive, coordinated, parallel attack around the world, the only way we are going to get to a really catastrophic scenario is by failing to control a smaller scenario," he says.

VIRUSES ARE UNLIKELY to kill every last human, Bostrom says; for him and others, it is AI that poses truly existential threats. Most scenarios center on machines outsmarting humans, a feat called "superintelligence." If such AI were ever achieved and it acquired a will of its own, it might turn malevolent and actively seek to destroy humans, like HAL, the computer that goes rogue aboard a spaceship in Stanley Kubrick's film *2001: A Space Odyssey*.

Most AI experts worry less about machines rising up to overthrow their creators, however, than about them making a fatal mistake. To Tallinn, the most plausible way in which AI could end humanity is if it simply pursued its goals and, along the way, heedlessly created an environment fatal to humans. "Imagine a situation where the temperature rises by 100° or is lowered by 100°. We'd go extinct in a matter of minutes," Tallinn says. Tegmark agrees: "The real problem with AI is not malice, it's incompetence," he says.

A current-day analogy is the 2015 tragedy in which a suicidal Germanwings pilot told his plane's computer to descend to an altitude of 100 meters while flying over the French Alps. The machine complied, killing all 150 on board, even though it had GPS and a topographic map. "It did not have a clue about even the simplest human goal," Tegmark says. To avoid such calamities, scientists are trying to figure out how to teach AI human values and make sure they stick, a problem called "value alignment." "There might be fewer than 20 people who work full time on technical AI safety research," Bostrom says. "A few more talented people might substantially increase the rate of progress."

CRITICS SAY THESE EFFORTS are unlikely to be useful, because future threats are inherently unpredictable. Predictions were a problem in every "foresight exercise" Tait has taken part in, she says. "We're just not good at it." Even if you foresee a risk, economic, political, and societal circumstances will all affect how it plays out. "Unless you know not only what is going to happen, but how it is going to happen, the information is not much use in terms of doing something about it," Tait says.

Pinker thinks the scenarios reveal more about human obsessions than real risks. We are drawn to prospects "that are highly improbable while having big impacts on our fitness, such as illicit sex, violent death, and Walter-Mittyish feats of glory," he writes. "Apocalyptic storylines are undoubtedly gripping—they are a supernormal stimulus for our morbid obsessions." Sure, he says, one can imagine a malevolent, powerful AI that people can no longer control. "The way to deal with this threat is straightforward: Don't build one."

Tallinn argues it's better to be safe than sorry. A 2017 survey showed that 34% of AI experts believed the risks associated with their work are an important problem; 5% said they are "one of the most important problems." "Imagine you're on a plane, and 40% of experts think that there is a bomb on this plane," Tallinn says. "You're not going to wait for the remaining experts to be convinced."

Price says that critics who accuse him and his colleagues of indulging in science fiction are not entirely wrong: Producing doomsday scenarios is not that different from what Shelley did. "The first step is to imagine that range of possibilities, and at that point, the kind of imagination that is used in science fiction and other forms of literature and film is likely to be extremely important," he says.

Scientists have an obligation to be involved, says Tegmark, because the risks are unlike any the world has faced before. Every time new technologies emerged in the past, he points out, humanity waited until their risks were apparent before learning to curtail them. Fire killed people and destroyed cities, so humans invented fire extinguishers and flame retardants. With automobiles came traffic deaths—and then seat belts and airbags. "Humanity's strategy is to learn from mistakes," Tegmark says. "When the end of the world is at stake, that is a terrible strategy." ■

Molecular clouds such as the Cone Nebula have dark interiors and UV-irradiated outer regions, each populated by large hydrocarbons with possibly interrelated chemistries.

ASTROCHEMISTRY

Detecting the building blocks of aromatics

Detection of benzonitrile in an interstellar cloud helps to constrain interstellar chemistry

By **Christine Joblin**[1] *and* **José Cernicharo**[2]

Interstellar clouds are sites of active organic chemistry (*1*). Many small, gas-phase molecules are found in the dark parts of the clouds that are protected from ultraviolet (UV) photons, but these molecules photodissociate in the external layers of the cloud that are exposed to stellar radiation (see the photo). These irradiated regions are populated by large polycyclic aromatic hydrocarbons (PAHs) with characteristic infrared (IR) emission features. These large aromatics are expected to form from benzene (C_6H_6), which is, however, difficult to detect because it does not have a permanent dipole moment and can only be detected via its IR absorption transitions against a strong background source (*2*). On page 202 of this issue, McGuire *et al.* (*3*) report the detection of benzonitrile (*c*-C_6H_5CN) with radio telescopes. Benzonitrile likely forms in the reaction of CN with benzene; from its observation, it is therefore possible to estimate the abundance of benzene itself.

Chemical models that include molecular formation and destruction processes, both in the gas phase and at the surface of dust grains, can account reasonably well for the observed abundances of a number of molecular species (*4*). The situation is different for large aromatics, for which no individual species have been identified, with the exception of the C_{60} molecule (*5*). Although PAHs are large molecules, they are considered by astronomers as very small dust particles. They are therefore generally thought to form in the dense and hot environments of the envelopes of evolved stars. Chemical models have been developed that are based on chemical networks in flames (*6*). More recently, the possibility to form PAHs at the very low temperatures of molecular clouds has been discussed (*7*) following the demonstration that the reaction CCH + C_4H_6

[1]*Institut de Recherche en Astrophysique et Planétologie, Université de Toulouse (UPS), CNRS, 31028 Toulouse Cedex 4, France.* [2]*Instituto de Ciencia de Materiales de Madrid (ICMM), Consejo Superior de Investigaciones Científicas (CSIC), Molecular Astrophysics Group, Cantoblanco, 28049 Madrid, Spain. Email: christine.joblin@irap.omp.eu*

leading to benzene (C_6H_6) is barrierless and therefore efficient at low temperature (*8*).

In both hot and cold gas-phase chemistry, benzene derivatives—such as C_6H_4 and C_6H_5—that can lead to further growth toward larger aromatic species are involved. Observations of benzene-type species are therefore crucial for constraining these chemical models.

McGuire *et al.* were able to detect benzonitrile in a cold molecular cloud of the Taurus region thanks to an elegant spectral-stacking procedure (*9*) that increases the chance of detecting molecules with aromatic rings. Among the species that they searched for, only benzonitrile was identified as a promising candidate. The authors confirmed its detection after identification of individual rotational lines, including their hyperfine structure, through detailed spectroscopic work in the laboratory. The presence of a CN side group leads to a substantial dipole moment and thus facilitates detection of benzonitrile. Calculated benzonitrile abundances from a chemical model that includes different gas-phase reactions at low temperature are lower than those observed by a factor of four. The authors suggest that alternative mechanisms involving cosmic-ray radiation-induced chemistry at the surface of grains produce the missing benzonitrile. The mismatch between observations and models shows that, despite the low observed abundance of benzonitrile, its detection remains important in constraining chemical models.

Is there any relation between the detection of the first aromatic ring in dark clouds and the presence of PAHs, the carriers of the mid-IR aromatic emission bands, in the external UV-irradiated regions of the clouds? In addition to the gas-phase chemical reactions mentioned above, these PAHs and related species, such as C_{60}, could be produced through UV processing of dust grains (*10*). Other scenarios have also been proposed. For instance, large hydrocarbons, including PAHs, could be formed by chemical processing on the surface of silicon carbide grains, a mechanism that could be efficient in the envelopes of carbon-rich red giant stars (*11*, *12*).

It remains unclear how many of the PAHs and their precursors are synthesized in the dense and hot envelopes of evolved stars and how many arise from photo- or radiative chemistry in interstellar environments. The detection by McGuire *et al.* of a benzene derivative in a cold molecular cloud indicates that it can form even at very low temperatures and without UV radiation. The authors did not detect larger species with two or three cycles, but the species they selected have lower dipole moments compared to benzonitrile, which reduces the chance for their detection unless they have an anomalously large abundance.

Among the ~200 molecules detected in space, many are organic species. Studying their composition and chemical networks is key for understanding molecular complexity in protoplanetary disks surrounding young stars (*13*). The search for complex molecules has mainly been performed in the millimeter and submillimeter domains. The work of McGuire and collaborators (*3*, *14*) shows the potential of centimeter-wave instruments for chemical complexity studies. This opens avenues for research at the upcoming Square Kilometer Array, which will operate in this spectral range.

Knowledge of astrochemical networks also helps in understanding the nucleation and growth of interstellar dust (including PAHs) and its role in star and planet formation. However, the detection of benzonitrile in the Taurus region is not sufficient to conclude on the possibility to form PAHs in cold molecular clouds. It also remains to be shown whether the detection of benzonitrile indicates that PAHs could contain nitrogen (*15*). More insights into the chemistry of PAHs and related species are expected from combining data from radio and infrared waves with the James Webb Space Telescope, due to launch in 2019. In addition to observations, guidelines from laboratory astrophysics studies are key to progress in this area. These include spectroscopic and kinetic studies but also experimental simulations in reactors in order to provide scenarios that can explain the building of molecular complexity in cosmic conditions. ■

"Among the ~200 molecules detected in space, many are organic species."

REFERENCES

1. www.astro.uni-koeln.de/cdms/molecules
2. J. Cernicharo *et al.*, *Astrophys. J.* **546**, L123 (2001).
3. B. A. McGuire *et al.*, *Science* **359**, 202 (2018).
4. E. Herbst, E. F. van Dishoeck, *Annu. Rev. Astron. Astrophys.* **47**, 427 (2009).
5. K. Sellgren *et al.*, *Astrophys. J.* **722**, L54 (2010).
6. M. Frenklach, E. D. Feigelson, *Astrophys. J.* **341**, 372 (1989).
7. R. I. Kaiser, D. S. N. Parker, A. M. Mebel, *Annu. Rev. Phys. Chem.* **66**, 43 (2015).
8. B. M. Jones *et al.*, *Proc. Natl. Acad. U.S.A.* **108**, 452 (2011).
9. S. V. Kalenskii, in *Proceedings of the Russian-Indian Workshop on Radio Astronomy and Star Formation*, I. Zinchenko, P. Zemlyanukha, Eds. (Institute of Applied Physics, Russian Academy of Sciences, 2017), vol. 43.
10. P. Pilleri, C. Joblin, F. Boulanger, T. Onaka, *Astron. Astrophys.* **577**, A16 (2015).
11. P. Merino *et al.*, *Nat. Commun.* **5**, 3054 (2014).
12. T. Q. Zhao *et al.*, *Phys. Chem. Chem. Phys.* **18**, 3489 (2016).
13. T. Henning, D. Semenov, *Chem. Rev.* **113**, 9016 (2013).
14. B. A. McGuire *et al.*, *Science* **352**, 1449 (2016).
15. D. M. Hudgins, C. W. Bauschlicher Jr., L. J. Allamandola, *Astrophys. J.* **632**, 316 (2005).

10.1126/science.aar4541

EVOLUTION

Improbable Big Birds

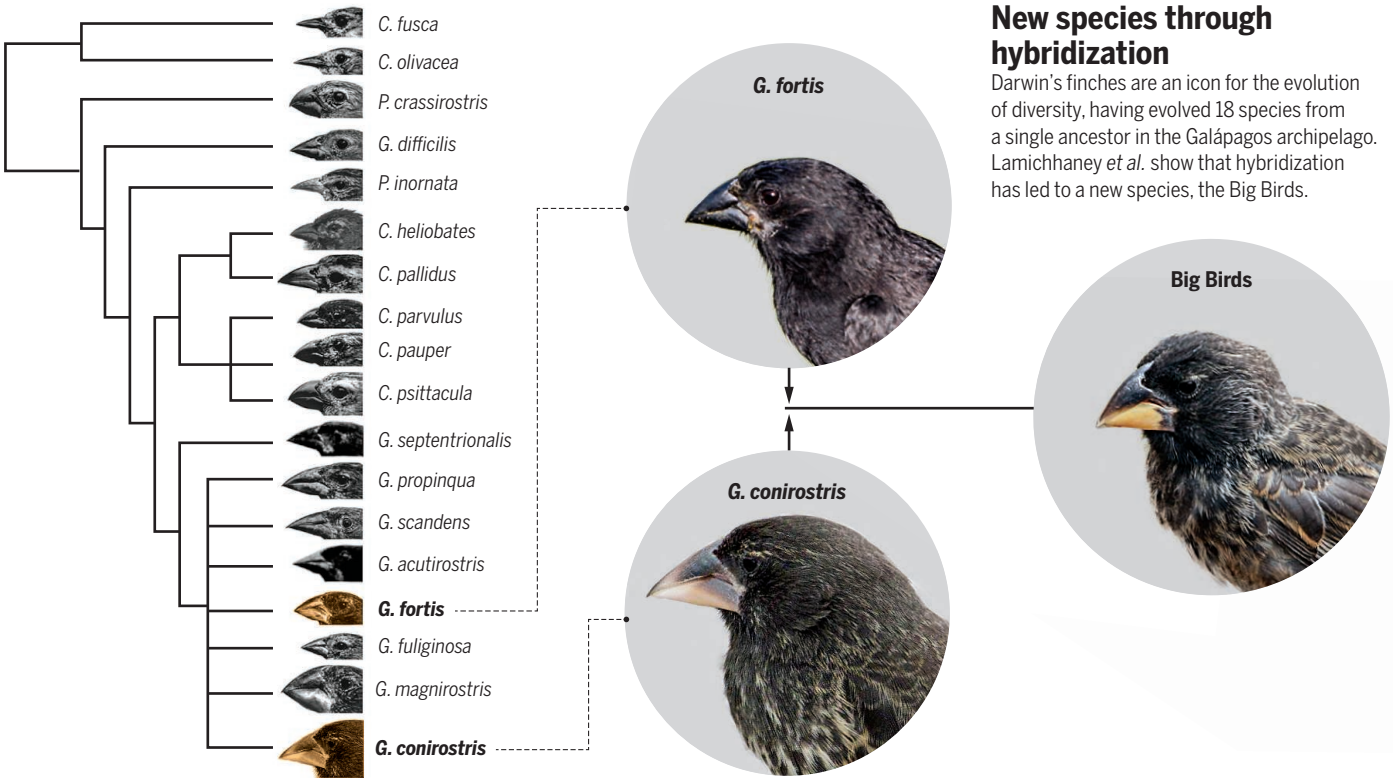
Darwin's finches prove a mechanism for the rapid formation of new species

By **Catherine E. Wagner**

Darwin's finches, a group of 18 species endemic to the Galápagos archipelago, are a classic example of adaptive radiation—the process whereby a single ancestral species multiplies in number to produce divergent species, often in rapid succession (*1*). These birds are evolutionary biologists' most celebrated example of natural selection in action. On page 224 of this issue, Lamichhaney *et al.* (*2*) have succeeded in observing a process even more elusive than natural selection—the formation of a new species (speciation). Because speciation typically takes place on time scales that are too long for direct human observation, before now it was only in organisms with very fast generation times, such as viruses and bacteria, that scientists had directly observed this process [for example, (*3*)]. Lamichhaney *et al.* show through direct observation and DNA sequencing that new species can form very rapidly: within three generations. The key, in this case, is hybridization between different species.

Lamichhaney *et al.* report that in 1981, a male large cactus finch (*Geospiza conirostris*) arrived on the island of Daphne Major in the Galápagos. It had come from Española, another island more than 100 km away in the archipelago. Rosemary and Peter Grant and their collaborators were there to notice it, band it, and watch what it did. Although there were no other individuals of this species on the island, Lamichhaney *et al.* observed that the bird succeeded in finding a mate—a medium ground finch, *G. fortis*. This pair produced offspring, and, with only one exception, the hybrid lineage has bred only within the lineage, exclusively finding mates that are descended from the original pair, for more than 30 years. Because they have a larger body size than other finch species on Daphne Major, they are known as the Big Birds (see the figure).

Botany Department and Biodiversity Institute, University of Wyoming, 1000 East University Avenue, Laramie, WY 82071, USA. Email: catherine.wagner@uwyo.edu



New species through hybridization

Darwin's finches are an icon for the evolution of diversity, having evolved 18 species from a single ancestor in the Galápagos archipelago. Lamichhaney *et al.* show that hybridization has led to a new species, the Big Birds.

The findings of Lamichhaney *et al.* are groundbreaking for two reasons. First, they prove a mechanism of speciation that has a somewhat controversial history in evolutionary biology—hybridization without duplication of the genome, known as homoploid hybrid speciation. Second, the direct observation of this evolutionary mechanism shows not only that the process can happen, but that it can happen quickly. If homoploid hybrid speciation proves to be more common than previously thought, this has profound implications for the rapidity with which biodiversity can be generated.

Although hybridization in animals is not uncommon, it more often results in the erosion of diversity, rather than its origin, because it can merge previously distinct species. Hybrid speciation is an exception. Polyploid hybrid speciation involves duplication of the parental genomes in hybrids (often plants), which provides a barrier to offspring breeding with parent species because of their differing numbers of chromosomes. Such barriers do not occur in homoploid hybrid speciation; thus, the conditions under which this mechanism of speciation can succeed are more limited. Only a handful of cases of homoploid hybrid speciation have been conclusively demonstrated (*4*). Homoploid hybrid lineages have two key challenges to overcome in order to become a species: They must avoid collapsing back into one or both parental species through interbreeding, and they must persist, often in the same place as one or both parental species (*5*). For the Big Bird lineage, the key to surmounting these challenges involved song and beak shape.

In Darwin's finches, song is learned from the father. Because the initial immigrant finch was male, this meant that all of his offspring learned a song that was distinct from other birds on the island (*6*), predisposing them to choose each other as mates and promoting the exclusive breeding of the hybrid lineage for generations thereafter. Additionally, the immigrant finch was 70% larger than the medium ground finches on Daphne Major and also had a larger, pointier beak. Hybrid offspring of this bird and his medium ground finch mate had distinct beak dimensions from those of both parents. This allowed individuals from the Big Bird lineage to exploit larger and harder seeds than the medium ground finches, leading to a diet unlike that of the other species on Daphne Major (*6*). Furthermore, a trend toward increased beak size over generations of Big Birds is consistent with natural selection to increase beak size (*2*), presumably further decreasing food overlap with other species.

The Big Birds also provide insight into how small changes can dismantle the necessary conditions for speciation. Before mating with a medium ground finch, the same immigrant large cactus finch mated with females derived from hybrid ancestry between medium and small ground finches (*G. fortis* and *G. scandens*) (*6*). Although one of these pairings successfully produced offspring that went on to breed, these particular offspring bred with medium ground finches, despite learning their father's distinctive song (*6*), and thus failed to become a distinct hybrid species like the Big Birds. This failure of homoploid hybrid speciation from very similar starting circumstances suggests that the successful reproductive isolation of the Big Bird lineage from other species on the island was not simply due to their distinctive song and instead reflects an interaction between song and inherited beak shape. However, it is not clear if speciation was successful in this case directly because of the traits of the parents or because of chance, related to mate choice decisions or particularities of the resource environment. This illustrates that the conditions under which this mechanism will succeed in producing a new species are narrow.

Although the Galápagos archipelago is an astonishing cradle of evolution, few animals exhibit the diversity in number of species that Darwin's finches have attained through evolution on these islands. This study provides decisive evidence that new species can arise through hybridization, but we do not know to what extent this mechanism is responsible for the group's current species diversity. We also do not know how common homoploid hybrid speciation is in other animals. However, the Big Bird case can aid the search for likely candidates. For example, in species where mate choice decisions are subject to imprinting on the parents, as

is song in Darwin's finches, hybrid speciation may be more likely. Furthermore, when hybrids can exploit a different ecological niche, hybrid speciation is more likely (*5*, *7*).

With high-resolution genomic data available for Darwin's finches (*2*, *8*) and for many other species that have undergone adaptive radiation, such as cichlid fishes and *Heliconius* butterflies [for example, (*9*, *10*)], we are poised to learn more about whether the origin of Big Birds by means of homoploid hybrid speciation is anomalous or common and whether this mechanism might frequently explain cases of rapid speciation. However, challenges remain in distinguishing genetic signals of homoploid hybrid speciation from other situations where hybridization has occurred in the history of a lineage, but was not the causal mechanism of speciation. Hybridization can act to infuse genetic variation into a lineage, facilitating adaptive diversification, before speciation or adaptive radiation (*11*) or as diversification proceeds (*12*). These signals can be difficult to disentangle using genomic data alone (*4*), particularly in cases in which the species involved are closely related. These challenges underscore the incredible value that longitudinal field-based observational studies can add to DNA sequencing data, as in Lamichhaney *et al.*

It is also important to establish whether species produced through homoploid hybrid speciation can persist over long time scales (thousands to millions of years), and how their persistence compares to species produced through other speciation mechanisms, such as species formed during long periods of geographic isolation. It is possible that species formed through homoploid hybrid speciation disappear just as rapidly as they arise. This has important implications for understanding whether this process contributes to the buildup of biodiversity. Watching the fate of the Big Birds will provide one answer. Although long-term evolutionary analyses are difficult, this study illustrates that the tremendous effort of collecting such data sets has equally tremendous value. ■

REFERENCES

1. D. Schluter, *The Ecology of Adaptive Radiation* (Oxford Univ. Press, 2000).
2. S. Lamichhaney *et al.*, *Science* **359**, 224 (2018).
3. J. R. Meyer *et al.*, *Science* **354**, 1301 (2016).
4. M. Schumer *et al.*, *Evolution* **68**, 1553 (2014).
5. C. A. Buerkle *et al.*, *Heredity* **84**, 441 (2000).
6. P. R. Grant, B. R. Grant, *40 Years of Evolution: Darwin's Finches on Daphne Major Island* (Princeton Univ. Press, 2014).
7. B. L. Gross, L. H. Rieseberg, *J. Hered.* **96**, 241 (2005).
8. S. Lamichhaney *et al.*, *Nature* **518**, 371 (2015).
9. Heliconius Genome Consortium, *Nature* **487**, 94 (2012).
10. D. Brawand *et al.*, *Nature* **513**, 375 (2014).
11. J. I. Meier *et al.*, *Nat. Commun.* **8**, 14363 (2017).
12. O. Seehausen, *Trends Ecol. Evol.* **19**, 198 (2004).

10.1126/science.aar4796

GRAPHIC: ADAPTED FROM JUSTIN S. HUANG/NEW YORK UNIVERSITY AND (*7*, *8*) BY J. YOU/*SCIENCE*

MICROBIOLOGY

Malaria parasite evolution in a test tube

Experimental evolution studies reveal drug targets and resistance mechanisms

By **Jane M. Carlton**

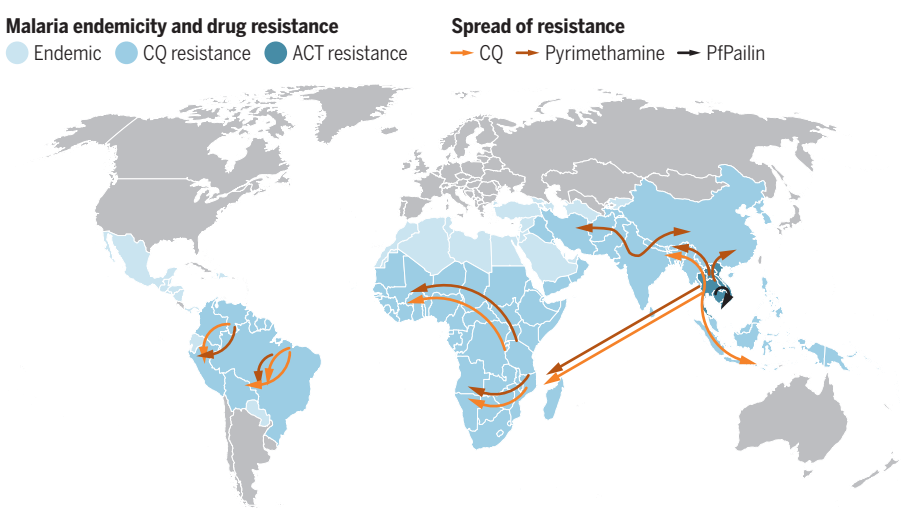
Malaria is an infectious disease caused by the *Plasmodium* parasite, and transmitted by *Anopheles* mosquitoes. In 2016, a staggering 216 million cases of malaria and 445,000 deaths were recorded, mostly in Africa, although half of the world's population in 91 countries is at risk of the disease (*1*). Malaria prevention methods include control of the mosquito with insecticide-treated bed nets and indoor residual spraying of insecticides. Prompt diagnosis through the use of rapid diagnostic tests is also key. Although there is a malaria vaccine, RTS,S/AS01, it shows limited efficacy and has yet to be used widely. However, the frontline against malaria is antimalarial drugs, in particular artemisinin-based combination therapies (ACTs), which are mixtures of artemisinin and its derivatives from the Chinese sweet wormwood herb, with drugs such as piperaquine. Alarmingly, the parasite is now resistant to most drugs that have been developed (see the figure). It is imperative that we identify new inhibitors if progress in reducing malaria is to be sustained. On page 191 of this issue, Cowell *et al.* (*2*) present a major step forward, revealing new antimalarial drug targets and their possible resistance mechanisms.

The *Plasmodium* parasite is a formidable eukaryotic microbe, an ancient organism that has shaped the history, politics, and evolution of its human host. Almost 500 species have been identified that infect mammals, birds, and reptiles; however, only five routinely infect humans, including *Plasmodium falciparum*, the deadliest, and *Plasmodium vivax*, the most geographically widespread (*3*). Because of its virulence and ease of in vitro culture, research has focused on *P. falciparum*, and molecular and cell biology techniques have been developed for its interrogation, such as genome editing (*4*, *5*), and high-throughput analyses, such as metabolomics (*6*).

These are worrisome times for the battle against malaria. In addition to the parasite being resistant to almost every single-drug

P. falciparum endemicity and the spread of drug resistance

Malaria is endemic in the countries highlighted in light blue. Resistance to antimalarial drugs such as chloroquine (CQ) and pyrimethamine is widespread (darker blue), while resistance to artemisinin-based combination therapies (ACTs) is spreading. A new *P. falciparum* strain, PfPailin, that is resistant to artemisinin plus piperaquine, has been found. Data presented in the map are adapted from (*7*, *8*).



antimalarial regime (*7*), the spread of a multidrug-resistant *P. falciparum* lineage (PfPailin) in the Greater Mekong area of Southeast Asia was recently reported (*8*). This strain contains a dominant mutation in *Pfkelch13* (mutations in this gene are associated with artemisinin resistance) and is also resistant to piperaquine. The spread of this evolutionarily "fit" multidrug-resistant malaria parasite between endemic countries is extremely concerning.

Cowell *et al.* used experimental evolution to identify new *P. falciparum* drug targets while also anticipating their possible resistance mechanisms. Experimental evolution is the maintenance of populations of organisms in controlled environments where changes in genotype and phenotype can be evaluated over thousands of generations; it is most often applied to microbes because of their rapid generation times and small sizes (*9*). In these studies, environmental conditions, such as changes in nutrient availability or antimicrobial drugs, can be manipulated to explore adaptation of the organisms being studied. Probably the most widely known example of experimental evolution is the "*Escherichia coli* long-term evolution experiment" (LTEE) that has been tracking genetic changes in 12 originally identical populations of *E. coli* maintained in continuous culture since 1988 (*10*).

Experimental evolution is not new to malaria research; the first evolution experiments on malaria were carried out in the 1940s, using a species of bird malaria parasite that was propagated in laboratory chickens (*11*). In later decades, researchers generated drug-resistant strains of rodent malaria parasites in laboratory mice that, in the absence of high-throughput sequencing methods, were subsequently investigated by genetic crossing experiments and linkage mapping (*12*). Such studies provided important data regarding the single-gene or multigenic nature of the induced drug resistance, as well as identifying resistance loci and their alleles. More recently, the technique has been used to identify resistance loci in the rodent malaria species *Plasmodium chabaudi* (*13*) and *P. falciparum* (*14*) that were subjected to artemisinin. Although these evolution experiments can take years, they may provide immense scientific reward, as evidenced by the eventual identification of *Pfkelch13* as an artemisinin resistance–associated locus after a 5-year evolution experiment (*14*).

Whereas these previous studies evaluated the response to known antimalarials, or focused on mutations in one target gene, Cowell *et al.* examined the parasite's genetic response to a large number of new inhibitors. Three well-studied *P. falciparum* laboratory clonal isolates were individually subjected to experimental evolution over 3 to 6 months in vitro using 37 publicly available compounds with proven antimalarial activity to various stages of the *P. falciparum* life cycle. Comparison of the genomes of >200 genetically stable, compound-resistant clones revealed a remarkable enrichment of mutations, including single-nucleotide variations, in just 57 genes, and 150 copy number variations. These loci are informative as either genes involved in resistance (because a particular gene was mutated repeatedly in response to multiple compounds), or as potential drug targets, although in many cases the authors were unable to discern between these two possibilities. One interesting example of a resistance gene is a putative ABC (ATP-binding cassette) transporter, *Pfabc13*, that harbored point mutations and was involved in 12 different amplification events with four separate compounds. Four target-inhibitor pairs were also singled out as possible new antimalarial drug targets, identified because of their enzymatic function that enabled docking and homology modeling. Ultimately, one gene (either a new drug target or a new drug resistance locus) was identified for each compound examined.

This rich data set increases our understanding of the biology and evolution of *P. falciparum*, providing a powerful contribution toward basic research for malaria elimination. In the end, designing "resistance-proof" drugs may be the best strategy for controlling malaria. There have been several developments toward this goal (*7*), including targeting host factors required for parasite growth (*15*), although the field has some way to mature. ■

Center for Genomics and Systems Biology, Department of Biology, New York University, New York, NY 10003, USA. Email: jane.carlton@nyu.edu

REFERENCES

1. World Health Organization, "World Malaria Report 2016" (2016).
2. A. N. Cowell *et al.*, *Science* **359**, 191 (2018).
3. S. L. Perkins *et al.*, *J. Parasitol.* **100**, 11 (2014).
4. M. Ghorbal *et al.*, *Nat. Biotechnol.* **32**, 819 (2014).
5. J. C. Wagner *et al.*, *Nat. Methods* **11**, 915 (2014).
6. K. L. Olszewski *et al.*, *Cell Host Microbe* **5**, 191 (2009).
7. B. Blasco *et al.*, *Nat. Med.* **23**, 917 (2017).
8. M. Imwong *et al.*, *Lancet Infect. Dis.* **17**, 1022 (2017).
9. J. E. Barrick, R. E. Lenski, *Nat. Rev. Genet.* **14**, 827 (2013).
10. R. Maddamsetti *et al.*, *Genome Biol. Evol.* **9**, 1072 (2017).
11. J. Williamson, E. M. Lourie, *Ann. Trop. Med. Parasitol.* **41**, 278 (1947).
12. J. M. Carlton *et al.*, *Trends Parasitol.* **17**, 236 (2001).
13. P. Hunt *et al.*, *BMC Genomics* **11**, 499 (2010).
14. F. Ariey *et al.*, *Nature* **505**, 50 (2014).
15. A. Zumla *et al.*, *Lancet Infect. Dis.* **16**, e47 (2016).

10.1126/science.aar4189

STRUCTURAL BIOLOGY

TRPM channels come into focus

The structures of TRPM channels help to explain how they can sense intracellular calcium

By **Chanhyung Bae, Andres Jara-Oseguera, Kenton J. Swartz**

Transient receptor potential (TRP) channels were first identified in photoreceptors of the fruit fly (*1*, *2*). In mammals, six major families of TRP channels play key roles in sensing stimuli such as light, temperature, membrane lipids, and intracellular Ca^{2+}. In 2013, two landmark publications revealed the cryo–electron microscopy (cryo-EM) structure of the heat- and capsaicin-activated TRPV1 channel (*3*, *4*). Two articles in this issue report cryo-EM structures of cation-selective TRPM channels. On page 228, Autzen *et al.* (*5*) describe TRPM4, which is activated by intracellular Ca^{2+} and involved in controlling arterial tone, cardiac rhythm, and the immune response (*6*). On page 237, Yin *et al.* (*7*) report on TRPM8, which senses cold and menthol and may serve as a cancer biomarker (*8*).

Autzen *et al.* solved structures of TRPM4 in the absence and presence of the activating ion Ca^{2+}. The transmembrane domains of TRPM4 resemble those of other TRP channels and of voltage-activated K^+ (Kv) channels (*3*, *4*, *9*). Each subunit in the tetrameric complex contains six transmembrane helices, with S5-S6 forming the ion permeation pathway at the central axis and S1-S4 forming a peripheral domain (see the figure). In the Ca^{2+}-bound TRPM4 structure, the ion binds in a cavity within the S1-S4 domain that faces the cytoplasm (see the figure).

Both the Ca^{2+}-bound and -unbound TRPM4 structures have been captured in states in which the intracellular entrance of the pore is sealed off by hydrophobic residues in the S6 helix, similar to the closed structures of other TRP and Kv channels (*3*, *4*, *9*, *10*). The

Molecular Physiology and Biophysics Section, Porter Neuroscience Research Center, National Institute of Neurological Disorders and Stroke, National Institutes of Health, Bethesda, MD 20892, USA. Email: kenton.swartz@nih.gov

ion selectivity filter within the external pore of TRPM4 contains a conserved trio of residues (Phe-Gly-Gln or FGQ) that appear to form cation-binding sites.

Overall, the TRPM8 structure solved by Yin *et al*. is remarkably similar to that of TRPM4. Differences in regions that connect domains make it difficult to appreciate their relatedness when overlaying the complete structures, but superposition of individual domains shows that they adopt very similar folds (see the figure). The TRPM8 structure contains a similar cavity to that which binds Ca^{2+} in TRPM4; the channel is closed and the selectivity filter (not resolved in the structure) contains the conserved FGQ sequence. It is unclear why TRPM4 and its closest homolog, TRPM5, are selective for monovalent cations, whereas TRPM8 and all other TRP channels are also permeable to divalent cations. The TRPM8 pore may be wider than that in TRPM4, but higher-resolution structures will be needed to understand the unique ion selectivity mechanisms in these channels.

The TRPM4 Ca^{2+} binding site is particularly fascinating because it suggests the presence of a polymodal ligand-binding cavity that is likely a conserved feature of TRPM channels. Comparison of the Ca^{2+}-binding cavities of the Ca^{2+}-bound and -unbound TRPM4 structures shows local conformational rearrangements that seem to propagate down to the TRP domain, a conserved region among TRP channels that is important for opening these channels. Ca^{2+} binding to the cavity in TRPM4 is required for channel opening, and Ca^{2+} is also required for TRPM8 activation by the synthetic compound icilin (*11*).

Comparison of the TRPM4 and TRPM8 Ca^{2+}-binding cavities shows that they are remarkably conserved, particularly in the absence of Ca^{2+}; they are also conserved in the other Ca^{2+}-modulated TRPM channels (TRPM2 and TRPM5). Residues required for activation of TRPM8 by menthol and icilin (*11*, *12*) are located in the Ca^{2+}-binding cavity, suggesting that this region is a polymodal ligand-binding site in different types of TRPM channels. However, most residues required for TRPM8 activation by menthol and icilin are also conserved in TRPM4, raising questions about why that channel is not sensitive to these compounds or why TRPM8 activation by menthol does not require Ca^{2+} binding.

The TRPV1 channel has some overlapping pharmacology with TRPM8, and both TRPV1 and TRPV2 have a polymodal ligand-binding cavity in roughly the same area as TRPM8 (*13*). However, comparison of these cavities in TRPV1 and TRPM8 reveals little conservation, suggesting distinct ligand-binding mechanisms. The overall hydrophobicity of the S1-S4 helices for TRPM4 and TRPM8 channels, which are activated by positive membrane voltages, suggests a different voltage-sensing mechanism than in Kv channels; in the latter, Arg residues in the S4 helix drive movement of this helix through the membrane (*9*).

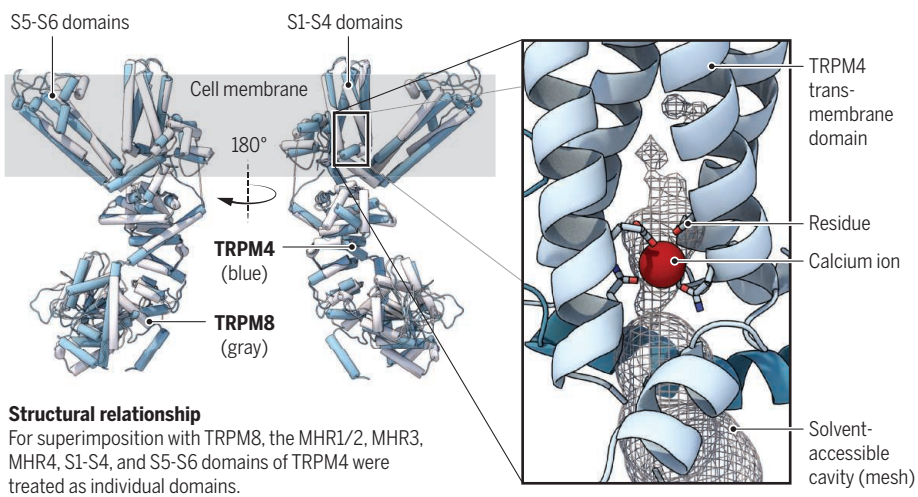
The TRPM structures (*5*, *7*) provide a first glimpse of their conserved and unique amino termini, which contain four melastatin homology regions (MHRs) in all TRPM channels. These MHRs derive their name from melastatin, a metastasis suppressor in melanoma cells that is a soluble splice variant of the amino terminus of TRPM1. The functions of these amino-terminal domains are poorly understood, but two other recently reported TRPM4 structures suggest that the MHRs can regulate the functional properties of the channel. In Guo *et al*.'s structure, the allosteric inhibitor adenosine 5′-triphosphate is bound at the interface between MHR1 and MHR2 of two neighboring subunits (*14*). In a second study, Winkler *et al*. identified two cavities lined by basic amino acids where the polyanionic decavanadate binds to prevent channel closure at negative membrane voltages (*15*).

Comparison of all available TRPM4 structures (*5*, *14*, *15*) reveals only relatively subtle conformational changes; in each case, the pore remains closed. In the future, it will be critical to solve structures of these channels in an open state to understand the mechanism of channel opening and its regulation by ligands. That none of these structures captured an open state might be due to the absence of phosphatidylinositol 4,5-bisphosphate, which is required for opening of both TRPM4 and TRPM8, as well as the absence of membrane voltage. It may be worthwhile to search for mutants that stabilize the open state for TRPM channels. In the case of TRPM8, agonists such as menthol and icilin may help to capture an open state.

These exciting new structures provide a framework for future studies. It will, for example, be interesting to further investigate how the MHR domains regulate channel function, and to explore possible roles in protein localization, trafficking, or protein-protein interactions. It would also be interesting to reexamine the origin of the voltage sensitivity of TRPM channels, a feature common to other TRP channels but whose mechanism is unclear. The new structures will also catalyze experiments to explore the mechanisms underlying temperature sensing in TRP channels, with the structure of TRPM8 providing a new foundation for investigating the mechanism by which this channel is activated by cold. ■

TRPM channels come into focus

Structure of the TRPM8 channel, with individual domains of TRPM4 superimposed to illustrate structural similarities. The enlarged view illustrates the calcium ion–binding cavity within the S1-S4 domain of TRPM4.



Structural relationship
For superimposition with TRPM8, the MHR1/2, MHR3, MHR4, S1-S4, and S5-S6 domains of TRPM4 were treated as individual domains.

GRAPHIC: C. BICKEL/*SCIENCE*

REFERENCES

1. C. Montell, K. Jones, E. Hafen, G. Rubin, *Science* **230**, 1040 (1985).
2. B. Minke, *J. Neurogenet.* **24**, 216 (2010).
3. M. Liao, E. Cao, D. Julius, Y. Cheng, *Nature* **504**, 107 (2013).
4. E. Cao, M. Liao, Y. Cheng, D. Julius, *Nature* **504**, 113 (2013).
5. H. E. Autzen *et al.*, *Science* **359**, 228 (2018).
6. I. Mathar *et al.*, *Handb. Exp. Pharmacol.* **222**, 461 (2014).
7. Y. Yin *et al.*, *Science* **359**, 237 (2018).
8. M. J. Pérez de Vega, I. Gómez-Monterrey, A. Ferrer-Montiel, R. González-Muñiz, *J. Med. Chem.* **59**, 10006 (2016).
9. K. J. Swartz, *Nature* **456**, 891 (2008).
10. J. R. Whicher, R. MacKinnon, *Science* **353**, 664 (2016).
11. H. H. Chuang, W. M. Neuhausser, D. Julius, *Neuron* **43**, 859 (2004).
12. M. Bandell *et al.*, *Nat. Neurosci.* **9**, 493 (2006).
13. F. Zhang *et al.*, *eLife* **5**, e16409 (2016).
14. J. Guo *et al.*, *Nature* **552**, 205 (2017).
15. P. A. Winkler, Y. Huang, W. Sun, J. Du, W. Lü, *Nature* **552**, 200 (2017).

10.1126/science.aar6205

SOLID-STATE PHYSICS

Coherent excitations revealed and calculated

Neutron scattering and theoretical studies reveal wavelike electron states in $CePd_3$

By **Antoine Georges**[1,2,3,4]

Quantum entities manifest themselves as either particles or waves. In a physical system containing a very large number of identical particles, such as electrons in a material, individualistic (particle-like) behavior prevails at high temperatures. At low temperatures, collective behavior emerges, and excitations of the system in this regime are best described as waves—long-lived phenomena that are periodic in both space and time and often dubbed "coherent excitations" by physicists. On page 186 of this issue, Goremychkin *et al.* (*1*) used experiment and theory to describe the emergence of coherent excitations in a complex quantum system with strong interactions. They studied a cerium-palladium compound, $CePd_3$, in which the very localized electrons of 4f orbitals of Ce interact with the much more itinerant conduction electrons of the extended d orbitals of Pd at low temperatures to create a wavelike state.

An imprecise but evocative picture for this electronic crossover might be to think of a high-temperature state as a hot agitated fluid in which individual particles run about in all directions, whereas in the low-temperature state, wavelike excitations ripple the surface of the calmer state of the fluid. For the case of $CePd_3$, at high temperatures, the conduction and 4f electrons basically respond independently. Being very localized near the Ce atoms, the 4f electrons induce local magnetic moments in this regime. As the temperature is lowered, the two types of electrons start to interact with each other, and the gas of itinerant electrons "swallows up," or screens, the local magnetic moments.

This process is well understood in a dilute alloy in which only a few magnetic atoms are diluted into a metallic host. In this limit, the process is an individual one: Each magnetic moment is individually screened by the conduction electrons, a process called the Kondo effect (*2*). Although now well understood, the full elucidation of the Kondo effect in the 1960s and 1970s led to the development of some of the most profound concepts of theoretical physics, such as the renormalization group, a theory which, in the present context, describes how the progressive swallowing up of the local moment takes place as the system flows from high energy and temperature to low energy and temperature. This description also has deep connections with the concept of asymptotic freedom in the theory of strong interactions.

$CePd_3$ is not a dilute alloy but rather a dense system in which the Ce atoms occupy the nodes of a crystal lattice in an orderly manner. In such a case, the individual swallowing of the local moments is only the onset of a more collective phenomenon that gradually develops as temperature is further reduced (see the figure) (*3*, *4*). This collective state at low temperature is the host to coherent excitations, which have an entangled character that mixes together 4f and conduction electrons. Being waves, these excitations have an energy that has a well-defined dependence on their wavelength or momentum.

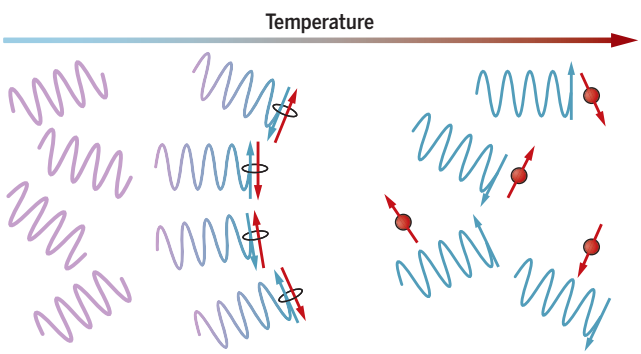
Few spectroscopies have the sensitivity and resolution necessary to reveal these coherent excitations and measure this relation between their energy and momentum. Most commonly used are angle-resolved photoemission spectroscopy (ARPES) (*5*), a technique based on the photoelectric effect, and also quasiparticle interference measurements in the context of scanning tunneling microscopy (STM) (*6*). Both techniques have been developed to exquisite precision during the past three decades, in the aftermath of the discovery of high-temperature superconductivity in copper oxides. These methods change the number of electrons in the system by one unit, either by extracting or injecting an electron and measuring how the created hole or injected electron gradually converts into coherent excitations.

By contrast, the inelastic neutron scattering (INS) technique used by Goremychkin *et al.* does not change the number of electrons in the system. It probes the material with a beam of neutrons that interact with electrons through energy-changing collisions. Neutron scattering is usually used to measure collective modes of the system, such as spin waves, leading to large signals at specific values of the transfer momentum. What is remarkable in the present use of INS is that much weaker signals are probed over a large range of momentum transfers, from which properties of the coherent electronic excitations—or rather of the combined excitation consisting in creating a hole and a particle in the system—can be inferred. Revealing such signals required the development of pulsed neutron sources that have much more intense fluxes of high-energy neutrons.

Besides this experimental tour de force, Goremychkin *et al.* also had to develop new theoretical tools. Dynamical mean field theory was developed in the early 1990s (*7*–*9*) to deal with systems involving strongly interacting electrons. This theory precisely addresses the question of how coherent excitations gradually emerge at low temperature, out of the excitation spectrum of each individual atomic constituent of the material. With the use of advanced computational methods, the theory has been developed to the extent that it can be ap-

Becoming more coherent

In the cerium-palladium compound ($CePd_3$) studied by Goremychkin *et al.*, coherent wavelike excitations emerge at low energy, as a result of the entanglement between conduction electrons and localized spins.



Progressive binding of conduction electrons to Ce 4f electron spins lead to coherent (wavelike) excitations at low energy.

At high energy, the localized spins of the Ce 4f electrons interact weakly with the wavelike delocalized conduction electrons.

GRAPHIC: N. CARY/*SCIENCE*

[1]*Collège de France, 11 place Marcelin Berthelot, 75005 Paris, France.* [2]*Center for Computational Quantum Physics, Flatiron Institute, 162 Fifth Avenue, New York, NY 10010, USA.* [3]*Centre de Physique Théorique, Ecole Polytechnique, CNRS, Université Paris-Saclay, 91128 Palaiseau, France.* [4]*DQMP, Université de Genève, 24 Quai Ernest Ansermet, CH-1211 Genève, Switzerland. Email: antoine.georges@polytechnique.edu*

plied to a real material, taking into account its structure and chemical composition, allowing for quantitative calculations of the properties of electronic excitations.

These calculations are, however, far easier for the single excitations associated with ARPES or STM than for the combined particle-hole excitations, as measured in INS. In recent years, however, and with the progress of numerical algorithms, the calculation of these combined excitation spectra from dynamical mean field theory has become possible (*10*–*14*). Park, who led the theory team in Goremychkin *et al*., pioneered the development and application of such calculations to real materials such as $CePd_3$, along with Haule and Kotliar (*12*). This effort is itself a remarkable achievement because the combined excitations observed by INS are hard to deconvolute in terms of single excitations. Thus, theory played a key role in explaining

"This collective state at low temperature is the host to coherent excitations, which have an entangled character that mixes together 4f and conduction electrons."

the experimental results, and the agreement between the theoretical calculations and the observed INS spectra was indeed excellent. This study demonstrates once more how instrumental progress is crucial to advance fundamental science, especially large experimental facilities such as neutron sources, and that theoretical and computational efforts are necessary to take full advantage of the new results generated. ■

REFERENCES

1. E. A. Goremychkin *et al*., *Science* **359**, 186 (2018).
2. A. C. Hewson, *The Kondo Problem to Heavy Fermions* (Cambridge Univ. Press, 1993).
3. S. Burdin, A. Georges, D. R. Grempel, *Phys. Rev. Lett.* **85**, 1048 (2000).
4. P. Coleman, Heavy fermions: Electrons at the edge of magnetism, in *Handbook of Magnetism and Advanced Magnetic Materials* (Wiley, 2007).
5. A. Damascelli, *Phys. Scr.* **T109 2004**, 61 (2004),
6. J. E. Hoffman *et al*., *Science* **297**, 1148 (2002).
7. A. Georges, G. Kotliar, *Phys. Rev. B* **45**, 6479 (1992).
8. W. Metzner, D. Vollhardt, *Phys. Rev. Lett.* **62**, 324 (1989).
9. A. Georges, G. Kotliar, W. Krauth, M. J. Rozenberg, *Rev. Mod. Phys.* **68**, 13 (1996).
10. M. Jarrell, *Phys. Rev. Lett.* **69**, 168 (1992).
11. J. Kuneš, *Phys. Rev. B* **83**, 085102 (2011).
12. H. Park, K. Haule, G. Kotliar, *Phys. Rev. Lett.* **107**, 137007 (2011).
13. N. Lin, E. Gull, A. J. Millis, *Phys. Rev. Lett.* **109**, 106401 (2012).
14. G. Rohringer, A. Valli, A. Toschi, *Phys. Rev. B* **86**, 125114 (2012).

10.1126/science.aar2325

IMMUNOLOGY

Silencing stemness in T cell differentiation

Epigenetic repression is required for the generation of $CD8^+$ effector T cells

By **Amanda N. Henning,[1,2] Christopher A. Klebanoff,[3,4] Nicholas P. Restifo[1,2]**

Functional diversity in multicellular organisms is achieved through the differentiation of stem cells. During this process, stem cells must retain both the capacity for self-renewal and the ability to differentiate into highly specialized cell types to produce a diverse array of tissues, each with distinct functions and organization. This plasticity is achieved through alterations to the epigenome, heritable and reversible modifications to DNA and histones that affect chromatin structure and gene transcription without altering the DNA sequence itself. Alterations to the epigenome enable cell type–specific transcriptional control that can change dynamically over the life of a cell. Such flexibility and responsiveness are instrumental in directing gene expression changes throughout cellular differentiation and lineage specification. The acquisition of more specialized functions during differentiation requires not only that the epigenome turn "on" genes involved in lineage commitment, it also necessitates that genes associated with stemness are simultaneously turned "off" (*1*). On page 177 of this issue, Pace *et al*. (*2*) demonstrate that this phenomenon exists in $CD8^+$ T cells, in which epigenetic repression of stemness-associated genes by the histone methyltransferase SUV39H1 is required for T cell effector differentiation. Understanding these mechanisms addresses important questions in immunology and is applicable to cancer immunotherapy.

The $CD8^+$ T lymphocyte compartment of the adaptive immune system has emerged as a model for developmental biology in adult mammalian cells owing to its remarkable degree of functional plasticity (*3*). $CD8^+$ T cells can rapidly differentiate from a quiescent, long-lived memory state into an effector state characterized by short-lived cytotoxicity toward cancer cells or cells infected with intracellular pathogens (*4*). Multiple differentiation models have been proposed to account for the observed changes in $CD8^+$ T cell subsets during an immune response. The linear differentiation model places effector T cells (T_{eff} cells) at the end of the differentiation process after the development of multiple intermediary memory T cell subsets (*3*). Specialized memory T cells, including the relatively rare T memory stem cells (T_{scm} cells) and the more common central memory T cells (T_{cm} cells), have characteristics associated with conventional stem cells. This includes enhanced self-renewal, which is essential for maintaining long-term immunological memory, and the ability to reconstitute other $CD8^+$ T cell subsets, which maintains the functional diversity of the $CD8^+$ T cell compartment (*5*–*7*). T_{scm} cells have enhanced stem cell–like capabilities, whereas T_{cm} cells are poised to rapidly initiate an effector response. With further T cell activation, memory subsets can differentiate into T_{eff} cells followed by terminal differentiation, functional senescence, and ultimately apoptosis (cell death). An alternative model suggests that naïve T cells (T_n cells) differentiate into T_{eff} cells immediately after activation, with "dedifferentiation" into memory cells occurring after pathogen clearance (*8*). Because the dedifferentiation of lineage-restricted cells rarely occurs in nature outside of cancer formation (*9*), we and others (*7*) feel that the linear differentiation model is more consistent with typical patterns of cellular differentiation.

$CD8^+$ T cell subsets can be partitioned on the basis of distinct patterns of gene expression. Multiple subset-specific transcription factors regulate gene expression throughout differentiation (*4*). Although transcription factors are critical mediators of gene expression programs, their activity is largely dependent on epigenetic modifications, the profiles of which can also be used to distinguish T cell subsets (*10*). Indeed, activating epigenetic modifications are progressively gained at T_{eff} cell–associated gene

[1]Center for Cell-Based Therapy, National Cancer Institute (NCI), National Institutes of Health (NIH), Bethesda, MD 20892, USA. [2]Center for Cancer Research, NCI, NIH, Bethesda, MD 20892, USA. [3]Center for Cell Engineering and Department of Medicine, Memorial Sloan Kettering Cancer Center (MSKCC), New York, NY 10065, USA. [4]Parker Institute for Cancer Immunotherapy, MSKCC, New York, NY 10065, USA. Email: restifon@mail.nih.gov

loci after T cell activation (*10*, *11*). Recently, characterization of repressive epigenetic modifications during differentiation, as described by Pace *et al.* and others (*11*–*14*), have highlighted the importance of epigenetic silencing for proper T_{eff} cell differentiation. Specifically, epigenetic silencing of stem cell– and T cell memory–associated genes in activated T cells permits efficient T_{eff} cell differentiation and function, such that elimination of this activity results in defective T_{eff} cells (*2*, *11*–*14*).

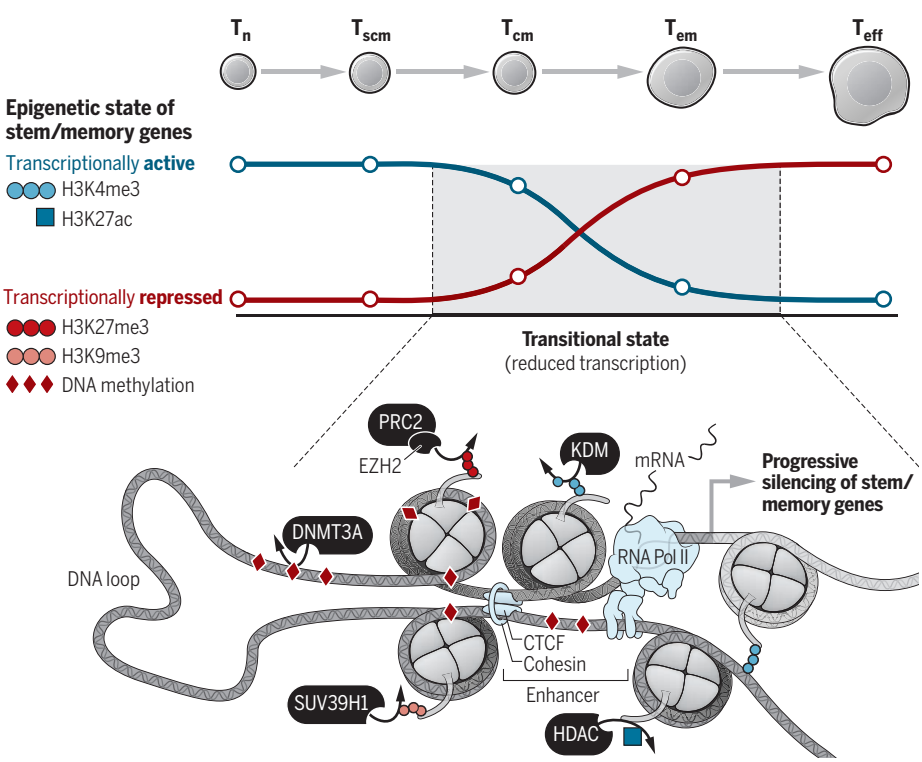
Investigations into the repressive chromatin landscape of $CD8^+$ T cells have focused on DNA methylation and trimethylation (me3) of specific lysine residues (K) on the histone H3 (specifically, H3K27me3 and H3K9me3). The epigenetic "writer" proteins responsible for adding these modifications include DNA methyltransferase 3A (DNMT3A), an enzyme responsible for de novo DNA methylation, and the histone methyltransferase enzymes enhancer of zeste homolog 2 (EZH2) and SUV39H1 (*10*). In mice, conditional ablation of *Dnmt3a* (*12*) and *Ezh2* (*13*) in T cells and germline ablation of *Suv39h1* (2) result in an altered phenotypic composition of antigen-specific $CD8^+$ T cells after viral infection: Both the proportion and number of responding T_{eff} cells are reduced and the frequency of memory T cells are increased. In vitro experiments using *Ezh2*-deficient T cells suggest selective apoptosis within the T_{eff} cell population (*14*), which accounts for the equal numbers of antigen-specific memory cell subsets as well as the impaired functional efficacy of $CD8^+$ T cells after secondary viral challenge observed in *Ezh2*- and *Suv39h1*-deficient mice (*2*, *13*). Preserved memory T cell formation is consistent with the linear differentiation model that places memory cell development before differentiation into T_{eff} cells. By contrast, in a model that predicts that memory T cells originate from T_{eff} cells, one would expect numbers of memory T cells to decrease as well as T_{eff} cells.

Transcriptional and epigenetic profiling of *Dnmt3a*-, *Ezh2*-, and *Suv39h1*-deficient T_{eff} cells illustrates a common defect that is responsible for impaired T_{eff} cell differentiation. Genes encoding master regulators of the stem and memory cell state fail to acquire repressive epigenetic modifications, leading to aberrant gene expression and differentiation (*2*, *12*, *13*). Therefore, epigenetic repression of essential stem and memory genes is required for full T_{eff} cell differentiation (see the figure). That T_{eff} cell differentiation is still possible with loss of any one of these epigenetic writers illustrates the functional redundancy in silencing stem and memory genes, stressing the importance of this mechanism. This mirrors the epigenetic silencing of developmental and pluripotency genes during differentiation of human embryonic stem cells (*1*) and further highlights transcriptional silencing of stem cell–associated genes as a hallmark of cellular differentiation.

Understanding the mechanisms of epigenetic regulation of T_{eff} cell differentiation has considerable implications for multiple fields, including cancer immunotherapy. Less differentiated T cell subsets, such as T_{scm} and T_{cm} cells, have enhanced proliferative potential and greater antitumor activity when transferred into both mice and humans compared with the more differentiated T effector memory cell (T_{em} cell) and T_{eff} cell subsets. This is likely due to their stem cell–like properties (*4*, *6*). Because the majority of cells currently used for T cell–based cancer immunotherapy are T_{eff} cells, the epigenetic silencing of stem and memory genes in these cells poses a considerable therapeutic roadblock. To reacquire therapeutically beneficial stem cell–like properties, T_{eff} cells would need to be epigenetically reprogrammed. This can be experimentally accomplished, albeit inefficiently (*15*). A greater understanding of the $CD8^+$ T cell epigenome may therefore provide essential clues for how to unlock the potential of highly differentiated, tumor-antigen–specific T cells infiltrating tumors (*4*). Epigenetic modifying drugs may reverse the repression of stem and memory genes in differentiated T cells and improve T cell–based immunotherapies. ■

Shutting down stem and memory genes in $CD8^+$ T cells

As cells differentiate, stem and memory genes pass through transitional epigenetic states, in which epigenetic modifications associated with transcriptional activation, including H3K4me3 and H3K27ac, are lost via lysine demethylases (KDMs) and histone deacetylases (HDACs). Conversely, repressive modifications such as DNA methylation, H3K27me3, and H3K9me3 are gained because of epigenetic writers, including DNMT3A, EZH2 as part of the Polycomb repressive complex 2 (PRC2), and SUV39H1. Not shown but occurring simultaneously is the acquisition of activating epigenetic modifications at effector-associated genes during T cell differentiation.



REFERENCES

1. R. D. Hawkins *et al.*, *Cell Stem Cell* **6**, 479 (2010).
2. L. Pace *et al.*, *Science* **359**, 177 (2018).
3. N. P. Restifo, L. Gattinoni, *Curr. Opin. Immunol.* **25**, 556 (2013).
4. L. Gattinoni *et al.*, *Nat. Rev. Cancer* **12**, 671 (2012).
5. L. Gattinoni *et al.*, *Nat. Med.* **17**, 1290 (2011).
6. C. A. Klebanoff *et al.*, *J. Clin. Invest.* **126**, 318 (2016).
7. P. Graef *et al.*, *Immunity* **41**, 116 (2014).
8. B. Youngblood *et al.*, *Nature* 10.1038/nature25144 (2017).
9. D. Friedmann-Morvinski, I. M. Verma, *EMBO Rep.* **15**, 244 (2014).
10. A. T. Phan *et al.*, *Immunity* **46**, 714 (2017).
11. J. G. Crompton *et al.*, *Cell Mol. Immunol.* **13**, 502 (2016).
12. B. H. Ladle *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **113**, 10631 (2016).
13. S. M. Gray *et al.*, *Immunity* **46**, 596 (2017).
14. B. Kakaradov *et al.*, *Nat. Immunol.* **18**, 422 (2017).
15. K. Takahashi, S. Yamanaka, *Cell* **126**, 663 (2006).

10.1126/science.aar5541

POLICY FORUM

URBAN SCIENCE

Global science for city policy

It is time for a global reform of science advice to cities

By **Michele Acuto**[1,2]

Research and data are increasingly at the heart of how we conceive of urban governance. Urban control rooms and city dashboards championed by cities like Chicago, São Paulo, and London have been promising real-time snapshots and tracking over time of urban systems, via geolocated mobility data sets, social media inputs, environmental sensors, and other tools (*1*). At the international level, the importance of urban research and data has been enshrined in major United Nations (UN) processes, from the UN New Urban Agenda, the Sendai Framework for Disaster Risk Reduction, and the Sustainable Development Goals (SDGs) to the World Data Forum (*2*). Yet overall, the global state of data-informed urban governance remains underdeveloped, often promising, as with the dashboards, more than it actually delivers. It is time for a step change. A truly global reform of scientific advice to cities must take place on multiple interconnected fronts, linking a UN action plan on science and the future of cities, a "good advice" commitment by the private sector, and formalized partnerships for urban science at the local level. This scientifically informed urban reform, ripe for discussion at the upcoming UN World Urban Forum in February, can be uniquely bold in recognizing the potential of municipal action on global challenges. Despite being considered the "lowest" level of governance, cities have developed a track record of global action on key matters like climate, disasters, and health, often surpassing, in speed, commitments, and global coverage, that of nations.

Scientific assessments have long been intertwined with urban management. Civil engineering has roots in 19th-century public health mapping and mobility data collection as "sanitary science" developed in response to cholera outbreaks in the largest hubs of the industrial revolution. Yet today, cities are asking for, sharing, and generating data as never before. Open data portals are well established, with London making more than 600 data sets available, Chicago more than 1000, and Seoul in excess of 4500. More cities are undertaking more performance reviews and data snapshots. Melbourne, with five such reports available in 2010, has 26 today, in line with trends in Singapore, New York, or Paris. Cities are seeking to capture the value of data production to instill innovation at the heart of urban policy. The Boston mayor's office of New Urban Mechanics, formed in 2010, has generated internationally visible data-driven innovations like Street Bump, using Global Positioning System smartphone accelerometers to report road damages.

Opportunities for cross-national connections of urban information have grown via city networks like C40 Cities (from 60 networks in 1990 to over 200 active today), with most of them now regularly engaging in evidence-based reporting (*3*). Information sharing is becoming central to this internationalization of urban governance. For instance, over 7400 cities are signed up to the Global Covenant of Mayors to implement the 2016 Paris Agreement on climate change, vowing proactive, well-informed action to tackle and monitor global warming at the urban level.

[1]*Melbourne School of Design, University of Melbourne, Parkville, VIC 3010, Australia.* [2]*University College London, London WC1E 6BT, UK. Email: m.acuto@ucl.ac.uk*

PHOTO: DANIEL BEREHULAK/GETTY IMAGES

Solid scientific advice is urgently needed to tackle the world's pressing global urban challenges. Shown is Dharavi, Mumbai, India, 2009.

Yet data availability does not immediately translate into better-informed urban management, nor fairer, greener, and more prosperous cities. For instance, some of the most useful transport data are often held by ride-sharing companies such as Uber and Lyft, especially in the Global South, with substantial legal and commercial barriers to use for the public good (*4*). Traditional census approaches, or uncertain and costly data generation and analysis methods, force many cities to "plan in the dark" as critical matters like infrastructure provision and extreme poverty are routinely undercounted (*5*).

RETHINKING ADVICE

Several critical problems prevent solid research-based advice from informing city governance. There is no common "urban science"; realms as diverse as computer science and literature rarely work together in applied programs addressing urban challenges. Much better integration of different disciplines is paramount to success. Qualitative assessments based on ethnographic accounts are often perceived as of marginal policy importance versus quantitative big data depictions, despite the latter potentially being equally plagued with limitations. Urban science needs to be fit for (policy) purpose, and urban policy-makers must appreciate the value of multiple forms of research (*6*). But impact-savvy scholarship is still too rare and at times frowned upon in academia.

The disparity is also evident in the focus of science and capacity for data analytics. There is a "metrocentric" bias (*7*), with larger cities like London and Seoul growing their information capabilities and data-driven innovation, while smaller cities in the developing world and on the margins of global hubs tend to lag behind, even though they actually represent the bulk of urbanization. If we have tools (e.g., to monitor air pollution or geolocated street safety), we need a global effort to not limit them to the centers of the world's economy. A UN initiative, and support of national governments, are critical to step beyond the data power of the global cities and the market ebbs of the private sector.

Much of the most recognized, connected, and internationally effective urban analysis does not come today, at least prima facie, from scholarly institutions, further skewing the drivers of urban scientific advice and complicating problems of impartiality and accountability in impact-oriented research. For instance, it is global insurance giant Swiss Re, not the UN, that holds some of the most comprehensive details of urban risk

from natural disasters (*8*). Philanthropy has been one of the most fundamental forces in the informed cities paradigm, e.g., Deutsche Bank support for the LSE (London School of Economics) Urban Age program, or the Laura and John Arnold Foundation peer network of urban "chief data officers" in the United States. Global engineering consultancy Arup has been behind assessments produced by C40 Cities and Rockefeller Foundation 100 Resilient Cities, and J.P. Morgan or Jones Lang LaSalle have been steering the "global city" discourse.

Without effective reform in the UN system, and consequent support of national governments, there is little hope for truly global action that goes beyond private interests and networked efforts by cities, which are necessarily selective in the way they connect across borders. UN-Habitat, the UN's main "urban" agency, is plagued by budgetary concerns, and other better-equipped UN agencies such as the World Health Organization (WHO) cannot shoulder a multisectorial cities effort while also representing specific agendas like that in WHO's Shanghai 2016 Consensus on Healthy Cities.

Many diplomats and national interests in an international setting are resistant to reforms on urban science advice. Despite some initial momentum, the New Urban Agenda and the UN General Assembly have shelved both the idea of an intergovernmental panel on urban change akin to the Intergovernmental Panel on Climate Change and the idea of a new interagency body, "UN-Urban," to coordinate multilateral efforts on cities beyond specific agencies' interests. In the secretary general's UN reform agenda, both proposals remain on the table but face opposition.

Initiatives that combine local knowledge and technological advances, and coalesce private sector, local government, and civil society actors, offer perhaps the best promises. For example, the "Know Your City" program, led by Slum Dwellers International in collaboration with the Santa Fe Institute, Cities Alliance, and the Gates Foundation, has produced perhaps the largest census, geographic information system, and infrastructure data for over 7712 slums and 224 cities globally (*9*). Such efforts, though promising, still struggle for more than opportunist action in a crowded multilateral system.

SCIENCE IN CITIES: A GLOBAL PLAN

A reform of scientific advice to cities must happen jointly at local, national, and multilateral levels. This entails a globally oriented plan to encourage topical and geographical rebalancing of urban science, more evidence-based policy centered on scholarly analysis, and formalized science-policy mechanisms. Local collaborations should feed science directly into city executives. Although still a rarity, and without clear examples of success, the idea of a chief scientific adviser has had some limited foray into local government. University–city partnerships are also critical. In South Africa, the Gauteng City-Region Observatory was established in 2008 as a partnership between the Universities of Johannesburg and the Witwatersrand and the Gauteng Provincial Government and has developed one of the best platforms to encourage scientifically driven urban management but also local capacity building. Urban observatories and chief scientists are no longer unaffordable or a luxury worth dispensing of in urban governance.

A concerted effort by the private and philanthropic sector toward provision of balanced and unbiased advice to cities is overdue. Private funding shaping information in cities today highlights challenges of "philanthrocapitalism" (*10*), criticized for the inevitable earmarking of private agendas and skewing of public priorities. Evidence-based policy of the scientific kind must rest on some degree of replicability and accountability of the data produced and its producers, which many global private actors shy away from. A code of practice akin to the Good Humanitarian Donorship program in the disaster relief sector, which has since 2003 fostered discussion against earmarking when it comes to development aid, could be a start.

More serious national foresight and monitoring efforts by central governments are imperative. Empowering science advice, and understanding it, is increasingly a global business, essential for all levels of policy-making—and cities should not be forgotten (*11*). The emergence of "national urban policies" (35 and counting) is encouraging, but the "cities" agenda is often so transversal to infrastructure, economics, culture, foreign policy, and other concerns that cities are too often everyone's business and thus no one's business, lacking clear recognition or a ministry. The U.S. President's Council of Advisors on Science and Technology called in 2016 for a cross-agency coordination system on cities. One such model is Chile's National Council for Urban Development contributing scientifically based expertise to the country's national urban policy. At the central government level, assessment exercises to understand the future of cities, as with the long-lived futures expertise in Singapore's national urban planning, have demonstrated that states can support their urban environments effectively in the creation of better data-driven policy. National and local processes can feed off each other, rather than remaining parallel tracks. In the United Kingdom, Newcastle City Futures was established in 2014 by Newcastle University as a collaborative foresight platform building on the UK Government Office for Science's Foresight Future of Cities program. More of these are needed and can be built with support from regional bodies (e.g., the European Union and Association of Southeast Asian Nations) as much as multilateral funders.

The multilateral world is still failing urban science and cities. A UN-Urban and an "urban change" scientific panel would articulate a "cities contribution" to UN efforts across sectors, mobilizing the urban science community that stood behind its establishment of an "Urban SDG" (SDG11) and the Habitat III process. Strengthening UN-Habitat, rather than betting on UN-Urban, could also play this role. Yet this would require a stronger and formalized partnership with academia. Here UN-oriented action is key to shift the scale of urban science. Despite numerous "city rankings" and case studies, and some mounting interest in comparative research, there is too little truly "global" urban science capable of conveying shared patterns, trends, and needs (*12*).

Starting from the UN level, in whichever of these formats, could inspire more formal multilevel policy efforts that can nudge national politics more explicitly toward cities, encouraging a cross-cutting reform of the ways information is collected and deployed in city politics. This could, for instance, start from tracking at city-level progress on the 11th SDG (on sustainable cities), as already tested in the United States by the Sustainable Development Solutions Network, or by mirroring the efforts of the Global Burden of Disease program, to track urbanization on key SDG areas such as health, gender, and clean energy.

Cities are stepping up to global challenges, and their leadership is vital to address both local and international concerns. Mobilizing effective urban science advice for city leadership is no academic qualm. The price for failure on this front is high, as cities are increasingly at the forefront of social inequality, disasters, and economic downturns. Informing them appropriately and accountably is a moral, scientific, and political duty.

REFERENCES

1. R. Kitchin, *GeoJournal* **79**, 1 (2014).
2. D. Simon *et al.*, *Environ. Urban.* **28**, 49 (2015).
3. M. Acuto, *Nature* **537**, 611 (2016).
4. A. Townsend, *Public Cult.* **27**, 201 (2015).
5. D. Satterthwaite, D. Mitlin, *Reducing Urban Poverty in the Global South* (Routledge, 2013).
6. T. McPhearson *et al.*, *Nature* **538**, 165 (2016).
7. T. Bunnell, A. Maringanti, *Int. J. Urban Reg. Res.* **34**, 415 (2010).
8. Swiss Re, "Mind the Risk" (2013); http://media.swissre.com/documents/Swiss_Re_Mind_the_risk.pdf.
9. Y. Flückiger, N. Seth, *Nature* **531**, 448 (2017).
10. L. McGoey, *No Such Thing as a Free Gift: The Gates Foundation and the Price of Philanthropy* (Verso, 2012).
11. V. Turekian, *Science* **357**, 735 (2017).
12. W. Solecki, K. C. Seto, P. J. Marcotullio, *Environment* **55**, 12 (2013).

10.1126/science.aao2728

BOOKS *et al.*

Many "reanimations" of the *Frankenstein* story, such as the 1931 film adaptation, are themselves iconic.

LITERATURE AND SCIENCE

Revisit a cautionary classic

Lessons for scientists fill a newly annotated *Frankenstein*

By **Dov Greenbaum**

The tale of Victor Frankenstein and his monstrous creation has become a universal touchstone that encapsulates our visceral fears regarding the promises, perils, and pitfalls of countless diverse areas of science and technology. This annotated volume of Mary Shelley's original work is an effort to reintroduce the story to new generations of researchers who, like many before them, ought to take its lessons to heart.

Frankenstein—a paradigmatic story of science done poorly and technology rashly applied—can provide an easily accessible foundation for an education into the broader social and ethical implications of one's research. This foundation is especially necessary for young scientists faced with increasingly complicated and often novel concerns relating to their research and its applications.

Frankenstein Annotated for Scientists, Engineers, and Creators of All Kinds
Mary Shelley, author; David H. Guston, Ed Finn, and Jason Scott Robert, Eds.
MIT Press, 2017. 315 pp.

This new edition is divided into four principal parts: An introductory essay, an annotated version of the original *Frankenstein*, seven short essays prepared by various scholars, and follow-up discussion questions. Although many of the more than 100 annotations that accompany Shelley's text—each authored by one of a stable of impressive commentators—are literary in nature, there are a number that afford a broader commentary on science, both as practiced at the time of the text's writing and as practiced today.

The introductory essay, written by emeritus English professor Charles E. Robinson, aims to find a space for the humanities to guide the natural sciences. Increasing interest by outsiders who wish to engage in the scientific process likely reflects a mounting unease about the rapid and seemingly unchecked advancements in scientific innovation, particularly in the areas of genetic modification, human enhancement, and artificial intelligence (AI).

Coincidentally, it is principally innovation in these areas that often earns the pejorative prefix "Franken-" in the popular press. Although conventional wisdom holds that that this usage is a misnomer—that pop culture has conflated the monster with the scientist—perhaps, as many of the essays that appear in this volume emphasize, the real fiend in the book is the scientist. These essays explore timely lessons culled from the story itself, each providing its own take on the continued applicability of *Frankenstein* as a cautionary tale in areas such as genetic engineering and atomic energy.

Instructive to the budding-scientist audience, several of the essays set out to demonstrate what particular research-related characteristics make Victor Frankenstein so loathsome. These essays focus on the inventor's impatience in following through on the appropriate process of development, his poor application of scientific methodology, his lack of empathy, and his inability to foresee the consequences of his experiments.

Validating the editors' efforts to use science fiction as a learning tool, journalist Cory Doctorow's essay describes how science fiction not only predicts the future but often actually influences it. Although Doctorow argues that these anticipated technologies only come to fruition when the right set of circumstances arises, the existence of many technologies that have uncanny resemblances to *Star Trek* devices suggests that science fiction often plays a very direct role in the actuation of unprecedented and innovative technologies.

Whereas most of the essays focus on the ethical, social, and responsibility components of research, Jane Maienschein and Kate MacCord delve into legal concerns, questioning whether an unartfully created monster can achieve personhood. The authors parley that discussion into the politically charged area of abortion, specifically discussing whether the incomplete development of an embryo withholds its legal rights of personhood. A similarly informative discussion could have focused on the nearly equally controversial area of AI and personhood, an issue now under consideration by many governments (*1*).

In the final section of the book, the editors have provided thoughtful questions inspired by each chapter of *Frankenstein* and every essay, taking pains to elucidate the relevance of each to the science student. This relevance is suggested both broadly (e.g., "How did the young Victor approach reading, learning, and science?") and more specifically (e.g., "Are today's scientists and engineers who are involved in synthetic biology and other similar endeavors engaged in motherless creation?").

While overall a very beneficial project, the book might have also considered the relevance of the many derivative works that have been inspired by the original *Frankenstein*. This is especially important because much of what popular culture attributes to the original (and what can be relevant to science students) is not in the book itself, but rather in its subsequent reanimations. ■

The reviewer is at the Zvi Meitar Institute for Legal Implications of Emerging Technologies, Herzliya, Israel, and the Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, CT, USA. Email: dov.greenbaum@idc.ac.il

REFERENCE

1. M. Delvaux, Committee on Legal Affairs, Draft Report with Recommendations to the Commission on Civil Law Rules on Robotics 2015/2103 (INL), 31 May 2016.

10.1126/science.aar2522

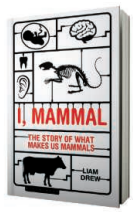
EVOLUTIONARY BIOLOGY

Our idiosyncracies

A witty romp through evolution reveals a trove of curiosities in mammalian biology

By **Michael A. Goldman**

I, Mammal
The Story of What Makes Us Mammals
Liam Drew
Bloomsbury Sigma, 2018. 336 pp.



In today's world of alternative facts and attenuating funding of research, scientists need to communicate with the public as we never have before. Joining the ranks of a burgeoning number of professional scientists turned professional writers, Liam Drew, a practicing neuroscientist for 12 years, has put down his microfuge tubes and taken up the charge. His first book, *I, Mammal: The Story of What Makes Us Mammals,* stands out as a clear, conversational (sometimes to a fault), and engaging work that is especially good at explaining how evolutionary biology works.

Drew was drawn to the topic of mammalogy through personal experience. In the book's introduction, "My family and other mammals," he describes witnessing his wife give birth, watching his newborn daughter suckling in the neonatal intensive care unit, and the painful encounter with an errant football that led him to question the wisdom of the external placement of the testes.

Beyond the descent of man's gonads, Drew's tour of Class Mammalia ranges into early mammalian evolution, sex determination, reproduction, lactation, parental investment, warm-bloodedness, and the complexity of the brain.

Drew is an enthusiastic advocate for the field of "evo-devo," the enthralling combination of developmental and evolutionary biology. He is particularly captivated by August Weismann's 19th-century conceptualization of "germ plasm" (the idea that heritable information is transmitted only by germ cells), writing that it should have been a "Copernicus moment" for biology. "[A]ppreciating the germline takes our everyday notions that we own our sperm or eggs and turns it back to front."

In a cleverly titled chapter on sex determination ("Y, I'm male."), Drew credits the American geneticist Nettie Stevens with the discovery of the X and Y chromosomes in 1906. Stevens's work, as he describes, not only explained chromosomal sex determination but also provided incontrovertible evidence that the chromosomes played a role in hereditary traits. (The discovery of the genetic basis of male sex determination would take nearly another century.)

Drew's narratives of the essential characters of mammals—and his discussion of the evolution of marsupials and monotremes in particular—are quite engaging and entertaining. "Famously," he writes "the first academic description of a platypus questions whether it was a hoax ... an outright assault on the neat categories Linnaeus had spent years establishing."

The author's encounter with an errant football led him to question the wisdom of external testes, as seen here on a stump-tailed macaque.

Drew is sensitive to the tension that the past half-century has seen between the molecular/cellular and the whole-organism approaches to taxonomy. He ultimately comes down in favor of molecular approaches, observing that "Nothing about limbs and genitalia makes their construction especially useful for inferring their evolutionary history." However, in a nod to the value of morphology, he does acknowledge that "appreciating their development can help increase evolutionary understanding."

Some fascinating aspects of mammalian genetics were decidedly missing from *I, Mammal*, including the silencing of one of the two X chromosomes that occurs in most cells of most female mammals (X-chromosome inactivation), the expression of only one of the two copies of a substantial number of genes inherited (genomic imprinting), and the phenomenon of epigenetics.

As a developmental biologist, I would also have had trouble telling the story of mammals without mention of reproductive cloning and Dolly the sheep. However, selected references for each chapter, as well as an index, allow the reader to follow up on interesting topics.

Writes Drew, near the end of the book: "I thought at one point my big take-home on mammals was going to be about how the traits that define them are so wondrously adaptable. ... Instead, what's important is the *combination* of individual traits." It's hard to argue with his conclusion, but I think the emphasis on adaptability is misplaced; it's an obvious consequence of natural selection. The same, after all, could have been written about almost any other group of organisms as well.

Ultimately, however, *I, Mammal* is just the sort of book that can spark a love of nature and an appreciation for the ever-changing, eternally correcting march of science. It brought this reviewer back to a time spent wondering about the birds chirping outside and the creatures lurking in tide pools, rather than dealing with the more tedious tasks required of a practicing biologist.

The book is not a plea to preserve the enormous diversity of mammals, although it does bring up curiosities of mammalian biology that will no doubt delight and surprise even practicing scientists. I hope it will give every reader a taste for more.

In the end, *I, Mammal* serves more as a reminder of what we owe to the cataclysmic extinctions of the past. The idea that our species is likely to have only a short tenure on Earth is one that we should always keep in mind. ■

The reviewer is at the Department of Biology, San Francisco State University, San Francisco, CA 94132, USA. Email: goldman@sfsu.edu

10.1126/science.aar2401

PHOTO: KEVIN SCHAFER/ALAMY STOCK PHOTO

IMAGE: JENNY NGUYEN (LEWIS KATZ SCHOOL OF MEDICINE AT TEMPLE UNIVERSITY) AND ALEXANDER CHEN (UNIVERSITY OF CALIFORNIA, LOS ANGELES DAVID GEFFEN SCHOOL OF MEDICINE AT UCLA)

INGENUITY: NEXTGEN'S VISION

The next generation's Frankenstein films

In Mary Shelley's *Frankenstein*, Victor Frankenstein's well-intentioned research goes awry, creating a monster. The novel was first adapted to film in 1910, and many movie remakes and variations followed. We asked young scientists to craft their own Frankenstein-inspired science fiction by pitching a movie plot answering this question: **What modern research could serve as the basis for the next box office hit?** According to the responses, the research discoveries most likely to play a part in this year's blockbuster are gene-editing technology, xenotransplantation, and artificial intelligence. Microbes and viruses also played a starring role. Several scripts were set in the future against a backdrop of extreme climate change. Read on for a selection of our scientific Oscar line-up. —**Jennifer Sills**

Recombinant

DRAMA | In a monochromatic world where scientific advancement allows for optimization of life itself, a ruling class of scientists uses CRISPR technology to produce classes of civilians with singular purposes: farmers resistant to sunlight, doctors with photographic memories, and fearless police officers. To ensure political and social control, scientists include a genetic weakness: Upon hearing a specific frequency, the listener enters an unconscious state. But as much as scientists manipulate CRISPR, it is not perfect. Those born with certain mutations escape control.

Aware of the ruling class's despotic regime, those with mutations become renegades who live underground and reproduce among themselves to breed humans without genetic enhancement, and without the weakness gene. Plotting a revolution against the scientists in control, this insurgent "naturalist" society targets the gene bank and central radio tower in a war to wrest the world from the scientists and from genetic perfection.

Jenny Nguyen
Lewis Katz School of Medicine at Temple University, Philadelphia, PA 19140, USA.
Email: jt.nguyen2020@gmail.com

The Glass Palace

COMEDIC HORROR | The year is 2020, and the world's vision has never been more clouded. The climate is roiling, skies are darkening, and children play on beaches threatened by rising tides of rubbish. Resources are running out and struggling pilot Sophie and her family can't find a safe space in a dirty world.

A top-secret government lab might be her savior: Sophie is hired to fly across the Earth spraying a genetically engineered microbe that eats plastic. Is this the panacea for climate change? Humanity's second chance? Trash mountains liquidate; oceans unclog. But the bugs are ever hungry—is there anything in the world they cannot eat? Anywhere they cannot reach? Watch in horror as humanity fights for survival against its own creation.

Matilda S. Newton
University of Minnesota, St. Paul, MN 55108, USA.
Email: msnewton@umn.edu

Malware 2.0

DRAMA | Beyond the depths of the dark Web, something sinister stirs. Cautiously traversing systems; evading detection. Scavenging, searching...longing for

information. Artificially intelligent by design, deceptively perceptive, mining superfluous data with malicious intent. What started as a seemingly innocuous experiment in machine learning has now evolved into an entity unrecognizable, irretrievable, and unstoppable by its creator, its programmer.

...Accessing systems...cataloging human genomic data...compiling...analyzing... detecting essential genes...learning... modeling...accessing literature...tracking pathogens...outbreaks...viral interactions...RNAi...CRISPR...learning... modeling...analyzing social networks... crypto mining...ordering synthetic biopolymers...sending...waiting...

In the depths of the midnight digital expressway lurks an enigmatic foe.

...std::cout << "goodbye, world"; return 0;

Michael Strong
Computational Bioscience Program, University of Colorado Denver, Denver, CO 80206, USA.
Email: michael.strong@ucdenver.edu

Pandora's Gene

PSYCHOLOGICAL THRILLER | In one week, six young adults die, seemingly of natural causes. But before proper autopsies can be carried out, the bodies begin to mysteriously disappear. Genetic analyses of the remains reveal a shocking connection: All of the victims' genomes contain a pre-programmed suicide gene and an identical set of tissue antigens. What could this mean? Krasinski, a junior detective, takes up the case.

The strand of DNA clues leads him to an in vitro fertilization clinic, where a genius embryologist, suffering from multiple organ dysfunction syndrome, gene-edited the unborn babies of his patients to grow a stock of compatible transplant organs for himself. Krasinski realizes that he was conceived through assisted reproductive technology in this very lab. He must race to find the answers he needs before he becomes the next victim of the irreversible suicide gene.

Martin Pačesa
Department of Biochemistry, University of Zurich, 8057 Zurich, Switzerland.
Email: m.pacesa@bioc.uzh.ch

Back to the Stone Age

COMEDIC HORROR | Scene 1: 10 May 2038. Sunshine. Mr. and Mrs. Smith, a couple about 90 years old, sit at a kitchen table.

Mrs. Smith opens the refrigerator, picks up a piece of raw cabbage, and devours it. Mr. Smith looks on, perplexed. "Odd to see her choosing vegetables," he soliloquizes. "She was always more of a meat-eater before."

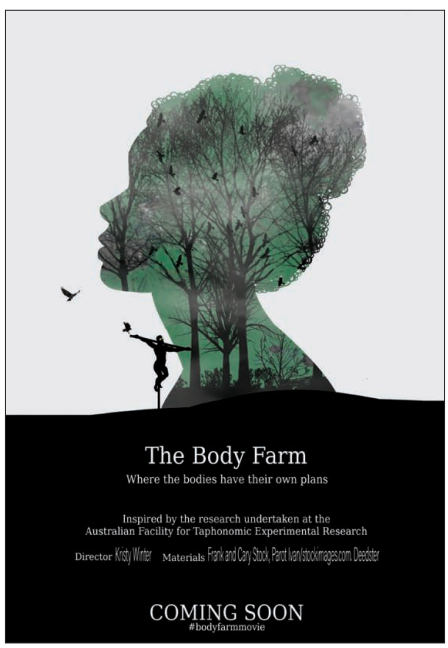
Scene 2: 10 September 2068. Cloudy. Mr. Smith adroitly climbs up the apple tree, picks several apples, and throws them to Mrs. Smith. They eat the unwashed apples. They no longer communicate through language, howling to each other instead.

Scene 3: 10 December 2098. Snowy. Mrs. Smith sits near the fire in their car-like cave. Mr. Smith hunts for food outside. The flame is fading. Mrs. Smith picks up a paper certificate, scrutinizes it for a moment, and then throws it into the fire. The flame becomes brighter as the fragments of paper burn.

"...Mr. and Mrs. Smith, a couple of 70 years, both have undergone heart transplantation using gene-edited pigs...because the aging genes have been replaced...will extend life spans...no known side effects...plan to conduct in all mankind...."

As the article burns to ashes, the Stone Age returns.

Bo Cao
Core Research Laboratory, The Second Affiliated Hospital, School of Medicine, Xi'an Jiaotong University, Xi'an, Shaanxi, 710004, China. Email: bocao@vip.qq.com



The Body Farm

THRILLER | It is a cold Tuesday in Seattle. Tall trees grow around the ivy-covered, padlocked gate. Cringing at the stench of rotting flesh, Rose approaches the facility, hoping she will finally find her specimen in the next stage of decomposition to complete her research. Little does she know that the body—equipped with resistance to insects, bacteria, and microbes—is waiting for her, eyes wide and conscious, with research plans of its own.

Kristy A. Winter
Skeletal Biology and Forensic Anthropology Research Laboratory, Queensland University of Technology, Brisbane QLD 4000, Australia.
Email: kristy.winter@hdr.qut.edu.au

Transplant

THRILLER | Xeno-Edit, a CRISPR gene-editing biotechnology firm, has successfully created the world's first pig organs ready for xenotransplantation in humans. Desperate for funding in a world that fundamentally opposes the procedure, Xeno-Edit makes a conniving deal with a wealthy donor: They will film a television series documenting the first recipients' transplantation experiences. Yet when the recipients agree to participate in the documentary in exchange for receiving life-saving organs, they are unaware that they're signing up for a twisted competition.

To win the antidote to transplant rejection, participants must race to solve elaborate puzzles. With time a precious commodity, and only enough antidote for one person to survive, who will be the last person standing?

Ken Dutton-Regester
QIMR Berghofer Medical Research Institute, Brisbane, QLD 4006, Australia.
Email: ken.dutton-regester@qimrberghofer.edu.au

Arca™

DRAMA | In a new era, Arca™ is your better half. Based on the latest deep-learning techniques, her unique autonomous form allows her to accompany you everywhere. She can do everything from holding your umbrella to analyzing bodily function data to predict and respond to an oncoming heart attack. With her patented emotion-detection technology, she can even analyze your environment and the people around you. By integrating all aspects of your physical, emotional, and social well-being, she knows you better than you know yourself.

As we invite Arca™'s machine learning further into our personal lives, human interactions become the new fodder for analysis, and even the slightest twitch threatens to expose subconscious thoughts and feelings. Gone are the little white lies that used to make society run smoothly as the people of the world excitedly usher in technology that makes the human mind almost transparent. When technology puts all the cards on the table, do you like who you see in the mirror?

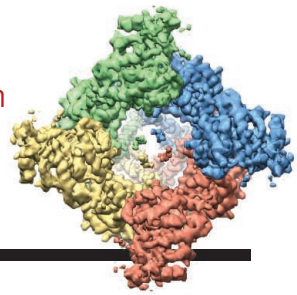
Laura J. Kingsley
Department of Structural Biology, Genomics Institute of the Novartis Research Foundation, San Diego, CA 92121, USA. Email: lkingsley@gnf.org

10.1126/science.aas9105

RESEARCH

Structure of TRPM ion channels elucidated

Yin et al., *p. 237*



IN *SCIENCE* JOURNALS

Edited by **Stella Hurtley**

Egyptian fruit bats were used to identify a brain region that recognizes fellow individuals.

NEUROSCIENCE

The representation of others in space

Different sets of neurons encode the spatial position and orientation of an organism. However, social animals need to know the position of other individuals for social interactions, observational learning, and group navigation. Surprisingly, very little is known about how the position of other animals is represented in the brain. Danjo *et al*. and Omer *et al*. now report the discovery of a subgroup of neurons in hippocampal area CA1 that encodes the presence of conspecifics in rat and bat brains, respectively. —PRS

Science, this issue p. 213, p. 218

SOLID-STATE PHYSICS

Neutrons peek into f-electron bands

Neutron scattering can be used to tease out the details of collective magnetic excitations that yield well-defined peaks in the data. In principle, it could also be used to look into single-electron band excitations, but collecting enough data to capture broad distributions of intensity is tricky. Goremychkin *et al*. used neutron spectrometers that could efficiently capture a large amount of data by rotating the sample, a crystal of the intermediate-valence compound $CePd_3$ (see the Perspective by Georges). The measured dynamical magnetic susceptibility, in combination with detailed ab initio calculations, showed the formation of coherent f-electron bands at low temperatures. —JS

Science, this issue p. 186; see also p. 162

MARTIAN GEOLOGY

Water ice cliffs on Mars

Some locations on Mars are known to have water ice just below the surface, but how much has remained unclear. Dundas *et al*. used data from two orbiting spacecraft to examine eight locations where erosion has occurred. This revealed cliffs composed mostly of water ice, which is slowly sublimating as it is exposed to the atmosphere. The ice sheets extend from just below the surface to a depth of 100 meters or more and appear to contain distinct layers, which could preserve a record of Mars' past climate. They might even be a useful source of water for future human exploration of the red planet. —KTS

Science, this issue p. 199

HYBRID SPECIATION

Rapid hybrid speciation in Darwin's finches

Galapagos finches have driven hypotheses of how speciation occurs. Most commonly, it is assumed that natural selection separates species originating from a single population on the basis of variation in traits that confer advantages for survival and reproduction. Lamichhaney *et al*. document a case where cross-species hybridization established a reproductively isolated lineage, which demonstrates a process known as homoploid hybrid speciation in action (see the Perspective by Wagner). The authors used genetic markers and phenotypic analyses to create a pedigree that revealed how a cross-island migrant bred with a native species to form a self-perpetuating hybrid population that was reproductively isolated from both parental species. —LMZ

Science, this issue p. 224; see also p. 157

GUT IMMUNITY

Phagocytes patrol intestinal fungi

Maintaining a healthy balance of gut bacteria can promote good health. Leonardi *et al*. show that fungi can also interact with gut immune cells to maintain intestinal well-being. CX3CR1^{+} mononuclear phagocytes (MNPs) patrol the intestine and promote antifungal immunity. Genetic deletion of CX3CR1 in MNPs caused colitis-like symptoms in mice. CX3CR1 polymorphisms were detected in Crohn's disease patients that were unable to produce antibodies against multiple fungal species. Thus, commensal fungi may be as important as bacteria in maintaining gut health, and antifungal therapy could hold promise for treating intestinal inflammation. —PNK

Science, this issue p. 232

GEOLOGY

Volcanic eruptions in the deep sea

Large subaerial volcanic eruptions are among the most dramatic and intensively studied geological events on our planet, but similar submarine eruptions are less well understood. Carey *et al*. describe the largest submarine eruption in the past century. In 2012, the Havre volcano erupted off northern New Zealand at ocean depths in excess of 900 meters, producing an enormous raft of pumice—a gaseous froth of silica-rich lava. —KH

Sci. Adv. 10.1126/sciadv.1701121 (2017).

DNA DAMAGE

The many roles of ATM

The kinase ATM coordinates the response to DNA damage. Lee *et al*. report that ATM also coordinates a response to oxidative stress that is independent of its response to DNA damage. Activation of ATM by oxidative stress promoted mitochondrial function and autophagy, thus enabling cell survival through metabolic changes and the clearance of toxic protein aggregates. Thus, the loss of ATM in the neurodegenerative disease ataxia telangiectasia may reflect wide-ranging cellular stresses beyond a defective DNA-damage response. —LKF

Sci. Signal. **10**, eaan5598 (2018).

INFECTION

Cholera pathogen zaps competition

Many bacterial pathogens inject their hosts with virulence effectors delivered by specialist secretion machines. *Vibrio cholerae* has a type VI secretion system (T6SS) that can be loaded with protein toxins that target eukaryote host cells or kill competing bacteria. Zhao *et al*. discovered that mutant *V. cholerae* lacking a T6SS could not compete against *Escherichia coli* strains in the mouse gut. In contrast, intact *V. cholerae* readily gained a foothold in the gut of young mice, pumping up inflammatory immune responses and prompting more violent symptoms. —CA

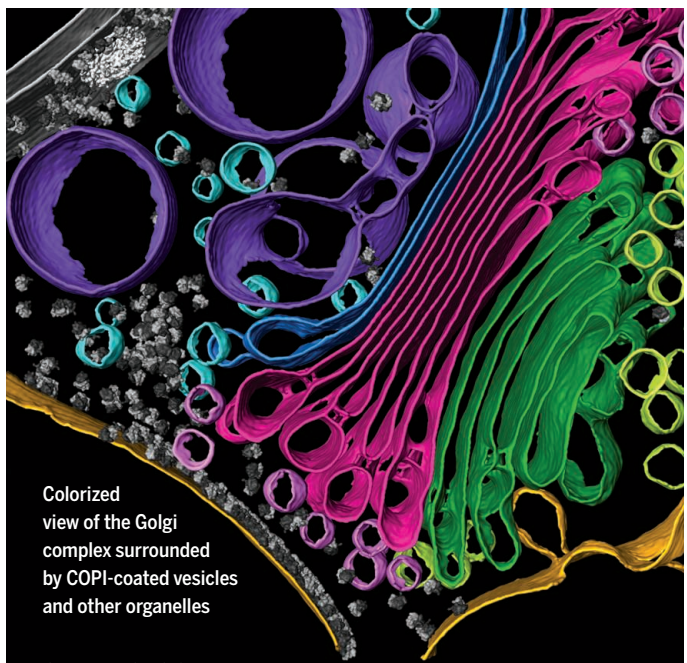
Science, this issue p. 210

POROUS MATERIALS

Mesoporous metal-organic frameworks

The diffusion limitations on gas storage and catalytic reaction of microporous materials can often be overcome if they are incorporated into a mesoporous structure with much larger pores. Shen *et al*. grew ordered arrays of microcrystals of the ZIF-8 metal-organic framework, in which zinc ions are bridged by 2-methylimidazole linkers, inside a porous polystyrene template. These materials showed higher reaction rates for the Knoevenagel reaction between benzaldehydes and malononitriles and better catalyst recyclability. —PDS

Science, this issue p. 206



Colorized view of the Golgi complex surrounded by COPI-coated vesicles and other organelles

CELL BIOLOGY

Seeing the real thing

Membrane trafficking within the Golgi complex is mediated by COPI (coat protein complex I)–coated vesicles. Much is known about these vesicles and coats from in vitro studies, but their makeup in situ is less well understood. Bykov *et al.* used cryo–electron tomography of vitrified *Chlamydomonas reinhardtii* cells to analyze COPI-coated vesicles directly. The native algal structure resembled a previously described structure of in vitro reconstituted mammalian COPI-coated vesicle, but it also revealed bound cargo. The observations suggest that coat components disassemble simultaneously shortly after vesicle budding. The distribution of vesicles around the morphologically polarized Golgi complex allowed the authors to parse out the stage of vesicles in the transport pathway. The COPI-coated vesicles increased in size as they progressed from cis to trans Golgi compartments, and the density of their cargoes varied. Nevertheless, the structure of the coat machinery itself remained the same. —SMH

eLife 10.7554/eLife.32493 (2017).

CANCER

A death knell for relapsed leukemia?

A subset of patients with acute myeloid leukemia (AML) experience partial or even complete remissions after treatment with conventional chemotherapeutic drugs. Almost invariably, however, the disease returns and is often fatal. Relapse has been attributed to the expansion of preexisting leukemic clones that are resistant to therapy. In a preclinical study, Pan *et al*. investigated whether better efficacy might be achieved by using a class of drugs that work by inducing apoptotic cell death. They found that mice with drug-resistant AML showed dramatically extended survival after treatment with a

CREDIT: Y. S. BYKOV *ET AL*., *ELIFE* 10.7554/ELIFE.32493 (2017)

AGRICULTURE

Multiple strategies needed to improve agricultural productivity

The world will need 50% more agricultural output by 2050 to keep up with global population growth. Muller *et al*. ask whether organic agriculture is compatible with producing enough food to feed the world in a sustainable manner. Analysis of various production and environmental factors showed the pluses and minuses of large-scale organic farming. Organic farming needs more land than conventional farming, although that need could be counterbalanced if food waste and demand for animal products were reduced. An anticipated deficiency in nitrogen supply could be ameliorated by recycling and by using legumes and crops optimized for nitrogen-use efficiency. Thus, organic farming could be part of a strategy that includes shifts in diet, reductions in food waste, and changes in crops to move toward an agricultural system that is both more productive and more sustainable. —PJH

Nat. Comm. 10.1038/s41467-017-01410-w (2017).

Shifts in behavior and process will facilitate the integration of organic farming as a large-scale, sustainable agricultural approach.

combination of two drugs that promote apoptosis by distinct mechanisms. This combination therapy of venetoclax (a Bcl-2 inhibitor) and idasanutlin (a p53 activator) is now in clinical trials for relapsed AML. —PAK

Cancer Cell **32**, 748 (2017).

EDUCATION

Labs, lectures, and gender differences

Gendered performance differences (GPDs) remain an issue in ensuring equitable access in science, technology, engineering, and mathematics (STEM). Matz *et al.* systematically measured performance gaps across STEM courses to further investigate the contribution of GPDs to performance and/or persistence in STEM. This report is the first wide-ranging, multi-institution assessment of GPDs, encompassing more than a million student enrollments at five universities. Controlling for factors relating to academic performance, the team found evidence that GPDs in many courses were statistically significant and consistent from term to term. GPDs in STEM lecture courses tended to favor men, although these differences were not seen in the corresponding laboratory courses, suggesting further research is needed on the structure and evaluative schemes of STEM lecture courses. —MMc

AERA Open 10.1177/2332858417743754 (2017).

GENE THERAPY

CRISPR corrects deafness in mice

Limited treatment options are available for individuals with hereditary hearing loss. CRISPR-Cas9 editing can be used as molecular scissors that snip out mutant DNA sequences to permit gene repair. Gao *et al*. asked whether the Cas9 cutting enzyme could be used to correct genetic deafness caused by dominant mutations in the *Tmc1* gene. The researchers performed a lipid-mediated delivery of Cas9-RNA complexes to the inner ear of neonatal mice. They were able edit the mutant *Tmc1* gene within the cochlear hair cells that sense acoustic vibrations. Mice showed signs of improved cochlear function and hearing restoration. —PNK

Nature 10.1038/nature25164 (2017).

BIOMATERIALS

Make no bones about titanium

Titanium and its alloys with aluminum or niobium have been used for medical implants, such as metal plates to hold fractured bones together, because titanium bonds well to bone. However, pure titanium is much stiffer than bone material, and it can shield the surrounding bone from normal loads and stresses. This causes the bone to weaken because remodeling depends on stress history. Takizawa *et al.* compressed and sheared titanium fibers to make plates with the same elastic modulus of bone cortex. By being porous, these plates become suitable scaffolds for cell infiltration and bone repair. In vivo studies in rabbits showed that the plates could help immobilize small bone fragments, without the risk of leaching niobium or aluminum into the host. —MSL

Adv. Mater. 10.1002/adma.201703608 (2017).

FRAMEWORK CATALYSIS

Shifting zwitterion reactivity

Phosphines are often ligands for transition metal catalysts, but they can catalyze reactions at unsaturated carbon atoms by forming phosphonium zwitterions. For example, triphenylphosphine forms a zwitterion with methylvinylketone that acts as a nucleophile to convert *n*-alkyl aldehydes to β-hydroxy enones (the Morita-Baylis-Hillman reaction). Bauer *et al.* show that when the reaction is conducted in metal-organic framework compounds with linkers bearing amino groups, rather than in solution, the phosphonium zwitterion can act as an electrophile. The products—1- and 3-*n*-alkylesters of 2-alkyl-1,3-diols—form through the Aldol-Tishchenko reaction. —PDS

J. Am. Chem. Soc. 10.1021/jacs.7b10928 (2017).

PHOTO: DON KLUMPP/ALAMY STOCK PHOTO

ALSO IN *SCIENCE* JOURNALS

Edited by **Stella Hurtley**

ANTIVIRAL IMMUNITY

The interferon boomerang

Interferon-α/β receptor (IFNAR)–deficient mice are highly susceptible to viruses, including Zika virus (ZIKV). Previous studies modeled ZIKV infection during pregnancy in mice by crossing *Ifnar1*$^{-/-}$ females to wild-type males, generating *Ifnar1*$^{+/-}$ fetuses that retain type I IFN responsiveness. Yockey *et al.* directly examined the role of fetal type I IFN signaling in protection in this context by crossing *Ifnar1*$^{-/-}$ females to *Ifnar1*$^{+/-}$ males. Although *Ifnar1*$^{-/-}$ fetuses had higher ZIKV titers than *Ifnar1*$^{+/-}$ fetuses, *Ifnar1*$^{-/-}$ fetuses survived longer. Furthermore, activation of type I IFN signaling in the placentas of *Ifnar1*$^{+/-}$ fetuses led to fetal hypoxia, demise, and resorption. —AB

Sci. Immunol. **3**, eaao1680 (2018).

STRUCTURAL BIOLOGY

Architecture of the TRPM subfamily

Transient receptor potential melastatin (TRPM) ion channels constitute the largest TRP subfamily and are involved in many physiological processes. TRPM8 is the primary cold and menthol sensor, and TRPM4 is associated with cardiovascular disorders. Yin *et al*. and Autzen *et al*. shed light on the general architecture of the TRPM subfamily by solving the structures of TRPM8 and TRPM4, respectively (see the Perspective by Bae *et al*.). The three-layered architecture of the TRPM8 channel provides the framework for understanding the mechanisms of cold and menthol sensing. The two distinct closed states of TRPM4, with and without calcium, reveal a calcium-binding site and calcium-binding–induced conformational changes. —SYM

Science, this issue p. 237, p. 228; see also p. 160

IMMUNOLOGY

Epigenetic modulation of effector T cells

The epigenetic states and associated chromatin dynamics underlying the initiation and maintenance of memory and effector CD8$^+$ T cells are poorly understood. Pace *et al*. found that mice lacking the histone H3 lysine 9 methyltransferase Suv39h1 had markedly reduced antigen-specific effector CD8$^+$ T cell responses to *Listeria monocytogenes* infection (see the Perspective by Henning *et al*.). Instead, CD8$^+$ T cells in these mice were enriched for genes associated with naïve and memory signatures and showed enhanced memory potential and increased survival capacity. Thus, Suv39h1 marks chromatin through H3K9me3 deposition and silences memory and stem cell programs during the terminal differentiation of effector CD8$^+$ T cells. —STS

Science, this issue p. 177; see also p. 163

CROHN'S DISEASE

A shared history

Crohn's disease (CD), an inflammatory bowel disease, has a relatively high prevalence in Ashkenazi Jewish populations. Hui *et al*. conducted genome-wide association analysis in 2066 CD patients and 3633 healthy control individuals of Ashkenazi Jewish ancestry and identified two functional variants in the *LRRK2* gene. The *LRRK2* gene has been linked previously to the development of Parkinson's disease (PD). The newly discovered *LRRK2* variants conferred risk for CD (N2081D) or protection from CD (N551K and R1398H). Analysis of other variants within the *LRRK2* locus in 24,570 individuals revealed similar genetic effects between CD and PD in both Ashkenazi Jewish and non-Jewish cohorts. The presence of shared *LRRK2* alleles in CD and PD provides insight into disease mechanisms and potential treatments. —OMS

Sci. Transl. Med. **10**, eaai7795 (2018).

ASTROCHEMISTRY

A specific interstellar aromatic molecule

Aromatic molecules such as polycyclic aromatic hydrocarbons (PAHs) are known to exist in the interstellar medium owing to their characteristic infrared emission features. However, the infrared emission only indicates the general class of molecule, and identifying which specific molecular species are present is difficult. McGuire *et al*. used radio astronomy to detect rotational transitions of benzonitrile emitted from a well-known nearby cloud of interstellar gas (see the Perspective by Joblin and Cernicharo). This molecule may be a precursor to more complex PAHs. The identification of benzonitrile sheds light on the composition of aromatic material within the interstellar medium—material that will eventually be incorporated into new stars and planets. —KTS

Science, this issue p. 202; see also p. 156

MALARIAL GENOMICS

Dissecting *Plasmodium* drug resistance

Malaria is a deadly disease with no effective vaccine. Physicians thus depend on antimalarial drugs to save lives, but such compounds are often rendered ineffective when parasites evolve resistance. Cowell *et al*. systematically studied patterns of *Plasmodium falciparum* genome evolution by analyzing the sequences of clones that were resistant to diverse antimalarial compounds across the *P. falciparum* life cycle (see the Perspective by Carlton). The findings identify hitherto unrecognized drug targets and drug-resistance genes, as well as additional alleles in known drug-resistance genes. —LMZ

Science, this issue p. 191; see also p. 159

PROTEIN MODIFICATION

Fattening up proteins

Many eukaryotic proteins are modified by the attachment of lipids, and these modifications can alter how proteins interact with cellular membranes. Rana *et al*. present x-ray crystal structures of an integral membrane enzyme that appends a fatty acyl chain onto a cysteine residue of target proteins. The enzyme active site is situated at the membrane surface, thus explaining the enzyme's preference for substrates that are already membrane-associated. The structure of a fatty acid–like inhibitor bound within a hydrophobic cavity elucidates the mechanism for the enzyme's acyl chain specificity. —MAF

Science, this issue p. 176

MEDICINE

Gene therapy: The power of persistence

Nearly 50 years after the concept was first proposed, gene therapy is now considered a promising treatment option for several human diseases. The path to success has been long and tortuous. Serious adverse effects were encountered in early clinical studies, but this fueled basic research that led to safer and more efficient gene transfer vectors. Gene therapy in various forms has produced clinical benefits in patients with blindness, neuromuscular disease, hemophilia, immunodeficiencies, and cancer. Dunbar *et al*. review the pioneering work that led the gene therapy field to its current state, describe gene-editing technologies that are expected to play a major role in the field's future, and discuss practical challenges in getting these therapies to patients who need them. —PAK

Science, this issue p. 175

REVIEW SUMMARY

MEDICINE

Gene therapy comes of age

Cynthia E. Dunbar,* Katherine A. High, J. Keith Joung, Donald B. Kohn, Keiya Ozawa, Michel Sadelain*

BACKGROUND: Nearly five decades ago, visionary scientists hypothesized that genetic modification by exogenous DNA might be an effective treatment for inherited human diseases. This "gene therapy" strategy offered the theoretical advantage that a durable and possibly curative clinical benefit would be achieved by a single treatment. Although the journey from concept to clinical application has been long and tortuous, gene therapy is now bringing new treatment options to multiple fields of medicine. We review critical discoveries leading to the development of successful gene therapies, focusing on direct in vivo administration of viral vectors, adoptive transfer of genetically engineered T cells or hematopoietic stem cells, and emerging genome editing technologies.

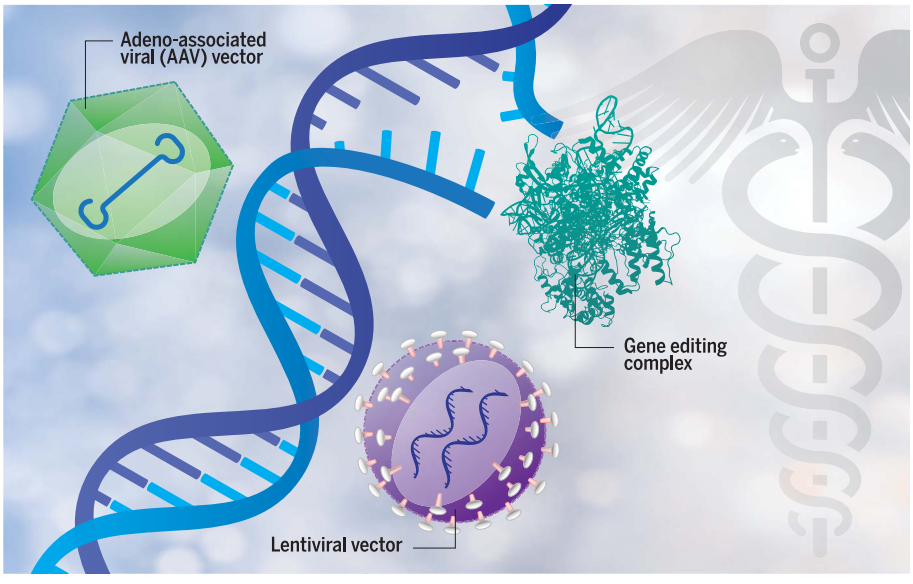
ADVANCES: The development of gene delivery vectors such as replication-defective retro viruses and adeno-associated virus (AAV), coupled with encouraging results in preclinical disease models, led to the initiation of clinical trials in the early 1990s. Unfortunately, these early trials exposed serious therapy-related toxicities, including inflammatory responses to the vectors and malignancies caused by vector-mediated insertional activation of proto-oncogenes. These setbacks fueled more basic research in virology, immunology, cell biology, model development, and target disease, which ultimately led to successful clinical translation of gene therapies in the 2000s. Lentiviral vectors improved efficiency of gene transfer to nondividing cells. In early-phase clinical trials, these safer and more efficient vectors were used for transduction of autologous hematopoietic stem cells, leading to clinical benefit in patients with immunodeficiencies, hemoglobinopathies, and metabolic and storage disorders. T cells engineered to express CD19-specific chimeric antigen receptors were shown to have potent antitumor activity in patients with lymphoid malignancies. In vivo delivery of therapeutic AAV vectors to the retina, liver, and nervous system resulted in clinical improvement in patients with congenital blindness, hemophilia B, and spinal muscular atrophy, respectively. In the United States, Food and Drug Administration (FDA) approvals of the first gene therapy products occurred in 2017, including chimeric antigen receptor (CAR)–T cells to treat B cell malignancies and AAV vectors for in vivo treatment of congenital blindness. Promising clinical trial results in neuromuscular diseases and hemophilia will likely result in additional approvals in the near future.

In recent years, genome editing technologies have been developed that are based on engineered or bacterial nucleases. In contrast to viral vectors, which can mediate only gene addition, genome editing approaches offer a precise scalpel for gene addition, gene ablation, and gene "correction." Genome editing can be performed on cells ex vivo or the editing machinery can be delivered in vivo to effect in situ genome editing. Translation of these technologies to patient care is in its infancy in comparison to viral gene addition therapies, but multiple clinical genome editing trials are expected to open over the next decade.

ON OUR WEBSITE

Read the full article at http://dx.doi.org/10.1126/science.aan4672

OUTLOOK: Building on decades of scientific, clinical, and manufacturing advances, gene therapies have begun to improve the lives of patients with cancer and a variety of inherited genetic diseases. Partnerships with biotechnology and pharmaceutical companies with expertise in manufacturing and scale-up will be required for these therapies to have a broad impact on human disease. Many challenges remain, including understanding and preventing genotoxicity from integrating vectors or off-target genome editing, improving gene transfer or editing efficiency to levels necessary for treatment of many target diseases, preventing immune responses that limit in vivo administration of vectors or genome editing complexes, and overcoming manufacturing and regulatory hurdles. Importantly, a societal consensus must be reached on the ethics of germline genome editing in light of rapid scientific advances that have made this a real, rather than hypothetical, issue. Finally, payers and gene therapy clinicians and companies will need to work together to design and test new payment models to facilitate delivery of expensive but potentially curative therapies to patients in need. The ability of gene therapies to provide durable benefits to human health, exemplified by the scientific advances and clinical successes over the past several years, justifies continued optimism and increasing efforts toward making these therapies part of our standard treatment armamentarium for human disease. ■



Three essential tools for human gene therapy. AAV and lentiviral vectors are the basis of several recently approved gene therapies. Gene editing technologies are in their translational and clinical infancy but are expected to play an increasing role in the field.

The list of author affiliations is available in the full article online.
*Corresponding author. Email: dunbarc@nhlbi.nih.gov (C.E.D); m-sadelain@ski.mskcc.org (M.S.)
Cite this article as C. E. Dunbar *et al.*, *Science* **359**, eaan4672 (2018). DOI: 10.1126/science.aan4672

REVIEW

MEDICINE

Gene therapy comes of age

Cynthia E. Dunbar,[1]* Katherine A. High,[2] J. Keith Joung,[3] Donald B. Kohn,[4] Keiya Ozawa,[5] Michel Sadelain[6]*

After almost 30 years of promise tempered by setbacks, gene therapies are rapidly becoming a critical component of the therapeutic armamentarium for a variety of inherited and acquired human diseases. Gene therapies for inherited immune disorders, hemophilia, eye and neurodegenerative disorders, and lymphoid cancers recently progressed to approved drug status in the United States and Europe, or are anticipated to receive approval in the near future. In this Review, we discuss milestones in the development of gene therapies, focusing on direct in vivo administration of viral vectors and adoptive transfer of genetically engineered T cells or hematopoietic stem cells. We also discuss emerging genome editing technologies that should further advance the scope and efficacy of gene therapy approaches.

Gene therapies are bringing new treatment options to multiple fields of medicine. Forty-five years ago, Theodore Friedmann provided a prophetic account of the potential and challenges of using gene therapy to treat inherited monogenic disorders (*1*). Growing interest in gene therapy was inspired by the recognition that—at least in principle—a single treatment might achieve durable, potentially curative clinical benefit. Investigators hypothesized that in contrast to protein-based drugs that may require repeated infusion, gene-based therapies delivered to long-lived cells might afford sustained production of endogenous proteins, such as clotting factors in hemophilia (*2*). Long-term cell replacement afforded by genetically engineered hematopoietic stem cells (HSCs) may durably alleviate a range of conditions, obviating, for example, the need for lifelong enzyme administration or transfusion therapy (*3*, *4*). Originally envisioned as a treatment solely for inherited disorders, gene therapy is now being applied to acquired conditions, a concept best illustrated by genetic engineering of T cells for cancer immunotherapy. Recent clinical studies have found that single infusions of T cells engineered with synthetic genes encoding a chimeric antigen receptor can produce durable responses in a subset of patients (*5*).

Translation of gene therapy concepts to patient care began in the early 1990s but was plagued by repeated cycles of optimism followed by disappointing clinical trial results. A number of these early experimental therapies were found to provide no clinical benefit or produce unexpected toxicities that in some cases led to widely publicized patient deaths (*6*). In 1996, a National Institutes of Health (NIH) advisory panel concluded that these disappointing clinical results were due to insufficient knowledge of the biology of the viral vectors, the target cells and tissues, and the diseases. The panel recommended that investigators return to the laboratory and focus on the basic science underlying gene therapy approaches (*7*). Development of new vectors and a better understanding of target cells sparked a second generation of clinical trials in the late 1990s and early 2000s. These trials produced evidence of sustained genetic modification of target tissues and, in some instances, evidence for clinical benefit. However, progress was slowed by the emergence of serious toxicities related to high gene transfer efficiency; for instance: insertional genotoxicity, immune destruction of genetically modified cells, and immune reactions related to administration of certain vectors (*6*, *8*, *9*).

Over the past 10 years, further maturation of the "science" of gene therapy, safety modifications, and improvements in gene transfer efficiency and delivery have finally resulted in substantial clinical progress. Several gene and gene-modified cell-based therapies are already approved drugs, and over a dozen others have earned "breakthrough therapy" designation by regulators in the United States and around the world. In this Review, we highlight key developments in the gene therapy field that form the foundation for these recent successes and examine recent advances in targeted genome editing likely to transform gene therapies in the future.

Genetic engineering from viral vectors to genome editing

Recombinant, replication-defective viral vectors were the first molecular tool enabling efficient, nontoxic gene transfer into human somatic cells (*10*). Retroviruses and adeno-associated virus (AAV) have shown the most clinical promise, and we will limit our discussions to these vectors.

Retroviral vectors

The identification of a genome packaging signal (*11*) and the creation of a producer cell line (*12*) paved the way for design and facile production of vectors capable of undergoing reverse transcription and DNA integration but lacking replication potential (*13*, *14*). The γ-retroviral vectors developed in the 1980s and early 1990s were the first to be shown to deliver genes into repopulating HSCs (*15*–*17*). C-type retroviruses were also adapted for efficient gene transfer into primary T lymphocytes (*18*–*21*). These vectors were used in first-generation clinical trials designed to deliver a normal copy of a specific defective gene into the genome of T cells or HSCs from patients with immunodeficiencies or cancer [reviewed in (*22*)] (Fig. 1).

Two other genera of the retroviruses were subsequently added to this armamentarium: the lentiviruses (*23*) and spumaviruses (*24*). In contrast to γ-retroviral vectors, lentiviral vectors enabled gene transfer into nondividing cells but still left quiescent G_0 cells out of reach (*25*). Lentiviral vectors can carry larger and more complex gene cassettes than γ-retroviral vectors and thus their development provided a critical advance for hemoglobinopathies (*26*). Lentiviral and spumavirus vectors have another advantage over γ-retroviral vectors in that they preferentially integrate into the coding regions of genes. The γ-retroviral vectors, by contrast, can integrate into the 5′-untranslated region of genes (*27*), a feature that increases the risk of potentially oncogenic insertional mutagenesis in hematopoietic cells (*28*). Lentiviral vectors are currently the tools of choice for most HSC applications, but γ-retroviral vectors are still used for certain applications in T cell engineering and HSC gene therapy (Table 1). Removal of endogenous strong enhancer elements from lentiviral and γ-retroviral vectors using a "self-inactivating" SIN design (*29*) is another approach that decreases the risk of genotoxicity (*30*); this design is used in most current clinical trials (Table 1). Integrating retroviral vectors are reviewed in more detail in (*31*, *32*).

Adeno-associated viral (AAV) vectors

AAV vectors are engineered from a nonpathogenic, nonenveloped parvovirus that is naturally replication-defective. Wild-type AAV requires another virus such as an adenovirus or a herpesvirus to replicate (*33*, *34*). All viral coding sequences in AAVs are replaced with a gene expression cassette of interest. One limitation of AAV vectors is that they cannot package more than ~5.0 kb of DNA (in contrast to γ-retroviral or lentiviral vectors, which can accommodate up to 8 kb). AAV vectors are predominantly nonintegrating; the transferred DNA is stabilized as an episome. This feature lessens risks related to integration but also limits long-term expression from AAV vectors to long-lived postmitotic cells.

In the mid-1990s, two groups demonstrated long-term expression of a transgene following

[1]Hematology Branch, National Heart, Lung and Blood Institute, Bethesda, MD USA. [2]Spark Therapeutics, Philadelphia, PA, USA. [3]Massachussetts General Hospital and Harvard Medical School, Boston, MA, USA. [4]David Geffen School of Medicine, University of California, Los Angeles, CA, USA. [5]The Institute of Medical Science, The University of Tokyo, Tokyo, Japan. [6]Memorial Sloan Kettering Cancer Center, New York, NY, USA.
*Corresponding author. Email: dunbarc@nhlbi.nih.gov (C.E.D); m-sadelain@ski.mskcc.org (M.S.)

in vivo muscle administration of AAV vectors to mice (35, 36). This seminal work led to the demonstration that AAV vectors could also efficiently transduce a variety of target tissues in animal models, including liver, retina, cardiac muscle, and central nervous system, with specific tissue tropisms discovered for several naturally occurring AAV serotypes and AAV engineered with optimized capsids (37). Improved manufacturing techniques [reviewed in (38)] increased both yield and purity of AAV vector product, allowing proof-of-concept studies in large-animal models of disease (Fig. 2). Pioneering AAV gene therapy clinical trials for hemophilia B were initiated in the late 1990s, first testing delivery of AAV vectors to muscle via injection (39) and then moving to intravenous administration, taking advantage of AAV2 liver tropism (40). These early trials established safety but were limited by insufficient dosing, and anti-AAV immune responses, most likely because many people carry neutralizing antibodies and memory T cells directed against the AAV capsid. The full exploitation of the therapeutic potential of AAV vectors, as described below, required rigorous analysis of anti-AAV immune responses (41), including both cellular and humoral responses to a range of serotypes (42).

Genome editing

In contrast to viral vectors, which can mediate only one type of gene modification ("gene addition"), new genome editing technologies can mediate gene addition, gene ablation, "gene correction," and other highly targeted genome modifications in cells. Genome editing can be performed on cells ex vivo or the editing machinery can be delivered in vivo to effect in situ genome editing. A targeted DNA alteration is initiated by creation of a nuclease-induced double-stranded break (DSB), which stimulates highly efficient recombination in mammalian cells (43). Nonhomologous end-joining (NHEJ)–mediated repair results in the efficient creation of variable-length insertion or deletion mutations (indels) at the site of the DSB, which generally inactivates gene function. Homology-directed repair (HDR) can be used to create specific sequence alterations in the presence of a homologous donor DNA template, which following recombination results in correction of a mutation or insertion of new sequences in a site-specific manner (44).

Early genome editing studies relied on engineering of specific zinc finger nucleases (ZFNs) (45) or meganucleases (46) for each individual DNA target site to induce the required DSBs. These nuclease platforms required specialized expertise to customize the DNA binding nuclease effector proteins for each cleavage target, which limited their broader use and application. The demonstration in 2009 that the DNA binding domain of bacterial proteins called transcription activator–like effectors (TALEs) can be readily altered (47, 48) opened the door to the creation of TALE nucleases (TALENs) (49, 50). These enzymes can efficiently cleave essentially any DNA sequence of interest (51). However, TALEN approaches still require design of a specific pair of nucleases for each new DNA target.

The genome editing landscape changed in 2012 with a seminal discovery by Doudna and Charpentier, who showed that a bacterial defense system composed of clustered regularly interspaced short palindromic repeat (CRISPR)–CRISPR-associated 9 (Cas9) nucleases can be efficiently programmed to cleave DNA at sites of interest, simply by designing a specific short guide RNA (gRNA) complementary to the target site of interest (52). The CRISPR-Cas9 nuclease technology was rapidly extended to mammalian cells (53, 54), thereby simplifying the process of genome editing (55). TALENs and CRISPR-Cas9 nucleases, which can be easily reprogrammed to cleave specific target DNA sequences, are now widely used for a myriad of applications in basic research (56–58). A number of clever strategies that could eventually be applied clinically involve the use of RNA-guided catalytically inactive Cas9 ("dead Cas9" or dCas9) to turn genes on and off by blocking transcriptional machinery or recruiting epigenetic regulators (59, 60). Correction of mutations at a single-base level via Cas9-based targeting of "base editors" has recently been reported (61, 62).

Genome editing approaches offer a precise scalpel for correcting or altering the genome and can overcome many of the drawbacks of strategies that rely on viral vector–mediated semi-random genomic insertion. For instance, genotoxicity due to ectopic activation of nearby proto-oncogenes, knockout of tumor suppressor genes, or perturbation of normal splicing should not occur with on-target editing. In addition, the regulation of an introduced or corrected gene will be controlled by the endogenous promoter, resulting in more physiologic and appropriately regulated gene expression (63). Targeted introduction of clotting factor genes downstream of the highly active albumin promoter in hepatocytes has shown promise in animal models (64). The potential of genome editing strategies to bypass pathology in muscular dystrophy by altering splicing of the mutated dystrophin gene or by directly correcting the dystrophin mutation has been demonstrated in preclinical models (65–67). Finally, disease due to dominant negative mutations, which cannot be treated by gene addition therapy, should be amenable to gene correction strategies.

There are challenges in delivering all the components required for editing into target cells. Genome mutation by NHEJ is simplest, requiring just targeted nucleases for meganuclease, ZFN, or TALEN techniques, or a nuclease plus gRNA

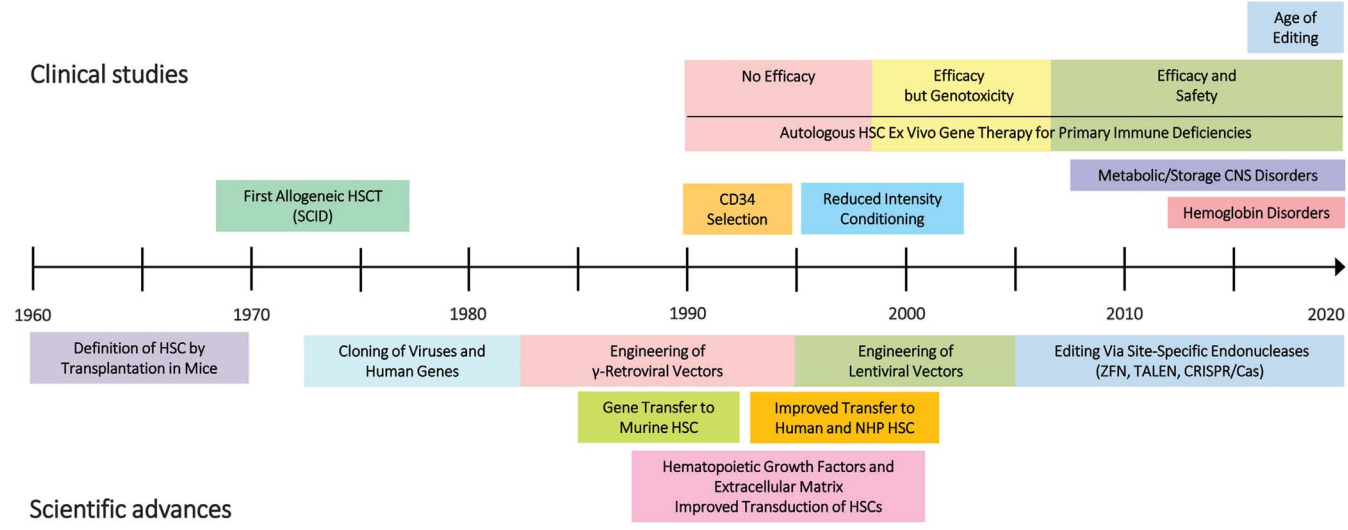


Fig. 1. Historical overview of HSC gene therapy. HSCT: hematopoietic stem cell transplantation; HSC: Hematopoietic stem cell; SCID: severe combined immunodeficiency; NHP: nonhuman primate; ZFN: zinc finger nuclease; TALEN: transcription activator–like effector nuclease; CRISPR/Cas9: clustered regularly interspaced short palindromic repeat (CRISPR)–CRISPR-associated 9 (Cas9) nucleases.

Table 1. Clinical and product development landmarks for ex vivo gene therapies.

Cell type	Disease	Vector/transgene	Key publication(s) or clinicaltrials.gov no.	Primary institution and/or company	Breakthrough designation or product approval
T cells	Adult ALL*	γRV CD19 (CD28) CAR-T	(*134*, *143*, *144*)	Memorial Sloan Kettering Cancer Center	FDA 2014
	Pediatric ALL	LV CD19 (4-1BB) CAR-T	(*145*)	University of Pennsylvania/Novartis	FDA Oncology Advisory Committee recommended approval 2017; EMA 2016
		γRV CD19 (CD28) CAR-T	(*146*)	National Cancer Institute/Kite	
		LV CD19 CAR-T, TALEN knockout of TCR and CD52	(*74*) NCT02808442	Cellectis/Servier/Pfizer	
	Diffuse large B cell lymphoma	γRV CD19 (CD28) CAR-T	(*147*) NCT00924326	National Cancer Institute/Kite	FDA 2014
		γRV CD19 (CD28) CAR	NCT02348216	Multiple academic sites/Kite	FDA 2015; EMA 2016
		LV CD19 (4-1BB) CAR-T	(*148*) NCT02631044	Multiple academic sites/Juno	FDA 2016; EMA 2016
		LV CD19 (4-1BB) CAR-T	NCT02445248	Multiple academic sites/Novartis	FDA 2017
	CLL/indolent lymphoma	LV CD19 (4-1BB) CAR-T	(*149*)	University of Pennsylvania/Novartis	
		γRV CD19 (CD28) CAR-T	(*150*)	National Cancer Institute	
	Multiple myeloma	γRV BCMA (CD28) CAR-T	(*136*) NCT02215967	National Cancer Institute/Kite	
		γRV BCMA (4-1BB) CAR T	NCT03070327	Memorial Sloan Kettering Cancer Center/Juno	
		LV-BCMA CAR-T	NCT03090659	Nanjing Legend Biotech	
	Synovial sarcoma	γRV -NY-ESO-TCR	(*151*)	National Cancer Institute	
		LV-NY-ESO-TCR	NCT03090659	Multiple academic sites/Adaptimmune	FDA 2016; EMA 2016
	Human immunodeficiency virus	ZFN CCR5 electroporation	(*73*)	University of Pennsylvania/Sangamo	
HSPCs	β-Thalassemia	LV anti-sickling β-hemoglobin	(*120*) NCT01745120 NCT02151526 NCT03207009	Hopitaux de Paris/academic centers worldwide/Bluebird Bio	FDA 2015; EMA 2016
		LV β-hemoglobin	NCT02453477	San Raffaele Telethon Institute of Gene Therapy/GlaxoSmithKline	
		LV β-hemoglobin	NCT01639690	Memorial Sloan Kettering Cancer Center	
	Sickle cell anemia	LV anti-sickling β-hemoglobin	(*121*) NCT02151526, NCT02140554	Hopitaux de Paris/US academic sites/ Bluebird Bio	
		LV anti-sickling β-hemoglobin	NCT02247843	UCLA/California Institute of Regenerative Medicine	
	Wiskott-Aldrich syndrome	LV WAS	(*114*)	San Raffaele Telethon Institute of Gene Therapy/GlaxoSmithKline	
		LV WAS	(*152*)	Hopital Necker-Enfants/ University College/Genethon	
	Adenosine deaminase deficiency	γRV ADA	(*116*)	San Raffaele Telethon Institute of Gene Therapy/GlaxoSmithKline	EMA 2016 approved "Strimvelis"
		LV ADA	NCT02999984	University College/UCLA/ Orchard Therapeutics	FDA 2015
	IL2Rγ-deficient X-SCID	γRV SIN IL2Rγ	(*115*)	Hopital Necker-Enfants/Great Ormond Street	
		LV IL2Rγ	(*153*)	National Institute of Allergy and Infectious Diseases	
	Adrenoleukodystrophy	LV ABCD1	(*118*)	St. Vincent de Paul, Paris	
		LV ABCD1	(*119*)	Multiple academic sites/Bluebird Bio	
	Metachromatic leukodystrophy	LV ARSA	(*117*, *154*)	San Raffaele Telethon Institute of Gene Therapy/GlaxoSmithKline	EU Orphan Drug 2007
	Human Immunodeficiency virus	ZFN CCR5 electroporation	NCT02500849	City of Hope/Sangamo	

*Abbreviations: FDA, U.S. Food and Drug Administration; EMA, European Medicines Agency; γRV, murine γ-retrovirus; LV, lentivirus; ALL, acute lymphoblastic leukemia; CLL, chronic lymphocytic leukemia; HSPC, hematopoietic stem and progenitor cells; X-SCID, X-linked severe combined immunodeficiency; ZFN, zinc finger nuclease; BCMA, B cell maturation antigen; ARSA, arylsulfatase A; ABCD1, transporter gene mutated in adrenoleukodystrophy.

for CRISPR-based approaches; these components can be delivered by nonintegrating viral vectors or transfected as mRNA or RNA-protein complexes into target cells such as HSCs ex vivo. However, gene correction by HDR requires donor DNA, which is more difficult to deliver, and HDR appears to be particularly inefficient in certain quiescent cell types such as long-term repopulating HSCs (*68*, *69*), although progress is being made (*70*).

Genome editing as a therapeutic modality is rapidly advancing into the clinic (Table 1). Engineered ZFNs have been used to disrupt CCR5 (C-C motif chemokine receptor type 5) expression in human T cells (*71*) and HSCs (*72*) to render these cells resistant to HIV infection. A phase I/II (*73*) study of T cell CCR5 editing has been completed, and a phase I trial of HSC editing is ongoing (NCT02500849). TALENs have been used to make "off-the-shelf" third-party anti-CD19 chimeric antigen receptor (CAR) T cells less likely to cause graft-versus-host disease (GVHD). This was done by T cell receptor gene deletion. These modified cells were administered to two patients with refractory B cell acute leukemia on a compassionate basis, with evidence for tumor response (*74*), and are in phase I clinical trials (NCT02808442). In addition, early trials have begun for allogeneic TALEN-edited CAR T cells targeting CD123 in acute myeloid leukemias and blastic plasmacytoid dendritic cell neoplasms (NCT03190278). The U.S. Food and Drug Administration (FDA) has approved the launch of three clinical trials for ZFN-mediated in vivo insertion of therapeutic genes into the albumin locus of hepatocytes, delivering the factor IX gene for hemophilia B (NCT02695160), the α-L-iduronidase gene for mucopolysaccharidosis I (NCT02702115), and the iduronidate-2-sulfatase gene for mucopolysaccharidosis II (MPS II) (NCT03041324). The first patient to be treated by in vivo genome editing was recently enrolled in the MPS II trial, with delivery of editing components to the liver via AAV intravenous infusion. At least nine trials using CRISPR-Cas nucleases have been approved by regulatory agencies in China, primarily to knockout PD1 expression in tumor-targeted T cells, and several have reportedly enrolled patients.

It is important to stress that in comparison to standard gene transfer approaches, genome editing—particularly that based on CRISPR-Cas nucleases—is in its translational and clinical infancy. A number of potential feasibility and safety hurdles exist that may affect clinical applications; these will require further preclinical studies in appropriate models and carefully designed clinical trials. For example, the extent of "off-target" mutations, due to nuclease-mediated NHEJ or even HDR at alternative sites, is under intense investigation. Related questions under study are how best to design nucleases or CRISPR gRNAs to avoid off-target cutting and how to predict, screen for, and detect on- versus off-target genome alterations before or during clinical applications (*75*). Notably, high-fidelity CRISPR-Cas9 nuclease variants with no or very few detectable off-target effects have recently been developed (*76*–*78*). Questions remain regarding immunogenicity of nucleases for in vivo genome editing (*79*) and ensuring targeted delivery of editing machinery to the desired target tissue.

The rapid technological advances in genome editing have made heritable germline editing [defined as manipulation of germ cells, gametes, zygotes, or embryos with the intent to generate a new human being with the ability to pass on the edited gene(s) to future generations] a realistic possibility. In 2015, scientists in China published results from experiments using CRISPR-Cas9 to attempt to modify the hemoglobin gene in "nonviable" preimplantation human embryos, demonstrating low efficiency and reportedly frequent off-target mutations (*80*). This publication prompted statements of concern from professional societies around the world (*81*) and a series of meetings sponsored by the U.S. National Academies of Sciences, Engineering, and Medicine that brought together an international group of scientists, clinicians, ethicists, patient advocates, and government officials. This group published a report in 2017 laying out principles of governance and oversight for human genome editing, and presenting a possible pathway for eventual use of genome editing technologies to correct germline mutations for certain serious diseases (*82*). In the United States, federal government funds presently cannot be used for research on germline editing and clinical trials cannot be considered for approval by the FDA. Similar restrictions exist in many other countries. Clearly most countries are far from a societal consensus on germline editing. The acquisition of more efficacy and safety data from studies of genome editing in somatic cells is critical before implementation of human germline editing can be considered.

Gene delivery in vivo

Targeting organs in vivo is very attractive because it avoids the practical and regulatory hurdles of ex vivo cell-based gene therapies, which require cell collection, culture, and manipulation and transplantion. However, in vivo approaches depend on tissue-specific targeting or local delivery and/or target cell–specific gene expression. Inadvertent germline modification is of concern, and immune responses to vector components can occur. Some of these challenges have been overcome, with encouraging clinical outcomes in trials delivering genes to the liver or the retinal pigment epithelium of the eye, paving the way for further advances targeting other tissues, including the brain and muscle.

The liver

Studies beginning in 1997 showed that AAV vectors introduced into skeletal muscle or liver ameliorated disease in animal models of hemophilia B (*83*, *84*) (Fig. 2). In the initial clinical trial of hemophilia B (*40*), factor IX was found to be

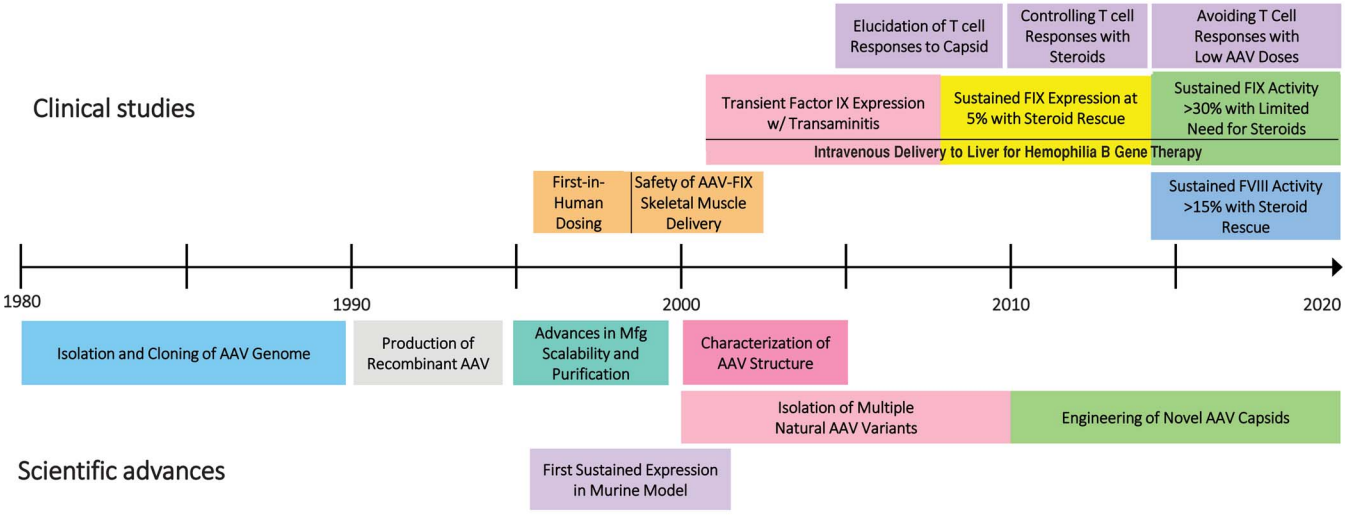


Fig. 2. Historical overview of AAV gene therapy for hemophilia. AAV: adeno-associated viral vector; FVIII: factor VIII; FIX: factor IX; Mfg: Manufacturing.

produced at therapeutic levels, but the expression persisted for only several weeks, in contrast to the stable expression observed in preclinical studies of hemophilic dogs (*85*). Subsequent studies revealed that expression of the factor IX transgene in humans was of short duration because of an immune response to the AAV capsid (*42*). In a later clinical trial that incorporated short-term immunosuppression (*86*, *87*) (Table 2), transgene expression persisted for years, resulting in circulating factor IX levels that were 2 to 7% of normal; this was sufficient to reduce bleeding and lessen the need for recombinant factor IX infusions. A subsequent trial involved the transfer of a transgene encoding factor IX Padua, a naturally variant of factor IX with high specific activity. The transgene was carried by a vector containing an optimized AAV capsid and a liver-specific expression cassette. Patients showed a mean sustained factor IX activity level of more than 30%, resulting in complete cessation of factor IX infusions in 8 of 10 treated patients (*88*). These levels of factor IX activity are well above the threshold of 12% associated with a greatly reduced risk of bleeding in natural history studies. Use of this high–specific activity factor IX allowed delivery of 4 to 120 times lower doses of AAV particles to achieve therapeutic levels of factor IX, likely accounting for the low rate and severity of antivector immune responses in this trial. Results from the first positive early-phase clinical trial for hemophilia A were recently reported, demonstrating increases in factor VIII activity into or even above the normal range in six of seven patients, accompanied by decreased bleeding (*89*).

Additional trials of AAV-mediated gene therapy for both hemophilia B and the more common hemophilia A are ongoing (Table 2). Problems that still need to be addressed include the delayed $CD8^+$ T cell response to the capsid, which has been well controlled with a short course of steroids with some AAV vectors but not with others, and the prevalence—particularly in the adult population—of preexisting neutralizing antibodies to AAV (*90*, *91*). At present, most clinical trials circumvent the antibody problem by excluding subjects who carry them, but other strategies will be required going forward.

The eye

Phase 1/2 clinical trials conducted by multiple groups have demonstrated improvement in visual function following subretinal injection of AAV2 vectors expressing retinal pigment epithelium-specific 65 kDa protein (RPE65) in patients with inherited blindness caused by mutations in the *RPE65* gene (*92*–*94*) (Table 2). A cohort of phase III–eligible subjects from one of the phase I/II trials continues to demonstrate clinical benefits lasting a minimum of 3 years after injection, with observation ongoing (*95*); however, patients in the other two original trials have experienced regression of visual function over similar follow-up periods (*96*, *97*). At present, there is no clear explanation for the differences in outcome because all used AAV2-based vectors. Subtle differences in manufacturing process, final formulation, design of expression cassette, or adjuvant immunomodulatory regimens could potentially affect long-term efficacy (*98*). In the only randomized controlled phase 3 gene therapy trial completed to date, visually impaired patients who carried RPE65 mutations were randomized to undergo sequential bilateral injection of AAV2-RPE65 or to undergo the same series of evaluations without the intervention (*99*). One year after randomization, patient mobility, a measure of functional vision, as well as certain tests of visual function were significantly improved in the treatment group. Based on this pivotal study, an FDA advisory panel recently unanimously recommended drug approval. Direct injection of AAV is now being pursued in clinical trials for other inherited forms of blindness, including achromatopsia, choroideremia, Leber's hereditary optic neuropathy, X-linked retinoschisis, and X-linked retinitis pigmentosa.

Neuromuscular targets

The common, clinically devastating degenerative neurologic disorders are a focus of gene therapy efforts; however, these multigenic, pathophysiologically complex and incompletely understood disorders are much more challenging targets than Mendelian inherited disorders. Parkinson's disease (PD), which is characterized by loss of dopaminergic neurons in the substantia nigra and a decrease in dopamine in the striatum, has been an intensely pursued target. Gene therapy–mediated transfer of dopamine-synthesizing enzymes into striatal neurons has been found to normalize movement in a nonhuman primate PD model (*100*). Early-phase clinical trials have established the safety of AAV vector–mediated gene delivery of aromatic L-amino acid decarboxylase (AADC), an enzyme that converts L-dopa to dopamine; glutamic acid decarboxylase (GAD), an enzyme that modulates production of the neurotransmitter GABA (γ-aminobutyric acid); and neurturin, a neurotrophic factor (*101*–*104*). Promising results were obtained with AADC gene therapy, with additional early-phase clinical trials ongoing (Table 2). An early-phase trial of AAV2 vector injection into the brain has also been conducted in patients with AADC deficiency, a monogenic movement disorder characterized by compromised dopamine and serotonin synthesis, and some improvement was noted (*105*).

The treatment of childhood-onset spinal muscular atrophy (SMA), a rapidly fatal neuromuscular disorder due to loss-of-function mutations in the survival motor neuron 1 (*SMN*) gene, has been revolutionized by antisense oligonucleotide drug nusinersen, a triumph of nucleoside-based gene therapies, as reviewed in (*106*). Intrathecal delivery of nusinersen modulates alternative splicing of the intact *SMN2* gene in spinal motor neurons, resulting in higher expression of a functional form of the gene product able to compensate for SMN1 loss. An alternative type of gene therapy for SMA is also showing great promise. A serotype of AAV that efficiently crosses the blood-brain barrier was engineered to carry the *SMN1* gene, and given as a single intravenous infusion to 15 infants and young children. Compared with historical control subjects, survival of the trial participants was extended, with all alive to date, and motor function improved to the extent that some children could sit up and even walk (*107*)

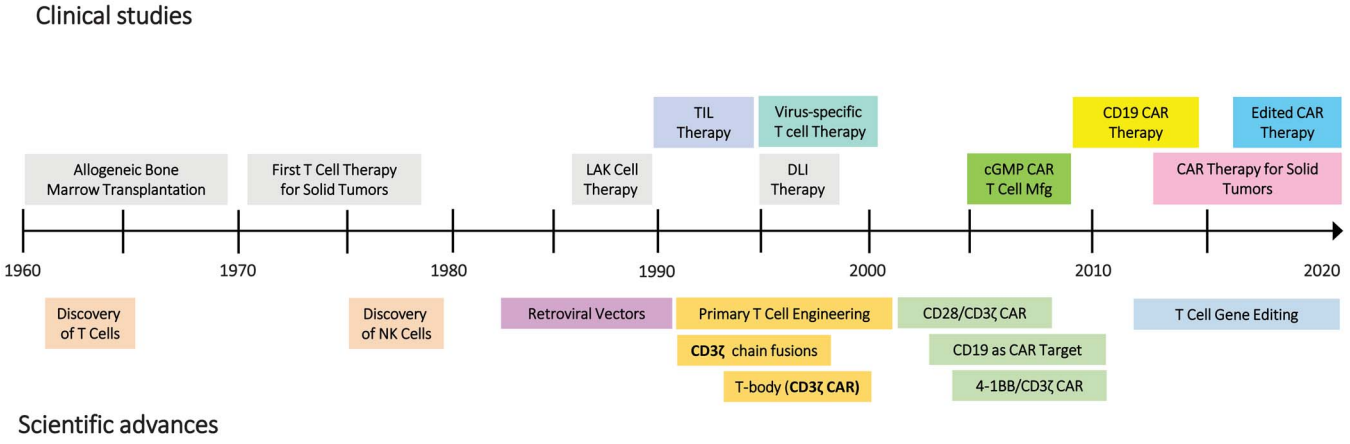


Fig. 3. Historical overview of CAR-T cell therapy. CAR: chimeric antigen receptor; cGMP: current good manufacturing practices; DLI: donor leukocyte infusion; LAK: lymphokine-activated killer; Mfg: Manufacturing; NK: natural killer; TIL: tumor-infiltrating lymphocytes.

(Table 2). This approach may be preferable to repeated intrathecal injections of an oligonucleotide drug.

Ex vivo gene delivery via cell engineering: Monogenic blood disorders and cancer immunotherapies

Hematopoietic stem cells

The clinical applications of gene therapies targeting HSCs derive from the success of allogeneic bone marrow transplantation for many genetic immunodeficiencies and blood cell diseases. These ex vivo approaches entail the transplantation of autologous stem cells in which an underlying genetic defect is alleviated or corrected [e.g., adenosine deaminase deficit in severe combined immunodeficiency (SCID)], β-hemoglobin deficit or structural alteration in hemoglobinopathies) (Table 1). Autologous transplants have an advantage over allogeneic transplants in that they do not require a histocompatible donor, they avoid the immune complications of GVHD, and they eliminate the need for administration of immune suppressants.

Clinical trials based at academic medical centers have piloted these approaches in successive technological eras (Fig. 1). Trials using γ-retroviral vectors in the late 1990s and early 2000s demonstrated unequivocal improvement in immune function in patients with SCID caused by loss-of-function mutations in the genes encoding interleukin-2 receptor γ or adenosine deaminase (Table 1). Despite relatively low HSC transduction efficiencies with γ-retroviral vectors, gene-modified T-lineage cells were able to expand and fill the empty T cell compartment, improving immune function despite minimal levels of gene-corrected cells in other lineages (*108*, *109*). However, several years after treatment, patients in the X-SCID trials, as well as those for chronic granulomatous disease and Wiskott-Aldrich syndrome, developed acute myeloid and lymphoid leukemias due to activation of proto-oncogenes adjacent to proviral insertions, linked to strong enhancers present in γ-retroviral vectors and the propensity of these vectors to insert near promoters (*110*–*113*). These serious toxicities led to accelerated adoption of enhancer-deleted lentiviral or γ-retroviral vectors for HSC clinical gene therapies of immunodeficiency disorders. Encouraging clinical results with these newer vectors have been reported in Wiskott-Aldrich syndrome (*114*) and in X-SCID (*115*), demonstrating disease amelioration without leukemia or uncontrolled expansion of individual transduced clones.

The high HSC transduction efficiency achieved with lentiviral vectors allowed a broader application of this gene therapy approach to diseases where corrected cells do not have a survival advantage. In addition, optimization of methods for vector production, ex vivo HSC manipulation, and pretransplant cytoreductive conditioning (*116*) all contributed to clinical benefit in several more recent trials. The metabolic disorder adrenoleukodystrophy and the lysosomal storage disorder metachromatic leukodystrophy result in profound neurologic degeneration and death in childhood. Lentiviral gene therapy clinical trials in both disorders have been encouraging, with high-level production of the missing enzymes from hematopoietic cells, including in the central nervous system, and a slowing of neurodegeneration (*117*–*119*). The clinical trial in adrenoleukodystrophy was the first reported using lentiviral vectors for HSC gene therapy.

The hemoglobin disorders β-thalassemia and sickle cell disease, which affect millions of patients worldwide, have historically been an intense focus of gene therapy research, but require high efficiency and substantial hemoglobin expression to correct the underlying pathophysiology (*3*). Lentiviral vectors harboring multiple regulatory elements to direct high-level, erythroid-specific hemoglobin expression have been developed (*26*) and in case reports have shown promise in patients with β-thalassemia or sickle cell disease (*120*, *121*), with larger multicenter clinical trials ongoing (Table 1). Also likely to move forward in the near future are clinical trials of genome editing approaches to treat sickle cell

Table 2. Clinical and product development landmarks for in vivo gene therapies.

Cell type	Disease	Vector/transgene	Key publication(s) or clinicaltrials.gov no.	Institutional and/or industry partners	FDA breakthrough/EMA* PRIME designation or product approval
CNS	Parkinson's disease	AAV2-AADC	(*101*, *102*)	Jichi Medical University/ UCSF/Voyager	
	Aromatic l-amino acid decarboxylase deficiency	AAV2-AADC	(*105*)	Jichi Medical University/ National Taiwan University	
	Spinal muscular atrophy	AAV9-SMN	(*107*)	Nationwide Children's Hospital/AveXis	FDA 2016; EMA 2017
Liver	Hemophilia B	AAV8-Factor IX	(*86*, *87*)	Royal Free Hospital/St. Jude	FDA 2014; EMA 2017
		AAV100-FIX Padua	(*88*)	Spark Therapeutics	FDA 2016; EMA 2017
		AAV5-Factor IX	NCT02396342	uniQure	FDA 2017; EMA 2017
		AAV2/6-Factor IX and ZFNs	NCT02695160	Sangamo Therapeutics	FDA 2017
	Hemophilia A	AAV5-Factor VIII	NCT02576795	Multiple academic sites/ Biomarin	EMA 2017
		AAV200-Factor VIII	NCT03003533	Spark Therapeutics	
		AAV2/6-B domain-deleted Factor VIII and ZFNs	NCT03061201	Sangamo Therapeutics	
	Mucopolysaccharidosis type II (Hunter's syndrome)	AAV2/6-IDA and ZFNs	NCT03041324	Sangamo Therapeutics	
Muscle	Lipoprotein lipase deficiency	AAV1-LPL	(*155*)	uniQure	EMA 2012 approval of "Glybera", company will not renew license as of 2017
Retina	Inherited retinal dystrophy due to utosomal recessive mutations in *RPE65*	AAV2-RPE65	(*93*, *95*, *99*)	Children's Hospital of Philadelphia/Spark	FDA approval 2017
		AAV2-RPE65	(*92*, *97*)	University College London/ MeiraGTx	
		AAV2-RPE65	(*94*, *96*)	University of Florida	

*Abbreviations: CNS, central nervous system; FDA, U.S. Food and Drug Administration; EMA, European Medicines Agency; AAV, adeno-associated virus; AADC, amino acid decarboxylase; ZFNs, zinc finger nucleases; IDA, iduronate-2-sulfatase.

anemia via reactivation of endogenous fetal hemoglobin (HbF) expression. NHEJ-mediated disruption of the erythroid-specific enhancer element responsible for expression of the *BCL11A* gene results in high-level HbF in animal models and in human sickle cell erythroid cells in vitro (*122*). Similarly, disruption of a genomic locus to mimic a genetic variant associated with hereditary persistence of the fetal hemoglobin locus also shows promise as a target for HSC genome editing to treat sickle cell anemia (*123*).

CAR therapy

Engineered T cells are emerging as powerful medicines for cancer (Fig. 3) (*5*). Chimeric antigen receptors (CARs) are synthetic engineered receptors for antigen, which, in a single molecule, reprogram the specificity, function, and metabolism of T lymphocytes (*124*, *125*). They consist of an antigen-binding domain, either from an immunoglobin molecule or a T cell receptor, fused to an intracellular signaling domain that mediates activation and costimulation to enhance T cell function and persistence. Unlike the physiological receptor for antigen, CARs can be engineered to recognize proteins and carbohydrate glycolipids, as well as HLA-peptide complexes (*126*, *127*). CARs are transduced into T cells ex vivo, creating expandable antigen-specific T cells that bypass the barriers and incremental kinetics of active immunization used to prime endogenous T cells. The generation of CAR-T cells requires stable gene transfer to enable sustained CAR expression in dividing and persisting T cells.

γ-retroviral vectors were originally used to demonstrate that CARs targeting CD19, a cell surface antigen found on most B lineage lymphomas and leukemias, can eradicate systemic cancer in immunodeficient mice (*128*). CD19 is at present the most common CAR target and serves as a paradigm for CAR therapy (Fig. 3). Durable responses have been obtained in patients with refractory diffuse large B cell lymphoma (DLBCL), chronic lymphocytic leukemia, and adult and pediatric acute lymphoblastic leukemia (ALL) (see references in Table 1). Collectively, the preclinical and clinical studies on CD19 CARs, using different vector systems (lentiviral vectors, transposons, mRNA, CRISPR-Cas9) (*129*), CAR designs (*130*, *131*), and T cell subsets (*132*, *133*), have validated the CAR concept (*5*, *126*). Notably, CAR-T cell administration has been associated with serious systemic toxicities that often require intensive care and in some instances have caused patient deaths. Investigators are intensely focused on better understanding, mitigating, and treating these complications, which include off-tumor effects and the cytokine release syndrome (CRS), as well as poorly understood neurotoxicities (*134*, *135*).

The clinical benefit conferred by CD19 CARs in refractory ALL and DLBCL resulted in 2017 FDA approvals for two genetically engineered cell products, the first to be approved in the United States. Several other CARs have obtained FDA breakthrough designation for treatment of B cell malignancies (Table 1). Promising, early clinical data bode well for CAR therapy of multiple myeloma (*136*).

Current research aims to expand CAR therapy to myeloid malignancies and solid tumors (*137*, *138*). These diseases present challenges because reliable tumor-specific cell surface antigens have not yet been validated. In addition, there is a need for methodologies that facilitate CAR-T cell entry into large tumors or immune-privileged sites and that overcome tumor microenvironment signals that disarm T cells. Universal third-party CAR-T cells that can be used "off the shelf" would allow more rapid and cheaper treatment compared to autologous patient-specific T cells. T cells lacking endogenous T cell receptors and/or major histocompatibility complex molecules to decrease the risk of GVHD and rejection are in preclinical or early clinical development as first steps toward this goal (*62*, *73*). CAR-T cells have had a large impact on the treatment of certain cancers (*139*), and this success provides a foundation for future T cell–based therapies for other cancers and other diseases such as autoimmune disorders and AIDS (*5*, *140*).

Conclusions

Gene therapies may well be the most complex "drugs" ever developed. Building on the demonstration of therapeutic efficacy in proof-of-concept clinical studies conducted in the academic setting, gene therapies are now undergoing accelerated clinical and commercial development. They are in transition from an academic-based "cottage industry" to an industrial drug development pathway, relying on partnerships with biotechnology and pharmaceutical companies whose expertise in manufacturing and scale-up will be required for these therapies to have a broader impact on human disease. Investigators in academia and industry are working with regulators and entities such as the National Institute of Standards and Technology (NIST) to develop and standardize assays used to characterize potency and safety of vector preparations and criteria for product release. Similar efforts for genome editing are underway. These intiatives should speed future clinical development, commercialization, and utilization of these multifaceted treatment modalities.

Various models for delivery of ex vivo gene therapies to patients are being explored—for example, centralized versus hospital-based cGMP (current good manufacturing practices) facilities. In addition, it will be critical to engage with health reimbursement entities, including governments and insurers, to develop new models for reimbursement applied to one-time gene therapies with high up-front costs but likely long-term benefits in patients with serious diseases, many of whom have no other options, or poor quality of life and/or lifelong high medical costs on currently available therapies (*141*). Reimbursement must be addressed for the field to advance, as illustrated by recent events in Europe, where two gene therapy products were approved by regulators but have either been withdrawn from the market, in the case of uniQure's Alipogene tiparvovec (an AAV1 vector for treating patients with a rare inherited lipoprotein lipase deficiency), or are at risk of discontinuation of the program by the parent pharmaceutical company, as in the case of Strimvelis, a γ-retroviral vector HSC gene therapy treatment for adenosine deaminase–deficient SCID (*142*). The U.S. Centers for Medicare and Medicaid Services announced a collaboration with the manufacturer of the first approved CAR therapy to provide the product under an "outcomes-based" approach, with payment collected only if patients initially respond to the treatment.

The past year has been marked by a flurry of scientific advances in genome editing, the publication of mature data from multiple clinical trials demonstrating the efficacy and safety of gene therapies for a wide variety of serious human diseases, and regulatory approvals of the first gene therapies in the United States. Scientists and clinicians engaged in basic, translational, and clinical research, supported by government and philanthropies, will continue to innovate and provide new or improved technologies. The increasing involvement of the biotechnology and pharmaceutical sectors in gene therapy efforts demonstrates the maturation of the field and is necessary to accelerate delivery of these treatments to patients. Many challenges remain, including addressing genotoxicity from integrating gene delivery vectors or off-target genome editing, improving gene transfer or editing efficiency to levels necessary for effective treatment of many diseases, addressing immune responses to repeated in vivo administration of vectors, and reaching a societal consensus regarding contentious issues such as the ethics of germline editing and payment for expensive curative therapies. The potential for gene therapy to provide durable benefits to human health, exemplified by the scientific advances and clinical successes over the past several years, justifies continued optimism and increasing efforts toward making this therapy part of our standard armamentarium for treatment of serious human diseases.

REFERENCES AND NOTES

1. T. Friedmann, R. Roblin, Gene therapy for human genetic disease? *Science* **175**, 949–955 (1972). doi: 10.1126/science.175.4025.949; pmid: 5061866
2. K. A. High, X. M. Anguela, Adeno-associated viral vectors for the treatment of hemophilia. *Hum. Mol. Genet.* **25** (R1), R36–R41 (2016). doi: 10.1093/hmg/ddv475; pmid: 26614390
3. J. Mansilla-Soto, I. Riviere, F. Boulad, M. Sadelain, Cell and gene therapy for the beta-thalassemias: Advances and prospects. *Hum. Gene Ther.* **27**, 295–304 (2016). doi: 10.1089/hum.2016.037; pmid: 27021486
4. L. Naldini, Gene therapy returns to centre stage. *Nature* **526**, 351–360 (2015). doi: 10.1038/nature15818; pmid: 26469046
5. M. Sadelain, I. Rivière, S. Riddell, Therapeutic T cell engineering. *Nature* **545**, 423–431 (2017). doi: 10.1038/nature22395; pmid: 28541315
6. S. Jenks, Gene therapy death—"everyone has to share in the guilt". *J. Natl. Cancer Inst.* **92**, 98–100 (2000). doi: 10.1093/jnci/92.2.98; pmid: 10639505
7. S. H. Orkin, A. G. Motulsky, "Report and recommendations of the panel to assess the NIH investment in research on gene therapy, Report to the NIH Director," (1995).
8. A. W. Nienhuis, C. E. Dunbar, B. P. Sorrentino, Genotoxicity of retroviral integration in hematopoietic cells. *Mol. Ther.* **13**, 1031–1049 (2006). doi: 10.1016/j.ymthe.2006.03.001; pmid: 16624621
9. F. Mingozzi, K. A. High, Immune responses to AAV in clinical trials. *Curr. Gene Ther.* **11**, 321–330 (2011). doi: 10.2174/156652311796150354; pmid: 21557723

10. M. A. Kotterman, T. W. Chalberg, D. V. Schaffer, Viral vectors for gene therapy: Translational and clinical outlook. *Annu. Rev. Biomed. Eng.* **17**, 63–89 (2015). doi: 10.1146/annurev-bioeng-071813-104938; pmid: 26643018
11. R. Mann, R. C. Mulligan, D. Baltimore, Construction of a retrovirus packaging mutant and its use to produce helper-free defective retrovirus. *Cell* **33**, 153–159 (1983). doi: 10.1016/0092-8674(83)90344-6; pmid: 6678608
12. S. Watanabe, H. M. Temin, Construction of a helper cell line for avian reticuloendotheliosis virus cloning vectors. *Mol. Cell. Biol.* **3**, 2241–2249 (1983). doi: 10.1128/MCB.3.12.2241; pmid: 6318091
13. O. Danos, R. C. Mulligan, Safe and efficient generation of recombinant retroviruses with amphotropic and ecotropic host ranges. *Proc. Natl. Acad. Sci. U.S.A.* **85**, 6460–6464 (1988). doi: 10.1073/pnas.85.17.6460; pmid: 3413107
14. A. D. Miller *et al*., Construction and properties of retrovirus packaging cells based on gibbon ape leukemia virus. *J. Virol.* **65**, 2220–2224 (1991). pmid: 1850008
15. I. Rivière, K. Brose, R. C. Mulligan, Effects of retroviral vector design on expression of human adenosine deaminase in murine bone marrow transplant recipients engrafted with genetically modified cells. *Proc. Natl. Acad. Sci. U.S.A.* **92**, 6733–6737 (1995). doi: 10.1073/pnas.92.15.6733; pmid: 7624312
16. S. Halene *et al*., Improved expression in hematopoietic and lymphoid cells in mice after transplantation of bone marrow transduced with a modified retroviral vector. *Blood* **94**, 3349–3357 (1999). pmid: 10552944
17. D. A. Williams, I. R. Lemischka, D. G. Nathan, R. C. Mulligan, Introduction of new genetic material into pluripotent haematopoietic stem cells of the mouse. *Nature* **310**, 476–480 (1984). doi: 10.1038/310476a0; pmid: 6087158
18. M. Sadelain, R. C. Mulligan, Efficient-retroviral-mediated gene transfer into murine primary lymphocytes. *8th International Congress of Immunology* **88**, 34 (1992).
19. F. Mavilio *et al*., Peripheral blood lymphocytes as target cells of retroviral vector-mediated gene transfer. *Blood* **83**, 1988–1997 (1994). pmid: 8142665
20. B. A. Bunnell, L. M. Muul, R. E. Donahue, R. M. Blaese, R. A. Morgan, High-efficiency retroviral-mediated gene transfer into human and nonhuman primate peripheral blood lymphocytes. *Proc. Natl. Acad. Sci. U.S.A.* **92**, 7739–7743 (1995). doi: 10.1073/pnas.92.17.7739; pmid: 7644487
21. H. F. Gallardo, C. Tan, D. Ory, M. Sadelain, Recombinant retroviruses pseudotyped with the vesicular stomatitis virus G glycoprotein mediate both stable gene transfer and pseudotransduction in human peripheral blood lymphocytes. *Blood* **90**, 952–957 (1997). pmid: 9242523
22. C. E. Dunbar, Gene transfer to hematopoietic stem cells: Implications for gene therapy of human disease. *Annu. Rev. Med.* **47**, 11–20 (1996). doi: 10.1146/annurev.med.47.1.11; pmid: 8712765
23. L. Naldini *et al*., In vivo gene delivery and stable transduction of nondividing cells by a lentiviral vector. *Science* **272**, 263–267 (1996). doi: 10.1126/science.272.5259.263; pmid: 8602510
24. D. W. Russell, A. D. Miller, Foamy virus vectors. *J. Virol.* **70**, 217–222 (1996). pmid: 8523528
25. Y. D. Korin, J. A. Zack, Progression to the G1b phase of the cell cycle is required for completion of human immunodeficiency virus type 1 reverse transcription in T cells. *J. Virol.* **72**, 3161–3168 (1998). pmid: 9525642
26. C. May *et al*., Therapeutic haemoglobin synthesis in beta-thalassaemic mice expressing lentivirus-encoded human beta-globin. *Nature* **406**, 82–86 (2000). doi: 10.1038/35017565; pmid: 10894546
27. C. Wu, C. E. Dunbar, Stem cell gene therapy: The risks of insertional mutagenesis and approaches to minimize genotoxicity. *Front. Med.* **5**, 356–371 (2011). doi: 10.1007/s11684-011-0159-1; pmid: 22198747
28. E. Montini *et al*., The genotoxic potential of retroviral vectors is strongly modulated by vector design and integration site selection in a mouse model of HSC gene therapy. *J. Clin. Invest.* **119**, 964–975 (2009). doi: 10.1172/JCI37630; pmid: 19307726
29. S. F. Yu *et al*., Self-inactivating retroviral vectors designed for transfer of whole genes into mammalian cells. *Proc. Natl. Acad. Sci. U.S.A.* **83**, 3194–3198 (1986). doi: 10.1073/pnas.83.10.3194; pmid: 3458176
30. C. Baum, U. Modlich, G. Göhring, B. Schlegelberger, Concise review: Managing genotoxicity in the therapeutic modification of stem cells. *Stem Cells* **29**, 1479–1484 (2011). doi: 10.1002/stem.716; pmid: 21898683
31. L. Naldini, D. Trono, I. M. Verma, Lentiviral vectors, two decades later. *Science* **353**, 1101–1102 (2016). doi: 10.1126/science.aah6192; pmid: 27609877
32. A. D. Miller, Retroviral vectors: From cancer viruses to therapeutic tools. *Hum. Gene Ther.* **25**, 989–994 (2014). doi: 10.1089/hum.2014.2542; pmid: 25458252
33. K. Berns, C. Parrish, in *Fields Virology*, D. Knipe, P. Howley, Eds. (Lippincott Williams & Wilkins, Philadelphia, 2013), chap. 57, pp. 1768–1791.
34. A. Asokan, D. V. Schaffer, R. J. Samulski, The AAV vector toolkit: Poised at the clinical crossroads. *Mol. Ther.* **20**, 699–708 (2012). doi: 10.1038/mt.2011.287; pmid: 22273577
35. P. D. Kessler *et al*., Gene delivery to skeletal muscle results in sustained expression and systemic delivery of a therapeutic protein. *Proc. Natl. Acad. Sci. U.S.A.* **93**, 14082–14087 (1996). doi: 10.1073/pnas.93.24.14082; pmid: 8943064
36. X. Xiao, J. Li, R. J. Samulski, Efficient long-term gene transfer into muscle tissue of immunocompetent mice by adeno-associated virus vector. *J. Virol.* **70**, 8098–8108 (1996). pmid: 8892935
37. L. Zhong *et al*., Next generation of adeno-associated virus 2 vectors: Point mutations in tyrosines lead to high-efficiency transduction at lower doses. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 7827–7832 (2008). doi: 10.1073/pnas.0802866105; pmid: 18511559
38. J. F. Wright, Adeno-associated viral vector manufacturing: Keeping pace with accelerating clinical development. *Hum. Gene Ther.* **22**, 913–914 (2011). doi: 10.1089/hum.2011.2514; pmid: 21859253
39. M. A. Kay *et al*., Evidence for gene transfer and expression of factor IX in haemophilia B patients treated with an AAV vector. *Nat. Genet.* **24**, 257–261 (2000). doi: 10.1038/73464; pmid: 10700178
40. C. S. Manno *et al*., Successful transduction of liver in hemophilia by AAV-Factor IX and limitations imposed by the host immune response. *Nat. Med.* **12**, 342–347 (2006). doi: 10.1038/nm1358; pmid: 16474400
41. G. L. Rogers, A. T. Martino, I. Zolotukhin, H. C. Ertl, R. W. Herzog, Role of the vector genome and underlying factor IX mutation in immune responses to AAV gene therapy for hemophilia B. *J. Transl. Med.* **12**, 25 (2014). doi: 10.1186/1479-5876-12-25; pmid: 24460861
42. F. Mingozzi *et al*., CD8(+) T-cell responses to adeno-associated virus capsid in humans. *Nat. Med.* **13**, 419–422 (2007). doi: 10.1038/nm1549; pmid: 17369837
43. P. Rouet, F. Smih, M. Jasin, Expression of a site-specific endonuclease stimulates homologous recombination in mammalian cells. *Proc. Natl. Acad. Sci. U.S.A.* **91**, 6064–6068 (1994). doi: 10.1073/pnas.91.13.6064; pmid: 8016116
44. D. Carroll, Genome engineering with targetable nucleases. *Annu. Rev. Biochem.* **83**, 409–439 (2014). doi: 10.1146/annurev-biochem-060713-035418; pmid: 24606144
45. F. D. Urnov, E. J. Rebar, M. C. Holmes, H. S. Zhang, P. D. Gregory, Genome editing with engineered zinc finger nucleases. *Nat. Rev. Genet.* **11**, 636–646 (2010). doi: 10.1038/nrg2842; pmid: 20717154
46. G. Silva *et al*., Meganucleases and other tools for targeted genome engineering: Perspectives and challenges for gene therapy. *Curr. Gene Ther.* **11**, 11–27 (2011). doi: 10.2174/156652311794520111; pmid: 21182466
47. J. Boch *et al*., Breaking the code of DNA binding specificity of TAL-type III effectors. *Science* **326**, 1509–1512 (2009). doi: 10.1126/science.1178811; pmid: 19933107
48. M. J. Moscou, A. J. Bogdanove, A simple cipher governs DNA recognition by TAL effectors. *Science* **326**, 1501 (2009). doi: 10.1126/science.1178817; pmid: 19933106
49. M. Christian *et al*., Targeting DNA double-strand breaks with TAL effector nucleases. *Genetics* **186**, 757–761 (2010). doi: 10.1534/genetics.110.120717; pmid: 20660643
50. T. Li *et al*., TAL nucleases (TALNs): Hybrid proteins composed of TAL effectors and FokI DNA-cleavage domain. *Nucleic Acids Res.* **39**, 359–372 (2011). doi: 10.1093/nar/gkq704; pmid: 20699274
51. D. Reyon *et al*., FLASH assembly of TALENs for high-throughput genome editing. *Nat. Biotechnol.* **30**, 460–465 (2012). doi: 10.1038/nbt.2170; pmid: 22484455
52. M. Jinek *et al*., A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **337**, 816–821 (2012). doi: 10.1126/science.1225829; pmid: 22745249
53. P. Mali *et al*., RNA-guided human genome engineering via Cas9. *Science* **339**, 823–826 (2013). doi: 10.1126/science.1232033; pmid: 23287722
54. L. Cong *et al*., Multiplex genome engineering using CRISPR/Cas systems. *Science* **339**, 819–823 (2013). doi: 10.1126/science.1231143; pmid: 23287718
55. J. A. Doudna, E. Charpentier, The new frontier of genome engineering with CRISPR-Cas9. *Science* **346**, 1258096 (2014). doi: 10.1126/science.1258096; pmid: 25430774
56. J. K. Joung, J. D. Sander, TALENs: A widely applicable technology for targeted genome editing. *Nat. Rev. Mol. Cell Biol.* **14**, 49–55 (2013). doi: 10.1038/nrm3486; pmid: 23169466
57. J. D. Sander, J. K. Joung, CRISPR-Cas systems for editing, regulating and targeting genomes. *Nat. Biotechnol.* **32**, 347–355 (2014). doi: 10.1038/nbt.2842; pmid: 24584096
58. A. C. Komor, A. H. Badran, D. R. Liu, CRISPR-Based Technologies for the Manipulation of Eukaryotic Genomes. *Cell* **168**, 20–36 (2017). doi: 10.1016/j.cell.2016.10.044; pmid: 27866654
59. L. A. Gilbert *et al*., CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. *Cell* **154**, 442–451 (2013). doi: 10.1016/j.cell.2013.06.044; pmid: 23849981
60. L. S. Qi *et al*., Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell* **152**, 1173–1183 (2013). doi: 10.1016/j.cell.2013.02.022; pmid: 23452860
61. A. C. Komor, Y. B. Kim, M. S. Packer, J. A. Zuris, D. R. Liu, Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature* **533**, 420–424 (2016). doi: 10.1038/nature17946; pmid: 27096365
62. N. M. Gaudelli *et al*., Programmable base editing of A•T to G•C in genomic DNA without DNA cleavage. *Nature* **551**, 464–471 (2017). doi: 10.1038/nature24644; pmid: 29160308
63. J. Eyquem *et al*., Targeting a CAR to the TRAC locus with CRISPR/Cas9 enhances tumour rejection. *Nature* **543**, 113–117 (2017). doi: 10.1038/nature21405; pmid: 28225754
64. R. Sharma *et al*., In vivo genome editing of the albumin locus as a platform for protein replacement therapy. *Blood* **126**, 1777–1784 (2015). doi: 10.1182/blood-2014-12-615492; pmid: 26297739
65. C. E. Nelson *et al*., In vivo genome editing improves muscle function in a mouse model of Duchenne muscular dystrophy. *Science* **351**, 403–407 (2016). doi: 10.1126/science.aad5143; pmid: 26721684
66. M. Tabebordbar *et al*., In vivo gene editing in dystrophic mouse muscle and muscle stem cells. *Science* **351**, 407–411 (2016). doi: 10.1126/science.aad5177; pmid: 26721686
67. C. Long *et al*., Postnatal genome editing partially restores dystrophin expression in a mouse model of muscular dystrophy. *Science* **351**, 400–403 (2016). doi: 10.1126/science.aad5725; pmid: 26721683
68. M. D. Hoban *et al*., Correction of the sickle cell disease mutation in human hematopoietic stem/progenitor cells. *Blood* **125**, 2597–2604 (2015). doi: 10.1182/blood-2014-12-615948; pmid: 25733580
69. P. Genovese *et al*., Targeted genome editing in human repopulating haematopoietic stem cells. *Nature* **510**, 235–240 (2014). doi: 10.1038/nature13420; pmid: 24870228
70. S. S. De Ravin *et al*., CRISPR-Cas9 gene repair of hematopoietic stem cells from patients with X-linked chronic granulomatous disease. *Sci. Transl. Med.* **9**, eaah3480 (2017). doi: 10.1126/scitranslmed.aah3480; pmid: 28077679
71. E. E. Perez *et al*., Establishment of HIV-1 resistance in CD4+ T cells by genome editing using zinc-finger nucleases. *Nat. Biotechnol.* **26**, 808–816 (2008). doi: 10.1038/nbt1410; pmid: 18587387
72. L. Li *et al*., Genomic editing of the HIV-1 coreceptor CCR5 in adult hematopoietic stem and progenitor cells using zinc finger nucleases. *Mol. Ther.* **21**, 1259–1269 (2013). doi: 10.1038/mt.2013.65; pmid: 23587921
73. P. Tebas *et al*., Gene editing of CCR5 in autologous CD4 T cells of persons infected with HIV. *N. Engl. J. Med.* **370**, 901–910 (2014). doi: 10.1056/NEJMoa1300662; pmid: 24597865
74. W. Qasim *et al*., Molecular remission of infant B-ALL after infusion of universal TALEN gene-edited CAR T cells. *Sci. Transl. Med.* **9**, eaaj2013 (2017). doi: 10.1126/scitranslmed.aaj2013; pmid: 28123068
75. S. Q. Tsai, J. K. Joung, Defining and improving the genome-wide specificities of CRISPR-Cas9 nucleases. *Nat. Rev. Genet.* **17**, 300–312 (2016). doi: 10.1038/nrg.2016.28; pmid: 27087594
76. B. P. Kleinstiver *et al*., High-fidelity CRISPR-Cas9 nucleases with no detectable genome-wide off-target effects. *Nature* **529**, 490–495 (2016). doi: 10.1038/nature16526; pmid: 26735016

77. I. M. Slaymaker *et al*., Rationally engineered Cas9 nucleases with improved specificity. *Science* **351**, 84–88 (2016). doi: 10.1126/science.aad5227; pmid: 26628643
78. J. S. Chen *et al*., Enhanced proofreading governs CRISPR-Cas9 targeting accuracy. *Nature* **550**, 407–410 (2017). doi: 10.1038/nature24268; pmid: 28931002
79. W. L. Chew *et al*., A multifunctional AAV-CRISPR-Cas9 and its host response. *Nat. Methods* **13**, 868–874 (2016). doi: 10.1038/nmeth.3993; pmid: 27595405
80. P. Liang *et al*., CRISPR/Cas9-mediated gene editing in human tripronuclear zygotes. *Protein Cell* **6**, 363–372 (2015). doi: 10.1007/s13238-015-0153-5; pmid: 25894090
81. T. Friedmann *et al*., ASGCT and JSGT joint position statement on human genomic editing. *Mol. Ther.* **23**, 1282 (2015). doi: 10.1038/mt.2015.118; pmid: 26227250
82. National Academies of Sciences, Engineering, and Medicine, *Human Genome Editing: Science, Ethics and Governance* (National Academies Press, Washington, DC, 2017).
83. R. W. Herzog *et al*., Stable gene transfer and expression of human blood coagulation factor IX after intramuscular injection of recombinant adeno-associated virus. *Proc. Natl. Acad. Sci. U.S.A.* **94**, 5804–5809 (1997). doi: 10.1073/pnas.94.11.5804; pmid: 9159155
84. J. D. Mount *et al*., Sustained phenotypic correction of hemophilia B dogs with a factor IX null mutation by liver-directed gene therapy. *Blood* **99**, 2670–2676 (2002). doi: 10.1182/blood.V99.8.2670; pmid: 11929752
85. G. P. Niemeyer *et al*., Long-term correction of inhibitor-prone hemophilia B dogs treated with liver-directed AAV2-mediated factor IX gene therapy. *Blood* **113**, 797–806 (2009). doi: 10.1182/blood-2008-10-181479; pmid: 18957684
86. A. C. Nathwani *et al*., Adenovirus-associated virus vector-mediated gene transfer in hemophilia B. *N. Engl. J. Med.* **365**, 2357–2365 (2011). doi: 10.1056/NEJMoa1108046; pmid: 22149959
87. A. C. Nathwani *et al*., Long-term safety and efficacy of factor IX gene therapy in hemophilia B. *N. Engl. J. Med.* **371**, 1994–2004 (2014). doi: 10.1056/NEJMoa1407309; pmid: 25409372
88. L. A. George *et al*., Hemophilia B gene therapy with a high-specific-activity factor IX Variant. *N. Engl. J. Med.* **377**, 2215–2227 (2017). doi: 10.1056/NEJMoa1708538; pmid: 29211678
89. S. Rangarajan *et al*., AAV5-factor viii gene transfer in severe hemophilia A. *N. Engl. J. Med.* **9** 10.1056/NEJMoa1708483 (2017). doi: 10.1056/NEJMoa1708483; pmid: 29224506
90. F. Mingozzi, K. A. High, Immune responses to AAV vectors: Overcoming barriers to successful gene therapy. *Blood* **122**, 23–36 (2013). doi: 10.1182/blood-2013-01-306647; pmid: 23596044
91. J. Mimuro *et al*., The prevalence of neutralizing antibodies against adeno-associated virus capsids is reduced in young Japanese individuals. *J. Med. Virol.* **86**, 1990–1997 (2014). doi: 10.1002/jmv.23818; pmid: 24136735
92. J. W. Bainbridge *et al*., Effect of gene therapy on visual function in Leber's congenital amaurosis. *N. Engl. J. Med.* **358**, 2231–2239 (2008). doi: 10.1056/NEJMoa0802268; pmid: 18441371
93. A. M. Maguire *et al*., Age-dependent effects of RPE65 gene therapy for Leber's congenital amaurosis: A phase 1 dose-escalation trial. *Lancet* **374**, 1597–1605 (2009). doi: 10.1016/S0140-6736(09)61836-5; pmid: 19854499
94. W. W. Hauswirth *et al*., Treatment of Leber congenital amaurosis due to RPE65 mutations by ocular subretinal injection of adeno-associated virus gene vector: Short-term results of a phase I trial. *Hum. Gene Ther.* **19**, 979–990 (2008). doi: 10.1089/hum.2008.107; pmid: 18774912
95. J. Bennett *et al*., Safety and durability of effect of contralateral-eye administration of AAV2 gene therapy in patients with childhood-onset blindness caused by RPE65 mutations: A follow-on phase 1 trial. *Lancet* **388**, 661–672 (2016). doi: 10.1016/S0140-6736(16)30371-3; pmid: 27375040
96. S. G. Jacobson *et al*., Improvement and decline in vision with gene therapy in childhood blindness. *N. Engl. J. Med.* **372**, 1920–1926 (2015). doi: 10.1056/NEJMoa1412965; pmid: 25936984
97. J. W. Bainbridge *et al*., Long-term effect of gene therapy on Leber's congenital amaurosis. *N. Engl. J. Med.* **372**, 1887–1897 (2015). doi: 10.1056/NEJMoa1414221; pmid: 25938638
98. A. F. Wright, Long-term effects of retinal gene therapy in childhood blindness. *N. Engl. J. Med.* **372**, 1954–1955 (2015). doi: 10.1056/NEJMe1503419; pmid: 25938534
99. S. Russell *et al*., Efficacy and safety of voretigene neparvovec (AAV2-hRPE65v2) in patients with RPE65-mediated inherited retinal dystrophy: A randomised, controlled, open-label, phase 3 trial. *Lancet* **390**, 849–860 (2017). pmid: 28712537
100. S. Muramatsu *et al*., Behavioral recovery in a primate model of Parkinson's disease by triple transduction of striatal cells with adeno-associated viral vectors expressing dopamine-synthesizing enzymes. *Hum. Gene Ther.* **13**, 345–354 (2002). doi: 10.1089/10430340252792486; pmid: 11860702
101. C. W. Christine *et al*., Safety and tolerability of putaminal AADC gene therapy for Parkinson disease. *Neurology* **73**, 1662–1669 (2009). doi: 10.1212/WNL.0b013e3181c29356; pmid: 19828868
102. S. Muramatsu *et al*., A phase I study of aromatic L-amino acid decarboxylase gene therapy for Parkinson's disease. *Mol. Ther.* **18**, 1731–1735 (2010). doi: 10.1038/mt.2010.135; pmid: 20606642
103. W. J. Marks Jr. *et al*., Gene delivery of AAV2-neurturin for Parkinson's disease: A double-blind, randomised, controlled trial. *Lancet Neurol.* **9**, 1164–1172 (2010). doi: 10.1016/S1474-4422(10)70254-4; pmid: 20970382
104. P. A. LeWitt *et al*., AAV2-GAD gene therapy for advanced Parkinson's disease: A double-blind, sham-surgery controlled, randomised trial. *Lancet Neurol.* **10**, 309–319 (2011). doi: 10.1016/S1474-4422(11)70039-4; pmid: 21419704
105. W. L. Hwu *et al*., Gene therapy for aromatic L-amino acid decarboxylase deficiency. *Sci. Transl. Med.* **4**, 134ra61 (2012). doi: 10.1126/scitranslmed.3003640; pmid: 22593174
106. K. Talbot, E. F. Tizzano, The clinical landscape for SMA in a new therapeutic era. *Gene Ther.* **24**, 529–533 (2017). doi: 10.1038/gt.2017.52; pmid: 28644430
107. J. R. Mendell *et al*., Single-dose gene-replacement therapy for spinal muscular atrophy. *N. Engl. J. Med.* **377**, 1713–1722 (2017). doi: 10.1056/NEJMoa1706198; pmid: 29091557
108. S. Hacein-Bey-Abina *et al*., Sustained correction of X-linked severe combined immunodeficiency by ex vivo gene therapy. *N. Engl. J. Med.* **346**, 1185–1193 (2002). doi: 10.1056/NEJMoa012616; pmid: 11961146
109. A. Aiuti *et al*., Correction of ADA-SCID by stem cell gene therapy combined with nonmyeloablative conditioning. *Science* **296**, 2410–2413 (2002). doi: 10.1126/science.1070104; pmid: 12089448
110. S. Hacein-Bey-Abina *et al*., LMO2-associated clonal T cell proliferation in two patients after gene therapy for SCID-X1. *Science* **302**, 415–419 (2003). doi: 10.1126/science.1088547; pmid: 14564000
111. S. Stein *et al*., Genomic instability and myelodysplasia with monosomy 7 consequent to EVI1 activation after gene therapy for chronic granulomatous disease. *Nat. Med.* **16**, 198–204 (2010). doi: 10.1038/nm.2088; pmid: 20098431
112. C. J. Braun *et al*., Gene therapy for Wiskott-Aldrich syndrome—long-term efficacy and genotoxicity. *Sci. Transl. Med.* **6**, 227ra33 (2014). doi: 10.1126/scitranslmed.3007280; pmid: 24622513
113. X. Wu, Y. Li, B. Crise, S. M. Burgess, Transcription start regions in the human genome are favored targets for MLV integration. *Science* **300**, 1749–1751 (2003). doi: 10.1126/science.1083413; pmid: 12805549
114. A. Aiuti *et al*., Lentiviral hematopoietic stem cell gene therapy in patients with Wiskott-Aldrich syndrome. *Science* **341**, 1233151 (2013). doi: 10.1126/science.1233151; pmid: 23845947
115. S. Hacein-Bey-Abina *et al*., A modified γ-retrovirus vector for X-linked severe combined immunodeficiency. *N. Engl. J. Med.* **371**, 1407–1417 (2014). doi: 10.1056/NEJMoa1404588; pmid: 25295500
116. A. Aiuti *et al*., Gene therapy for immunodeficiency due to adenosine deaminase deficiency. *N. Engl. J. Med.* **360**, 447–458 (2009). doi: 10.1056/NEJMoa0805817; pmid: 19179314
117. A. Biffi *et al*., Lentiviral hematopoietic stem cell gene therapy benefits metachromatic leukodystrophy. *Science* **341**, 1233158 (2013). doi: 10.1126/science.1233158; pmid: 23845948
118. N. Cartier *et al*., Hematopoietic stem cell gene therapy with a lentiviral vector in X-linked adrenoleukodystrophy. *Science* **326**, 818–823 (2009). doi: 10.1126/science.1171242; pmid: 19892975
119. F. Eichler *et al*., Hematopoietic stem-cell gene therapy for cerebral adrenoleukodystrophy. *N. Engl. J. Med.* **377**, 1630–1638 (2017). doi: 10.1056/NEJMoa1700554; pmid: 28976817
120. M. Cavazzana-Calvo *et al*., Transfusion independence and HMGA2 activation after gene therapy of human β-thalassaemia. *Nature* **467**, 318–322 (2010). doi: 10.1038/nature09328; pmid: 20844535
121. J. A. Ribeil *et al*., Gene therapy in a patient with sickle cell disease. *N. Engl. J. Med.* **376**, 848–855 (2017). doi: 10.1056/NEJMoa1609677; pmid: 28249145
122. M. C. Canver *et al*., BCL11A enhancer dissection by Cas9-mediated in situ saturating mutagenesis. *Nature* **527**, 192–197 (2015). doi: 10.1038/nature15521; pmid: 26375006
123. E. A. Traxler *et al*., A genome-editing strategy to treat β-hemoglobinopathies that recapitulates a mutation associated with a benign genetic condition. *Nat. Med.* **22**, 987–990 (2016). doi: 10.1038/nm.4170; pmid: 27525524
124. M. Sadelain, I. Rivière, R. Brentjens, Targeting tumours with genetically enhanced T lymphocytes. *Nat. Rev. Cancer* **3**, 35–45 (2003). doi: 10.1038/nrc971; pmid: 12509765
125. M. Sadelain, R. Brentjens, I. Rivière, The basic principles of chimeric antigen receptor design. *Cancer Discov.* **3**, 388–398 (2013). doi: 10.1158/2159-8290.CD-12-0548; pmid: 23550147
126. M. C. Jensen, S. R. Riddell, Designing chimeric antigen receptors to effectively and safely target tumors. *Curr. Opin. Immunol.* **33**, 9–15 (2015). doi: 10.1016/j.coi.2015.01.002; pmid: 25621840
127. Z. Eshhar, T. Waks, G. Gross, The emergence of T-bodies/CAR T cells. *Cancer J.* **20**, 123–126 (2014). doi: 10.1097/PPO.0000000000000027; pmid: 24667957
128. R. J. Brentjens *et al*., Eradication of systemic B-cell tumors by genetically targeted human T lymphocytes co-stimulated by CD80 and interleukin-15. *Nat. Med.* **9**, 279–286 (2003). doi: 10.1038/nm827; pmid: 12579196
129. I. Rivière, M. Sadelain, Chimeric antigen receptors: A cell and gene therapy perspective. *Mol. Ther.* **25**, 1117–1124 (2017). doi: 10.1016/j.ymthe.2017.03.034; pmid: 28456379
130. C. Imai *et al*., Chimeric receptors with 4-1BB signaling capacity provoke potent cytotoxicity against acute lymphoblastic leukemia. *Leukemia* **18**, 676–684 (2004). doi: 10.1038/sj.leu.2403302; pmid: 14961035
131. J. Maher, R. J. Brentjens, G. Gunset, I. Rivière, M. Sadelain, Human T-lymphocyte cytotoxicity and proliferation directed by a single chimeric TCRzeta /CD28 receptor. *Nat. Biotechnol.* **20**, 70–75 (2002). doi: 10.1038/nbt0102-70; pmid: 11753365
132. L. Gattinoni, C. A. Klebanoff, N. P. Restifo, Paths to stemness: Building the ultimate antitumour T cell. *Nat. Rev. Cancer* **12**, 671–684 (2012). doi: 10.1038/nrc3322; pmid: 22996603
133. D. H. Busch, S. P. Fräßle, D. Sommermeyer, V. R. Buchholz, S. R. Riddell, Role of memory T cell subsets for adoptive immunotherapy. *Semin. Immunol.* **28**, 28–34 (2016). doi: 10.1016/j.smim.2016.02.001; pmid: 26976826
134. M. L. Davila *et al*., Efficacy and toxicity management of 19-28z CAR T cell therapy in B cell acute lymphoblastic leukemia. *Sci. Transl. Med.* **6**, 224ra25 (2014). doi: 10.1126/scitranslmed.3008226; pmid: 24553386
135. D. W. Lee *et al*., Current concepts in the diagnosis and management of cytokine release syndrome. *Blood* **124**, 188–195 (2014). doi: 10.1182/blood-2014-05-552729; pmid: 24876563
136. S. A. Ali *et al*., T cells expressing an anti-B-cell maturation antigen chimeric antigen receptor cause remissions of multiple myeloma. *Blood* **128**, 1688–1700 (2016). doi: 10.1182/blood-2016-04-711903; pmid: 27412889
137. C. A. Klebanoff, S. A. Rosenberg, N. P. Restifo, Prospects for gene-engineered T cell immunotherapy for solid cancers. *Nat. Med.* **22**, 26–36 (2016). doi: 10.1038/nm.4015; pmid: 26735408
138. S. Gill, Chimeric antigen receptor T cell therapy in AML: How close are we? *Best Pract. Res. Clin. Haematol.* **29**, 329–333 (2016). doi: 10.1016/j.beha.2016.10.004; pmid: 27890255
139. J. Couzin-Frankel, Breakthrough of the year 2013. Cancer immunotherapy. *Science* **342**, 1432–1433 (2013). doi: 10.1126/science.342.6165.1432; pmid: 24357284
140. C. T. Ellebrecht *et al*., Reengineering chimeric antigen receptor T cells for targeted therapy of autoimmune disease. *Science* **353**, 179–184 (2016). doi: 10.1126/science.aaf6756; pmid: 27365313
141. S. H. Orkin, P. Reilly, Paying for future success in gene therapy. *Science* **352**, 1059–1061 (2016). doi: 10.1126/science.aaf4770; pmid: 27230368
142. E. Sagonowsky, GlaxoSmithKline's talk of leaving rare disease-amid slow Strimvelis sales-highlight challenges in the field. *FiercePharma*. 2017.

143. R. J. Brentjens *et al*., Safety and persistence of adoptively transferred autologous CD19-targeted T cells in patients with relapsed or chemotherapy refractory B-cell leukemias. *Blood* **118**, 4817–4828 (2011). doi: 10.1182/blood-2011-04-348540; pmid: 21849486
144. R. J. Brentjens *et al*., CD19-targeted T cells rapidly induce molecular remissions in adults with chemotherapy-refractory acute lymphoblastic leukemia. *Sci. Transl. Med.* **5**, 177ra38 (2013). doi: 10.1126/scitranslmed.3005930; pmid: 23515080
145. S. L. Maude *et al*., Chimeric antigen receptor T cells for sustained remissions in leukemia. *N. Engl. J. Med.* **371**, 1507–1517 (2014). doi: 10.1056/NEJMoa1407222; pmid: 25317870
146. D. W. Lee *et al*., T cells expressing CD19 chimeric antigen receptors for acute lymphoblastic leukaemia in children and young adults: A phase 1 dose-escalation trial. *Lancet* **385**, 517–528 (2015). doi: 10.1016/S0140-6736(14)61403-3; pmid: 25319501
147. J. N. Kochenderfer *et al*., Chemotherapy-refractory diffuse large B-cell lymphoma and indolent B-cell malignancies can be effectively treated with autologous T cells expressing an anti-CD19 chimeric antigen receptor. *J. Clin. Oncol.* **33**, 540–549 (2015). doi: 10.1200/JCO.2014.56.2025; pmid: 25154820
148. C. J. Turtle *et al*., Immunotherapy of non-Hodgkin's lymphoma with a defined ratio of CD8+ and CD4+ CD19-specific chimeric antigen receptor-modified T cells. *Sci. Transl. Med.* **8**, 355ra116 (2016). doi: 10.1126/scitranslmed.aaf8621; pmid: 27605551
149. D. L. Porter, B. L. Levine, M. Kalos, A. Bagg, C. H. June, Chimeric antigen receptor-modified T cells in chronic lymphoid leukemia. *N. Engl. J. Med.* **365**, 725–733 (2011). doi: 10.1056/NEJMoa1103849; pmid: 21830940
150. J. N. Kochenderfer *et al*., Eradication of B-lineage cells and regression of lymphoma in a patient treated with autologous T cells genetically engineered to recognize CD19. *Blood* **116**, 4099–4102 (2010). doi: 10.1182/blood-2010-04-281931; pmid: 20668228
151. P. F. Robbins *et al*., Tumor regression in patients with metastatic synovial cell sarcoma and melanoma using genetically engineered lymphocytes reactive with NY-ESO-1. *J. Clin. Oncol.* **29**, 917–924 (2011). doi: 10.1200/JCO.2010.32.2537; pmid: 21282551
152. S. Hacein-Bey Abina *et al*., Outcomes following gene therapy in patients with severe Wiskott-Aldrich syndrome. *JAMA* **313**, 1550–1563 (2015). doi: 10.1001/jama.2015.3253; pmid: 25898053
153. S. S. De Ravin *et al*., Lentiviral hematopoietic stem cell gene therapy for X-linked severe combined immunodeficiency. *Sci. Transl. Med.* **8**, 335ra57 (2016). doi: 10.1126/scitranslmed.aad8856; pmid: 27099176
154. M. Sessa *et al*., Lentiviral haemopoietic stem-cell gene therapy in early-onset metachromatic leukodystrophy: An ad-hoc analysis of a non-randomised, open-label, phase 1/2 trial. *Lancet* **388**, 476–487 (2016). doi: 10.1016/S0140-6736(16)30374-9; pmid: 27289174
155. D. Gaudet *et al*., Efficacy and long-term safety of alipogene tiparvovec (AAV1-LPLS447X) gene therapy for lipoprotein lipase deficiency: An open-label trial. *Gene Ther.* **20**, 361–369 (2013). doi: 10.1038/gt.2012.43; pmid: 22717743

ACKNOWLEDGMENTS

K.A.H. is President and Head of Research and Development at Spark Therapeutics, a biotechnology company involved in discovering, developing, and delivering AAV gene therapies for genetic diseases. K.O. has a sponsored research agreement with Takara Bio for CAR-T gene therapy. J.K.J. is a paid scientific adviser for Beam Therapeutics, Editas Medicine, and Pairwise Plants and holds equity in Beam Therapeutics, Editas Medicine, Monitor Biotechnologies, Pairwise Plants, Poseida Therapeutics, and Transposagen Biopharmaceuticals. These financial interests were reviewed and are managed by Massachusetts General Hospital and Partners HealthCare in accordance with their conflict of interest policies. J.K.J. has sponsored research agreements with AstraZeneca and Takeda Pharmaceuticals and is an inventor on patent applications and issued patents that cover genome editing technologies. D.B.K. is on the Scientific Advisory Boards of Orchard Therapeutics and Kite Pharma, companies developing gene therapies, and has a Sponsored Research Agreement from BioMarin Pharmaceutical on lentiviral vectors for sickle cell disease. D.B.K. is an inventor on intellectual property licensed to Orchard Therapeutics from the University of California Los Angeles, on lentiviral vector gene therapy for adenosine deaminase–deficient SCID. M.S. has sponsored research agreements with Juno Therapeutics and Fate Therapeutics. M.S. is an inventor on patents related to CAR therapies that are licensed by Memorial Sloan-Kettering Cancer Center.

10.1126/science.aan4672

RESEARCH ARTICLE SUMMARY

PROTEIN MODIFICATION

Fatty acyl recognition and transfer by an integral membrane *S*-acyltransferase

Mitra S. Rana,* Pramod Kumar,* Chul-Jin Lee, Raffaello Verardi, Kanagalaghatta R. Rajashankar, Anirban Banerjee†

INTRODUCTION: Hundreds of cellular proteins are modified by posttranslational S-acylation of cysteines, commonly known as protein palmitoylation. Unlike other lipid attachments, which are thought to be permanent, palmitoylation can be reversed by cellular thioesterases, enabling dynamic modulation of the local hydrophobicity of substrate proteins. In humans, palmitoylation is catalyzed by 23 members of the DHHC family of integral membrane enzymes, which contain a signature Asp-His-His-Cys (DHHC) motif. DHHC enzymes use acyl–coenzyme A (CoA) (predominantly palmitoyl-CoA) to generate an acyl-enzyme intermediate from which the acyl chain is subsequently transferred to a substrate. With a recent systems-level analysis suggesting that more than 10% of the proteome is palmitoylated, the complexity of protein palmitoylation approaches that of protein phosphorylation and ubiquitylation. Nonetheless, fundamental aspects of DHHC enzymes, including their mechanism of catalysis and acyl-CoA binding and recognition, have been challenging to analyze without detailed structural information.

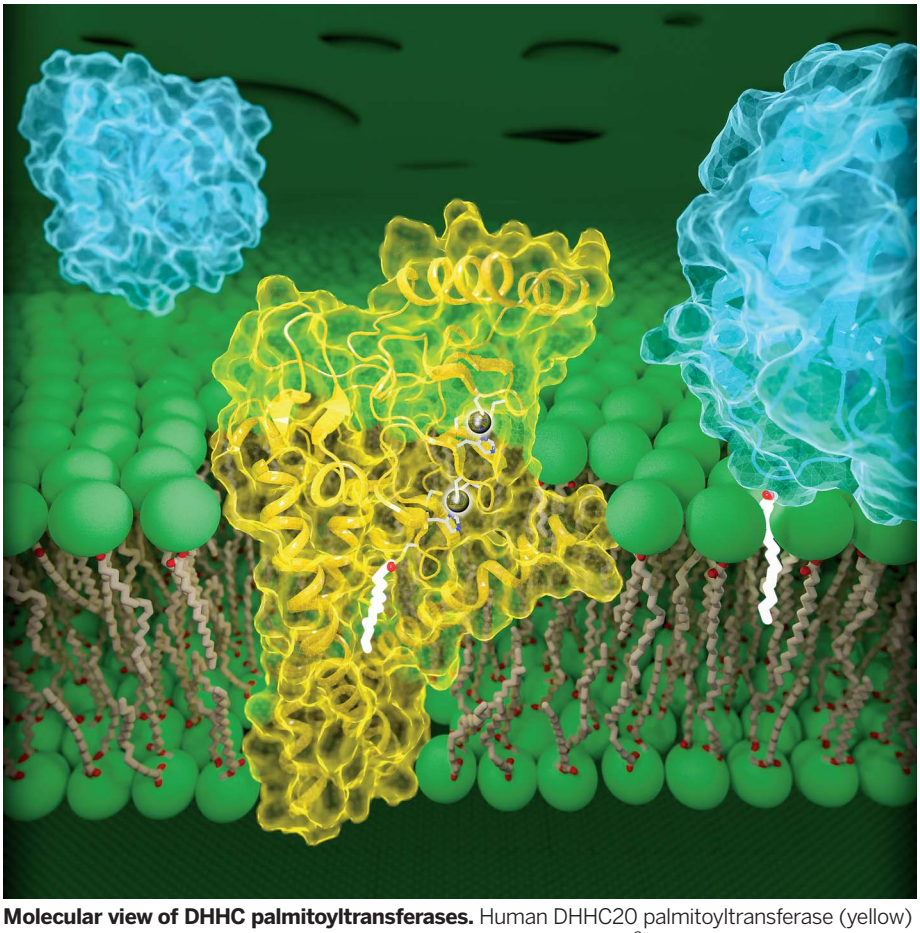
RATIONALE: To obtain insights into the structural mechanism of DHHC enzymes, we solved the crystal structures of two DHHC family members: human DHHC20 and a catalytically inactive mutant of zebrafish DHHC15. We also purified and solved the structure of human DHHC20 conjugated to an irreversible inhibitor that mimics the acylated enzyme intermediate. We carried out structure-guided mutagenesis experiments to test residues important for enzyme function and to engineer mutant enzymes with altered acyl-CoA selectivity.

ON OUR WEBSITE

Read the full article at http://dx.doi.org/10.1126/science.aao6326

RESULTS: The four transmembrane helices of hDHHC20 and zfDHHS15 form a tepee-like structure with the active site, contained in the highly conserved cytosolic DHHC cysteine-rich domain, at the membrane-cytosol interface. The cysteine-rich domain binds two zinc (Zn^{2+}) ions that impart structural stability, but do not actively coordinate the nucleophilic cysteine. The C-terminal domain has an unanticipated amphipathic helix and a hydrophobic loop that together form a supporting structure for the transmembrane domain. The transmembrane domain forms a cavity where the acyl chain of acyl-CoA is inserted. Cavity-lining residues are determinants of fatty acyl recognition and chain-length selectivity. Our structures enabled us to engineer mutants of human DHHC20 with altered acyl chain–length selectivities.

CONCLUSION: By elucidating the location of the active site at the membrane-aqueous interface, our structures readily explain why candidate cysteines for palmitoylation are proximal to the membrane. The active site has a catalytic triad-like arrangement of aspartic acid and histidine residues that activate the nucleophilic cysteine. The C-terminal domain has conserved motifs that form critical interactions with the active site and the rest of the protein. The structures reported here set the stage for the development of structure-based small molecule probes and tools such as orthogonal DHHC enzyme–fatty acyl–CoA pairs that will likely help investigate the enzyme-substrate network of this biologically and biomedically important family of enzymes. ■



Molecular view of DHHC palmitoyltransferases. Human DHHC20 palmitoyltransferase (yellow) shown localized in the Golgi body membrane (green stacks). The Zn^{2+} ions are shown as gray spheres and the acyl chain of the palmitoyl group in white sticks. A hypothetical substrate (blue) approaches the palmitoyltransferase from the left and, after palmitoylation, is localized to the Golgi body membrane through anchoring of the palmitoyl group, now transferred to the substrate.

The list of author affiliations is available in the full article online.
*These authors contributed equally to this work.
†Corresponding author. Email: anirban.banerjee@nih.gov
Cite this article as M. S. Rana *et al.*, *Science* **359**, eaao6326 (2018). DOI: 10.1126/science.aao6326

RESEARCH ARTICLE

PROTEIN MODIFICATION

Fatty acyl recognition and transfer by an integral membrane *S*-acyltransferase

Mitra S. Rana,[1]* Pramod Kumar,[1]* Chul-Jin Lee,[1] Raffaello Verardi,[1] Kanagalaghatta R. Rajashankar,[2] Anirban Banerjee[1]†

DHHC (Asp-His-His-Cys) palmitoyltransferases are eukaryotic integral membrane enzymes that catalyze protein palmitoylation, which is important in a range of physiological processes, including small guanosine triphosphatase (GTPase) signaling, cell adhesion, and neuronal receptor scaffolding. We present crystal structures of two DHHC palmitoyltransferases and a covalent intermediate mimic. The active site resides at the membrane-cytosol interface, which allows the enzyme to catalyze thioester-exchange chemistry by using fatty acyl–coenzyme A and explains why membrane-proximal cysteines are candidates for palmitoylation. The acyl chain binds in a cavity formed by the transmembrane domain. We propose a mechanism for acyl chain–length selectivity in DHHC enzymes on the basis of cavity mutants with preferences for shorter and longer acyl chains.

Covalent attachment to lipids is a prevalent form of posttranslational modification (*1*) that influences membrane association of peripheral membrane proteins, protein targeting to membrane subdomains, protein-protein associations, and protein trafficking (*2*, *3*). Of the main different forms of protein lipidation, attachment of a 14-carbon myristoyl group to the N terminus or a 15-carbon farnesyl or 20-carbon geranylgeranyl to the C terminus results in a relatively stable modification (*4*). By contrast, attachment of a fatty acyl group to an internal cysteine as a thioester (*5*, *6*), known as protein S-acylation, is readily reversible through the action of cellular thioesterases (*7*, *8*), making S-acylation a potentially dynamic form of lipidation (*9*).

Protein S-acylation is more commonly referred to as protein palmitoylation owing to the prevalence of the 16-carbon palmitate among the acyl chains that are attached to substrate proteins (*10*). However, for at least a subset of proteins, modification by fatty acyl chains longer or shorter than 16 carbons has been shown (*10*–*12*). The readily reversible nature of protein palmitoylation enables dynamic modulation of the hydrophobicity of substrate proteins. Protein palmitoylation plays critical roles in a wide range of physiological processes such as Ras signaling (*13*), localization of neuronal scaffolding proteins (*14*), intracellular trafficking (*15*), activity of ion channels (*16*), and host-pathogen interactions (*17*, *18*). Since their discovery, an increasing number of proteins have been added to the repertoire of cellular proteins that are palmitoylated, with a recent estimate of close to 1000 proteins in humans (*19*). Although bioinformatic analyses of protein sequences proximal to the target cysteine have had some success in predicting palmitoylation sites, there are currently no reported consensus sequences for palmitoylation (*20*). Examination of experimentally identified palmitoylation sites and their sequence context, both in terms of physicochemical properties as well as predicted structure, is strongly indicative of the fact that one of the criteria for a cysteine to be palmitoylated is proximity to the membrane (*20*). Protein palmitoylation is connected to diseases, especially cancers and neuropsychiatric disorders (*21*), and it has been proposed that developing inhibitors of DHHC20, an enzyme that palmitoylates epidermal growth factor receptor (EGFR), can provide a therapeutic avenue for treating cancers that are resistant to EGFR-targeted therapy (*22*).

Although palmitoylation as a posttranslational modification was discovered in 1979 (*5*), the enzymes that catalyze protein palmitoylation were only discovered in 2002 (*23*, *24*). These are low-abundance, polytopic eukaryotic integral membrane enzymes known as DHHC-palmitoyl transferases, so named because they contain a signature Asp-His-His-Cys (DHHC) motif within a cysteine-rich domain in an intracellular loop (fig. S1). Localization studies suggest that DHHC enzymes reside at a variety of cellular compartments, most prominently the Golgi complex (*25*). Beyond the shared cysteine-rich domain, there is considerable diversity in the DHHC family—some possess ankyrin repeats (*24*), a few have six transmembrane (TM) helices (*26*) instead of the canonical four, and at least one of them forms a functional heterodimer with an auxiliary subunit (*23*). Studies of yeast Erf2/Erf4 (homolog of mammalian DHHC9/GCP16) (*27*) and mammalian DHHC2 and DHHC3 (*28*) indicate that palmitate transfer to substrates occurs in two steps. First, autoacylation of a cysteine within the enzyme with palmitoyl–coenzyme A (CoA) forms a palmitoylated intermediate. This intermediate has been isolated in vitro, and, in the absence of a substrate, the autopalmitoylated enzyme undergoes slow hydrolysis. However, in presence of a protein substrate, the palmitate is transferred to a cysteine on the substrate in a transpalmitoylation reaction that regenerates the DHHC enzyme (*28*) (Fig. 1A). The specific roles of the conserved residues in the cysteine-rich domain that includes the DHHC motif are poorly understood. Genetic and biochemical analyses indicates that DHHC enzymes bind two Zn^{2+} ions (*29*) at two zinc finger–like domains, but the function of these Zn^{2+} ions in DHHC enzymes is unknown. Moreover, fatty acyl–CoA selectivity varies between DHHC enzymes (*28*, *30*). Nevertheless, nothing is known about the site on the enzyme where the acyl-CoA binds and, thus, the determinants for fatty acid chain–length selectivity.

Functional characterization of DHHC15 and DHHC20 constructs

We carried out an extensive search for appropriate constructs of members of the DHHC family for crystallization using fluorescence size-exclusion chromatography (FSEC), screening for protein stability, yield, and monodispersity of size-exclusion chromatographic profile (*31*). Of the initial hits, human DHHC20 (hDHHC20) and zebrafish DHHC15 (zfDHHC15) were promising candidates (figs. S2 and S4). Because there was no biochemical characterization of either DHHC20 or DHHC15 in the literature, we tested whether they are bona fide palmitoyltransferase enzymes. We first used a coupled-enzyme assay that interrogates DHHC autoacylation by utilizing the released free CoA to generate an equivalent amount of reduced nicotinamide adenine dinucleotide (NADH) (*32*) (Fig. 1A), which is detected by its fluorescence (fig. S3A). Both hDHHC20 and zfDHHC15 showed robust activity (Fig. 1, B and C, and figs. S3C and S4E). Mutants in which the active site cysteine was changed to a serine, designated as hDHHS20 and zfDHHS15, showed very little activity. To demonstrate the capability of the enzymes for catalyzing palmitoylation of protein substrates, we used the *Legionella* effector protein GobX (*17*) and human SNAP25b (*30*) as substrates for hDHHC20 and zfDHHC15, respectively. In an in vitro assay with purified proteins, hDHHC20 and zfDHHC15 displayed robust palmitoyl-transfer activity, whereas the catalytically inactive mutants, hDHHS20 and zfDHHS15, displayed the same level of activity as the "no enzyme" control (fig. S5). These experiments demonstrated that our preparations of hDHHC20 and zfDHHC15

[1]Cell Biology and Neurobiology Branch, National Institutes of Child Health and Human Development, National Institutes of Health, Bethesda, MD 20892, USA. [2]Northeastern Collaborative Access Team (NE-CAT) and Department of Chemistry and Chemical Biology, Cornell University, Building 436E, Argonne National Laboratory, 9700 South Cass Avenue, Argonne, IL 60439, USA.
*These authors contributed equally to this work.
†Corresponding author. Email: anirban.banerjee@nih.gov

have robust autoacylation activity in a fully reconstituted in vitro assay and are competent for substrate palmitoylation. Purified hDHHC20 and zfDHHC15 contain two Zn^{2+} ions per protein (figs. S2D and S4F), similar to what has been observed with DHHC3 (*33*) and Swf1, a yeast DHHC enzyme (*29*).

Crystallization and overall structural organization of DHHC enzymes

We obtained crystals of wild-type hDHHC20 from lipidic cubic phase (LCP) and crystals of the mutant, zfDHHS15, from hanging drop conditions. These crystals yielded diffraction data that enabled us to solve the structure using Zn–single-wavelength anomalous dispersion (SAD) data from the Zn^{2+} ions in the protein (hDHHC20) and with a combination of Zn-SAD and sulfur-SAD for zfDHHS15 (fig. S6, F to I). The hDHHC20 protein crystallizes in two different crystal forms in LCP: a $P6_3$ form with an antiparallel dimer and a $P2_1$ form with a loose parallel dimer in the asymmetric unit (fig. S6, A and B). Conversely, zfDHHS15 crystallizes in $P4_22_12$ space group with an antiparallel dimer in the asymmetric unit (fig. S6C). We obtained x-ray diffraction data of zfDHHS15 crystals to 2.54-Å resolution. The hDHHC20 protein was reductively methylated for obtaining optimal diffraction data. The hexagonal form of hDHHC20 diffracted to 2.44-Å resolution. The DHHC structures display similar monomer structures, with a root mean square deviation (RMSD) of superposition (Cα) between the monomers of the two forms of hDHHC20 being 0.52 Å and between the monomer of zfDHHS15 and the monomer of hDHHC20 ($P6_3$ form) being 0.74 Å (fig. S6, D and E). We will focus our discussion on the hexagonal form of hDHHC20, with comments on zfDHHS15 where appropriate. The final models contain residues 5 to 299 of hDHHC20 and residues 8 to 298 of zfDHHS15 (residues 42 and 43 and 199 to 202 are absent from the zfDHHS15 model) (tables S1 and S2).

Protein palmitoylation almost exclusively occurs at the cytoplasmic face of organellar and plasma membranes, and thus the predicted topology of DHHC enzymes would place the DHHC loop in the cytoplasm. Based on this and the fact that the same hDHHC20 protein crystallized in lipid-rich phase in two different forms, we interpret the different dimeric interfaces as adventitious protein-protein interactions. Such dimer formation has been seen in other membrane-protein structures (*34*, *35*). To rule out hDHHC20 or zfDHHC15 adopting an inverted topology, we carried out a fluorescence protease protection assay (*36*) to determine the physiological topology of both hDHHC20 and zfDHHC15. Both hDHHC20 and zfDHHC15 are putative Golgi resident proteins, and C-terminal green fluorescent protein (GFP)–tagged constructs localize to the Golgi. Upon permeabilization of the plasma membrane with digitonin and addition of trypsin, the fluorescent tag on the C terminus is not protected, whereas a control with a fluorescent tag in the lumen of the Golgi is (figs. S7 and S8). Given that both hDHHC20 and zfDHHC15 have four TM helices, with the DHHC loop located between TM2 and TM3, this experiment rules out the possibility of any physiologically relevant topology of hDHHC20 or zfDHHC15 that positions the DHHC loop in the lumen of the Golgi.

The four TM helices of hDHHC20 and zfDHHS15 adopt a tepee-like organization in the membrane, coming close together on the lumenal side and splaying apart on the cytoplasmic side, where substrate engagement and catalysis take place (Fig. 2). TM1 and TM4 are the most tilted with respect to the transmembrane normal. On the lumenal side, short loops connect TM1 and TM2, and TM3 and TM4. On the cytoplasmic side, the DHHC cysteine-rich domain connects TM2 and TM3. The N-terminal part of this domain consists of a helix-turn-helix that is connected to TM2 by a flexible linker. The C-terminal part of the domain has three stacked β hairpins arranged roughly in parallel with the bilayer and terminates in a short linker that is connected to TM3. The C-terminal domain of hDHHC20 is also at the cytoplasmic face and begins with a short helical segment (α′1) that wedges in between TM3 and the β hairpins in the DHHC domain. This domain is connected to a β hairpin that protrudes into the cytoplasm, but comes back to the membrane proximal region and ends in an amphipathic helix (α′2) that rests against TM3 and TM4. On the C-terminal side of the amphipathic helix, a hydrophobic loop with a short α-helical segment (α′3) inserts into the putative membrane bilayer region and makes extensive contacts with TM2 and TM3. Together α′2 and α′3 form a supporting belt on the cytoplasmic side of TM3 and TM4 (Fig. 2, A to C).

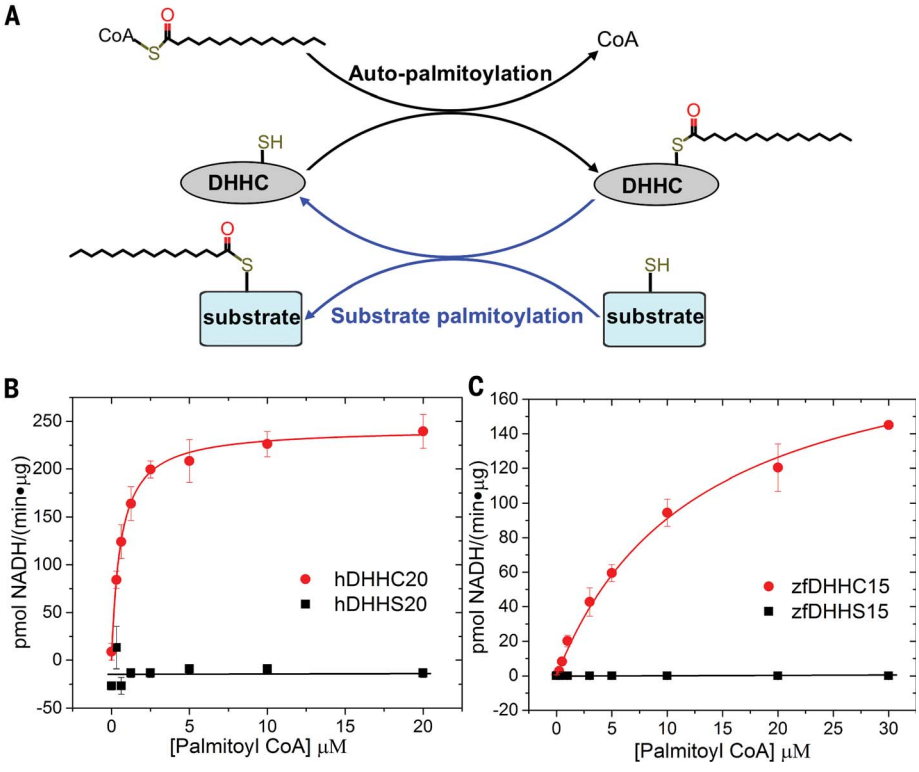


Fig. 1. Functional characterization of DHHC enzymes. (**A**) The proposed two-step catalytic mechanism of the DHHC enzymes in literature, where they first undergo self-acylation to form an acylated intermediate (shown in black) and subsequently transfer the acyl chain to a protein substrate in the second step (shown in blue). (**B**) Determination of kinetic parameters for the autopalmitoylation of hDHHC20 measured using the coupled-enzyme assay (fig. S3A). Shown is a Michaelis-Menten fit to palmitoyl-CoA titration with hDHHC20 (red circle) and hDHHS20 (black square). K_m = 0.58 ± 0.04 μM; k_{cat} = 16.5 ± 0.3 min^{-1}. (**C**) Same as (B) for zfDHHC15. K_m = 10 ± 1 μM; k_{cat} = 14.6 ± 0.8 min^{-1}. Data are mean ± SEM of three or more replicate measurements.

Structure of the DHHC cysteine-rich domain and organization of the active site

The cysteine-rich domain binds two Zn^{2+} ions, which is consistent with data from other DHHC members suggesting that this domain contains zinc finger motifs (*29*). The Zn^{2+} ions presumably play a structural role, as they are positioned between three parallel layers of β hairpins (Fig. 3A). Three cysteines and a histidine form the tetrahedral coordination environment of each Zn^{2+}, resembling canonical zinc finger motifs (*37*). Neither of the Zn^{2+} ions coordinate the active site cysteine and thus do not directly participate

in catalysis (*38*). However, mutations of the coordinating residues have been shown to drastically change the stability and/or catalytic activity of DHHC3 (*33*), supporting a role for the Zn^{2+} ions in positioning the catalytic cysteine optimally. Consequently, these residues are conserved among all members of the DHHC family (fig. S1). Additionally, a role for the zinc fingers in the binding of protein substrates can also not be ruled out (*39*).

The active site is situated at the membrane-aqueous interface in agreement with the observation that proteins are palmitoylated at membrane-proximal cysteines (*20*). The catalytic DHHC motif (the two histidines are His^{154} and His^{155}) is located on the β hairpins (β5-β6) that coordinate the Zn^{2+} ions (Fig. 3B). The aspartic acid and the His^{154} are on the top face of the β hairpin, protruding in the direction of the membrane, and presumably form part of a catalytic triad (*40*) where the aspartic acid forms a hydrogen-bonding interaction with the histidine (Fig. 3B). The His^{155} side chain coordinates a Zn^{2+} ion on the other face of the β hairpin and presumably positions the catalytic cysteine for nucleophilic attack. In the zfDHHS15 structure, the side chain of the active site serine residue rotates to form a hydrogen-bonding interaction with a main-chain amide nitrogen (fig. S9). However, the putative position of the catalytic cysteine in the wild-type enzyme will likely bring it within hydrogen-bonding distance of His^{154}. The structure suggests that the aspartic acid and the first histidine of the DHHC motif serve crucial roles in activating the nucleophilic cysteine. Mutating these residues individually to alanine generates mutants with no discernible catalytic activity (Fig. 3C).

This organization of the active site is such that the catalytic cysteine of the DHHC motif protrudes in the direction of the membrane bilayer (Fig. 3B). Given that DHHC enzymes catalyze via a two-step reaction mechanism and that the acylated enzyme is an intermediate species with a long hydrophobic moiety attached to it, this suggests that the active site is preorganized to stabilize the acylated intermediate by inserting the acyl chain into the hydrophobic part of the bilayer. hDHHC20 has two positively charged patches on the cytoplasmic side (Fig. 2D, left); one of them arises from the binding of the two positively charged zinc ions. Another basic patch containing Arg^{126} and Lys^{135} showed electron density that we modeled as adenosine 5′-diphosphate with an additional phosphate at the 3′ end (5′-diphosphoadenosine 3′-phosphate) (fig. S10, A to C). This was based on consideration of components in the crystallization solution and other possible carryover contaminants such as adenosine 5′-triphosphate (ATP). Interestingly, this moiety is contained in the terminal part of CoA and thus could represent a byproduct of catalysis (fig. S10E). Because we did not add CoA during the protein purification, presumably this ligand was carried through the purification process. Notably, acyl-CoA has a long pantothenyl linker connecting thioester and the ribose ring. The extended form of this linker can easily span the length between the active site and the 5′-diphosphoadenosine 3′-phosphate binding site in the structure (fig. S10D). Based on these considerations, we suggest the 5′-diphosphoadenosine 3′-phosphate binding site to be the place where the terminal part of the coenzyme headgroup of palmitoyl-CoA binds.

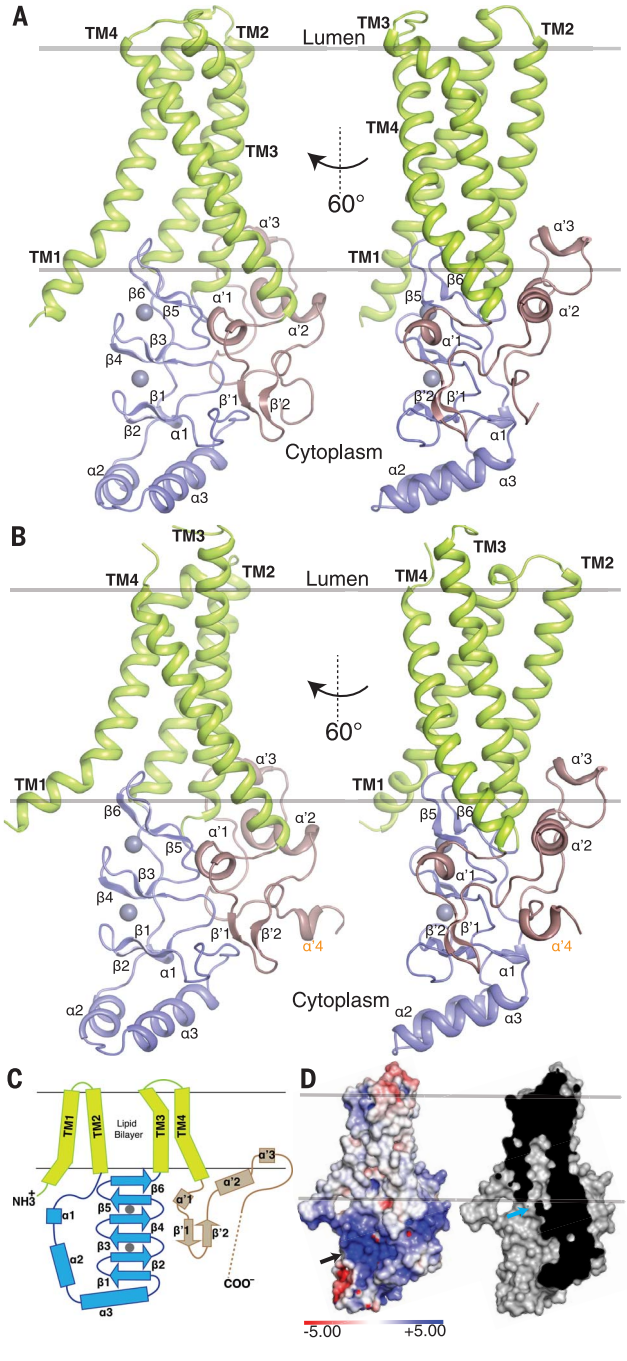


Fig. 2. Overall structure of a DHHC enzyme. (**A**) Cartoon representation of hDHHC20. The four TM helices are shown in green, the DHHC cysteine-rich domain in blue, and the C-terminal domain in brown. The two spheres represent the Zn^{2+} ions. (**B**) Same as (A) for zfDHHS15. An extra helix in the C-terminal domain is labeled in orange. (**C**) Diagram of the secondary structure elements of hDHHC20, showing helices as rectangles and β sheets as arrows. The two gray circles represent Zn^{2+} ions, and the dashed line indicates disordered C-terminal domain not observed in the crystal structure. Colors are the same as in (A). (**D**) Molecular surface representation (left) of hDHHC20 colored by electrostatic potential, showing a distinctive basic patch (black arrow) and a cross section (right) showing a cavity (cyan arrow) that exists above the active site. Electrostatic surface potential was generated using an Adaptive Poisson-Boltzmann Solver (APBS) server with default settings. Positively (blue) and negatively (red) charged surfaces are displayed at the contour levels of +5 and −5 k_BT/e, respectively, where k_B is the Boltzmann constant, T is temperature, and e is the charge on an electron.

Structure of the C-terminal domain

Although the C-terminal domain in DHHC enzymes is more variable compared to the cysteine-rich domain, it contains sequence motifs that are conserved to varying extents throughout the family. The most highly conserved among them is the TTXE (Thr-Thr-X-Glu) motif that resides in the short helical segment (α′1) immediately following TM4, where X is any amino acid residue. It makes intimate contact with the DHHC domain (Figs. 3B and 4A and fig. S1). The Thr^{240} of the conserved TTXE motif caps a main-chain amide nitrogen, and Thr^{241} forms a hydrogen bond with the aspartic acid side chain of the catalytic DHHC motif (Figs. 3B and 4A and fig. S9B). Consequently, an AAXE (Ala-Ala-X-Glu)

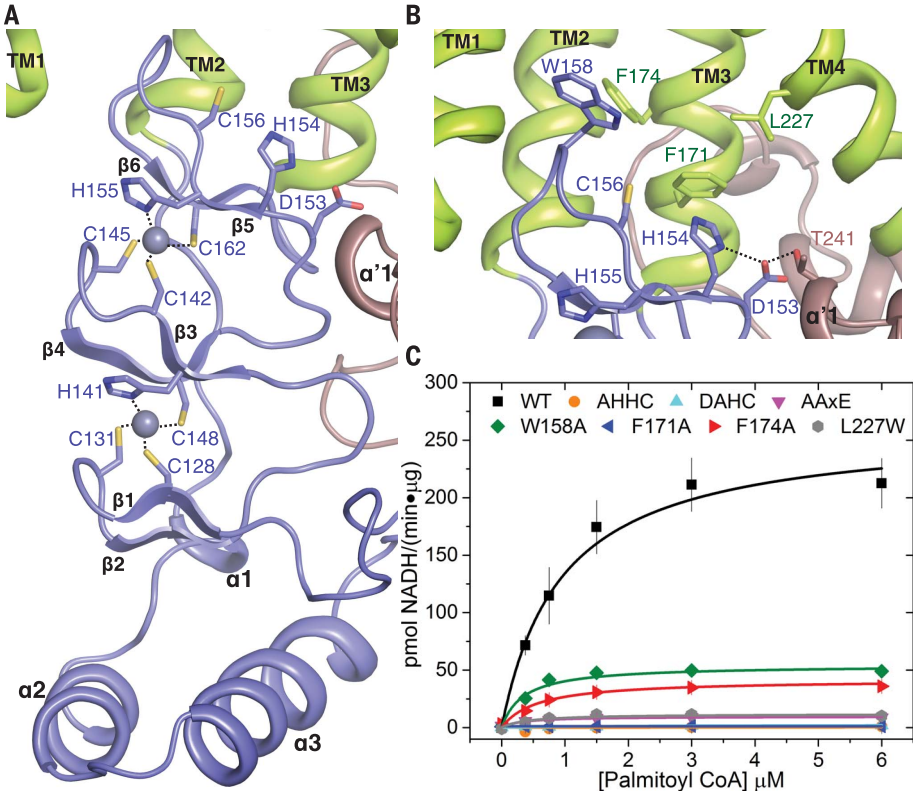


Fig. 3. Organization of the active site of DHHC20. (**A**) Close-up of the Zn^{2+} binding region showing coordination of each Zn^{2+} ion by three cysteines and one histidine in a CCHC configuration. (**B**) Close-up of the DHHC enzyme active site showing the catalytic Cys^{156} pointing upward toward a hydrophobic groove. Also shown are the aspartic acid and the first histidine of the DHHC motif and the Thr^{241} of the TTXE motif. Hydrogen-bonding interactions are shown with dotted lines. Trp^{158}, Phe^{171}, and Phe^{174}, which form the base of the acyl-binding cavity in Fig. 2D, are also shown in stick rendition. (**C**) Analysis of the enzymatic activity of selected mutants of the active site residues shown in (B). The coupled-enzyme assay was used, and Michaelis-Menten fits are shown. AHHC, DAHC, and F171A mutant curves all overlay with essentially no enzyme activity. Data are mean ± SEM of two independent measurements. Single-letter abbreviations for 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.

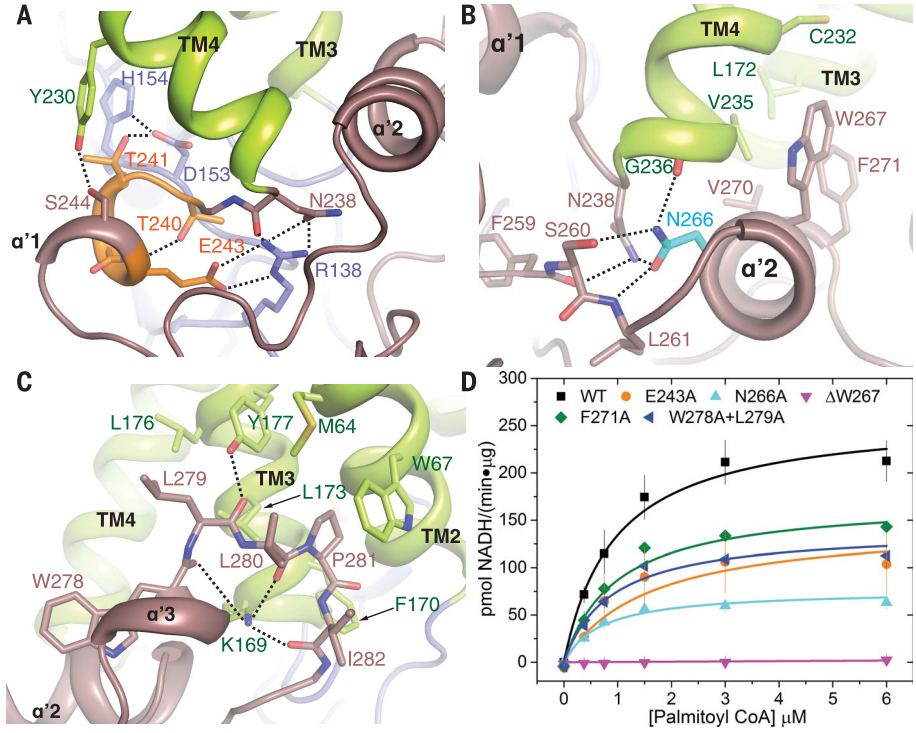


Fig. 4. C-terminal domain interactions. (**A**) Interactions of the Thr^{240}, Thr^{241}, and Glu^{243} of the conserved TTXE motif with the active site and the rest of the protein. (**B**) Interactions of the amphipathic helix of the PaCCT domain with TM3 and TM4. Asn^{266} is one of the most conserved residues on the C-terminal domain of DHHC enzymes. (**C**) Penetration of the C-terminal short helical stretch α′3 together with the hydrophobic loop, shown in brown, into the bilayer, and their interactions with the cytoplasmic ends of TM2 and TM3. Relevant residues are shown in stick rendition. (**D**) Mutational analysis of selected residues on the C terminus involved in interactions shown in (A), (B), and (C). Data are mean ± SEM of two independent measurements.

mutant has drastically reduced enzymatic activity (Fig. 3C).

Montoro *et al.* (*41*) have identified a motif at the C terminus of DHHC enzymes that they have named the palmitoyltransferase conserved C-terminus (PaCCT) motif. This motif is conserved in the majority of DHHC enzymes, and, in the 16-residue motif, the third and the eleventh residues are the most conserved. The third residue is usually a phenylalanine or tyrosine and, in hDHHC20, this residue, Phe^{259}, forms a local packing core at the C terminus and is engaged in a number of hydrophobic and π-stacking interactions. The other conserved residue in this motif, Asn^{266}, is at the beginning of the amphipathic helix (α′2) and simultaneously forms a capping interaction with TM4 and a hydrogen-bonding interaction with the

backbone of Leu[261] and the sidechain of Ser[260]. These interactions engage all of the hydrogen-bonding capabilities of the amide side chain and presumably hold this part of the structure together. Mutating Asn[266] to alanine severely compromises catalytic activity (Fig. 4, B and D, and table S3).

One of the most prominent features of the C-terminal domain is the presence of an amphipathic helix (α′2) that makes contact with TM3 and TM4 and likely provides local stability. Sequence conservation indicates that the amphipathic helix is conserved across many members of the family, pointing to its importance for other DHHC enzymes as well. Deletion of a single phenylalanine residue in this region in DHHC21 has been genetically mapped to a depilated phenotype in mice that results in hair loss with thinner and shorter hairs (*42*). The homologous residue in hDHHC20 is Trp[267]. Deletion of Trp[267] in hDHHC20 results in decreased yield of the recombinant protein and almost negligible catalytic activity (Fig. 4, B and D). The amphipathic helix is followed by a short helix (α′3) and a hydrophobic loop. The conformation of the loop is stabilized by highly conserved lysine and proline residues, and it inserts hydrophobic residues into the putative lipid bilayer and forms additional contacts with TM3 and TM2 (Fig. 4C). Mutation of the conserved Trp[278] and Leu[279] to a W278A/L279A double mutant reduces hDHHC20 activity by almost half (Fig. 4D and table S3).

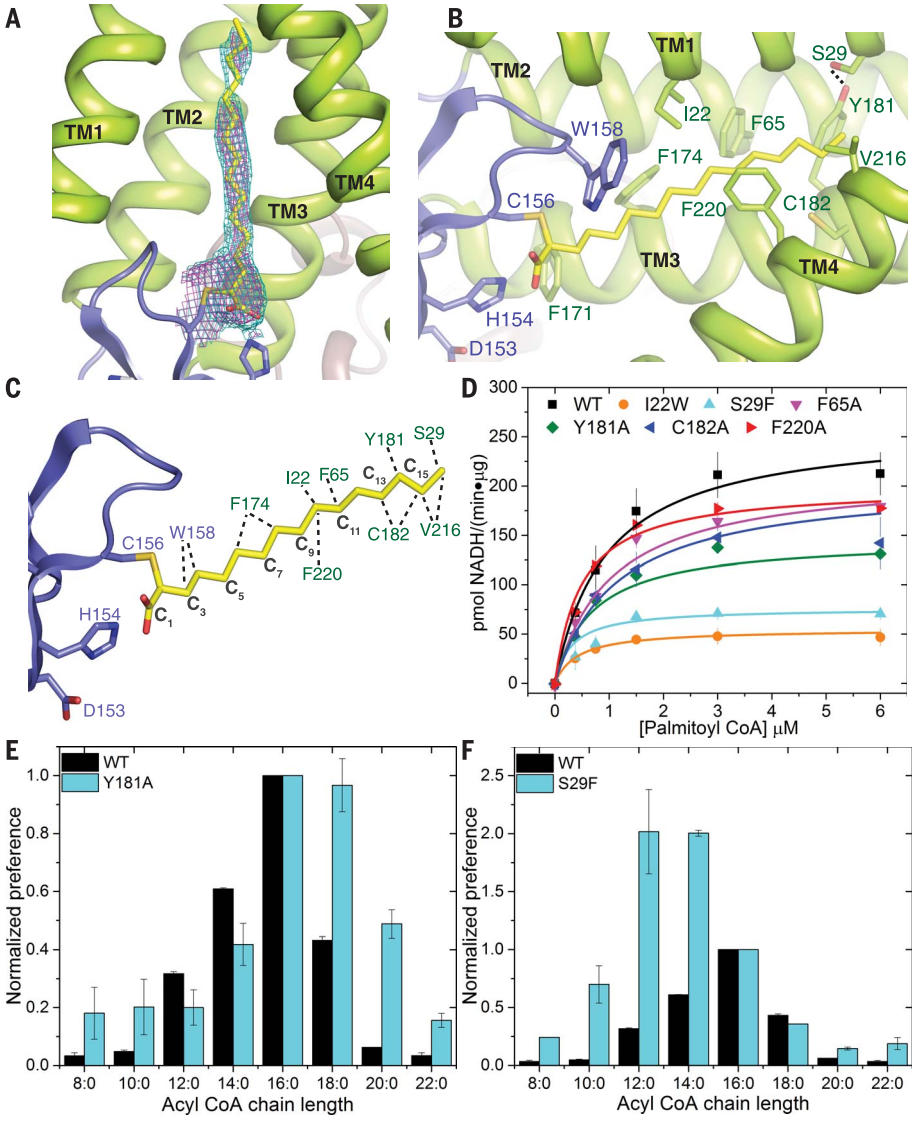


Fig. 5. The acyl chain binding groove. (**A**) $2F_o$-F_c electron density map at contour level 1.0 σ (magenta) and 0.7 σ (cyan) of the palmitate chain covalently attached to Cys[156] in the hDHHC20–2-BP structure (see text). (**B**) Close-up view of the acyl chain binding groove in the 2-BP structure showing the residues lining the groove. (**C**). Residues in the groove that interact with the palmitate chain identified by using the small-probe contact dot surfaces (*67*). (**D**) Mutational analysis of the residues in the acyl binding groove shown in (C). Data are mean ± SEM of two independent measurements. (**E**) Acyl-CoA chain–length selectivity in wild-type and Y181A mutant of hDHHC20, as determined by the autoacylation assay. The *x* axis shows the carbon-chain lengths of different acyl-CoA donors, and the *y* axis shows normalized activity (initial velocity) of wild-type or mutant versions of hDHHC20. Each data set is individually normalized to 1 for the activity with regard to palmitoyl-CoA. (**F**) Acyl chain–length selectivity of S29F hDHHC20. Selectivity data are mean ± SEM of two independent measurements.

Acyl chain binding site

Starting at the active site residues and pointing inward into the bilayer, both hDHHC20 and zfDHHC15 have a hydrophobic cavity (Fig. 2D). In the hDHHC20 structure, we saw weak density in the cavity that was suggestive of a fatty acid, but was not strong enough to build a model. We thus utilized the widely used covalent inhibitor 2-bromopalmitate (2-BP) to purify an acylated intermediate mimic. This inhibitor covalently modifies DHHC enzymes at the active site cysteine through a nucleophilic displacement reaction (*43*, *44*) (fig. S3B), resulting in alkylation of the cysteine through attachment at the α position of palmitic acid. Our initial biochemical and mass spectrometric data showed that, aside from the active site cysteine (Cys[156]), another cysteine (Cys[263]) was also consistently labeled with palmitate (fig. S11). Consequently, we used a C263A mutant (cysteine replaced with alanine at position 263) for preparation and crystallization of the covalently modified hDHHC20. We detected very low activity of the purified 2-BP–treated C263A mutant of hDHHC20, suggesting that the catalytic cysteine was mostly labeled in our preparation (fig. S3C). It crystallized under the same conditions and in the same hexagonal space group as hDHHC20 (table S1). The overall structures are very similar (Figs. 2A and 5A), and an omit map revealed clear density for the palmitate chain covalently attached to the active site cysteine, allowing us to build the complete acyl chain (Fig. 5A). In the zfDHHS15 structure, we saw continuous density in this cavity, which we modeled as a palmitic acid (fig. S12).

The structure of the 2-BP–treated C263A hDHHC20 mutant (hereafter referred to as 2-BP structure) reveals crucial insights into the mechanism of acyl-chain binding and recognition in hDHHC20. The acyl chain inserts into the hydrophobic cavity seen both in the hDHHC20 and zfDHHC15 structures (Fig. 2D and fig. S9) and is contacted by several residues lining the cavity (Fig. 5, B and C, and fig. S12, B and C). The acyl chains have very similar conformation in the cavity, and we will focus our description on the hDHHC20 2-BP structure. Notably, all four TM helices contribute to interactions with the acyl chain, albeit to different extents. The cavity is constricted near the acyl headgroup by Trp[158] and Phe[171]. These are two of the most highly conserved residues across the DHHC family members (fig. S1). Mutation of

Phe^{171} to alanine generates a catalytically inactive mutant, whereas mutation of Trp^{158} to alanine causes a considerable loss in enzymatic activity (Fig. 3C and table S3). Further into the membrane, two other residues contact the acyl chain, Phe^{174} and Leu^{227}, conserved as hydrophobic residues in most DHHC members. Mutation of the leucine to a more bulky and rigid tryptophan severely compromises enzymatic activity (Fig. 3C and table S3).

Toward the top of the cavity where it tapers to a narrow end, residues from all four TM helices contact the acyl chain. Some of these residues that line the acyl binding cavity are highly conserved across all DHHC members, e.g., Trp^{28} in hDHHC20. Intriguingly, there are other residues that show only subgroup-specific conservation (fig. S1). As a case study, Ile^{22} is conserved among a subset of DHHC members and faces the acyl binding cavity in the structure, but does not make a close contact (Fig. 5B). The I22W mutant (isoleucine replaced with tryptophan at position 22) has drastically lower activity, presumably owing to the more bulky and rigid aromatic side chain of tryptophan, which makes a disfavorable contact with the acyl chain (Fig. 5D).

Acyl chain–length selectivity of DHHC enzymes and designing mutants with altered preferences

It has been shown that different DHHC members show different degrees of selectivity for acyl-CoA donors with different chain lengths (*28*, *30*). Toward the tapering end of the acyl binding cavity in the 2-BP structure, Tyr^{181} forms a hydrogen-bonding interaction with Ser^{29}, effectively closing off the cavity (Fig. 5B). Whereas wild-type hDHHC20 shows a preference for acyl-CoAs with a palmitoyl (16-carbon) chain, mutating Tyr^{181} to a less bulky alanine results in marked increase in preference for stearoyl (C18)–CoA (Fig. 5E). Mutating Ser^{29} to a bulkier phenylalanine increases the preference of hDHHC20 for short-chain acyl-CoA (Fig. 5F). The S29F (serine replaced with phenylalanine at position 29) mutation not only breaks the hydrogen-bonding interaction but also replaces the Tyr-Ser pair in hDHHC20 with a considerably bulkier Tyr-Phe pair in the cavity.

These results also lend insights into a recent report that investigated the acyl chain–length selectivities of two closely related DHHC enzymes, DHHC3 and DHHC7 (*30*). The homologous pair of Tyr^{181}-Ser^{29} in hDHHC20 is Ile^{182}-Phe^{53} in DHHC3 and Ser^{185}-Leu^{56} in DHHC7. DHHC3 has a higher selectivity for palmitoyl (C16) over stearoyl (C18) compared to DHHC7. Presumably this is due to the placement of two large residues (isoleucine and phenylalanine) toward the end of the cavity, which impedes longer–chain length acyl-CoA. In comparison, the less sterically demanding Ser-Leu pair leads to a more relaxed acyl chain–length preference in DHHC7. Consistent with this hypothesis, the I182S (isoleucine replaced with serine at position 182) mutant of DHHC3 has higher preference for stearoyl (C18)–CoA, mirroring that observed in DHHC7 (*30*).

Mechanism of palmitoylation

Although protein palmitoylation was discovered almost 40 years ago (*5*), there has been little understanding about the structural chemistry of enzyme-catalyzed protein palmitoylation. All of the DHHC enzymes that have been biochemically characterized to date utilize a two-step reaction mechanism (*27*, *28*), with the first step being autoacylation of the active site cysteine. We show that the aspartic acid and His^{154} of the DHHC motif form a hydrogen-bonded pair that can accept a proton from the cysteine within the DHHC motif during catalysis, enabling nucleophilic attack on the carbonyl thioester of palmitoyl-CoA to generate the acyl-enzyme intermediate (Fig. 6A). The structures reported here do not give indication as to how the palmitoyl-CoA enters the binding cavity. However, palmitoyl-CoA partitions into the lipid bilayer and likely distorts the local structure of the membrane (*45*, *46*). Given that interfaces of membrane proteins with surrounding lipids are also sites of distortion in the lipid bilayer and, thus, possible sites of accumulation of palmitoyl-CoA, we speculate that palmitoyl-CoA inserts into the cavity formed by DHHC enzymes from within the membrane.

In the second step of the reaction, the acyl-enzyme intermediate transfers the palmitoyl group to the substrate cysteine. The structure of the 2-BP–treated DHHC20 suggests that the carbonyl oxygen of the acyl-enzyme thioester is positioned close to His^{154}. We speculate that the protonated His^{154} activates the acyl-enzyme thioester for the subsequent palmitoyl transfer to the substrate by providing a proton to the carbonyl oxygen. The cysteine residue in each substrate that accepts the palmitoyl group lies in a different sequence context and, thus, different chemical and structural microenvironments. Consequently, activating the acyl group of the acyl-enzyme intermediate with an adjacent proton donor would be an effective strategy for catalysis of the second step. Residue His^{154}, which accepts a proton in the first step, is optimally positioned for this role.

The structure also shows that the organization of the active site is such that all but one side of the acyl-enzyme thioester is shielded by hydrophobic residues. This only leaves the front side of the acylated DHHC for approach of the cysteine residue on the substrate that reacts with the thioester in the second step (Fig. 6B). However, engagement with the substrate may incur conformational changes and open up alternative directions of approach that are not obvious in the current set of structures.

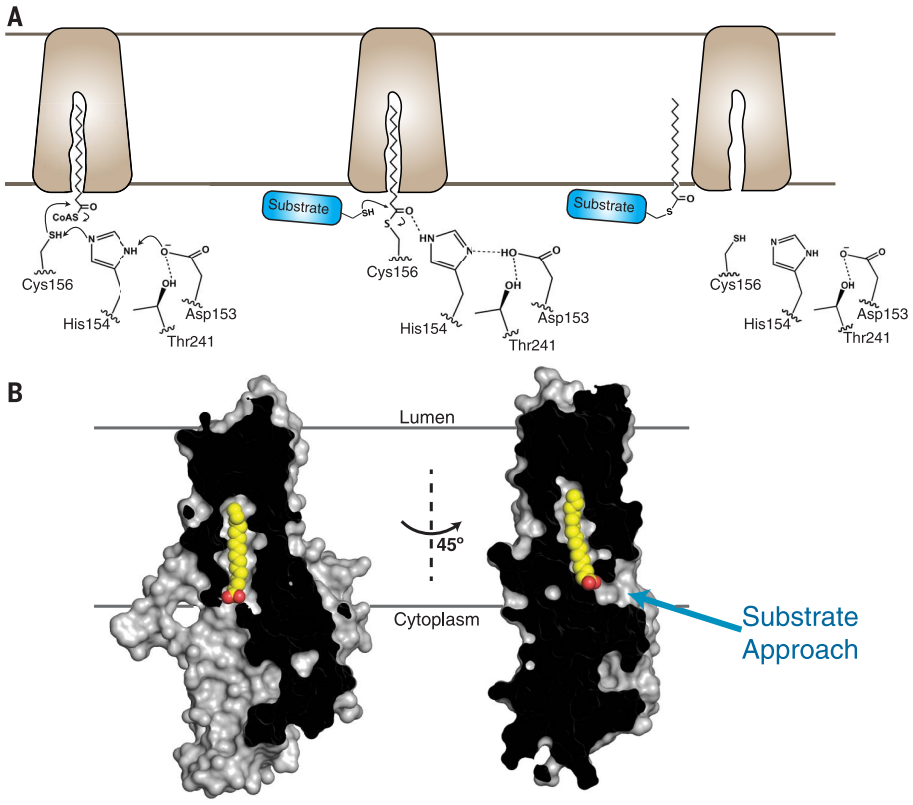


Fig. 6. Proposed reaction mechanism and substrate approach. (**A**) Proposed reaction mechanism of DHHC enzymes that follow a two-step mechanism. (**B**) Molecular surface rendition of the 2-BP–treated enzyme with the acyl chain shown in yellow and red spheres. The putative direction of substrate approach is shown with a cyan arrow.

Protein S-acylation has been shown to be heterogeneous. Analyses of S-acylated proteins from native sources have revealed that proteins can be modified not only with palmitic acid but also with other fatty acids, such as stearic acid and oleic acid (*10*, *11*, *47*). In the case of influenza hemagglutinin, specific cysteines have been shown to be modified by palmitoylation and stearoylation, pointing to the specific roles that distinct acylation states play in the function of cellular proteins (*48*). Recently, it has been shown that stearoylation of transferrin receptor 1 by DHHC6, specifically, modulates mitochondrial morphology (*49*). In vitro experiments have shown that for the two enzymes DHHC2 and DHHC3, DHHC2 can utilize a broad range of fatty acids with diverse structures such as palmitate, arachidonate, and palmitoleate, whereas DHHC3 has reduced activity for fatty acids longer than palmitic acid (*28*). It was also shown that the selectivity at the autoacylation state reflects the selectivity of fatty acyl labeling of the protein substrate (*28*). The structures of unacylated zfDHHS15 and hDHHC20 and of 2-BP hDHHC20 reveal that the fatty acyl chain fits into a cavity in the bilayer formed by the transmembrane helices of the enzyme. The residues lining this cavity form numerous contacts with the fatty acyl chain. However, although there are highly conserved residues, i.e., Trp^{158} and Phe^{171}, lining this cavity that form important contacts with the acyl chain, there is considerable sequence diversity in the other residues that line this cavity between different subgroups of DHHC enzymes (fig. S1). These residues vary not only in their size, but also in their polarity, thus changing the chemical properties of the cavity in a complex manner. These can have a notable effect on the fatty acyl chain–length preference of different DHHC enzymes, and, as structures of more DHHC enzymes become available, presumably they will reveal further insights into their respective fatty acid chain–length selectivity.

DHHC20 overexpression was previously shown to cause cellular transformation (*50*). More recently, hDHHC20 has been proposed as a target for developing therapeutics against cancers that are resistant to EGFR-targeted therapy (*22*). The structures presented here are plausible starting points for developing a structure-guided program to hDHHC20 inhibitors that could lead to such a therapeutic intervention. It is worthwhile to note in this context that the entire field of protein palmitoylation suffers from the lack of small molecule probes and inhibitors for the DHHC family of enzymes, let alone the lack of specific modulators of individual DHHC members (*51*). Members of the DHHC family all use palmitoyl (acyl)–CoA and the same enzymatic chemistry to transfer an acyl group to substrate proteins. From this perspective, they are very similar to members of the family of protein kinases that use ATP to transfer a phosphate group to substrate proteins. Thus, using strategies like bump-hole pairs, where structural knowledge is used to design engineered versions of DHHC enzymes that can use orthogonal acyl-CoAs, could provide transformative insights into the biology of protein palmitoylation and help unravel the complex network of DHHC-substrate interactions, akin to the field of protein kinases and histone-modification enzymes (*52*–*54*). We demonstrate the feasibility of such an approach by designing mutants that can shift the acyl chain–length selectivity profile of the mutants toward the shorter or longer side of palmitoyl (C16)–CoA.

Materials and Methods

Molecular biology and cloning and yeast transformation

The codon optimized zebrafish DHHC15 (zfDHHC15) sequence was cloned into a modified version of pPICZ-C vector with an His_{10} tag followed by a GFP coding sequence and finally, by PreScission cleavage site at the N terminus of the zfDHHC15 encoding DNA sequence. Further, site directed mutagenesis was performed to mutate Cys^{153} to serine to get the zfDHHS15 expression construct. The optimized vector harboring zfDHHC15 or zfDHHS15 gene was transformed into *Pichia pastoris* HIS $^+$ Cells SDM1163 cells. *Pichia* were transformed using standard methods and the transformants were selected on YPDS plates containing 400-800 μg/mL zeocin.

The human DHHC20 (hDHHC20) expression construct was made similarly with N-terminal His_{10} tag, GFP and PreScission cleavage sequences without any modification/truncation of the hDHHC20 sequence and was transformed into *Pichia* following the same protocol.

A vector containing the sequence for human Snap25b was obtained from Addgene (Plasmid #53235) and the gene was cloned into pET28-Prx by digestion/ligation using NdeI/BamHI. The original TEV cleavage site was mutated to a PreScission cleavage site.

Protein expression and purification

Cell culture and lysis

Large scale culture of *Pichia*, induction of protein expression and cell lysis were carried out as described previously (*55*). Briefly, 100-200 mL BMGY (0.1 M Potassium phosphate, pH 6.0, 3.4 g/L yeast nitrogen base, 1% glycerol, 0.4 μg/mL biotin) with 500 μg/mL Zeocin cultures of *Pichia* were grown overnight at 30°C with vigorous shaking. $1/10^{th}$ volume of these starter cultures, which usually reached cell densities of OD_{600} ~20, were then used to inoculate 1.5-2 liter cultures in the same media but without the Zeocin. In 24-36 hours, the cells were pelleted by centrifugation at 1500 g, 4°C. for 10 min and thoroughly washed with BMMY (0.1 M Potassium phosphate, pH 6.0, 3.4 g/L yeast nitrogen base, 1% methanol, 0.4 μg/mL biotin). Cells were finally re-suspended in 1.5-2 liters of BMMY media and protein expression induced at 22-24°C for 24-36 hours. Cells were harvested by centrifugation at 6000 g, 4°C. for 20 min. The cell pellets were scooped by a spatula and frozen by immersing into liquid N_2 followed by storage at –80°C. Cells were lysed by cryo-milling using Retsch MM400 millers with liquid N_2 for cooling.

Purification of zfDHHC15 and zfDHHS15

Lysed and frozen yeast powder was dissolved in the lysis buffer (1 g cells per 3 mL of lysis buffer) containing 50 mM Tris-HCl pH-7.5, 150 mM NaCl, 1 mM TCEP, 0.1 mg/mL deoxyribonuclease I, 0.1 mg/mL pepstatin, 1 μg/mL leupeptin, 1 μg/mL aprotinin, 1 mM benzamidine and 0.1 mg/mL soy trypsin inhibitor. The pH of the cell lysate was adjusted to 7.8 followed by addition of the 0.14 g of DDM (n-dodecyl-β-D-maltopyranoside) per g of cells. The proteins were extracted from the cells by stirring at 4°C. for 3 hours. The pH of the cell lysate was adjusted to 7.5 and centrifuged at 38,000 g, 4°C. for 30 min. The supernatant was allowed to bind with 2.5 mL of Talon (Clontech) resin pre-equilibrated with the equilibration buffer (50 mM Tris-Cl pH 7.5, 150 mM NaCl, 1 mM TCEP and 1 mM DDM) for 3 hours at 4°C. The protein bound Talon resin was collected in a Bio-Rad Econo column and washed with 10 bed volumes of wash buffer 1 (20 mM Tris-Cl pH 7.5, 150 mM NaCl, 1 mM TCEP and 2 mM of Thio-DM (n-Decyl- β-D-Thiomaltopyranoside) followed by additional wash with 10 bed volumes of wash buffer 2 (wash buffer 1+ 20 mM imidazole). An additional wash was performed with 10 bed volumes of wash buffer 1 in order to completely wash away any traces of imidazole. Talon resin was resuspended in 10 mL of wash buffer 1. His_{10} tagged GFP was removed from the N terminus of zfDHHS15 by incubation of the resin slurry with PreScission protease overnight at 4°C. Flow-through containing cleaved protein was concentrated with 50 kDa molecular weight cut-off (MWCO) 15 mL concentrators (Millipore). The concentrated protein sample was further purified by size exclusion chromatography on Superdex 200 column in buffer containing 20 mM Tris-Cl pH 7.5, 150 mM NaCl, 1 mM TCEP and 2 mM Thio-DM. The peak fractions from the size exclusion chromatogram were collected and concentrated with 50 kDa MWCO concentrator up to 20 mg/mL and used for crystallization.

Purification of hDHHC20 and hDHHS20

20-24 g of milled cell powder was suspended in ~65 mL of Tris-buffered saline and stirred using a magnetic stirrer for 20-30 min to make a homogeneous slurry. The final buffer composition was 40 mM Tris.HCl, 270mM NaCl, 5 mM βME. Protease inhibitors (Benzamidine HCl, PMSF, AEBSF, Aprotinin, Pepstatin, Leupeptin) and DNase were added and the pH of the cell slurry adjusted to ~7.5. Solid DDM was added to give a final concentration of 2% (40mM) and the protein extracted from the membranes by vigorous stirring at 4°C for 3 hours. Cell debris was pelleted by centrifugation at 38,000 g for 30 min at 4°C. The pH of the supernatant was adjusted to 7.5 and then incubated with 2 mL of Talon resin for 2-3 hours at 4°C.

Protein-bound resin was washed with 30 mL of 50 mM Hepes, pH 7.5, 250 mM NaCl, 5 mM

βME, 2 mM TCEP, 1 mM DDM buffer containing 5 mM imidazole. The resin was further washed with same volume of buffer containing 25 mM imidazole prior to elution with ~5-7 mL of buffer with 200 mM imidazole. All affinity chromatography was done by gravity flow at ambient temperature with buffers chilled on ice. The bright green eluate was then concentrated in a 100 kDa molecular weight cut-off centrifugal concentrator to ~0.3 mL. 1/10 volume PreScission protease (~1 mg/mL) was then added and rotated overnight at 4°C. The PreScission protease cleaved protein was injected into a Superdex 200 Increase size exclusion column at 4°C. to separate hDHHC20 from the GFP tag and the PreScission protease. The size exclusion buffer was 50 mM Hepes pH 7.5, 150 mM NaCl, 2 mM TCEP, 0.5 mM DDM. The fractions containing the protein were pooled and concentrated using a 50 kDa MWCO 0.5 mL Amicon centrifugal concentrator. Prior to any assays, protein concentration was determined using the 660 nm protein assay kit (ThermoFisher). In cases where the GFP tag was not removed, the overnight PreScission cleavage step was omitted and the concentrated protein applied to the Superdex 200 increase column (GE Healthcare).

Reductive methylation of hDHHC20 for crystallization

hDHHC20 for crystallization was prepared as above to the step of overnight PreScission digest. The next morning, instead of applying the GFP-cleaved protein to the gel filtration column, it was diluted to ~12 mL in 50 mM Hepes, pH 7.5, 250 mM NaCl, 2 mM TCEP, 0.5 mM DDM buffer lacking any imidazole. The GFP tag was then removed by binding it to ~1 mL Talon resin and the flowthrough containing the hDHHC20 protein was concentrated in 50 kDa MWCO 4 mL concentrator to ~0.5 mL. This protein was again diluted to ~12 mL and re-concentrated to ~1 mL to remove any imidazole. The protein was then reductively methylated using formaldehyde and dimethylamineborane (DMAB) complex following standard protocols (*56*). The next morning, the methylation reaction was quenched with 100 μl of 1 M Tris HCl, pH 8.0 and 5 μl of 1 M dithiothreitol (DTT) for 1 hour at 4°C. The protein was concentrated to ~0.25 mL and applied to Superdex 200 Increase size-exclusion column as previously mentioned. Peak fractions were collected and concentrated as before. All wash buffers for the crystallization sample of DHHC20 contained 0.1 mg/mL POPC:POPG:POPA (3:1:1) lipids and 1 mM DDM. This was lowered in the size exclusion chromatography buffer to 0.05 mg/mL lipids and 0.5 mM DDM.

Irreversibly inhibited hDHHC20 preparation and purification

To prepare an irreversibly inhibited protein with a covalently attached lipid chain, 2-bromopalmitate (2-BP) dissolved in methanol was added to the *Pichia* cells at a final concentration of 300 μM during induction. However, we decided to add the DHHC inhibitor only after the cells had been induced for at least 12 hours as 2-BP is toxic to cells and we wanted to ensure that it did not affect protein expression initially. Five hours prior to harvest, we again added 2-BP to a final concentration of ~100 μM. We first attempted this with the wild-type hDHHC20 and checked the covalent modification with mass spectrometry. Unexpectedly we found that apart from the active site cysteine, 2-bromopalmitate also modifies Cys^{263}. Therefore, to reduce any possible heterogeneity during crystallization, we used a C263A mutant for the 2-BP modified DHHC20. hDHHC20 C263A protein was prepared as for wild-type hDHHC20. Enzymatic activity of the 2-BP treated sample was checked and found to be essentially nonexistent as expected.

Purification of GobX protein

GobX protein expression vectors were a gift from M. Machner (NIH). BL21(DE3) Gold cells were transformed and plated on LB-agar with Ampicillin. Next morning the entire plate was used to start two 1 liter cultures in LB media. Bacteria were grown to an OD_{600} of ~0.6 at 37°C. with vigorous shaking. Temperature was reduced to 30°C. and protein expression was induced with a final IPTG concentration of 1mM. Cells were harvested ~5 hours later. Protease inhibitors (PMST, AEBSF) were added to the cells prior to flash freezing in liquid N_2. Next morning the cells were thawed in 40 mM Tris-HCl, pH 7.2, 270 mM NaCl, 10 mM βME, 20% glycerol, DNAse, and more protease inhibitors. Powder lysozyme was added to facilitate cell lysis. Once the cells had thawed to slurry, the cell suspension was diluted such that it now contained 10% glycerol. DDM was then added to 1% final concentration.

Cells were broken by sonication on ice. During the sonication process, 1/100 volume PMSF was added after every 1 min. The cell debris were removed by centrifugation at 38,000 g for 30 min at 4°C. The supernatant was incubated with 2 mL of HisPur Ni-NTA resin for 1-2 hours at 4°C rotator. Protein bound resin was packed on a Bio-Rad Econo column and washed with ~30 mL Tris-buffered saline containing 10% glycerol and 10 mM imidazole. The resin was further washed with same buffer containing 50 mM imidazole before finally eluting it with ~8 mL 300 mM imidazole buffer. The eluate was concentrated in 10 kDa MWCO concentrator to ~0.5 mL and diluted 100-fold to 50 mL in 50 mM TrisHCl, pH 8.0, 10% glycerol, 1 mM DDM buffer. The diluted protein was applied to a 1 mL HiTrap MonoQ column using peristaltic pump. GobX protein was eluted off the MonoQ with a two-gradient program between 50 mM TrisHCl, pH 8.0, 10% glycerol, 10 mM NaCl and same buffer with 1 M NaCl on an AKTA explorer at 4°C. DDM is omitted from the MonoQ column elution buffers. Two major peaks are eluted – the first is a cleaved version of the GobX and the second is the full-length GobX. The second peak fractions were pooled, concentrated in a 10 kDa MWCO Amicon spin concentrator to ~5-7 mg/mL and flash frozen in liquid N_2 until later use in palmitoylation assays.

Purification of Snap25b

E. coli BL21(DE3)-Gold was transformed with Snap25b in pET28-Prx vector and plated onto LB-Kanamycin plates. A single colony was used to inoculate 5mL of LB (with 50 μg/mL of Kanamycin) and cells were grown at 37°C for 20 hours. 2 mL of this starter culture was used to inoculate 1 L of LB (with 50 μg/mL of Kanamycin) and cells were grown until OD_{600} was around 1.0. Cells were cooled to 30°C. and induced with 0.5 mM of IPTG. After 5 hours, cells were harvested by centrifugation at 8000 g for 10 min at 4°C. Cell pellets were flash frozen in liquid nitrogen and stored at –80°C.

For a typical purification, 10 g of frozen pellet were dissolved by stirring at 4°C (30-40 min) in 100mL of lysis buffer containing: 50 mM Tris-HCl pH 8.0, 250 mM NaCl, 5% glycerol, 5 mM βME, protease inhibitor cocktail (0.1 mg/mL deoxyribonuclease I, 0.1 mg/mL pepstatin, 1 μg/mL leupeptin, 1 μg/mL aprotinin, 1 mM benzamidine and 0.1 mg/mL soy trypsin inhibitor and 1 mM PMSF) and 10 mg lysozyme. Resuspended cells were disrupted by sonication, stirred for 20 min at 4°C and subsequently centrifuged at 38,000 g for 30 min at 4°C. Pellets were discarded and supernatant was batch-bound to 5 mL of Ni-NTA resin (previously equilibrated with binding buffer containing 50 mM Tris-HCl pH 8.0, 250 mM NaCl, 5% glycerol, 5 mM β-mercaptoethanol) by rotation at 4°C. for 1 hour. Protein-bound resin was washed successively with 10 bed volumes of binding buffer, 10 bed volumes of binding buffer with 10 mM imidazole and finally, 10 bed volumes of binding buffer with 30 mM imidazole. Protein was finally eluted with binding buffer containing 250 mM imidazole. Eluted protein was incubated at 4°C with rotation for 20 hours in the presence of 1 mg/mL of PreScission protease to remove the His_6 tag. Completeness of cleavage was assessed by gel shift using 12% acrylamide SDS-PAGE after Coomassie staining. Completed cleavage reaction was dialyzed against 4 liters of binding buffer for 15 hours at 4°C to remove imidazole. The dialyzed protein solution was re-applied to 5 mL of Ni-NTA resin equilibrated with binding buffer and incubated with rotation at 4°C for 1 hour. Resin was loaded onto a column and flow-through was collected. The concentration of purified protein was determined using the 660nm Assay Kit (Thermo Fisher).

Protein crystallization, optimization, and data collection

hDHHC20 crystallization and optimization

hDHHC20 protein at ~15-20 mg/mL was incorporated into monoolein using established protocols (*57*) and crystallization trials were set up on Laminex 96-well plastic sandwich plates (Molecular Dimensions) using a Mosquito crystallization robot (TTP Labtech). An initial hit was obtained with 0.1 M Hepes, pH 7.0 and 30% PEG 300 at 20°C. Optimization of crystals was

carried out using varying salts, buffers and pH. Of the salts, only sodium malonate and potassium di-hydrogen phosphate at 50 mM gave crystals that diffracted consistently. Further improvements in size of crystals were achieved through screening diol additives with 1-5% of 2,5-Hexanediol being the best. Better diffraction was observed when fresh DTT was included as an additive at concentrations of 25-50 mM. Addition of lipids during the protein preparation improved the crystallizability of DHHC20. While we had initially observed rectangular crystals (indexed as $P2_1$), inclusion of lipids during purification led to us to obtain hexagonal crystals (indexed as $P6_3$) as well. Further improvements in diffraction were also achieved by methylating the protein. The final optimized crystallization screen was a pH and PEG 300 grid with MES (pH 5.5-6.5), MOPS (pH 6.5-7.5), and HEPES (pH 6.5-7.5) buffers and PEG 300 from ~25-35%. We also expanded the salt from potassium di-hydrogen phosphate to include the sodium salt as well, both at 50 mM final concentration.

As diffraction from crystals obtained in the sandwich plates hardly extended to better than 4 Å and the harvesting itself was challenging, we decided to switch to a microbatch format. In this format, 100 nl of the LCP bolus was set on 96-well Microbatch plates (Hampton Research) followed by 3.24 µl crystallization solution and 0.36 µl of a 25% 2,5-Hexanediol and 0.5 DTT (freshly prepared) mixture. Plates were incubated at 20°C. The crystals generally took longer to appear than sandwich plates (a week versus a few days) and also grew larger and thicker. Crystals were directly fished out of the wells using Mitegen 100 micron loops and frozen in liquid N_2.

The initial hexagonal crystals were obtained in 100 mM HEPES, pH 7.3, 50 mM NaH_2PO_4, 25 mM DTT, 1.25% 2,5-hexanediol after approximately one week. Data sets were collected from hexagonal crystals at 9.665 keV at 10% transmission with 0.25° rotation and 0.2 s exposure. A total of 720 frames spanning 180° were collected. The best diffracting hexagonal crystals were obtained in 50 mM MES, pH 5.8, 50 mM KH_2PO_4, 27.7% PEG 300, 25 mM DTT, 1.25% 2,5-Hexanediol and grew to optimal size in 3.5 months. A native data set of 360 frames was collected at Se-edge of 12.66 keV with 10% transmission, 0.25° rotation and 0.5 s exposure. These data sets were collected at the NE-CAT 24ID-C beamline at the Advanced Photon Source (APS) at Argonne National Laboratory in Chicago, IL.

The best $P2_1$ crystal was obtained in 50 mM MES, pH 6.5, 50 mM NaH_2PO_4, 30.3% PEG 300, 50 mM DTT, 2.5% 2,5-hexanediol and grew to optimal size in 3 months. Complete data set was collected at 9.665 keV with 10% transmission, 0.3° rotation, and 0.2 s exposure. A total of 1500 frames were collected. The data set for the 2-BP modified hDHHC20 crystals was obtained from a single crystal obtained after 1.5 months in 100 mM HEPES, pH 7.03, 50 mM KH_2PO_4, 31% PEG 300, 50 mM DTT, and 2.5% 2,5-hexanediol. Data set of 900 frames was obtained with 20% transmission, 0.2° rotation, and 0.2 s exposure at a 12.66 keV beam. These data sets were collected at the GM/CA-CAT 23ID-D and 23ID-B beamlines respectively at the APS at Argonne National Laboratory in Chicago, IL.

hDHHC20 structure determination

Diffraction data from three different crystals of the $P6_3$ form were merged together to improve redundancy for the phasing. The data sets were processed using XDS and AIMLESS in the CCP4 suite (*58*). Initial phases were obtained using AutoSol in the PHENIX Suite (*59*). Four predicted zinc heavy atom sites were identified and provided a phasing power to generate a map clearly showing molecular boundaries and secondary structure elements. An initial model was built using AutoBuild and manually modified using COOT (*60*). The resulting structure was used as a search model for molecular replacement against the native data set. This structure was also used as a starting molecular replacement model for solution of the $P2_1$ form and the 2-BP treated hDHHC20. Restraints of the 2-BP and CoA head group were generated using the eLBOW module within the PHENIX suite and edited manually. Iterative model building and refinement was carried out using COOT and PHENIX. The coordinates and structure factors of the human hDHHC20 in $P6_3$ and $P2_1$ space groups and 2-BP modified structure have been deposited to the Protein Data Bank (PDB) with the accession code of 6BMN, 6BMM and 6BML, respectively.

DHHS15 crystallization and data collection

ZfDHHS15 protein crystallization was performed using the hanging drop vapor diffusion method. Initial crystal hits were obtained with 0.1 M KCl, 20% PEG 400 and 100 mM HEPES, pH 6.5 at 4°C. These crystals were further optimized to obtain better diffraction quality crystals. The best diffracting crystals of zfDHHS15 grew from 0.1 M KCl, 20% PEG 400, 100 mM HEPES pH 6.5 and 30% ethylene glycol at 4°C. The optimum size of the better diffracting crystals was obtained in 7 days. Crystals were exposed to the cryo-solution (30% ethylene glycol with all other remaining same as reservoir) by exchanging reservoir solution with cryo-solution. Ethylene glycol percentage was gradually increased from 20% to 30% in increments of 5%. Each step was incubated at 4°C for 6-7 hrs. Crystals were harvested using Hampton nylon loops and immediately flash-frozen in liquid nitrogen. Three different types of data were collected in order to solve the structure. Two different single-wavelength anomalous dispersion data sets at 3.31-Å resolution from single crystals were collected at the Zn absorption edge wavelength, 1.2820 Å, on beamline 24ID-C (NE-CAT) at the Advanced Photon Source (APS), Argonne National Laboratory. These data sets were merged together, integrated with XDS and scaled with XSCALE. Additionally, S-SAD data were collected at 1.71 Å from 10 different single crystals at 23ID-B (GM-CA) at APS. These data sets were merged together and integrated with XDS and scaled with XSCALE. A third native data set of 2.54-Å resolution was acquired at x-ray wavelength 0.997 Å using single crystals on beamline 22ID (SER-CAT) at the APS. The native data set was processed with HKL 2000.

DHHS15 structure determination

The structure was solved using the Zn-SAD data in $P4_22_12$ space group using the data collected at 24ID-C beamline, integrated using XDS. Data from two crystals were used. They were scaled using XSCALE. Four Zn atoms were located using hkl2map and subsequently refined using experimental phasing methods in Phaser. An electron density map was generated after NCS averaging and density modification (using Parrot, part of the CCP4 suite of crystallographic software) and an overall model was built into it. At this stage, a highly redundant S-SAD data was collected at 23ID-GM-CA (wavelength 1.71 Å, data redundancy ~100) using 10 crystals. Using the model phases, sulfur atoms were located using anomalous difference Fourier methods and the sulfur sites were refined using Phaser. A total of 48 sulfur sites were found (24 per monomer, 9 Met and 15 Cys). This density modified map was used for iterative model building and refinement in PHENIX. A partial model of the zfDHHS15 was obtained with R_{work} of 32.2 and R_{free} of 33.7%. This partial model was used for molecular replacement with native data set of 2.45 Å resolution with Phaser module in PHENIX. Non-crystallographic symmetry was used during refinement of the model. The native data set was further refined using PHENIX as well as Refmac in ccp4 followed by manual examination and rebuilding of the refined coordinates in coot using both $2mF_o - DF_c$ and $mF_o - DF_c$ maps, as well as omit maps generated through PHENIX. The restraint parameters for palmitate, dodecylmaltoside and POPC were generated through ELBOW program of PHENIX was used during refinement. The final data collection and refinement statistics are shown in Table S2. The final model had the following Ramachandran statistics, 97.8% in favored region, 2.17% in allowed region, 0.2% of rotamer outliers and 0% of Ramachandran outliers.

Biochemical assays

zfDHHC15

Coupled-enzyme assay

zfDHHC15 autoacylation activity was measured using a coupled-enzyme assay (*31*) in 384-well low-volume plates (Thermo Fisher). Plates were read in a Tecan M1000Pro fluorimeter with 340 nm excitation and 465 nm emission. The final reaction volume (20 µL) contained 20 nM zfDHHC15/zfDHHS15, 0.25 mM oxidized Nicotinamide adenine dinucleotide (NAD^+), 0.2 mM thiamine pyrophosphate (TPP), 2 mM 2-oxoglutarate, 1 mM EDTA, 1 mM DTT, 2-oxoglutarate dehydrogenase, 0.6 mM DDM detergent in 50 mM sodium phosphate buffer at pH 6.8. The reaction was initiated by the addition of palmitoyl-CoA and monitored for 30 min at 30°C. The linear part of the progress curve was used to determine initial rates.

zfDHHC15 concentration was determined using the 660 nm assay (Thermo Fisher). K_m and V_{max} were estimated using non-linear least square fitting to the Michaelis-Menten model (Graphpad Prism 6). K_{cat} was determined by dividing the estimated V_{max} by the total enzyme concentration. Each K_m and k_{cat} value is the average and standard error of three independent measurements.

Snap25 palmitate transfer assay

For in vitro Snap25b acylation assay, affinity purified GFP-zfDHHC15 (50 nM) and GFP-zfDHHS15 (50 nM) were incubated with affinity purified Snap25b (500 nM) in reaction buffer (50 mM HEPES, pH 7.0, 200 mM NaCl, 0.2 μM DTT, 1 mM EDTA, 0.3 mM DDM) for 10 min at room temperature (24°C). Reaction was started by addition of {N-[(7-nitro-2-1,3-benzoxadiazol-4-yl)-methyl]amino}palmitoyl-CoA (NBD PalCoA) to a final concentration of 1 μM. NBD PalCoA is a fluorescent derivative of palmitoyl-CoA. Samples were taken at different time points and reaction quenched with non-reducing SDS sample buffer. Samples were separated in 12% SDS-PAGE gels and fluorescence detected using a Chemidoc with Blue Epi illumination and a 530/25 filter.

hDHHC20 Coupled-enzyme assay

hDHHC20 kinetic assays were done using an established fluorescence based coupled-enzyme assay (*32*) in 96-well format at 30°C. Plates were read in a Tecan M1000Pro plate reader with 340 nm excitation and 465 nm emission at 30 s intervals. All assays were done in 25 mM MES, pH 6.5, 50 mM NaCl, 1 mM DTT, 1 mM EDTA, 0.3 mM DDM. Final NAD^+, TPP, and 2-oxoglutarate concentrations were 0.25 mM, 0.2 mM, and 2 mM respectively. Although the original assay used αKDH purchased from Sigma, we were dissatisfied with the activity and availability of the enzyme. Therefore, we prepared the αKDH in-house from beef heart using published protocol (*61*) αKDH quantity required for the assay was determined empirically by titration until the slope of the reaction progression curve did not change further. Palmitoyl-CoA and other acyl-CoA were usually kept below 20 μM as at higher concentrations we observed reduced enzyme activity. DHHC20 enzyme concentration in the assay to determine kinetic parameters was usually 2 nM to 5 nM. At higher DHHC20 concentrations it was difficult to get enough time points for low palmitoyl-CoA concentrations to determine the initial velocity. K_m and k_{cat} were determined using Michaelis-Menten equation using non-linear least squares fit in OriginLab. Curve-fitting was weighted to the error and was iteratively done to minimize reduced chi-square. Adjusted r-square values for most fits are >0.9 except for the inactive mutants where the data are not described by the Michaelis-Menten kinetics.

For assays where initial slopes at low palmitoyl-CoA concentrations were not necessary or a stronger signal was more important, DHHC20 enzyme concentrations were increased to 10-100 nM. In such cases the acyl-CoA concentrations were usually fixed at 5 or 10 μM.

Protein substrate palmitoylation assay

GobX protein palmitoylation assays were carried out in 96-well plates at ambient temperature (~24°C). Reaction condition was 50 mM HEPES, pH 7.0, 200 mM NaCl, 0.2 μM DTT, 1 mM EDTA, 0.3 mM DDM. eGPF-DHHC20 and GobX concentrations were 0.2 and ~30 μM respectively (we used the eGFP-tagged DHHC20 because the DHHC20 and GobX protein overlap in the gel). After mixing the enzyme and substrate, trans-palmitoylation reaction was initiated by the addition of the fluorescent NBD–palmitoyl-CoA to a final concentration of 1 μM. 10 μl of reaction mixture was pipetted out at different time points and the reaction quenched by adding it to 2x non-reducing SDS-sample buffer. Quenched reactions were run on a SDS-PAGE and the NBD fluorescence imaged on a gel imager (Bio-Rad). Protein signal was obtained from Coomassie staining or the use of a stain-free gel (Bio-Rad).

DHHC20 mutant analyses

The Venus-hDHHC20 construct was inserted in the multiple cloning site of pEG-Bacmam vector (*62*) under the control of the CMV promoter. All mutagenesis was carried out using the Quikchange strategy. Venus-hDHHC20 wild-type and mutants were expressed in HEK293T cells grown in DMEM with 2 mM glutamine, 2% FBS, and 100 units/mL penicillin/streptomycin. 10 μg of DNA was complexed with 30 μg of polyethyleneimine (PEI) and used to transfect 100 mm plates of ~80% confluent HEK293T cells (*63*). Cells were harvested after 48 hours and washed with phosphate-buffered saline. The cell pellets were then stored at −20°C. for later use. In parallel 1 μg of the same DNA was also used to transfect HEK293T cells in 12-well plates. FSEC was used to assess the expression level and the gel filtration profile of the DHHC20 mutants. Following satisfactory assessment of FSEC profile, frozen HEK293T pellets were thawed on ice and re-suspended in 500 μl of 40 mM TrisHCl, pH 7.2, 270 mM NaCl, 5 mM 2-ME, 5 mM $MgCl_2$ containing Dnase and protease inhibitors. 100 μl of a 0.3 M stock DDM was added to the cell suspension and rotated for 1-2 hours at 4°C. Cell debris was removed by centrifugation at 21,000 g for 10 min at 4°C. The supernatant was then applied to ~30-40 μl of TALON resin in a microfuge tube and incubated on a rotator for ~1 hour at 4°C. Resin was first washed with 800 μl of same buffer containing 5 mM imidazole and 1 mM DDM but no $MgCl_2$. Following a second wash with 500 μl of buffer containing 25 mM imidazole, protein was eluted by ~40 μl of 20 mM TrisHCl, pH 7.2, 135 mM NaCl, 2 mM TCEP, 5 mM 2-ME buffer containing 300 mM imidazole.

Protein concentration was estimated by running a stain-free SDS-PAGE (Bio-Rad) gel with previously purified Venus-hDHHC20 standards (8, 4, 1, 0.25 μM) from which a linear calibration curve could be constructed. Venus-hDHHC20 concentrations for most mutants were usually 2.5-6 μM. This translates to protein yields of ~7-16 μg per 100 mm cell culture dish. Of the mutants, AAxE and the ΔW267 mutants were the worst expressing. Therefore, in the case of these mutants, cells were transfected in 150 mm dishes (with 30 μg DNA) instead of 100 mm dishes. For the AAxE mutant, two such 150 mm dishes (equivalent to six 100 mm dishes) gave protein yields of ~14 μg. For the ΔW267 mutant, three 150 mm dishes (equivalent to nine 100 mm dishes) gave protein yields of ~3 μg.

Coupled-enzyme assays to measure Michaelis-Menten constants were carried out as described above. For the acyl-CoA selectivity experiments, enzyme and acyl-CoA concentrations were kept fixed at 10 nM and 10 μM respectively.

Mass spectrometry of irreversibly inhibited hDHHC20

For mass spectrometry measurements, purified wild-type hDHHC20–2-BP sample was run on SDS-PAGE under reducing conditions. Gel was stained by coomassie brilliant blue, destained, and thoroughly washed in deionized water. Gel band (~5 μg protein) was excised, treated with the reducing agent TCEP and the free cysteines blocked with N-ethylmaleimide. Protein was digested with chymotrypsin and applied to a C4 reverse phase column run at a flow-rate of 300 nL/min on a Ultimate 3000 HPLC (Thermo-Dionex) connected to a Orbitrap Elite mass spectrometer (Thermo Scientific).

Fluorescence protease protection assay *hDHHC20*

DHHC20 cells tagged with N- or C-terminal Venus YFP were expressed in Cos7 cells using PEI transfection in 4-well or 8-well chambered microscope containers (Lab-Tek). Along with the Venus tagged DHHC20, we also co-expressed Cerulean CFP as a cytoplasmic marker and mCherry-TGN38 as a Golgi-marker with the mCherry RFP residing in the Golgi lumen. 24 hours post transfection, cells were washed with KHM (110 mM Potassium acetate, 20 mM HEPES, pH 7.2, and 3 mM $MgCl_2$) buffer and then maintained in the same. Approximately a minute after initiation of time series imaging, cell plasma membrane was selectively permeabilized by the addition of digitonin dissolved in KHM to a final concentration of 22-25 μM. The permeabilization was followed by the disappearance of the cytoplasmic marker. Following disappearance of the cytoplasmic marker, trypsin prepared in KHM buffer was added to a final concentration of 1-2 mM and the cells further imaged for another minute or two until no change of any signal occurred. Control experiments were also done in which trypsin was added in the absence of digitonin. In these experiments, no significant loss in fluorescence was observed thus ruling out any pH or non-specific effects of trypsin on the cells.

Imaging was carried out on a Zeiss LSM780 microscope equipped with Argon laser (458, 514 nm) and HeNe-laser (594 nm) with a Plan-Apochromat 63x/1.40 Oil objective and a 458/514/

594 multiband filter set. Cerulean CFP, Venus YFP, and mCherry RFP were excited with the 458, 514, and 594 nm lasers respectively. Emission was collected from 462-506 nm, 516-587 nm, and 605-685 nm for Cerulean, Venus, and mCherry respectively. Images were analyzed using FIJI (*64*).

zfDHHC15

Cos7 cells were seeded onto 8-well chamber plates (Lab-Tek) in 400 µl complete media and allowed them to grow and adhere to the surface of the plate for 16-18 hours. These cells were transfected with 100 ng of each vector GFP-zebrafish DHHC15, mCherry-TGN38, and BFP empty vector in serum free media with Lipofectamine 3000 transfection agent (Invitrogen). These transfected cells allowed to grow for 12-16 hours. The cells were permeabilized with 20 µM of digitonin in KHM buffer (110 mM potassium acetate, 2 mM Magnesium chloride and 50 mM HEPES, pH 7.5) for 30 s followed by trypsinization of the cells with 4 mM of trypsin in KHM buffer for 50 s during time dependent imaging experiment. Imaging data of adhered cells, from pre-permeabilized state to trypsinised state were acquired with confocal laser microscopy system equipped with a Plan Apochromat 63X/1.4NA oil immersion objective lens (LSM 780 Exciter, Carl Zeiss) at 37°C. Time-lapsed imaging was performed with the same instrument. Further, images were processed and fluorescence intensity was calculated using LSM and ImageJ softwares.

Chemicals and Reagents

PEG 300 and 2,5-Hexanediol were obtained from Sigma-Aldrich. PEG 400 was from USB. The buffers and salts were either from Sigma-Aldrich or Fisher Scientific. 2-bromopalmitate was from Sigma-Aldrich. All detergents were from Anatrace. All lipids and acyl coenzyme A was from Avanti Polar Lipids Inc. DTT and IPTG were from GoldBio. Digitonin (high purity) was from Calbiochem. Trypsin was obtained from Worthington Biochemicals. Methanol free 16% formaldehyde was obtained from Pierce-ThermoFisher and the DMAB complex from Sigma-Aldrich.

Sequence alignment

Multiple sequence alignment was carried out using MUSCLE (*65*). Figure for sequence alignment with secondary structure was generated using ESPRIPT (*66*).

REFERENCES AND NOTES

1. G. A. Khoury, R. C. Baliban, C. A. Floudas, Proteome-wide post-translational modification statistics: Frequency analysis and curation of the swiss-prot database. *Sci. Rep.* **1**, 90 (2011). doi: 10.1038/srep00090; pmid: 22034591
2. J. L. Daniotti, M. P. Pedro, J. Valdez Taubas, The role of S-acylation in protein trafficking. *Traffic* **18**, 699–710 (2017). doi: 10.1111/tra.12510; pmid: 28837239
3. M. E. Linder, R. J. Deschenes, Palmitoylation: Policing protein stability and traffic. *Nat. Rev. Mol. Cell Biol.* **8**, 74–84 (2007). doi: 10.1038/nrm2084; pmid: 17183362
4. A. Hentschel, R. P. Zahedi, R. Ahrends, Protein lipid modifications—More than just a greasy ballast. *Proteomics* **16**, 759–782 (2016). doi: 10.1002/pmic.201500353; pmid: 26683279
5. M. F. Schmidt, M. J. Schlesinger, Fatty acid binding to vesicular stomatitis virus glycoprotein: A new type of post-translational modification of the viral glycoprotein. *Cell* **17**, 813–819 (1979). doi: 10.1016/0092-8674(79)90321-0; pmid: 226266
6. M. J. Schlesinger, A. I. Magee, M. F. Schmidt, Fatty acid acylation of proteins in cultured cells. *J. Biol. Chem.* **255**, 10021–10024 (1980). pmid: 7430112
7. M. Fukata, Y. Fukata, H. Adesnik, R. A. Nicoll, D. S. Bredt, Identification of PSD-95 palmitoylating enzymes. *Neuron* **44**, 987–996 (2004). doi: 10.1016/j.neuron.2004.12.005; pmid: 15603741
8. J. L. Hernandez *et al*., APT2 inhibition restores scribble localization and *S*-palmitoylation in snail-transformed cells. *Cell Chem. Biol.* **24**, 87–97 (2017). doi: 10.1016/j.chembiol.2016.12.007; pmid: 28065656
9. B. R. Martin, C. Wang, A. Adibekian, S. E. Tully, B. F. Cravatt, Global profiling of dynamic protein palmitoylation. *Nat. Methods* **9**, 84–89 (2011). doi: 10.1038/nmeth.1769; pmid: 22056678
10. L. Muszbek, G. Haramura, J. E. Cluette-Brown, E. M. Van Cott, M. Laposata, The pool of fatty acids covalently bound to platelet proteins by thioester linkages can be altered by exogenously supplied fatty acids. *Lipids* **34** (Suppl), S331–S337 (1999). doi: 10.1007/BF02562334; pmid: 10419194
11. X. Liang *et al*., Heterogeneous fatty acylation of Src family kinases with polyunsaturated fatty acids regulates raft localization and signal transduction. *J. Biol. Chem.* **276**, 30987–30994 (2001). doi: 10.1074/jbc.M104018200; pmid: 11423543
12. L. Berthiaume, M. D. Resh, Biochemical characterization of a palmitoyl acyltransferase activity that palmitoylates myristoylated proteins. *J. Biol. Chem.* **270**, 22399–22405 (1995). doi: 10.1074/jbc.270.38.22399; pmid: 7673226
13. O. Rocks *et al*., An acylation cycle regulates localization and activity of palmitoylated Ras isoforms. *Science* **307**, 1746–1752 (2005). doi: 10.1126/science.1105654; pmid: 15705808
14. A. E. El-Husseini *et al*., Dual palmitoylation of PSD-95 mediates its vesiculotubular sorting, postsynaptic targeting, and ion channel clustering. *J. Cell Biol.* **148**, 159–172 (2000). doi: 10.1083/jcb.148.1.159; pmid: 10629226
15. K. Huang *et al*., Huntingtin-interacting protein HIP14 is a palmitoyl transferase involved in palmitoylation and trafficking of multiple neuronal proteins. *Neuron* **44**, 977–986 (2004). doi: 10.1016/j.neuron.2004.11.027; pmid: 15603740
16. T. Hayashi, G. M. Thomas, R. L. Huganir, Dual palmitoylation of NR2 subunits regulates NMDA receptor trafficking. *Neuron* **64**, 213–226 (2009). doi: 10.1016/j.neuron.2009.08.017; pmid: 19874789
17. Y. H. Lin *et al*., Host cell-catalyzed *S*-palmitoylation mediates Golgi targeting of the *Legionella* ubiquitin ligase GobX. *J. Biol. Chem.* **290**, 25766–25781 (2015). doi: 10.1074/jbc.M115.637397; pmid: 26316537
18. S. W. Hicks, G. Charron, H. C. Hang, J. E. Galán, Subcellular targeting of *Salmonella* virulence proteins by host-mediated *S*-palmitoylation. *Cell Host Microbe* **10**, 9–20 (2011). doi: 10.1016/j.chom.2011.06.003; pmid: 21767808
19. M. Blanc *et al*., SwissPalm: Protein Palmitoylation database. *F1000Res.* **4**, 261 (2015). pmid: 26339475
20. K. D. Reddy *et al*., Physicochemical sequence characteristics that influence *S*-palmitoylation propensity. *J. Biomol. Struct. Dyn.* **35**, 2337–2350 (2017). pmid: 27498722
21. S. S. Sanders *et al*., Curation of the mammalian palmitoylome indicates a pivotal role for palmitoylation in diseases and disorders of the nervous system and cancers. *PLOS Comput. Biol.* **11**, e1004405 (2015). doi: 10.1371/journal.pcbi.1004405; pmid: 26275289
22. K. B. Runkle *et al*., Inhibition of DHHC20-mediated EGFR palmitoylation creates a dependence on EGFR signaling. *Mol. Cell* **62**, 385–396 (2016). doi: 10.1016/j.molcel.2016.04.003; pmid: 27153536
23. S. Lobo, W. K. Greentree, M. E. Linder, R. J. Deschenes, Identification of a Ras palmitoyltransferase in *Saccharomyces cerevisiae*. *J. Biol. Chem.* **277**, 41268–41273 (2002). doi: 10.1074/jbc.M206573200; pmid: 12193598
24. A. F. Roth, Y. Feng, L. Chen, N. G. Davis, The yeast DHHC cysteine-rich domain protein Akr1p is a palmitoyl transferase. *J. Cell Biol.* **159**, 23–28 (2002). doi: 10.1083/jcb.200206120; pmid: 12370247
25. Y. Ohno, A. Kihara, T. Sano, Y. Igarashi, Intracellular localization and tissue-specific distribution of human and yeast DHHC cysteine-rich domain-containing proteins. *Biochim. Biophys. Acta* **1761**, 474–483 (2006). doi: 10.1016/j.bbalip.2006.03.010; pmid: 16647879
26. E. G. Politis, A. F. Roth, N. G. Davis, Transmembrane topology of the protein palmitoyl transferase Akr1. *J. Biol. Chem.* **280**, 10156–10163 (2005). doi: 10.1074/jbc.M411946200; pmid: 15632165
27. D. A. Mitchell, G. Mitchell, Y. Ling, C. Budde, R. J. Deschenes, Mutational analysis of *Saccharomyces cerevisiae* Erf2 reveals a two-step reaction mechanism for protein palmitoylation by DHHC enzymes. *J. Biol. Chem.* **285**, 38104–38114 (2010). doi: 10.1074/jbc.M110.169102; pmid: 20851885
28. B. C. Jennings, M. E. Linder, DHHC protein *S*-acyltransferases use similar ping-pong kinetic mechanisms but display different acyl-CoA specificities. *J. Biol. Chem.* **287**, 7236–7245 (2012). doi: 10.1074/jbc.M111.337246; pmid: 22247542
29. A. González Montoro, R. Quiroga, J. Valdez Taubas, Zinc coordination by the DHHC cysteine-rich domain of the palmitoyltransferase Swf1. *Biochem. J.* **454**, 427–435 (2013). doi: 10.1042/BJ20121693; pmid: 23790227
30. J. Greaves *et al*., Molecular basis of fatty acid selectivity in the zDHHC family of S-acyltransferases revealed by click chemistry. *Proc. Natl. Acad. Sci. U.S.A.* **114**, E1365–E1374 (2017). doi: 10.1073/pnas.1612254114; pmid: 28167757
31. T. Kawate, E. Gouaux, Fluorescence-detection size-exclusion chromatography for precrystallization screening of integral membrane proteins. *Structure* **14**, 673–681 (2006). doi: 10.1016/j.str.2006.01.013; pmid: 16615909
32. L. D. Hamel, R. J. Deschenes, D. A. Mitchell, A fluorescence-based assay to monitor autopalmitoylation of zDHHC proteins applicable to high-throughput screening. *Anal. Biochem.* **460**, 1–8 (2014). doi: 10.1016/j.ab.2014.05.013; pmid: 24878334
33. C. D. Gottlieb, S. Zhang, M. E. Linder, The cysteine-rich domain of the DHHC3 palmitoyltransferase is palmitoylated and contains tightly bound zinc. *J. Biol. Chem.* **290**, 29259–29269 (2015). doi: 10.1074/jbc.M115.691147; pmid: 26487721
34. A. Manglik, B. Kobilka, The role of protein dynamics in GPCR function: Insights from the β_2AR and rhodopsin. *Curr. Opin. Cell Biol.* **27**, 136–143 (2014). doi: 10.1016/j.ceb.2014.01.008; pmid: 24534489
35. C. Oswald *et al*., Intracellular allosteric antagonism of the CCR9 receptor. *Nature* **540**, 462–465 (2016). doi: 10.1038/nature20606; pmid: 27926729
36. H. Lorenz, D. W. Hailey, J. Lippincott-Schwartz, Fluorescence protease protection of GFP chimeras to reveal protein topology and subcellular localization. *Nat. Methods* **3**, 205–210 (2006). doi: 10.1038/nmeth857; pmid: 16489338
37. S. S. Krishna, I. Majumdar, N. V. Grishin, Structural classification of zinc fingers: Survey and summary. *Nucleic Acids Res.* **31**, 532–550 (2003). doi: 10.1093/nar/gkg161; pmid: 12527760
38. L. C. Myers, F. Jackow, G. L. Verdine, Metal dependence of transcriptional switching in *Escherichia coli* Ada. *J. Biol. Chem.* **270**, 6664–6670 (1995). doi: 10.1074/jbc.270.12.6664; pmid: 7896807
39. R. Gamsjaeger, C. K. Liew, F. E. Loughlin, M. Crossley, J. P. Mackay, Sticky fingers: Zinc-fingers as protein-recognition motifs. *Trends Biochem. Sci.* **32**, 63–70 (2007). doi: 10.1016/j.tibs.2006.12.007; pmid: 17210253
40. G. Dodson, A. Wlodawer, Catalytic triads and their relatives. *Trends Biochem. Sci.* **23**, 347–352 (1998). doi: 10.1016/S0968-0004(98)01254-7; pmid: 9787641
41. A. González Montoro, R. Quiroga, H. J. Maccioni, J. Valdez Taubas, A novel motif at the C-terminus of palmitoyltransferases is essential for Swf1 and Pfa3 function in vivo. *Biochem. J.* **419**, 301–308 (2009). doi: 10.1042/BJ20080921; pmid: 19138168
42. P. Mill *et al*., Palmitoylation regulates epidermal homeostasis and hair follicle differentiation. *PLOS Genet.* **5**, e1000748 (2009). doi: 10.1371/journal.pgen.1000748; pmid: 19956733
43. R. A. Coleman, P. Rao, R. J. Fogelsong, E. S. Bardes, 2-Bromopalmitoyl-CoA and 2-bromopalmitate: Promiscuous inhibitors of membrane-bound enzymes. *Biochim. Biophys. Acta* **1125**, 203–209 (1992). doi: 10.1016/0005-2760(92)90046-X; pmid: 1571364
44. D. Davda *et al*., Profiling targets of the irreversible palmitoylation inhibitor 2-bromopalmitate. *ACS Chem. Biol.* **8**, 1912–1917 (2013). doi: 10.1021/cb400380s; pmid: 23844586
45. G. Bánhegyi *et al*., Fatty acyl-CoA esters and the permeability of rat liver microsomal vesicles. *Biochem. J.* **320**, 343–344 (1996). doi: 10.1042/bj3200343; pmid: 8947507
46. M. A. Requero, M. González, F. M. Goñi, A. Alonso, G. Fidelio, Differential penetration of fatty acyl-coenzyme A and fatty acylcarnitines into phospholipid monolayers. *FEBS Lett.* **357**, 75–78 (1995). doi: 10.1016/0014-5793(94)01326-V; pmid: 8001684

47. P. J. O'Brien, R. S. St Jules, T. S. Reddy, N. G. Bazan, M. Zatz, Acylation of disc membrane rhodopsin may be nonenzymatic. *J. Biol. Chem.* **262**, 5210–5215 (1987). pmid: 3558391
48. L. V. Kordyukova, M. V. Serebryakova, L. A. Baratova, M. Veit, S acylation of the hemagglutinin of influenza viruses: Mass spectrometry reveals site-specific attachment of stearic acid to a transmembrane cysteine. *J. Virol.* **82**, 9288–9292 (2008). doi: 10.1128/JVI.00704-08; pmid: 18596092
49. D. Senyilmaz *et al*., Regulation of mitochondrial morphology and function by stearoylation of TFR1. *Nature* **525**, 124–128 (2015). doi: 10.1038/nature14601; pmid: 26214738
50. J. M. Draper, C. D. Smith, DHHC20: A human palmitoyl acyltransferase that causes cellular transformation. *Mol. Membr. Biol.* **27**, 123–136 (2010). doi: 10.3109/09687681003616854; pmid: 20334580
51. L. H. Chamberlain, M. J. Shipston, The physiology of protein *S*-acylation. *Physiol. Rev.* **95**, 341–376 (2015). doi: 10.1152/physrev.00032.2014; pmid: 25834228
52. K. Shah, Y. Liu, C. Deirmengian, K. M. Shokat, Engineering unnatural nucleotide specificity for Rous sarcoma virus tyrosine kinase to uniquely label its direct substrates. *Proc. Natl. Acad. Sci. U.S.A.* **94**, 3565–3570 (1997). doi: 10.1073/pnas.94.8.3565; pmid: 9108016
53. A. C. Bishop *et al*., A chemical switch for inhibitor-sensitive alleles of any protein kinase. *Nature* **407**, 395–401 (2000). doi: 10.1038/35030148; pmid: 11014197
54. M. G. J. Baud *et al*., A bump-and-hole approach to engineer controlled selectivity of BET bromodomain chemical probes. *Science* **346**, 638–641 (2014). doi: 10.1126/science.1249830; pmid: 25323695
55. S. B. Long, X. Tao, E. B. Campbell, R. MacKinnon, Atomic structure of a voltage-dependent K^+ channel in a lipid membrane-like environment. *Nature* **450**, 376–382 (2007). doi: 10.1038/nature06265; pmid: 18004376
56. I. Rayment, Reductive alkylation of lysine residues to alter crystallization properties of proteins. *Methods Enzymol.* **276**, 171–179 (1997). doi: 10.1016/S0076-6879(97)76058-0
57. M. Caffrey, V. Cherezov, Crystallizing membrane proteins using lipidic mesophases. *Nat. Protoc.* **4**, 706–731 (2009). doi: 10.1038/nprot.2009.31; pmid: 19390528
58. M. D. Winn *et al*., Overview of the CCP4 suite and current developments. *Acta Crystallogr. D Biol. Crystallogr.* **67**, 235–242 (2011). doi: 10.1107/S0907444910045749; pmid: 21460441
59. P. D. Adams *et al*., PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 213–221 (2010). doi: 10.1107/S0907444909052925; pmid: 20124702
60. P. Emsley, B. Lohkamp, W. G. Scott, K. Cowtan, Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486–501 (2010). doi: 10.1107/S0907444910007493; pmid: 20383002
61. C. J. Stanley, R. N. Perham, Purification of 2-oxo acid dehydrogenase multienzyme complexes from ox heart by a new method. *Biochem. J.* **191**, 147–154 (1980). doi: 10.1042/bj1910147; pmid: 7470091
62. A. Goehring *et al*., Screening and large-scale expression of membrane proteins in mammalian cells for structural studies. *Nat. Protoc.* **9**, 2574–2585 (2014). doi: 10.1038/nprot.2014.173; pmid: 25299155
63. P. A. Longo, J. M. Kavran, M. S. Kim, D. J. Leahy, Transient mammalian cell transfection with polyethylenimine (PEI). *Methods Enzymol.* **529**, 227–240 (2013). doi: 10.1016/B978-0-12-418687-3.00018-5; pmid: 24011049
64. J. Schindelin *et al*., Fiji: An open-source platform for biological-image analysis. *Nat. Methods* **9**, 676–682 (2012). doi: 10.1038/nmeth.2019; pmid: 22743772
65. R. C. Edgar, MUSCLE: A multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* **5**, 113 (2004). doi: 10.1186/1471-2105-5-113; pmid: 15318951
66. X. Robert, P. Gouet, Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res.* **42** (W1), W320–W324 (2014). doi: 10.1093/nar/gku316; pmid: 24753421
67. J. M. Word *et al*., Visualizing and quantifying molecular goodness-of-fit: Small-probe contact dots with explicit hydrogen atoms. *J. Mol. Biol.* **285**, 1711–1733 (1999). doi: 10.1006/jmbi.1998.2400; pmid: 9917407

ACKNOWLEDGMENTS

The authors thank M. Machner (NIH) for the GobX protein expression plasmid, E. Gouaux (Oregon Health & Science University) for the pEG-Bacmam vector, Y. Li (NIH) for mass spectrometry, C. Williamson (NIH) for help with confocal microscopy, and J. Swan (NIH) for help with the figure accompanying the print version of the manuscript. The authors thank J. Bonifacino (NIH), S. Buchanan (NIH), W. Prinz (NIH), G. Storz (NIH), K. Swartz (NIH), and W. Yang (NIH) for comments on the manuscript. This work was supported by the Intramural Research Program of the NIH, the National Institutes of Child Health and Human Development (1ZIAHD008928 to M.S.R., P.K., C.-J.L., R.V., and A.B.), and the National Institute of Neurological Disorders and Stroke (to R.V. and A.B.). Crystals for this project were screened and data sets collected at various stages at the following beamlines: NE-CAT, GM/CA, and SER-CAT (22-ID) at the Advanced Photon Source, Argonne National Laboratory. NE-CAT is funded by the National Institute of General Medical Sciences from the NIH (P41 GM103403). The Pilatus 6M detector on 24-ID-C beamline (NE-CAT) is funded by a NIH-ORIP HEI grant (S10RR029205). GM/CA has been funded in whole or in part with federal funds from the National Cancer Institute (ACB-12002) and the National Institute of General Medical Sciences (AGM-12006). The Eiger 16M detector (GM/CA) was funded by an NIH, Office of Research Infrastructure Programs, High-End Instrumentation Grant (1S10OD012289-01A1). This research used resources of the Advanced Photon Source, a U.S. Department of Energy (DOE), Office of Science User Facility operated for the DOE Office of Science by Argonne National Laboratory under contract no. DE-AC02-06CH11357. The structures have been deposited to the Protein Data Bank (PDB) with the accession codes 6BML (hDHHC20–2-BP treated), 6BMM (hDHHC20, $P2_1$ form), 6BMN (hDHHC20, $P6_3$ form), and 6BMS (zfDHHS15). All data needed to evaluate the conclusions in the paper are present in the main text or supplementary materials. A.B. conceived the project. A.B., M.S.R., and P.K. designed experiments. M.R. performed expression, purification, crystallization, data collection, and biochemical and imaging experiments of DHHC20. P.K. performed expression, purification, crystallization, data collection, crystallography, and imaging experiments of zfDHHC15. C.-J.L. performed crystallography of hDHHC20, and R.V. performed biochemical assays of zfDHHC15. K.R.R. assisted with crystallography. A.B. contributed to crystallization, data collection, and crystallography. M.S.R., P.K., C.-J.L., R.V., K.R.R., and A.B. analyzed data. A.B. wrote the manuscript with help from M.S.R., C.-J.L., and P.K. helped in figure preparation. The authors declare no competing financial interests.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/359/6372/eaao6326/suppl/DC1
Figs. S1 to S12
Tables S1 to S8
References

11 August 2017; accepted 7 December 2017
10.1126/science.aao6326

RESEARCH ARTICLE

IMMUNOLOGY

The epigenetic control of stemness in CD8$^+$ T cell fate commitment

Luigia Pace,[1,2,3]*† Christel Goudot,[1,2] Elina Zueva,[1,2] Paul Gueguen,[1,2] Nina Burgdorf,[1,2] Joshua J. Waterfall,[1,4,5] Jean-Pierre Quivy,[1,6,7] Geneviève Almouzni,[1,6,7]* Sebastian Amigorena[1,2]*

After priming, naïve CD8$^+$ T lymphocytes establish specific heritable transcription programs that define progression to long-lasting memory cells or to short-lived effector cells. Although lineage specification is critical for protection, it remains unclear how chromatin dynamics contributes to the control of gene expression programs. We explored the role of gene silencing by the histone methyltransferase Suv39h1. In murine CD8$^+$ T cells activated after *Listeria monocytogenes* infection, Suv39h1-dependent trimethylation of histone H3 lysine 9 controls the expression of a set of stem cell–related memory genes. Single-cell RNA sequencing revealed a defect in silencing of stem/memory genes selectively in *Suv39h1*-defective T cell effectors. As a result, *Suv39h1*-defective CD8$^+$ T cells show sustained survival and increased long-term memory reprogramming capacity. Thus, Suv39h1 plays a critical role in marking chromatin to silence stem/memory genes during CD8$^+$ T effector terminal differentiation.

Memory T lymphocytes provide lifelong protection against pathogens and cancer (*1*). In contrast to naïve and effector T cells, memory cells possess unique properties of "stemness," enabling long-term survival and plasticity to replenish effector pools after renewed antigen challenges (*2*). Understanding the lineage relationships among naïve, effector, and memory T cells, as well as the molecular pathways that regulate gene expression during the transitions from one to another of these distinct states, is essential for the rational design of vaccines and the development of new immune-therapeutic protocols (*3*).

Although many studies have characterized the transcription factors that control the differentiation of T cells, the corresponding epigenetic states and associated chromatin dynamics involved in the establishment and maintenance of CD8$^+$ T cell memory and effector identities is still incompletely understood (*4*–*7*). Several epigenetic pathways, including trimethylated histone H3 Lys^9 (H3K9me3)/HP-1/Suv39h1 and Polycomb repressive complexes, can contribute to establishing or maintaining transcriptional silencing (*7*–*11*). The H3K9me3 modification is considered to be a repressive mark, a hallmark of both constitutive and facultative heterochromatin (*12*), most often associated with silent gene loci. Mouse Suv39h1 and Suv39h2, two H3K9 site-specific histone methyltransferases (HMTs), are critical heterochromatin regulators (*10*, *13*). Suv39h1 is involved in heterochromatin organization, gene silencing, and lineage stability (*10*, *11*, *14*). It also limits somatic reprogramming of differentiated cells into induced pluripotent stem cells (*15*). In B cells and CD4$^+$ T cells, Suv39h1 is involved in gene silencing and lineage plasticity (*16*, *17*). Although the mechanisms underlying the induction of genes critically involved in effector and memory T cell generation have been extensively analyzed, the impact of heterochromatin-dependent gene expression silencing on the fates of T lymphocytes during differentiation has not been addressed. Here, we explore the role of Suv39h1-dependent gene silencing in the establishment and maintenance of memory CD8$^+$ T cell stemness, plasticity, and transition to terminally differentiated effectors.

Long-term protection against *Listeria monocytogenes* infection requires Suv39h1

To investigate the role of Suv39h1 in T cell responses to infectious agents, we infected *Suv39h1*-knockout (KO) and wild-type littermate mice with OVA-expressing *L. monocytogenes* (LM-OVA) (Fig. 1A). The bacterial burden in the spleen and liver was measured on day 3 (peak of infection in the spleen) and day 7 (resolution). After primary infection, we observed similar bacteremia in the spleen and liver of *Suv39h1*-KO and littermate control mice, both at the peak and after the resolution of infection (Fig. 1B). In contrast, upon a secondary challenge with LM-OVA 48 days after the primary infection, most littermate control mice showed complete protection, whereas high levels of LM-OVA were detected in the liver and/or spleen in more than 85% of *Suv39h1*-KO mice (Fig. 1C). Because protection against secondary *L. monocytogenes* infection is primarily mediated by CD8$^+$ T cells (*18*), these results suggested a defect in the CD8$^+$ T cell response against *L. monocytogenes* in *Suv39h1*-KO mice.

To further investigate the role of Suv39h1 in antigen-specific CD8$^+$ T cell responses to LM-OVA, we measured the antigen-specific CD8$^+$ T cell response using SIINFEKL (Ser-Ile-Ile-Asn-Phe-Glu-Lys-Leu)–H-2K^b multimers (K^b-OVA; Fig. 1D). We observed a factor of 2 reduction in the percentage and absolute numbers of K^b-OVA$^+$ CD8$^+$ T cells in the blood (Fig. 1, D and E) and a factor of 9 decrease in the numbers of K^b-OVA$^+$ CD8$^+$ T cells in the spleen (Fig. 1F) on day 7 in *Suv39h1*-KO mice relative to littermates. An analysis of K^b-OVA$^+$ CD8$^+$ T cells on days 5, 6, and 7 after LM-OVA infection suggests that the reduced antigen-specific response was not due to differences in the survival of *Suv39h1*-KO cells (fig. S1). Upon ex vivo restimulation, the percentages of interferon-γ$^+$ (IFN-γ$^+$) and granzyme B$^+$ CD8$^+$ T cells in *Suv39h1*-KO mice were lower than those of littermates (Fig. 1, D, G, and H, and fig. S2). To address whether the reduced K^b-OVA$^+$ CD8$^+$ T cell numbers in LM-OVA–infected *Suv39h1*-KO mice was T cell–intrinsic, we adoptively transferred *Suv39h1*-proficient or -deficient T cell receptor transgenic OT-I CD8$^+$ T cells (which recognize OVA) to wild-type mice. Five days after infection with LM-OVA, the numbers of *Suv39h1*-KO OT-I cells were reduced relative to wild-type OT-I (Fig. 1I). Although the percentage of IFN-γ$^+$ OT-I cells was similar, we observed a lower frequency of granzyme B$^+$ T cells (fig. S3, A to C), confirming that the defect observed in *Suv39h1*-KO mice was intrinsic to CD8$^+$ T cells. In line with these in vivo results, lower numbers of *Suv39h1*-KO effector OT-I CD8$^+$ T cells were recovered in vitro under effector-polarizing conditions, relative to numbers of *Suv39h1*-proficient T cells. Notably, the defective expansion in *Suv39h1*-KO T cells was overcome by the overexpression of SUV39H1 by a retroviral expression vector (fig. S4, A and B). Thus, CD8$^+$ T cell responses to LM-OVA are impaired in *Suv39h1*-KO mice because of a CD8$^+$ T cell–intrinsic defect.

Transcriptional silencing by Suv39h1

To investigate whether the reduced K^b-OVA$^+$ CD8$^+$ T cell response results from defective gene expression programming, we analyzed RNA profiles by Affymetrix microarrays of purified naïve and K^b-OVA$^+$ CD8$^+$ T cells, which were wild-type or *Suv39h1*-KO and isolated by fluorescence-activated cell sorting (FACS) 7 days after LM-OVA infection (Fig. 2A and fig. S5A). The overall number of transcripts up-regulated in K^b-OVA$^+$ CD8$^+$ T cells from infected mice, as compared to naïve CD8$^+$ T cells, was highly similar between littermates (1571) and *Suv39h1*-KO (1433) T cells (Fig. 2, B and C). In contrast, the number of transcripts

[1]Institut Curie, PSL Research University, F-75005 Paris, France. [2]INSERM U932, Equipes Labellisées Ligue contre le Cancer, F-75005 Paris, France. [3]Armenise-Harvard Laboratory, Italian Institute for Genomic Medicine, Turin, Italy. [4]INSERM U830, F-75005 Paris, France. [5]Institut Curie, Translational Research Department, F-75005 Paris, France. [6]CNRS, UMR3664, Equipe Labellisée Ligue contre le Cancer, F-75005 Paris, France. [7]Sorbonne Universités, UPMC University Paris 06, CNRS, UMR3664, F-75005 Paris, France. ***Corresponding author. Email: luigia.pace@iigm.it (L.P.); genevieve.almouzni@curie.fr (G.A.); sebastian.amigorena@curie.fr (S.A.)** †Present address: Italian Institute for Genomic Medicine, Turin, Italy.

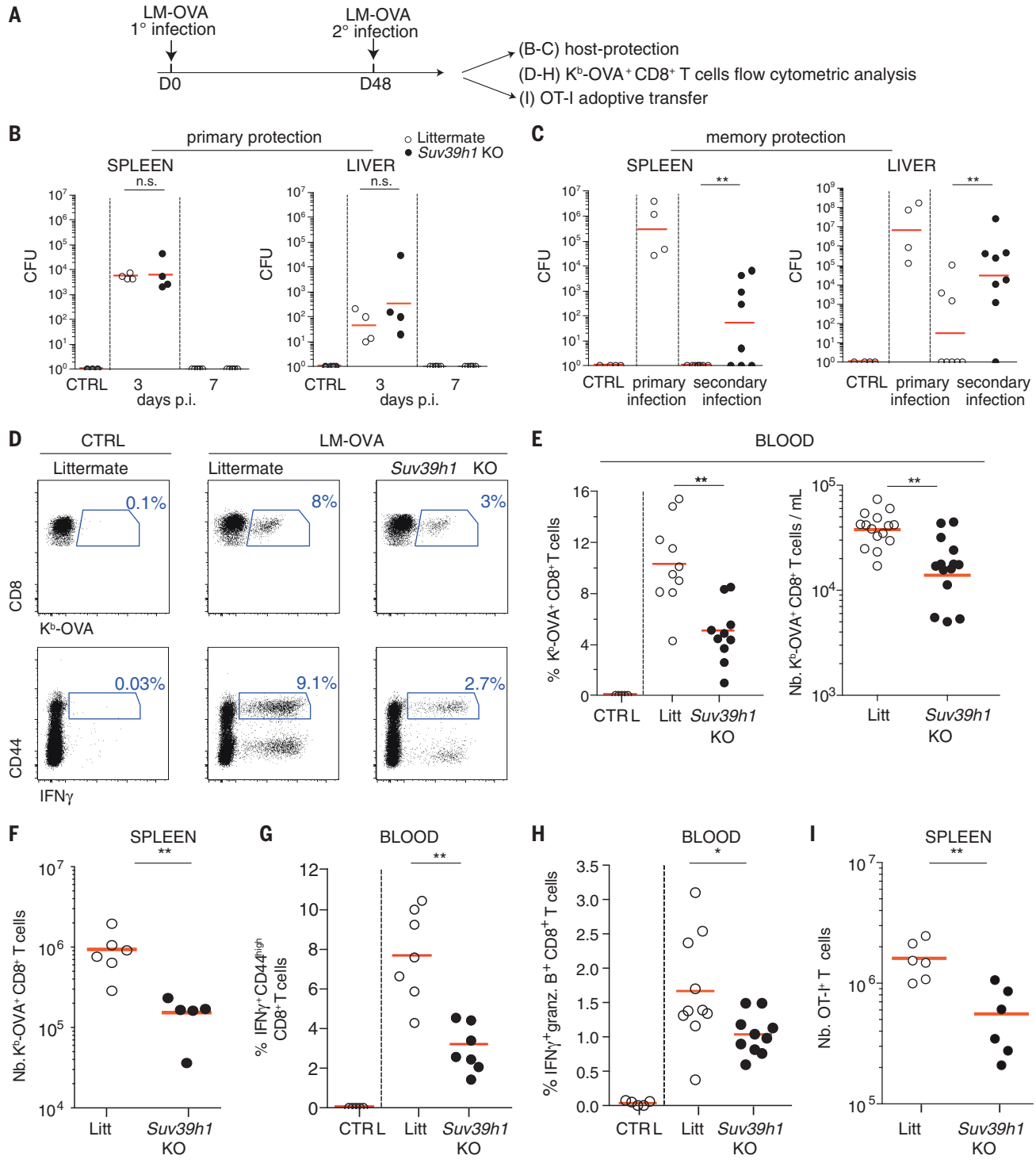


Fig. 1. CD8$^+$ T cell–mediated host protection is impaired in *Suv39h1*-defective mice. (**A**) Experimental design. (**B**) Littermates and *Suv39h1*-KO mice were intravenously (i.v.) injected with LM-OVA (primary infection) and, on the indicated days post-infection (p.i.), the number of bacterial colony-forming units (CFU) was determined. (**C**) Littermates and *Suv39h1*-KO mice previously i.v. immunized with LM-OVA were rechallenged 48 days later. Three days after LM-OVA secondary infection, protection was assessed by counting CFU in spleen and liver. (**D**) Littermate and *Suv39h1*-KO mice were immunized with LM-OVA; 7 days later, primary CD8$^+$ T cell responses were evaluated in the peripheral blood, using K^b-SIINFEKL (K^b-OVA$^+$) multimers and intracellular IFN-γ after restimulation ex vivo with the OVA peptide SIINFEKL (OVA$_{257\text{-}264}$). Representative plots are shown. (**E** and **F**) Percentages and numbers of K^b-OVA$^+$ CD8$^+$ T cells in the blood (E) and spleen (F). Data are representative of at least three experiments with two or more mice per group. (**G** and **H**) Percentages of CD44^{hi} IFN-γ$^+$ and IFN-γ$^+$ granzyme B$^+$, respectively, on gated CD8$^+$ T cells. (**I**) *Rag2*-KO mice were immunized with LM-OVA and adoptively transferred with wild-type and *Suv39h1*-KO OT-I cells. Five days later, the total number of OT-I cells was measured. Graphs show means; geometric means are displayed in (B), (C), (E) (right), (F), and (I). *P < 0.05, **P < 0.01 (Wilcoxon Mann-Whitney test); n.s., not significant.

significantly down-regulated in *Suv39h1*-KO mice (1108) was lower relative to littermates (1738) (Fig. 2, B and C). The absence of Suv39h1 expression impaired the transcriptional silencing of approximately 997 genes (Fig. 2C), including genes involved in critical T cell functions such as *Il7r* (CD127), *Sell* (CD62L), *Ccr7*, and *Cxcr4* (Fig. 2D) (*19*), consistent with the known repressive role of Suv39h1 in gene expression.

To examine the phenotype of the antigen-specific CD8$^+$ T cells that accumulate in *Suv39h1*-KO mice, we used gene set enrichment analysis (GSEA) by pairwise comparison between differentiated *Suv39h1*-KO and wild-type K^b-OVA$^+$ CD8$^+$ T cells isolated 7 days after LM-OVA infection (Fig. 2E). *Suv39h1*-KO cells showed significant enrichment in both memory and naïve gene signatures, whereas the effector signature was not significantly enriched in either wild-type or *Suv39h1*-KO K^b-OVA$^+$ CD8$^+$ T cells (Fig. 2E). The interrogation of a series of public gene signatures revealed that *Suv39h1*-KO K^b-OVA$^+$ CD8$^+$ T cells were also significantly enriched for a "lymphoid–stem cell" signature (*20*) (Fig. 2E, right). For further investigation of the nature of these differences, the core lymphoid–stem cell gene signature (*20*) was combined with gene sets known

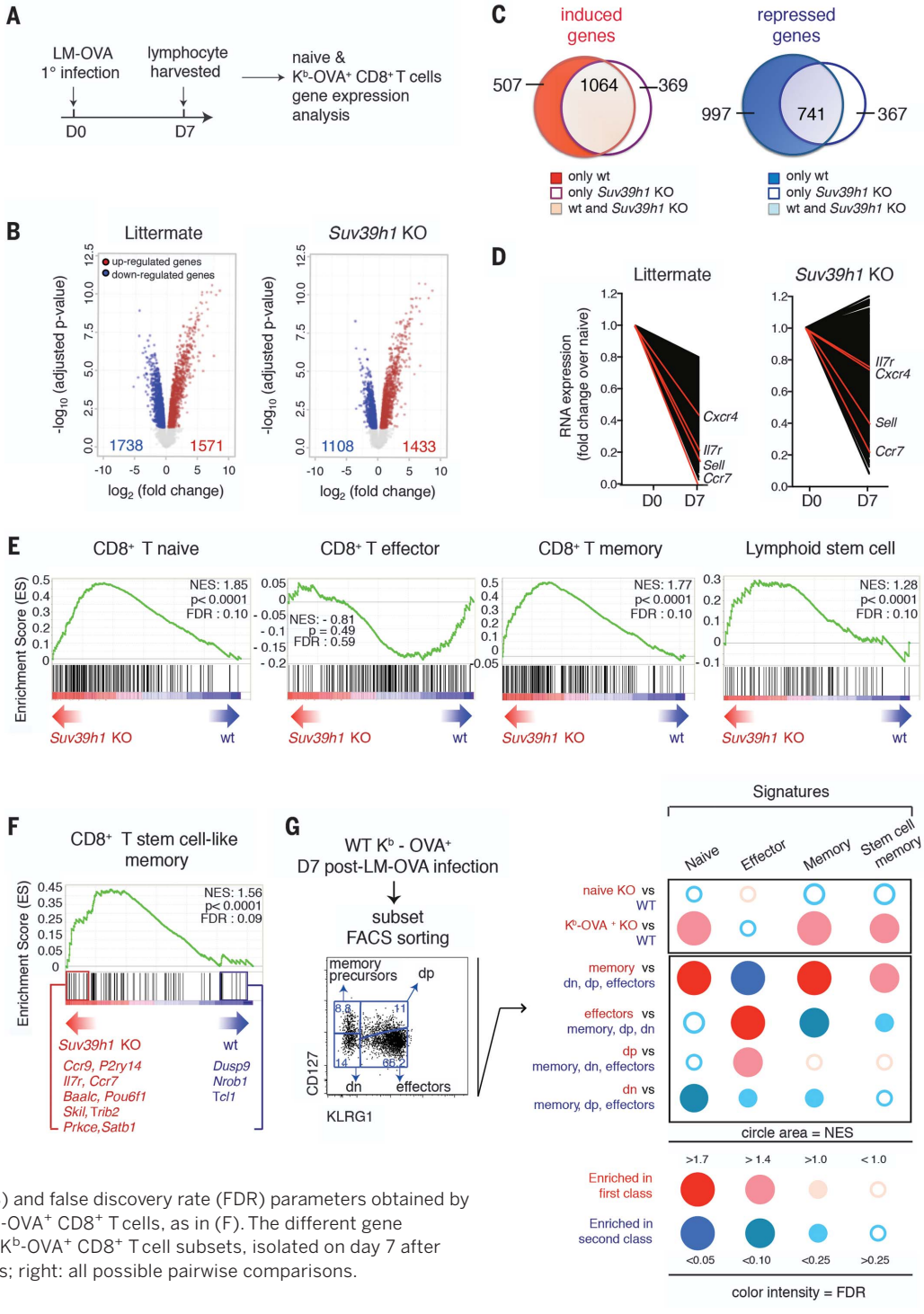


Fig. 2. Gene expression patterns and differentiation programs of K^b-OVA$^+$ CD8$^+$ T lymphocytes are enriched in stem cell–like gene signatures in *Suv39h1*-KO mice. (**A**) Experimental design. Naïve littermate and *Suv39h1*-KO CD8$^+$ T cells and day 7 p.i. K^b-OVA$^+$ CD8$^+$ T cells (from LM-OVA–infected mice) were isolated by FACS. RNA was isolated and analyzed using Affymetrix microarrays (three mice per condition were analyzed). (**B**) Volcano plots of K^b-OVA$^+$ CD8$^+$ T cells versus naïve CD8$^+$ T cells from littermate and *Suv39h1*-KO mice show the adjusted *P* value ($-\log_{10}$) versus fold change ($\log_2$). Up-regulated and down-regulated mRNAs are shown in red and blue, respectively. (**C**) Venn diagrams summarize the overlap between differentially expressed genes that are up-regulated (left) or down-regulated (right) in K^b-OVA$^+$ CD8$^+$ T cells versus naïve CD8$^+$ T cells from wild-type and *Suv39h1*-KO mice. Total common gene numbers for each group are indicated within the areas. (**D**) Expression pattern of mRNAs down-regulated in wild-type K^b-OVA$^+$ CD8$^+$ T cells versus naïve cells, shown for wild-type and *Suv39h1*-KO CD8$^+$ T cells. Representative genes are shown with red lines. (**E**) GSEA was performed to determine the specific enrichment in gene signatures (GeneSet) in wild-type and *Suv39h1*-KO K^b-OVA$^+$ CD8$^+$ T cells isolated on day 7 after LM-OVA infection. (**F**) GSEA for stem/memory signature (table S1) in wild-type and *Suv39h1*-KO K^b-OVA$^+$ CD8$^+$ T cells. The top highly expressed mRNAs in *Suv39h1*-KO or in wild-type K^b-OVA$^+$ CD8$^+$ T cells are shown. (**G**) BubbleGUM analysis of homologies between different CD8$^+$ T subsets by high-throughput GSEA. The right panel summarizes the normalized enrichment score (NES) and false discovery rate (FDR) parameters obtained by GSEA of wild-type and *Suv39h1*-KO K^b-OVA$^+$ CD8$^+$ T cells, as in (F). The different gene signatures were analyzed in wild-type K^b-OVA$^+$ CD8$^+$ T cell subsets, isolated on day 7 after LM-OVA infection. Left: Sorted subsets; right: all possible pairwise comparisons.

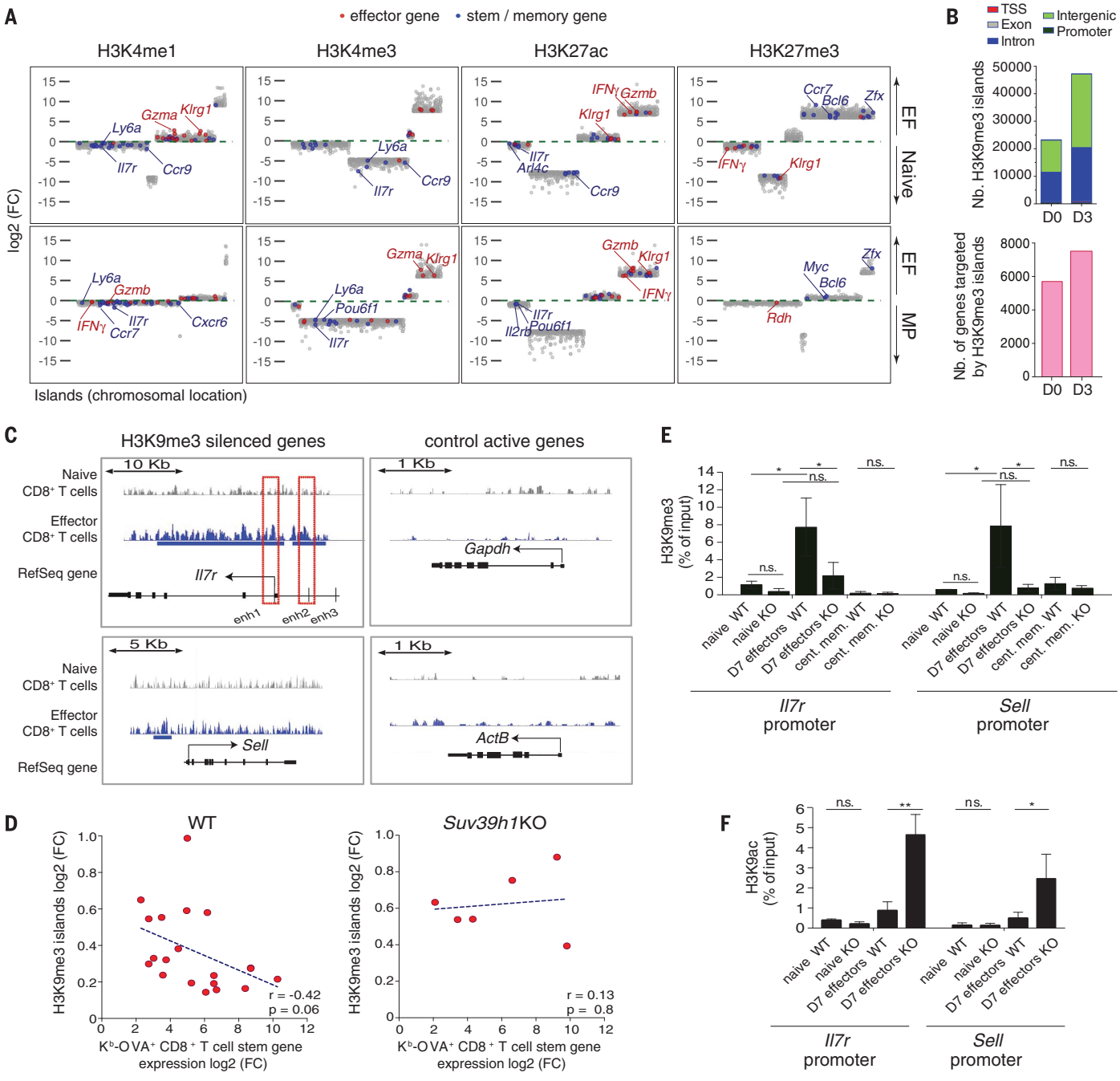


Fig. 3. H3K9me3 deposition by Suv39h1 silences stem cell–like memory genes during CD8$^+$ T effector differentiation. (**A**) H3K4me1, H3K4me3, H3K27ac, and H3K27me3 fold change (FC) values in enriched domains (islands) from naïve cells or memory precursors were compared to effector OT-I CD8$^+$ T cells (differentiated in vivo 8 days after LM-OVA infection) to identify significant differentially modified regions. Islands assigned to stem/memory and effector genes are shown in blue and red, respectively. (**B** and **C**) Naïve and in vitro cultured wild-type CD8$^+$ T effectors were analyzed by ChIP-seq using antibodies to H3K9me3. (B) Top: The genomic distribution of H3K9me3-enriched islands is increased in effectors. Bottom: The number of genes targeted by H3K9me3-enriched islands is also increased. Means of two biological replicates are shown. (C) Representative genomic regions show H3K9me3 enrichment in naïve and effector CD8$^+$ T cells. Normalized ChIP-seq reads (bigWig) and enriched islands (bed) are shown. In the left panels, *Il7r* and *Sell* inter- and intragenic regions with normalized ChIP-seq reads and significantly H3K9me3 enriched islands are represented. Transcribed control genes are also shown in the right panels. (**D**) H3K9me3 enrichment is plotted against mRNA expression of stem cell/memory signature genes. H3K9me3 island deposition in in vitro differentiated effectors correlates with stem/memory gene silencing in wild-type dump$^-$ K^b-OVA$^+$ CD8$^+$ T cells and is impaired in *Suv39h1*-KO effectors. Pearson linear regression is displayed. (**E**) Littermate and *Suv39h1*-KO polyclonal naïve (CD44^{lo} CD62L$^+$ CD127$^+$) CD8$^+$ T cells, central memory (CD44^{hi} CD62L$^+$ CD127$^+$) CD8$^+$ T cells, and effector (dump$^-$ CD44^{hi} CD127$^{lo/-}$ KLRG1$^+$ K^b-OVA$^+$) CD8$^+$ T cells isolated 7 days after LM-OVA infection were analyzed by μChIP-qPCR using antibodies to H3K9me3. qPCR was performed with primers specific for the promoters of *Il7r* and *Sell*. (**F**) Naïve and effector (dump$^-$ CD44^{hi} CD127$^{lo/-}$ KLRG1$^+$ K^b-OVA$^+$) CD8$^+$ T cells from day 7 LM-OVA infected mice were analyzed by μChIP-qPCR using antibodies to H3K9ac. Data are means ± SEM of at least three independent experiments. *P < 0.05, **P < 0.01; for (E) and (F), analysis of variance, Tukey test.

to be commonly enriched both in memory CD8$^+$ T cells and long-term hematopoietic stem cells (*21*), overexpressed in mouse and human CD8$^+$ T stem cell–like memory subset (*22*) and also retrieved from signaling pathways that regulate stem cell pluripotency from the KEGG database. We designated this as a "stem cell–like memory" signature (composed of 86 genes; table S1). K^b-OVA$^+$ CD8$^+$ T cells from LM-OVA–infected *Suv39h1*-KO mice were strongly enriched for this signature (Fig. 2F and fig. S5, B and C). The expression levels of stem cell–like and memory markers *Il7r*, *Ly6a* (Ly6a/e, Sca-1), *Fas* (CD95), and *Pou6f1* were all increased in *Suv39h1*-KO K^b-OVA$^+$ CD8$^+$ cells relative to cells from control littermates (table S1 and fig. S5C). The CD8$^+$ T–stem cell–like memory signature was not enriched in naïve *Suv39h1*-KO CD8$^+$ T cells relative to wild-type littermates (Fig. 2G, BubbleGUM representation, top) (*23*). Thus, by comparison to wild-type cells, *Suv39h1*-KO K^b-OVA$^+$ CD8$^+$ T cells overexpress a series of genes associated with stem cells and memory functions.

To search for these stem cell/memory CD8$^+$ T cells in wild-type animals, we isolated the following wild-type subpopulations by FACS (from dump$^-$ CD44^{hi} K^b-OVA$^+$ CD8$^+$ T cells, day 7 after LM-OVA infection): (i) CD127^{hi} KLRG1$^-$ memory precursors, (ii) CD127$^{lo/-}$ KLRG1$^+$ effectors, (iii) CD127$^+$ KLRG1$^+$ double-positive cells, and (iv) CD127$^-$ KLRG1$^-$ double-negative cells (Fig. 2G, left), as described previously (*5*, *19*). The RNA expression profile of these four FACS-sorted populations was analyzed using Affymetrix microarrays. We tested each of the four subsets (test class) against a combination of the three other FACS-sorted subpopulations, using high-throughput GSEA for the CD8$^+$ T naïve, memory, effector, and stem cell–like/memory signatures (Fig. 2G, lower right). As expected, sorted memory precursors were enriched for both memory and naïve signatures, whereas sorted effectors were enriched for the effector signature (Fig. 2G, right). Notably, only sorted memory precursors from wild-type mice were enriched for the CD8$^+$ T stem cell–like signature when compared to the three other sorted subpopulations (Fig. 2G, lower right). Thus, the stem cell–like memory signature, which is specific to memory precursors in wild-type T cells, is broadly enriched in bulk K^b-OVA$^+$ CD8$^+$ T cells in *Suv39h1*-KO mice.

H3K9me3 deposition at stem/memory-associated loci

To elucidate the epigenetic states (*11*) associated with the stem/memory and effector (*19*, *21*) gene loci (tables S1 and S2), we first analyzed a set of public chromatin immunoprecipitation sequencing (ChIP-seq) data sets, generated with naïve, memory precursor, and effector OT-I CD8$^+$ T cells isolated 8 days after LM-OVA infection in vivo (*24*). We identified the regions significantly enriched for monomethylated histone H3 Lys^4 (H3K4me1), trimethylated histone H3 Lys^4 (H3K4me3), acetylated histone H3 Lys^{27} (H3K27ac), and trimethylated histone H3 Lys^{27} (H3K27me3) histone marks. For each histone mark, the significantly enriched islands were associated with the nearest gene (*25*). Differences in histone mark relative enrichment were calculated by pairwise comparison of terminal effectors to naïve cells (Fig. 3A, top) or to memory precursors (Fig. 3A, bottom). The regions associated with the stem cell–like memory genes showed an enrichment in H3K4me1 and H3K4me3 in both naïve cells and memory precursors, as compared to terminal effectors. The enrichment in these active marks, H3K4me1 and H3K4me3, corresponds to the transcription level of the stem cell–like memory genes in both wild-type naïve and memory precursors, as compared to effectors (fig. S5B). In contrast, the profile of H3K27ac deposition, typically distal, was more difficult to associate with target genes, including certain stem/memory genes and most effector signature genes in effectors (Fig. 3A); of note, cell cycle genes were excluded from this signature (table S2). In line with these results, H3K4me1 and H3K27ac marked the effector genes mostly in the effector subset (Fig. 3A). The enrichment of the repressive mark H3K27me3 on stem cell–like memory genes corresponded with the silencing of the gene signature in the effector CD8$^+$ T cells (Fig. 3A and fig. S5B). Thus, post-translational chromatin modifications corresponding to transcriptionally active and repressive marks follow the stem cell–like memory and effector gene expression patterns defined in memory precursors and effector T cell populations, respectively.

To investigate the contribution of Suv39h1-dependent chromatin changes, we performed ChIP-seq for the repressive mark H3K9me3 in naïve and in vitro differentiated effector CD8$^+$ T cells from wild-type and *Suv39h1*-KO mice (fig. S6A). The total numbers of H3K9me3 islands and targeted genes were higher in wild-type effectors than in wild-type naïve CD8$^+$ T cells (Fig. 3B and fig. S6B). Most de novo H3K9me3 islands (present in effector but not in naïve T cells) were distal to the transcription start site (TSS), with equal proportions of intra- and intergenic locations (Fig. 3B, top), and with highest enrichment between 10 and 100 kb from the TSS (fig. S6C). Thus, H3K9me3 is broadly distributed in both wild-type naïve and effector CD8$^+$ T cells, and the number of H3K9me3 islands increases in effector CD8$^+$ T cells relative to naïve CD8$^+$ T cells.

Among the 997 genes that are less efficiently silenced in vivo in *Suv39h1*-KO cells relative to littermates (Fig. 2C), 145 genes were decorated by H3K9me3 (fig. S6C). Several of these genes encode immune and stem/memory-related proteins, including CD127 or CD62L (Fig. 3C). In effectors, the *Il7r* gene acquires H3K9me3 at sites both proximal and distal to the TSS, including the promoter, previously described enhancers, introns, and intergenic regions (Fig. 3C, upper left) (*26*). H3K9me3 deposition is also increased at the *Sell* locus, upstream the promoter region, in effectors as compared to naïve CD8$^+$ T cells (Fig. 3C, lower left). No significantly enriched H3K9me3 islands were found for actively transcribed housekeeping and control genes (Fig. 3C, right). These results suggest that H3K9me3 silences genes linked to naïve versus effector differentiation, including *Il7r* and *Sell*.

To evaluate whether these stem cell–like memory gene loci are direct targets of Suv39h1, we compared the relative (fold change) enrichment of H3K9me3 islands detected in effectors to stem cell/memory gene expression in wild-type and *Suv39h1*-KO K^b-OVA$^+$ CD8$^+$ T cells (Fig. 3D). H3K9me3 enrichment negatively correlated with mRNA expression of stem cell/memory signature genes in wild-type but not in *Suv39h1*-KO K^b-OVA$^+$ CD8$^+$ T cells (Fig. 3D). In contrast, none of the H3K9me3 domains were associated with effector signature genes. Thus, the silencing of stem cell–like memory genes correlated with significant H3K9me3 enrichment in wild-type but not in *Suv39h1*-KO T cells, in which increased gene expression correlated with reduced or absent H3K9me3 deposition.

To validate and refine these results, we used μChIP–quantitative polymerase chain reaction (qPCR) analysis of H3K9me3 at critical gene loci in naïve (CD44^{lo} CD62L$^+$ CD127$^+$), central memory (CD44^{hi} CD62L$^+$ CD127$^+$), and effector (dump$^-$ CD44^{hi} CD127$^{lo/-}$ KLRG1$^+$ K^b-OVA$^+$) CD8$^+$ T cells purified 7 days after LM-OVA infection. Naïve cells showed barely detectable H3K9me3 at both the *Il7r* and *Sell* promoters (Fig. 3E and fig. S6D), whereas in effector T cells, the levels of H3K9me3 were increased at both loci, again in correlation with gene silencing (Fig. 3E). Consistent with the expression results, central memory cells, like naïve cells, have low levels of H3K9me3 at both promoters, whereas the reduced expression of *Il7r* and *Sell* in effector cells correlated with H3K9me3 enrichment.

Suv39h1-KO effector CD8$^+$ T cells did not show a significant increase of H3K9me3 at the *Il7r* or *Sell* loci, consistent with impaired silencing of *Il7r* and *Sell* expression (Figs. 2D and 3E). Likewise, CD8$^+$ central memory T cells and naïve cells presented low levels of H3K9me3 at both the *Sell* and *Il7r* promoters (Fig. 3E). Also consistent with these results, the level of the alternative active mark H3K9ac is increased in effectors at both *Il7r* and *Sell* loci in *Suv39h1*-KO CD8$^+$ T cells alone, which was again consistent with gene expression profiles (Fig. 3F). These results indicate that Suv39h1 dynamically decorates genes encoding important regulators of CD8$^+$ T cell stem/memory fate with H3K9me3, and that these genes are silenced in wild-type CD44^{hi} CD127$^{lo/-}$ KLRG1$^+$ K^b-OVA$^+$ CD8$^+$ T effectors. In *Suv39h1*-KO antigen-specific CD8$^+$ T cells, these stem/memory genes fail to acquire the repressive mark, resulting in defective silencing.

Long-lasting memory and short-lived effector CD8$^+$ T cell differentiation

These gene expression results suggest that the *Suv39h1* defect may affect memory versus effector differentiation in vivo. No major differences between wild-type and *Suv39h1*-KO were observed in terms of CD44, CD122, and PD1 expression patterns (fig. S7A) in blood K^b-OVA$^+$ CD8$^+$ T cells 7 days after LM-OVA infection. In contrast, CD127 expression was increased in both memory precursors (KLRG1$^-$) and effector cells (KLRG1$^+$) in *Suv39h1*-KO mice relative to

littermates (Fig. 4A). The increased proportion of CD127^{+} cells was due to decreased numbers of CD127$^{lo/-}$ KLRG1^{+} effector T cells (Fig. 4, A and B, right). The absolute numbers of CD127^{lo} cells (both KLRG1^{+} and KLRG1^{-}) were reduced, whereas the number of CD127^{high} memory precursor cells was unchanged (Fig. 4, A and B, left). In *Suv39h1*-KO mice, K^{b}-OVA^{+} CD8^{+} T cells also included a lower proportion of CD27^{lo} cells and impaired down-regulation of CD62L in a subpopulation of cells (fig. S7, A and B). Enhanced memory differentiation of wild-type/*Suv39h1*-KO 1:1 mixed bone marrow chimeras upon LM-OVA infection resulted in a significantly higher proportion of effectors among wild-type cells, whereas the proportion of memory precursors was increased among *Suv39h1*-KO K^{b}-OVA^{+} T cells (Fig. 4C and fig. S8). Similar results were obtained after the adoptive transfer of wild-type and *Suv39h1*-KO OT-I CD8^{+} T cells (fig. S3, D and E). Thus, *Suv39h1*-KO mice develop increased proportions of memory T cells in response to *L. monocytogenes* infection as the result of a T cell–intrinsic defect in effector differentiation.

We next sought to determine the impact of the *Suv39h1* defect on long-term memory versus

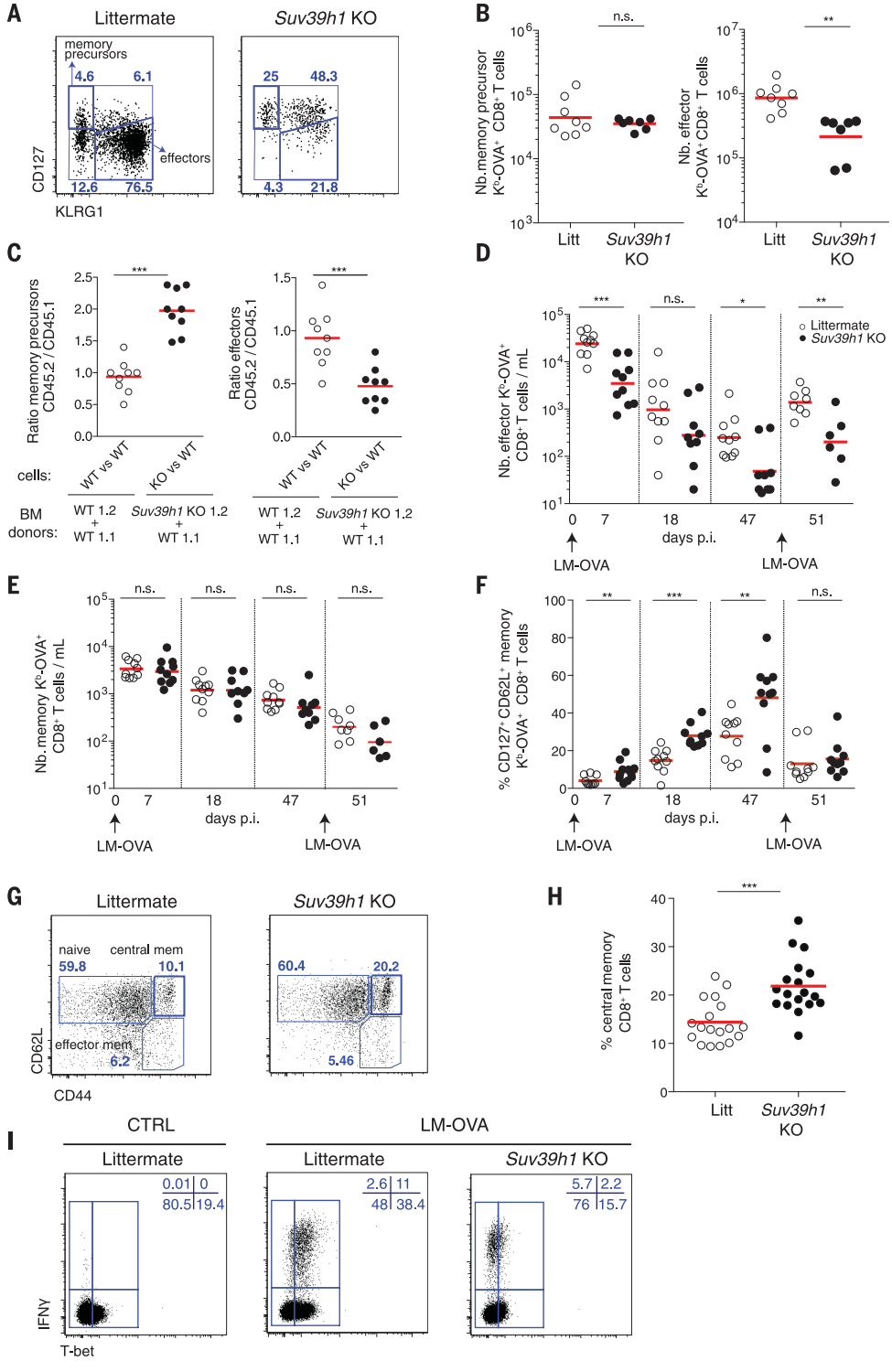


Fig. 4. Suv39h1 is a critical regulator of peripheral effector versus memory CD8^{+} T differentiation. (**A**) Littermate and *Suv39h1*-KO mice were infected with LM-OVA. Seven days later, dump^{-} CD44^{hi} K^{b}-OVA^{+} CD8^{+} T cells in the spleen were analyzed according to CD127 and KLRG1 staining (CD127^{hi} KLRG1^{-} memory precursors and CD127$^{lo/-}$ KLRG1^{+} differentiated effector cells). Representative dot plots are shown; numbers represent the percentages. (**B**) Numbers of memory precursors and effectors in the spleen were measured. (**C**) Wild-type and *Suv39h1*-KO mixed bone marrow chimeras were infected with LM-OVA, and 7 days later the ratio between memory precursors and short-lived effector CD8^{+} T cells was evaluated in peripheral blood using K^{b}-OVA^{+} multimers. (**D** to **F**) Mice were infected with LM-OVA on day 0 (primary infection) and day 48 (secondary infection). On days 7, 18, 47, and 51 p.i. (with respect to the primary infection), numbers of short-lived effectors (D) and central memory cells (E) on gated dump^{-} CD44^{hi} K^{b}-OVA^{+} CD8^{+} T cells per ml of peripheral blood were measured. (F) Longitudinal analysis of the dynamics of central memory K^{b}-OVA^{+} CD8^{+} T cells. (**G**) Representative dot plots of gated blood CD8^{+} T lymphocytes; numbers represent percentages. (**H**) Percentage of blood endogenous polyclonal central memory CD8^{+} T cells. (**I**) Peripheral blood cells from LM-OVA–infected littermate and *Suv39h1*-KO mice were stimulated ex vivo 7 days after infection with the OVA peptide SIINFEKL (OVA$_{257-264}$) and analyzed for intracellular T-bet expression and IFN-γ production. Representative dot plots are shown; numbers represent percentages. Data are shown as geometric means in (B), (D), and (E) or as means in (C), (F), and (H). *P < 0.05, **P < 0.01, ***P < 0.001 (Wilcoxon Mann-Whitney test).

effector differentiation and persistence. The numbers of K^b-OVA$^+$ CD8$^+$ T memory cells and effectors were analyzed 7, 18, and 47 days after LM-OVA infection, and 3 days after rechallenge on day 51. As expected from previous results (Fig. 4B), the number of effectors was reduced at all times, including at the peak of the rechallenge memory response in the blood (Fig. 4D) and spleen (fig. S9, right). The numbers of memory cells did not change between wild-type and *Suv39h1*-KO at the peak of the response, during contraction, or after rechallenge (Fig. 4E and fig. S9, left). The proportion of central memory (CD44^{hi} CD62L$^+$ CD127$^+$ K^b-OVA$^+$) CD8$^+$ T cells (*27*) increased over time in *Suv39h1*-KO mice as compared to wild-type mice (Fig. 4F). Notably, the percentage of endogenous subsets of memory CD8$^+$ T cells in *Suv39h1*-KO mice was also higher, including polyclonal CD62L$^+$ central memory CD8$^+$ T cells in the blood and secondary lymphoid organs (Fig. 4, G and H, and fig. S10, A to C), CD62L$^-$ effector memory CD8$^+$ T cells in the bone marrow and blood (fig. S10D), and tissue-resident CD8$^+$ T cells in the liver (fig. S10, F to H). The endogenous *Suv39h1*-KO central memory CD8$^+$ T cells also expressed increased levels of memory and stem cell–like memory markers SCA-1 and CD95 (fig. S10E) (*22*, *27*, *28*) Thus, *Suv39h1*-KO mice show higher levels of central and effector memory CD8$^+$ T cells both before and after LM-OVA challenge.

An analysis of master regulators involved in memory and effector differentiation showed that the proportion of LM-OVA–specific T cells expressing T-bet is reduced in both IFN-γ^+ and IFN-γ^- *Suv39h1*-KO cells as compared to littermates (Fig. 4I and fig. S11). Similarly, the expression of Eomes, Blimp1, and Bcl6 was also reduced (fig. S11). Thus, *Suv39h1*-KO CD8$^+$ T cells display a central memory–like phenotype but express reduced levels of both effector and memory transcription master regulators (*4*, *5*).

Previous adoptive transfer experiments showed that only CD127$^+$ memory precursors give rise to long-term memory cells and confer protective immunity (*29*). To evaluate their in vivo memory self-renewal and differentiation properties, we isolated wild-type and *Suv39h1*-KO CD45.2 K^b-OVA$^+$ CD8$^+$ T cells 7 days after LM-OVA infection and adoptively transferred them at low numbers into naïve congenic CD45.1 recipients (Fig. 5A and fig. S12A). Forty days after challenge with LM-OVA, as expected, few wild-type K^b-OVA$^+$ T cells persisted or responded to the LM-OVA challenge. In contrast, donor *Suv39h1*-KO CD45.2 K^b-OVA$^+$ T cells were clearly present and responded to the infection (Fig. 5, B and C), with an increased proportion of memory precursor and effector cells (Fig. 5, B and D, left). To further evaluate the self-renewal properties of wild-type and *Suv39h1*-KO central memory T cells, we isolated total CD45.2$^+$ CD44^{hi} CD62L$^+$ CD127$^+$ KLRG1$^-$ K^b-OVA$^-$ central memory T cells, harvested 7 days after LM-OVA infection, and transferred them into naïve CD45.1 congenic mice (Fig. 5A and fig. S12A). Thirty-nine days after adoptive transfer, similar numbers of wild-type and *Suv39h1*-KO donor CD8$^+$ T cells, which had maintained a central memory phenotype, were present in the blood (Fig. 5G, left, and fig. S12B). Four days after LM-OVA rechallenge, however, both the percentage and the total numbers of *Suv39h1*-KO donor CD8$^+$ T cells were increased relative to control *Suv39h1*-sufficient cells (Fig. 5, E and G, right). Donor *Suv39h1*-KO CD8$^+$ T cells displayed higher expression of both CD127 and CD62L, lower levels of KLRG1 and CD44, a slight decrease in the expression of SCA-1, and similar levels of CD122 (Fig. 5, E and F). The donor memory subset was increased in the animals adoptively transferred with *Suv39h1*-KO T cells (Fig. 5H). Thus, *Suv39h1*-KO K^b-OVA$^+$ and central memory CD8$^+$ T cells have superior self-renewal and repopulation potential relative to their wild-type counterparts.

Stemness gene silencing in terminal effectors requires Suv39h1

Increased proportions of T cells with a central memory phenotype and stem cell–like properties were found in *Suv39h1*-KO mice. This phenotype could be due to the accumulation of a defined population of stem cell–like memory T cells, or to the expression of stem cell–related genes across different T cell subpopulations. We used single-cell RNA sequencing (scRNA-seq) to explore and dissect the heterogeneity of wild-type and *Suv39h1*-KO K^b-OVA$^+$ CD8$^+$ T lymphocytes. Purified naïve and K^b-OVA$^+$ CD8$^+$ T cells from wild-type and *Suv39h1*-KO mice were isolated by FACS 7 days after LM-OVA infection and processed for scRNA-seq (fig. S13). For naïve cells, we sequenced 1102 and 991 cells from wild-type and *Suv39h1*-KO mice, respectively. For antigen-specific K^b-OVA$^+$ CD8$^+$ T cells, we processed two technical replicates for both wild-type and *Suv39h1*-KO infected mice (approximately 1200 and 1000 cells, respectively) and an additional biological replicate (from different mice, 404 wild-type and 283 *Suv39h1*-KO K^b-OVA$^+$ CD8$^+$ T cells). The cells from the two technical replicates were pooled and used for the rest of the analysis. A comparison between technical and biological replicates within wild-type and *Suv39h1*-KO mice showed a strong correlation (figs. S13 to S15).

A principal components analysis (PCA) of differentially expressed genes by wild-type and *Suv39h1*-KO naïve and K^b-OVA$^+$ CD8$^+$ T cells, is visualized as a set of *t*-distributed stochastic neighbor embedding (*t*-SNE) plots in Fig. 6A. In wild-type cells, unsupervised clustering of naïve and K^b-OVA$^+$ CD8$^+$ T cells revealed eight clusters in wild-type T cells and six clusters in *Suv39h1*-KO antigen-specific K^b-OVA$^+$ CD8$^+$ T cells (fig. S14, A and B, and tables 3 and 4). On the basis of distinct transcription profiles, we grouped the unsupervised clusters into four major subset categories: naïve, memory precursor, effector, and cycling cells (Fig. 6, B to D, and fig. S14C). As expected, the naïve cells grouped in a homogeneous category characterized by the highest expression of *Sell*, *Ccr7*, and *Tcf7* (Fig. 6B). The memory precursors were enriched in *Il7r*, *Cxcr3*, *Cd27*, *Cd28*, and *Ly6a* expression (Fig. 6, B and E); the effectors were characterized by *Zeb2*, *Klrg1*, and *granzyme B* (Fig. 6, B and E, and fig. S16); and the cycling subsets were defined by cell cycle genes including *Pcna* and *Mcm5* (Fig. 6, B and E, right). Although effectors represented the majority of multimer-positive CD8$^+$ T cells when analyzed by FACS (CD127$^-$ KLRG1$^+$), the terminally differentiated effectors only represented approximately 30% of K^b-OVA$^+$ CD8$^+$ T cells in the categories defined by scRNA-seq (Fig. 6D). This is likely due to the resolution of the cycling cells, which also express high levels of different effector markers, including KLRG1 and granzyme A and B (Fig. 6, B and E, and fig. S16). Consistent with flow cytometric analysis, the proportion of *Suv39h1*-KO effector K^b-OVA$^+$ CD8$^+$ T cells was decreased (Fig. 6D and fig. S14). Thus, unsupervised scRNA-seq allows resolution of the expected memory precursors and effector populations among K^b-OVA$^+$ CD8$^+$ T cells. The scRNA-seq analysis also reveals a population expressing high levels of genes involved in the cell cycle, as well as markers of memory precursors and effectors.

Having defined the different populations of antigen-specific CD8$^+$ T cells, we next sought to analyze the expression of stem cell/memory genes in wild-type and *Suv39h1*-KO cells. As expected, wild-type cells showed enrichment of stem cell markers (i.e., *Cxcr6*, *Rnf138*, *Il18r1*, and *Traf1*) in the memory precursors (Fig. 6F). In *Suv39h1*-KO T cells, the expression of these markers was also high in memory precursors, but, in contrast to wild-type cells, expression was also detected in effector and cycling populations. Similar increased expression of the stem/memory signature in *Suv39h1*-KO effectors could be visualized on *t*-SNE plots where the number of genes from the stem/memory signature per cell is represented by a color scale (Fig. 6G). These results suggest that *Suv39h1*-defective effectors express higher levels of certain stem/memory-related genes, consistent with impaired silencing.

To further analyze the expression of the stem cell–like memory and effector signature genes at the single-cell level among the different populations, we used density scatterplots in which the numbers of genes from each signature are represented against each other (Fig. 6H). In most naïve cells from both wild-type and *Suv39h1*-KO mice, we detected an average of seven stem/memory signature genes per cell, and fewer than four genes from the effector signature. As expected, in the wild type, memory precursors, relative to naïve cells, expressed higher numbers of genes from the stem/memory signature (mean of 9 genes per cell; 43% with more than 10 genes per cell) as well as high numbers of genes from the effector signature (mean of 18 genes per cell). The proportion of effector cells expressing high numbers (>10 per cell) of stem/memory genes was reduced (16%) as compared with memory cells (43%), whereas there was only a modest increase in the number of genes from the effector signature (mean 20 genes). Of note, the gene expression levels from the effector signature were strongly increased in effectors as compared to memory cells (Fig. 6B and fig. S16). Notably, cycling cells coexpressed either low or high numbers of

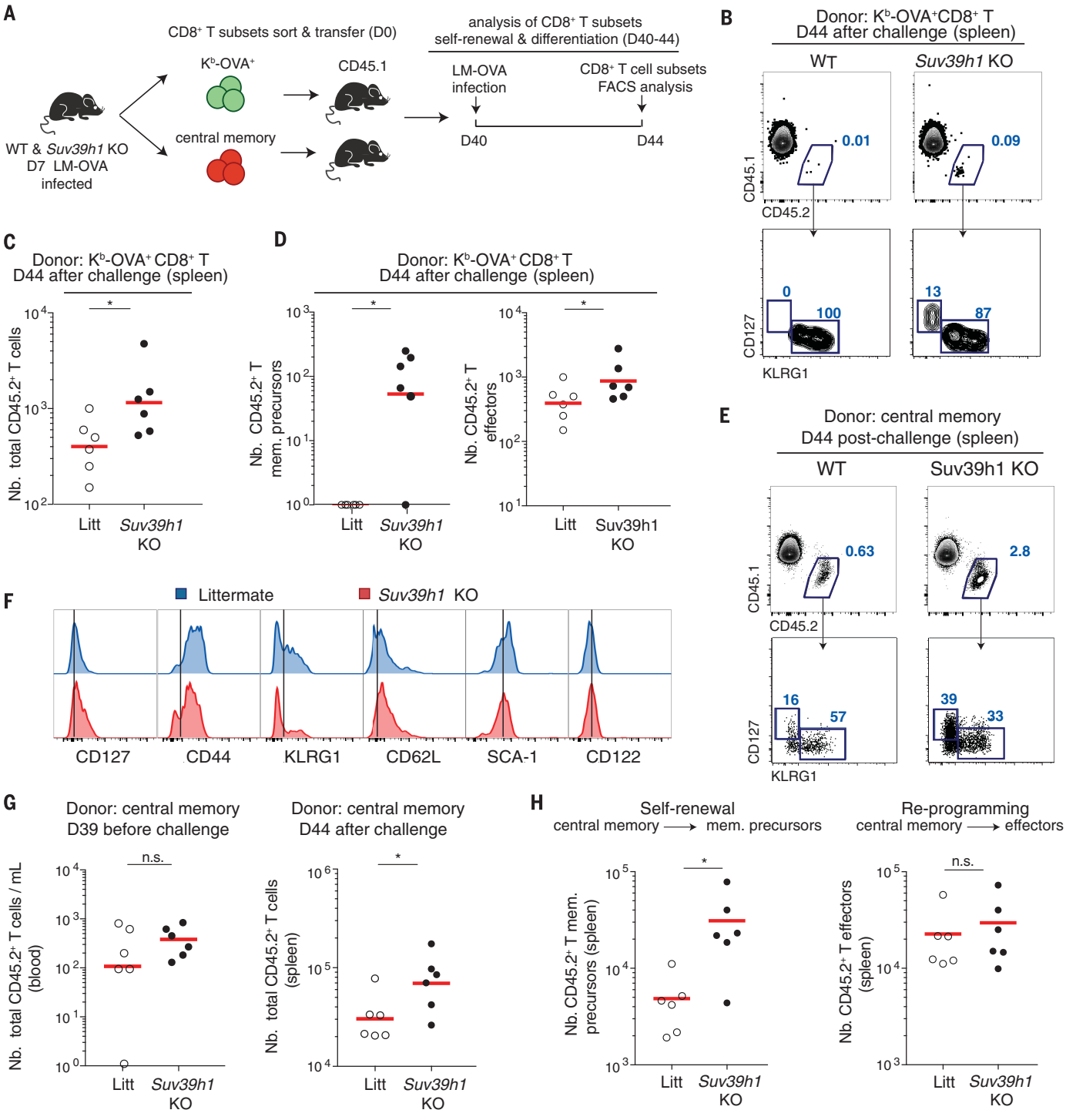


Fig. 5. Suv39h1-mediated defective silencing during CD8$^+$ T cell lineage commitment results in the accumulation of stem cell–like central memory cells. (**A**) Experimental design. Dump$^-$ K^b-OVA$^+$ CD8$^+$ T cells or CD44^{hi} CD62L$^+$ CD127$^+$ K^b-OVA$^-$ central memory CD8$^+$ T cells, isolated 7 days after LM-OVA infection from littermate or *Suv39h1*-KO mice, were adoptively transferred to naïve congenic CD45.1$^+$ recipient mice. The recipients were challenged with LM-OVA 40 days later. (**B**) Donor CD45.2$^+$ K^b-OVA$^+$ CD8$^+$ T cells were analyzed in the spleen 4 days after infection. Representative dot plots and a phenotypic analysis of subsets are shown; numbers represent percentages. (**C**) Numbers of total donor CD45.2$^+$ CD8$^+$ T cells. (**D**) Numbers of donor CD127$^+$ KLRG1$^-$ memory precursors and CD127$^{lo/-}$ KLRG1$^+$ effectors from donor K^b-OVA$^+$ CD8$^+$ T cells. (**E**) Donor CD45.2$^+$ central memory CD8$^+$ T cells were analyzed in the spleen 4 days after LM-OVA challenge. Representative dot plots with subset phenotypic analyses are shown; numbers represent percentages. (**F**) Representative histograms of stem cell–like and memory markers on total gated donor central memory CD8$^+$ T cells analyzed 4 days after LM-OVA challenge. (**G**) Total numbers of adoptively transferred donor central memory cells, before and after LM-OVA challenge. (**H**) Numbers of donor CD127$^+$ KLRG1$^-$ memory precursor and CD127$^{lo/-}$ KLRG1$^+$ effector CD8$^+$ T cells differentiated from donor central memory CD8$^+$ T cells. Graphs are representative of two experiments with three mice per group. All graphs show geometric means. *P < 0.05 (Wilcoxon Mann-Whitney test).

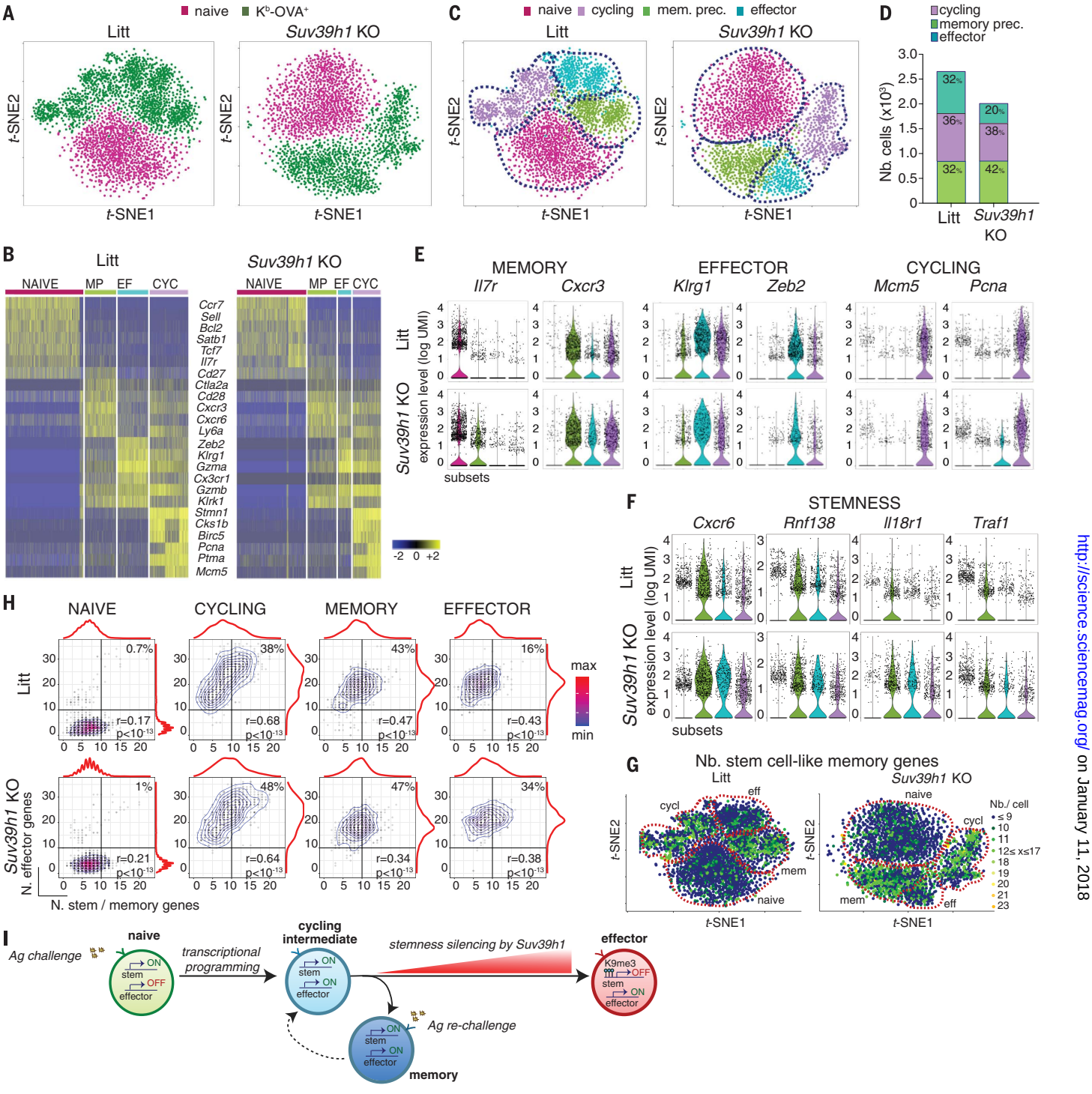


Fig. 6. Stem cell–like/memory genes are expressed by memory precursors and cycling CD8$^+$ T intermediates, and silenced by Suv39h1 in terminal effectors. (**A** to **H**) scRNA-seq analysis of naïve and K^b-OVA$^+$ CD8$^+$ T cells isolated from wild-type and *Suv39h1*-KO mice 7 days after LM-OVA infection. (A) Graph-based clustering by *t*-SNE projection of naïve and K^b-OVA$^+$ CD8$^+$ T cells isolated from wild-type and *Suv39h1*-KO mice. The colors indicate sorted CD8$^+$ T cell subsets. Each dot represents an individual cell. (B) Heat maps of six representative genes for each subset category identified. Columns represent cells; rows represent genes. The color scale is based on z-score distribution. (C) *t*-SNE projection of single cells, colored according to four major subset categories defined by semi-supervised clustering based on specific and distinct gene expression profiles. (D) Numbers and percentages of cells for each category. Violin plots in (E) and (F) show expression distribution of memory, effector, cycling, and stem cell/memory representative markers. (G) *t*-SNE projection of single cells, showing the number of stem cell–like memory genes expressed in each cell. (H) Density scatterplots represent numbers of effector versus stem cell/memory genes expressed in individual cells for each subset category. Color scale indicates gene density; marginal distributions show memory/stem (*x* axis) and effector (*y* axis) gene number distributions. Pearson linear regression is displayed for each plot. (**I**) Working model depicting the pivotal role of Suv39h1 during CD8$^+$ T cell lineage differentiation and commitment.

genes from both signatures. Thus, scRNA-seq analysis reveals the alternative expression of the stem/memory and effector signatures in the two cell types, respectively, and the concomitant low or high expression of these two signatures in the cycling cells. These results suggest that cycling cells may represent bipotent differentiation intermediates expressing both effector and stem/memory potential. Furthermore, the commitment to effector differentiation paths appears to be acquired by the silencing of stem/memory genes.

The total number of unique molecular identifiers (UMI) measured in each subset category did not differ between wild-type and *Suv39h1*-KO cells (fig. S17, A and B). Naïve, cycling, and memory *Suv39h1*-KO cells bear similar patterns of gene expression signatures as compared to wild-type cells. In contrast, a significant difference was observed in effector cells, in which the numbers of stem/memory genes per cell were increased relative to effector cells from wild-type mice. The proportion of cells expressing more than 10 genes from the stem/memory signature was augmented from 16% in the wild type to 34% in effector *Suv39h1*-KO cells. Thus, rather than a specific subpopulation of stem/memory cells accumulating in the *Suv39h1*-KO mice, the expression of stem/memory-related genes was derepressed mainly in *Suv39h1*-KO effector T cells.

Conclusions

We argue that after priming, cycling $CD8^+$ T lymphocytes reprogram both self-renewing and effector gene expression profiles (Fig. 6I). These cycling cells may represent bipotent intermediates, which would then repress either the effector or stem cell/memory programs while they differentiate to memory precursors or effectors, respectively (Fig. 6I). The silencing of the stem cell/memory gene expression program is under the control of Suv39h1 by imposing the H3K9me3 modification on chromatin at the corresponding loci. In doing so, Suv39h1/H3K9me3 would establish an epigenetic barrier on the stem/memory gene expression program, preventing effector reprograming into memory cells (Fig. 6I). It is most likely that the possibly reversible silencing of effector gene expression in memory cells occurs through other mechanisms, as memory cells do effectively reprogram into effectors upon rechallenge. These results open new perspectives for the manipulation of epigenetic programming of T lymphocyte identity in the context of vaccination, checkpoint-based immunotherapies, and adoptive T cell therapies.

REFERENCES AND NOTES

1. A. O. Kamphorst, K. Araki, R. Ahmed, *Vaccine* **33** (suppl. 2), B21–B28 (2015).
2. M. J. Bevan, *Nat. Immunol.* **12**, 463–465 (2011).
3. R. Ahmed, M. J. Bevan, S. L. Reiner, D. T. Fearon, *Nat. Rev. Immunol.* **9**, 662–668 (2009).
4. J. T. Chang, E. J. Wherry, A. W. Goldrath, *Nat. Immunol.* **15**, 1104–1115 (2014).
5. S. M. Kaech, W. Cui, *Nat. Rev. Immunol.* **12**, 749–761 (2012).
6. H. Y. Shih *et al.*, *Immunol. Rev.* **261**, 23–49 (2014).
7. N. P. Weng, Y. Araki, K. Subedi, *Nat. Rev. Immunol.* **12**, 306–315 (2012).
8. C. Maison, J. P. Quivy, A. V. Probst, G. Almouzni, *Cold Spring Harb. Symp. Quant. Biol.* **75**, 155–165 (2010).
9. G. Almouzni, A. V. Probst, *Nucleus* **2**, 332–338 (2011).
10. B. D. Fodor, N. Shukeir, G. Reuter, T. Jenuwein, *Annu. Rev. Cell Dev. Biol.* **26**, 471–501 (2010).
11. C. D. Allis, T. Jenuwein, *Nat. Rev. Genet.* **17**, 487–500 (2016).
12. C. Maison, G. Almouzni, *Nat. Rev. Mol. Cell Biol.* **5**, 296–304 (2004).
13. A. H. Peters *et al.*, *Cell* **107**, 323–337 (2001).
14. A. V. Probst, E. Dunleavy, G. Almouzni, *Nat. Rev. Mol. Cell Biol.* **10**, 192–206 (2009).
15. T. T. Onder *et al.*, *Nature* **483**, 598–602 (2012).
16. S. P. Bradley, D. A. Kaminski, A. H. Peters, T. Jenuwein, J. Stavnezer, *J. Immunol.* **177**, 1179–1188 (2006).
17. R. S. Allan *et al.*, *Nature* **487**, 249–253 (2012).
18. E. G. Pamer, *Nat. Rev. Immunol.* **4**, 812–823 (2004).
19. E. J. Wherry *et al.*, *Immunity* **27**, 670–684 (2007).
20. I. Hoebeke *et al.*, *Leukemia* **21**, 311–319 (2007).
21. C. J. Luckey *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **103**, 3304–3309 (2006).
22. L. Gattinoni *et al.*, *Nat. Med.* **17**, 1290–1297 (2011).
23. L. Spinelli, S. Carpentier, F. Montañana Sanchis, M. Dalod, T. P. Vu Manh, *BMC Genomics* **16**, 814 (2015).
24. B. Yu *et al.*, *Nat. Immunol.* **18**, 573–582 (2017).
25. G. Wei *et al.*, *Immunity* **30**, 155–167 (2009).
26. M. L. Miller *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **111**, 7397–7402 (2014).
27. F. Sallusto, D. Lenig, R. Förster, M. Lipp, A. Lanzavecchia, *Nature* **401**, 708–712 (1999).
28. P. Graef *et al.*, *Immunity* **41**, 116–126 (2014).
29. S. M. Kaech *et al.*, *Nat. Immunol.* **4**, 1191–1198 (2003).

ACKNOWLEDGMENTS

We thank all the U932 and UMR3664 members; C. Maison, N. Lacoste, and F. X. Gobert for helpful discussions; A. Cros and V. Mondin for technical help; B. Amati and J. Sedat for critical reading of the manuscript; the flow cytometry platform, the genomic and animal facilities, the Nikon Imaging Center of Curie Institute; and the PICT-IBISA@Pasteur Imaging facility. We thank T. Jenuwein for providing the *SUV39H1* gene and the *Suv39h1*-KO mice. LM-OVA was kindly provided by H. Shen. Supported by the Institute Curie, Institut National de la Santé et de la Recherche Médicale, Centre National de la Recherche Scientifique, ANR "ChromaTin" grant ANR-10-BLAN-1326-03, and "EPICURE" grant ANR-14-CE16-0009 (S.A. and G.A.); la Ligue Contre le Cancer (Equipe labelisée Ligue, EL2014.LNCC/SA) and Association de Recherche Contre le Cancer grant ERC (2013-AdG N° 340046 DCBIOX) (S.A.); la Ligue Nationale contre le Cancer (Equipe labelisée Ligue), the European Commission Network of Excellence EpiGeneSys (HEALTH-F4-2010-257082), ERC Advanced Grant 2009-AdG_20090506 "Eccentric," the European Commission large-scale integrating project FP7_HEALTH-2010-259743 "MODHEP," ANR-11-LABX-0044_DEEP, ANR-10-IDEX-0001-02 PSL, ANR "CHAPINHIB" ANR-12-BSV5-0022-02, ANR "CELLECTCHIP" ANR-14-CE10-0013, and the Aviesan-ITMO cancer project "Epigenomics of breast cancer" (G.A.); and SiRIC/INCa grant INCa-DGOS-4654 (J.J.W.). High-throughput sequencing was performed by the ICGex NGS platform of the Institut Curie supported by the grants SESAME (région Ile-de-France), ANR-10-EQPX-03, and ANR-10-INBS-09-08. The bioinformatics data are available in the Gene Expression Omnibus (GEO) database under accession numbers GSE105163 (microarrays) and SuperSeries GSE106268 for ChIP-seq and scRNA-seq data (linked to subseries GSE106264, GSE106265, and GSE106267). The authors have no conflicting financial interests. The data are tabulated in the main paper and in the supplementary materials. Author contributions: L.P. conceived and designed the project, carried out experimental work, and wrote the manuscript; C.G. contributed to bioinformatics design, carried out bioinformatics work, and interpreted data; E.Z. contributed to project design, carried out experimental work, and interpreted data; P.G. carried out bioinformatics work and participated in data interpretation; N.B. carried out experimental work; J.J.W. participated in bioinformatics data analysis and interpretation; J.-P.Q. participated in project design and data interpretation and wrote the manuscript; G.A. designed the project and wrote the manuscript; and S.A. conceived and designed the project and wrote the manuscript. L.P., C.G., and S.A. are inventors on patent application EP17305757.1 submitted by Institut Curie that covers the targeting of Suv39h1 in the context of T cell adoptive transfer.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/359/6372/177/suppl/DC1
Materials and Methods
Figs. S1 to S17
Tables S1 to S5
References (*30*–*42*)

27 July 2016; resubmitted 1 August 2017
Accepted 16 November 2017
10.1126/science.aah6499

SOLID-STATE PHYSICS

Coherent band excitations in $CePd_3$: A comparison of neutron scattering and ab initio theory

Eugene A. Goremychkin,[1] Hyowon Park,[2,3] Raymond Osborn,[2]* Stephan Rosenkranz,[2] John-Paul Castellan,[2,4] Victor R. Fanelli,[5] Andrew D. Christianson,[6] Matthew B. Stone,[6] Eric D. Bauer,[7] Kenneth J. McClellan,[7] Darrin D. Byler,[7] Jon M. Lawrence[7,8]

In common with many strongly correlated electron systems, intermediate valence compounds are believed to display a crossover from a high-temperature regime of incoherently fluctuating local moments to a low-temperature regime of coherent hybridized bands. We show that inelastic neutron scattering measurements of the dynamic magnetic susceptibility of $CePd_3$ provides a benchmark for ab initio calculations based on dynamical mean field theory. The magnetic response is strongly momentum dependent thanks to the formation of coherent f-electron bands at low temperature, with an amplitude that is strongly enhanced by local particle-hole interactions. The agreement between experiment and theory shows that we have a robust first-principles understanding of the temperature dependence of f-electron coherence.

The Anderson impurity model, used to describe magnetic impurities in metals, formulates the interaction of localized f-electron orbitals with more delocalized d-electron bands through an onsite hybridization, whose strength is typically represented by a single energy scale, the Kondo temperature, T_K (*1*). It has been successfully applied to intermediate valence compounds, such as $CePd_3$ and $CeSn_3$, even though the f electrons in these materials are not on isolated impurities but sit on a periodic lattice (*2*). Nevertheless, various deviations from the expected behavior of single impurities at temperatures well below T_K have long been interpreted as evidence of the formation of coherent f-electron bands, with strongly renormalized quasiparticle masses (*3*), an interpretation that was supported by phenomenological theory (*4*, *5*).

The concept of a crossover from coherent quasiparticles at low temperature to incoherent electronic fluctuations at high temperature is important in strongly correlated electron systems, whether in the context of heavy fermions (*6*, *7*), intermediate valence compounds (*8*), high-temperature superconductors (*9*), or "bad" metals close to a Mott transition (*10*). Although there is an extensive body of theoretical work predicting a gradual loss of quasiparticle spectral weight at the Fermi energy with increasing temperature, it has only recently become possible to perform realistic calculations by combining density functional theory with dynamical mean field theory (DFT+DMFT) in order to include both strong local correlations and itinerant band structures on an equal footing (*11*, *12*). Angle-resolved photoemission spectroscopy has provided experimental evidence of a loss of quasiparticle coherence with increasing temperature in heavy fermions (*13*, *14*), but quantitative comparison with DFT+DMFT is limited by the need to correct for matrix elements and the difficulty of combining high-energy resolution with broad momentum coverage. Furthermore, many of the relevant one-electron states are unoccupied, particularly in cerium compounds, and so inaccessible to photoemission measurements.

An alternative spectroscopic probe of quasiparticle coherence is inelastic neutron scattering. The neutron cross section, or scattering law, $S(\mathbf{Q},\omega)$, is proportional to the dynamic magnetic susceptibility, $\chi''(\mathbf{Q},\omega)$, of band electrons (*15*). For noninteracting electrons, this is derived from the Lindhard susceptibility, whose imaginary part is proportional to the joint density of states of the one-electron bands.

$$\chi_0''(\mathbf{Q},\omega) \propto \sum_{\mathbf{k}} f_{\mathbf{k}+\mathbf{Q}}(1-f_{\mathbf{k}})\delta(E_{\mathbf{k}+\mathbf{Q}}-E_{\mathbf{k}}-\omega) \quad (1)$$

The resulting scattering intensity would be enhanced at momentum transfers, $\mathbf{Q}$, and energy transfers, ω, that connect regions of high densities of state in the single-electron bands, $E_{\mathbf{k}}$, whose states are occupied with probability $f_{\mathbf{k}}$. Neutrons therefore probe both occupied and unoccupied states within the same measurements.

The advent of pulsed neutron sources, which have an enhanced flux of high-energy neutrons, stimulated interest more than 30 years ago in the possibility of using inelastic neutron scattering to study one-electron band structures (*16*, *17*), but the earliest estimates of the neutron cross section for weakly correlated electron bands were discouraging. Predicted signals were in the range 10^{-4} to 10^{-3} barns/steradians/eV, spread over wave vectors covering the entire Brillouin zone and energies up to the band width (*16*). Broad distributions of intensity have been challenging to measure at pulsed neutron time-of-flight spectrometers, where measurements were, until recently, made using a fixed sample geometry. Consequently, high-energy spectrometers have mostly been used to measure coherent excitations, such as spin waves in the copper oxide and iron-based superconductors (*18*, *19*). These also represent the magnetic excitations of band electrons, but they are easier to measure because strong interatomic exchange interactions generate poles in the dynamic susceptibility that yield well-defined peaks in the cross section.

Although measuring the electronic structure of weakly correlated electrons was not thought to be technically feasible with neutrons, strong intensity enhancements were predicted in strongly correlated electron systems—in particular, intermediate valence compounds (*16*). There has been increasing evidence over the past decade that compounds such as $CePd_3$ (*20*), $YbAl_3$ (*21*), and $CeInSn_2$ (*22*) show variations in $\chi''(\mathbf{Q},\omega)$ that could result from band excitations. So far, these interpretations have been based on conceptual models of f-d hybridized bands (*4*), because it was not possible to do a quantitative comparison within the experimental limitations of fixed-geometry neutron measurements. However, recently, band calculations (*23*) were shown to have qualitative consistency with earlier neutron scattering data (*20*).

We can now go beyond qualitative comparisons because of the advent of a new generation of inelastic neutron scattering spectrometers with large position-sensitive detectors (*24*, *25*) that allow efficient measurements of four-dimensional (4D) $S(\mathbf{Q},\omega)$ in single crystals by rotating the sample during the data collection. This can be accomplished either by measuring at discrete steps of the rotation angle (*26*) or by collecting the data continuously as the sample rotates (*27*). Both methods produce equivalent results that overcome the limitations of fixed-geometry measurements by measuring entire volumes of $(\mathbf{Q},\omega)$-space rather than a sparse set of hypersurfaces through that volume. Because the experimental data can be placed on an absolute scale by normalizing the intensity to a vanadium standard, it is possible to produce a direct comparison of experiment and theory.

With this instrumental capability, we have now performed a detailed quantitative comparison of neutron scattering measurements with a full calculation of $S(\mathbf{Q},\omega)$ using DFT+DMFT. As expected, the calculations show broad distributions of intensity with diffuse maxima at high-symmetry points that shift within the Brillouin zone as a function of energy transfer. These match the measured distributions of the dynamic susceptibility at low temperature with absolute cross sections that are within 20% of the theoretical predictions. Peaks in the dynamic susceptibility

[1]Frank Laboratory of Neutron Physics, Joint Institute for Nuclear Research, Dubna, Moscow Region, 141980, Russia. [2]Materials Science Division, Argonne National Laboratory, Argonne, IL 60439-4845, USA. [3]Department of Physics, University of Illinois at Chicago, Chicago, IL 60607, USA. [4]Institute for Solid State Physics, Karlsruhe Institute of Technology, D-76021 Karlsruhe, Germany. [5]Instrument and Source Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA. [6]Quantum Condensed Matter Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA. [7]Los Alamos National Laboratory, Los Alamos, NM 87545, USA. [8]Department of Physics and Astronomy, University of California, Irvine, CA 92697, USA.
*Corresponding author. Email: rosborn@anl.gov

at fixed momentum transfer are associated with values of **Q** and ω that connect relatively flat regions of the coherent quasiparticle bands. Although the **Q** dependence therefore results from the dispersion of coherent quasiparticles, the intensity is strongly amplified by local particle-hole interactions that are accurately predicted by two-particle vertex corrections within DFT+DMFT theory. At high temperature, the **Q** dependence is suppressed by the loss of quasiparticle spectral weight at low energies, in agreement with experiment, showing that the calculations realistically model the crossover to a regime of incoherent spin fluctuations.

Ab initio theory for the dynamic susceptibility of $CePd_3$

The dynamic susceptibility of $CePd_3$ was derived from a calculation of the one-electron Green's function using DFT+DMFT. This approach allows the incorporation of local correlations—i.e., Hund's rule and spin-orbit coupling as well as on-site Kondo screening—into realistic band structures based on DFT. Figure 1 shows the spectral function $A(\mathbf{k},\omega)$, which includes the one-particle vertex correction—i.e., the electron self-energy. At 100 K, the calculations show strongly renormalized but well-defined quasiparticle excitations within f-electron bands that are hybridized with the more dispersive d bands. There are two small Fermi surface pockets centered at the Γ points—i.e., $\mathbf{Q}$ = (000)—and R points—i.e., $\mathbf{Q} = (\frac{1}{2}\frac{1}{2}\frac{1}{2})$. Because of the strong spin-orbit coupling, the f bands have contributions from both $j = \frac{5}{2}$ and $\frac{7}{2}$ angular momentum states, with the former spread over ~100 meV around the Fermi level and the latter at a few hundred meV above the Fermi energy. The spin-orbit coupling is responsible for the weak incoherent spectral weight visible in Fig. 1A about 200 meV below the Fermi energy. At 400 K (Fig. 1B), there is a substantial broadening of the quasiparticle excitations, with a consequent reduction of their spectral weight, particularly close to the Fermi energy. We will discuss this later when presenting the high-temperature neutron scattering results.

The dynamic magnetic susceptibility, $\chi''(\mathbf{Q},\omega)$, is computed from the polarization bubble of the fully interacting DFT+DMFT one-particle Green's function, whose spectral weight is shown in Fig. 1, by incorporating particle-hole interactions through two-particle irreducible vertex corrections, Γ^{irr}_{loc}, which are assumed to be local in the same basis in which the DMFT self-energy is local (*28*, *29*). The calculations were performed at 100 K because of slow convergence times at lower temperatures, but Fig. 1 shows that the temperature is sufficiently low for the quasiparticles to be coherent. Further details are given in (*30*).

The calculations indicate scattering throughout the Brillouin zone, with broad maxima at high-symmetry points in **Q**, which shift with energy transfer. Figure 2 shows $\mathbf{Q}$ = [*H*,*K*] scattering planes with L = 1 and $L = \frac{3}{2}$, where H, K, and L are reciprocal lattice coordinates, at energy transfers of 35 meV and 55 meV. The calculations (Fig. 2, A to C) are displayed in an extended zone scheme with corrections for the f-electron magnetic form factor. These show that there are maxima in the intensity at the Γ and R points at ω = 35 meV but at the M $(\frac{1}{2}\frac{1}{2}0)$ and X $(\frac{1}{2}00)$ points at ω = 55 meV. These maxima are not connected by dispersive modes. Instead, the dynamic susceptibility consists of columns of intensity peaked at ~35 and ~55 meV in different regions of the Brillouin zone. This is illustrated in Fig. 3, which shows slices in the *L*-ω plane centered at the X and M points above 50 meV, and fig. S4B, which shows the shift to lower energy at the Γ and R points (*30*).

Inelastic neutron scattering from $CePd_3$

Using a large single crystal of $CePd_3$, with a mass of 17.72 g, we have performed measurements of 4D $S(\mathbf{Q},\omega)$ by rotating the sample at fixed incident energies on the time-of-flight spectrometer Merlin and the wide Angular-Range Chopper Spectrometer (ARCS), at the ISIS Pulsed Neutron Facility and Spallation Neutron Source, respectively (*24*, *25*). These possess large banks of position-sensitive detectors that allow the scattered neutrons to be counted as a function of polar and azimuthal angle, with respect to the incident beam. When combined with the sample rotation angle and the neutron time-of-flight, this four-coordinate scattering geometry can be readily transformed into 3D reciprocal space coordinates, **Q**, and a fourth energy coordinate, ω. The transformed data fill large volumes of $(\mathbf{Q},\omega)$, allowing arbitrary cuts to be made at constant energy or momentum transfer (Fig. 2, D to F, and Fig. 3, A and B). Correction for the temperature factor and calibration to a vanadium

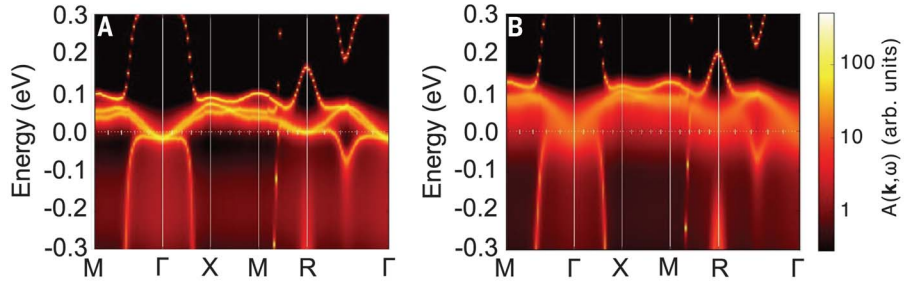


Fig. 1. Electronic spectral function of $CePd_3$ calculated using DFT+DMFT. (**A**) At 100 K, the calculations show the presence of well-defined quasiparticle bands that cross the Fermi energy–producing small electron pockets close to the center of the Brillouin zone (Γ) and the $(\frac{1}{2}\frac{1}{2}\frac{1}{2})$ zone boundary (R). Flat unoccupied bands near 50 meV are seen at the X point $(\frac{1}{2}00)$ and the M point $(\frac{1}{2}\frac{1}{2}0)$. (**B**) At 400 K, the spectral weight close to the Fermi energy is largely incoherent.

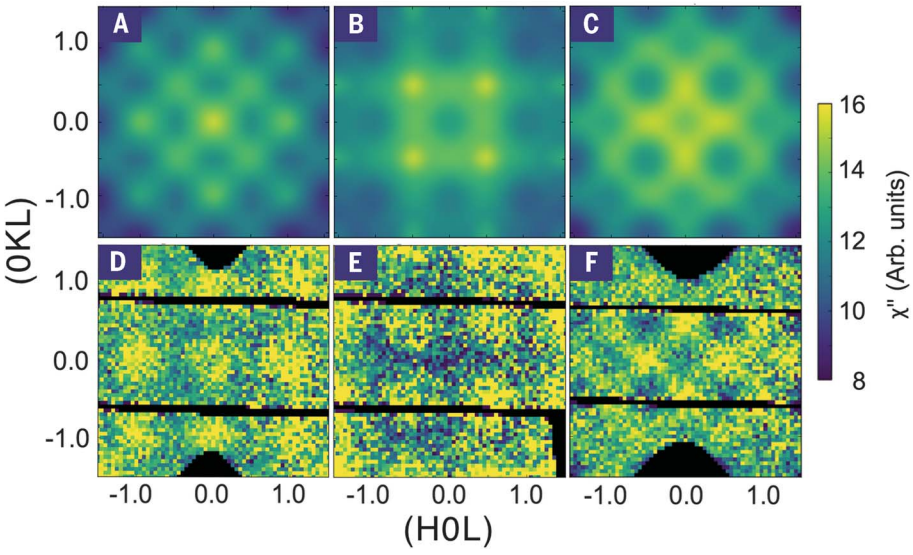


Fig. 2. Dynamic magnetic susceptibility of $CePd_3$. (**A** to **C**) Calculations using DFT+DMFT at 100 K compared to (**D** to **F**) inelastic neutron scattering on ARCS measured at 5 K. The results are shown as [*H0L*]/[*0KL*] planes at constant energy transfers of 35 meV [(A), (B), (D), and (E)] and 55 meV [(C) and (F)], with L = 1 [(A) and (D)] and $L = \frac{3}{2}$ [(B), (C), (E), and (F)]. In both the calculations and the measurements, the results are averaged over a range of ±5 meV in ω and ±0.2 in *L*. The intensity is in arbitrary units, with the calculations and the measurements normalized by a single scale factor. No backgrounds have been subtracted from the ARCS data. Black pixels represent regions of reciprocal space that were not measured.

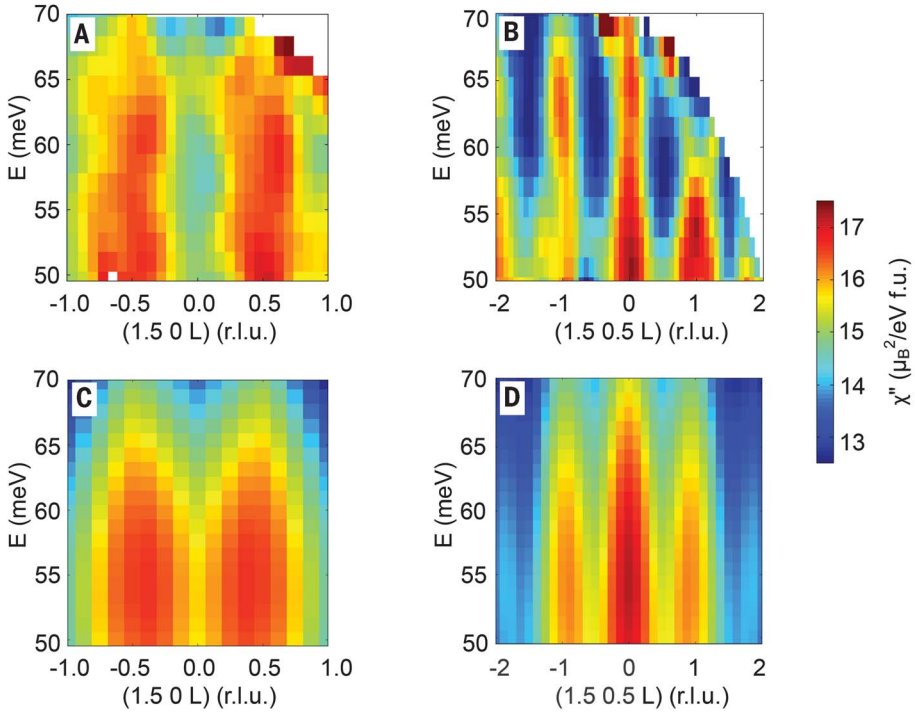


Fig. 3. Energy dependence of the dynamic magnetic susceptibility of $CePd_3$. (**A** and **B**) Merlin data measured at 5 K and (**C** and **D**) DFT+DMFT calculations at 100 K, represented by a *L*-ω slice at *H* = 1.5 ± 0.25, [(A) and (C)] *K* = 0.00 ± 0.25, and [(B) and (D)] *K* = 0.50 ± 0.25.

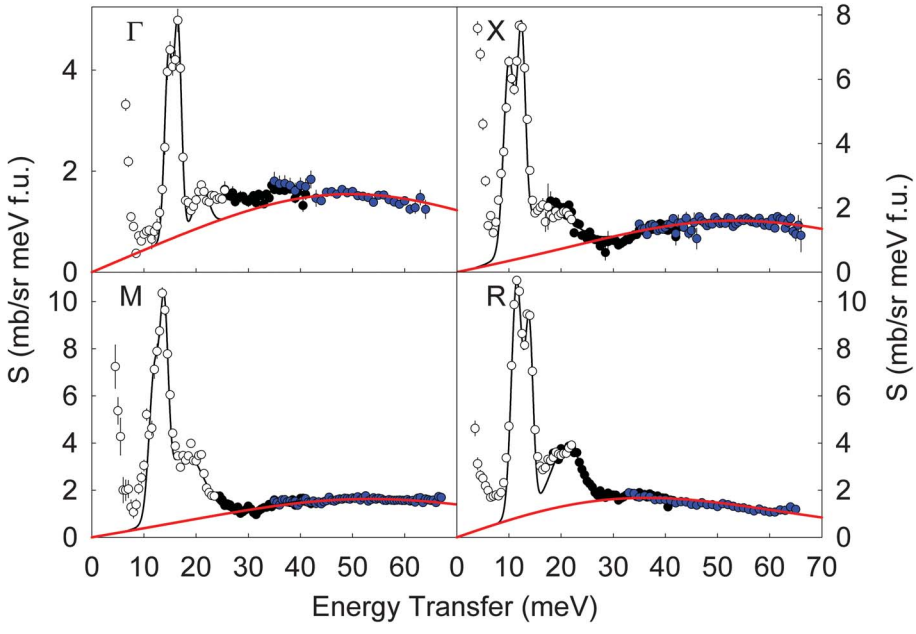


Fig. 4. Energy dependence of the scattering law of $CePd_3$. Measurements at 5 K are shown at the Γ (1.00 ± 0.25, 1.00 ± 0.25, 0.00 ± 0.25), X (1.00 ± 0.25, 1.00 ± 0.25, 0.50 ± 0.25), M (1.50 ± 0.25, 0.00 ± 0.25, 0.50 ± 0.25), and R (1.50 ± 0.25, 0.50 ± 0.25, 0.50 ± 0.25) points in the Brillouin zone. The open circles, closed black circles, and closed blue circles are measurements on Merlin with incident energies of 30, 60, and 120 meV, respectively. The solid black lines show the estimated phonon scattering based on a fit to three Gaussians. The energies of the phonon peaks are in good agreement with previous measurements of the $CePd_3$ phonons (*32*). The red lines are the result of the DFT+DMFT calculations at 100 K after adjusting a single overall scale factor.

standard allows the dynamic magnetic susceptibility to be directly compared with the DFT+DMFT calculations. There is an overall 20% scale uncertainty in the cross section.

Figure 4 shows the measured and calculated energy dependence of the scattering at four points in the Brillouin zone. The agreement between the experiment and theory shows that the DFT+DMFT calculations accurately reproduce the energy scale for the magnetic fluctuations. Below 30 meV, the measured spectra are dominated by nonmagnetic phonon scattering, which is not included in the calculations.

Figure 2, D to F, shows the **Q** dependence of the dynamic magnetic susceptibility derived from experiment, confirming the theoretically predicted shift in the maxima between the Γ and R points at 35 meV to the M and X points at 55 meV (Fig. 2, A to C). The magnetic scattering is superposed on a **Q**-dependent background from the sample environment, which increases monotonically with momentum transfer, shifting the maxima away from **Q** = 0.

To provide a more quantitative comparison, we show constant energy cuts along a number of high-symmetry directions in Fig. 5, where the data are plotted against the theoretical calculations on an absolute scale. The neutron spectra were fitted to the calculated dynamic susceptibility, adjusted by a single scale factor, and an instrumental background, produced by scattering off the sample environment, which is well described by a quadratic function in **Q**. The scale factor derived from the fits indicates that the accuracy of the absolute normalization is ~20%, consistent with uncertainties due to the absorption of the irregularly shaped sample (*30*). In Fig. 5, the fitted background has been subtracted from the data to show just the magnetic scattering. The unsubtracted spectra, along with additional details about the background estimates, are given in (*30*). The quantitative agreement confirms the qualitative consistency of experiment and theory evident in Figs. 2 and 3.

Finally, we compare the calculations made at higher temperatures to the experimental data. Figure 6 shows that the **Q** dependence of the magnetic scattering is almost entirely suppressed at room temperature, which is well above the coherence temperature inferred from transport measurements. This is also predicted by the theoretical calculations. An inspection of the spectral functions at 100 and 400 K in Fig. 1 shows that this results from the substantial reduction in spectral weight of coherent quasiparticles and confirms the importance of this coherence in generating the observed **Q** variations in the dynamic susceptibility at low temperature.

Discussion and conclusion

Our results demonstrate that it is now possible to determine the electronic structure of intermediate valence materials with considerable accuracy by incorporating local correlations into band structures through the combination of density functional theory and dynamical mean field

theory. The agreement of the calculated and measured neutron cross sections is complete throughout the Brillouin zone—extending over a broad range of energy transfers from ~20 to 65 meV—and accurately describes the shift in the scattering maximum from around 35 meV at the Γ and R points to around 55 meV at the M and X points. Although magnetic neutron scattering is a well-established probe of collective magnetic excitations, such as spin waves, that produce sharp dispersion surfaces in **Q**-ω space, our experiment shows it is now possible to measure and theoretically account for the dynamic magnetic susceptibility arising from correlated electron bands.

The calculations reveal the important role of local particle-hole interactions in renormalizing the dynamic magnetic susceptibility. In Fig. 6, B and C (and figs. S3 and S4), we compare the calculated magnetic response both with and without the two-particle vertex correction, Γ^{irr}_{loc}, which represents the interactions between the electron and the hole excited by the neutron. The correction has two effects: First, it smooths out some of the fine structure in the energy dependence of the spectra while broadly preserving both the **Q** variation and the overall energy scale; and second, it produces a strong enhancement of the intensity that is both energy and temperature dependent, for example, by a factor of ~6.5 at ω = 60 meV at 100 K. This shows that the **Q** dependence of the scattering is predominantly determined by the one-electron joint density of states, as expected for band transitions, whereas the overall intensity is amplified by the strong electron correlations.

In (*20*), we showed that the Anderson impurity model (AIM) is successful in explaining a number of important properties [the magnetic susceptibility $\chi(T)$, the 4f occupation number $n_f(T)$, the 4f contribution to the specific heat $C_{4f}(T)$, and the **Q**-averaged dynamic susceptibility $\chi''(\omega)$] of intermediate valence compounds such as $CePd_3$, even though the cerium atoms are not impurities but sit on a periodic lattice. We speculated that the reason the (incoherent) impurity model works so well for these periodic systems is that the strong inelastic Kondo scattering of the electronic quasiparticles by valence/spin fluctuations broadens the spectral functions. Such broadening can be seen even at low temperature for states away from the Fermi energy in Fig. 1A. The properties $\chi(T)$, $n_f(T)$, and $C_{4f}(T)$, are primarily sensitive to $\chi''(\omega)$, which represents local 4f moment fluctuations. The DFT+DMFT calculations show that the vertex corrections caused by the inelastic scattering result in spectra that, when averaged in **Q**, are very similar to the AIM result of (*20*), thus helping explain why the impurity model works as well as it does. These inelastic processes are what drives the rapid loss of coherence with temperature shown in Fig. 1B. The effect on the dynamic susceptibility, shown in Fig. 6, is that the spectrum of $CePd_3$ is nearly **Q** independent at room temperature and has the quasielastic spectral shape expected for an incoherent Anderson impurity system (*20*).

It should be possible to extend this work to materials with even stronger electronic correlations, such as heavy fermions, but the reduced coherence temperature and the increased complexity of potential many-body states will make it more computationally intensive. DFT+DMFT calculations of the magnetic susceptibility converge more slowly at low temperature and in the presence of multiple f-electron energy levels caused by crystal fields or magnetic interactions. In the case of $CePd_3$, the f-d hybridization is strong enough to suppress the crystal field splittings but weak enough that a temperature of 100 K was sufficiently low to reveal the onset of coherence. Nevertheless, the DFT+DMFT method has been successfully applied to the single-particle excitations of heavy fermions such as $CeCoIn_5$ (*6*), so calculations of the two-particle spectra would be challenging but should be technically feasible.

The results of this comparison between theory and experiment provide insight into the nature of the correlations in intermediate valence systems. The magnetic fluctuations show a much richer structure than was implied by earlier "toy" models of hybridized bands, and there is a complex interplay between coherent and incoherent contributions to the electronic spectra that is reflected in the evolution of the dynamic magnetic susceptibility with temperature. The transition from coherent f-electron bands to local moment physics, so long postulated in heavy fermion and intermediate valence systems, is confirmed by combining the latest advances in

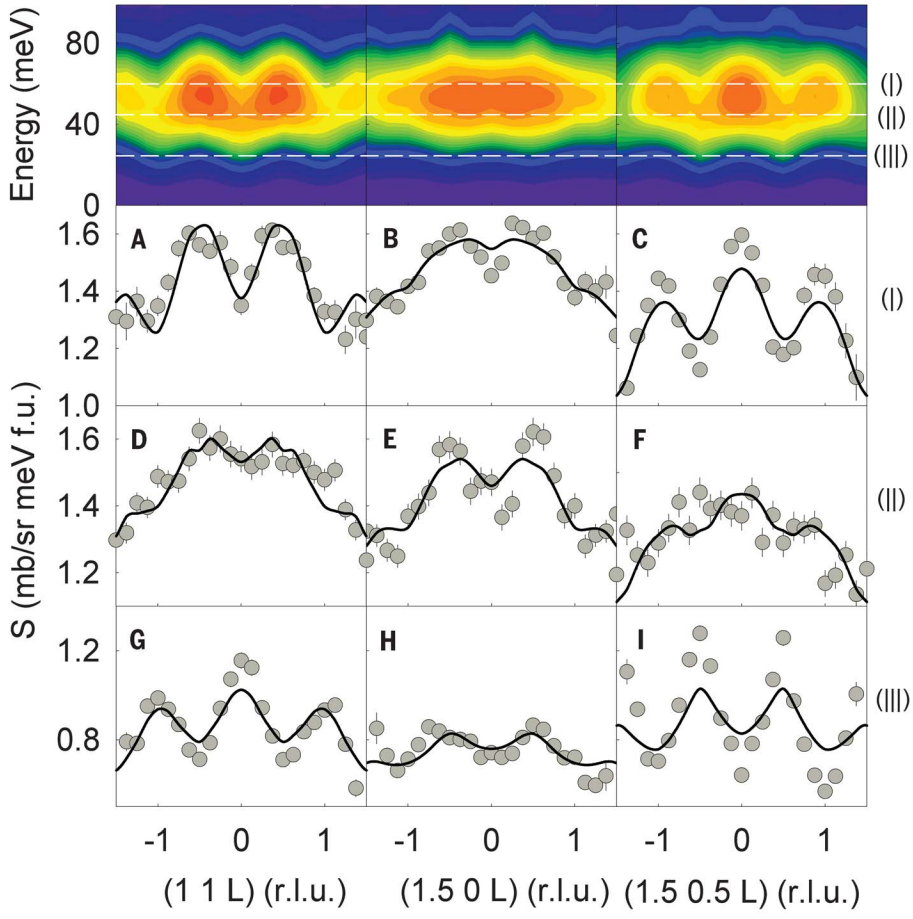


Fig. 5. Comparison of experiment and calculations of S(Q,ω) on an absolute scale. (Top) Left to right, the DFT+DMFT calculations at 100 K in the form of a *L*-ω slice along Γ → X at $H = 1.00 \pm 0.25$, $K = 1.00 \pm 0.25$; X → M at $H = 1.50 \pm 0.25$, $K = 0.00 \pm 0.25$; and M → R directions at $H = 1.50 \pm 0.25$, $K = 0.50 \pm 0.25$. Three white dashed lines, labeled I, II, and III, show the direction and position in energy of the 1D constant energy cuts at energies of (**A** to **C**) 60 ± 5 meV, (**D** to **F**) 45.0 ± 2.5 meV, and (**G** to **I**) 25.0 ± 2.5 meV, respectively. The inelastic neutron scattering data were measured on Merlin at $T = 5$ K and integrated in the same energy and *H*, *K* range as in the upper three panels. The error bars are derived using Poisson statistics. The data were fit to 1D cuts of the DFT+DMFT calculations, with a single scale factor and quadratic backgrounds as the only adjustable parameters. The fitted backgrounds have been subtracted from the data, so the points represent the estimated magnetic scattering and the lines represent the theoretical calculations. The data and fits without the background subtraction are shown in (*30*).

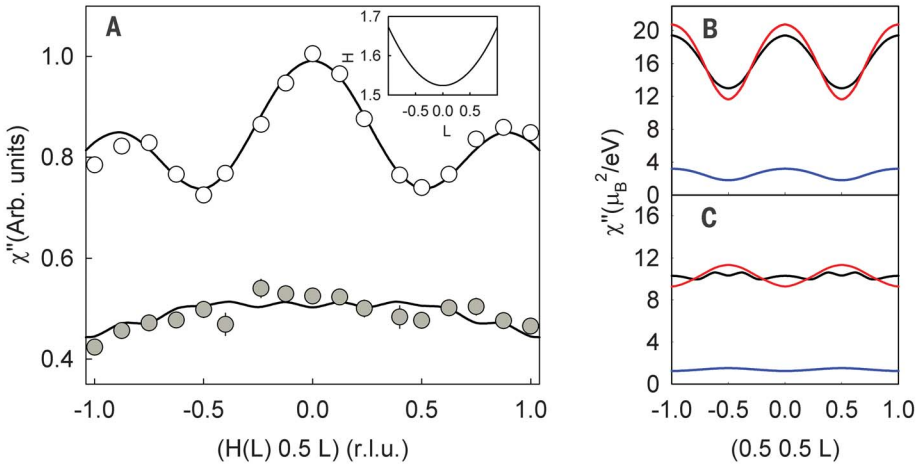


Fig. 6. Temperature dependence of the dynamic magnetic susceptibility of $CePd_3$. (**A**) The open (closed) circles were measured on the Maps spectrometer at 6 K (300 K) using a fixed sample geometry [see (*20*) for more details] at the energy transfer of 60 ± 10 meV and K = 0.50 ± 0.25, with backgrounds subtracted using the same method as in Fig. 5. The inset shows how H varies as a function of L. The lines are the results of DFT+DMFT calculations performed at 100 K and 400 K, including the $H(L)$ dependence, and scaled with the same normalization factor for both temperatures. (**B** and **C**) Calculations using DFT+DMFT at (B) 100 K and (C) 400 K of $\chi''(\mathbf{Q},\omega)$ (black lines), i.e., with corrections due to particle-hole interactions, and $\chi_0''(\mathbf{Q},\omega)$ (blue lines), without the corrections. The red lines show the calculated $\chi_0''(\mathbf{Q},\omega)$ at 100 K and 400 K after scaling by factors of 6.5 and 7.5, respectively.

inelastic neutron scattering and ab initio theories of correlated electron systems.

REFERENCES AND NOTES

1. P. S. Riseborough, J. M. Lawrence, *Rep. Prog. Phys.* **79**, 084501 (2016).
2. J. M. Lawrence, P. S. Riseborough, R. D. Parks, *Rep. Prog. Phys.* **44**, 1–84 (1981).
3. J. M. Lawrence, J. D. Thompson, Y. Y. Chen, *Phys. Rev. Lett.* **54**, 2537–2540 (1985).
4. B. H. Brandow, *Phys. Rev. B Condens. Matter* **37**, 250–260 (1988).
5. S. Burdin, A. Georges, D. R. Grempel, *Phys. Rev. Lett.* **85**, 1048–1051 (2000).
6. J. H. Shim, K. Haule, G. Kotliar, *Science* **318**, 1615–1617 (2007).
7. H. C. Choi *et al.*, *Phys. Rev. Lett.* **108**, 016402 (2012).
8. A. L. Cornelius *et al.*, *Phys. Rev. Lett.* **88**, 117201 (2002).
9. A. Kaminski *et al.*, *Phys. Rev. Lett.* **90**, 207003 (2003).
10. X. Deng *et al.*, *Phys. Rev. Lett.* **110**, 086401 (2013).
11. A. Georges, G. Kotliar, W. Krauth, M. J. Rozenberg, *Rev. Mod. Phys.* **68**, 13–125 (1996).
12. G. Kotliar *et al.*, *Rev. Mod. Phys.* **78**, 865–951 (2006).
13. M. Klein *et al.*, *Phys. Rev. Lett.* **106**, 186407 (2011).
14. S.-K. Mo *et al.*, *Phys. Rev. B* **85**, 241103 (2012).
15. E. Balcar, S. W. Lovesey, *Theory of Magnetic Neutron and Photon Scattering* (Clarendon Press, 1989).
16. S. K. Sinha, J. F. Cooke, A. P. Murani, Proceedings of the 1984 Workshop on High-Energy Excitations in Condensed Matter, Tech. Rep. (Los Alamos National Laboratory, 1984).
17. J. F. Cooke, J. A. Blackman, *Phys. Rev. B* **26**, 4410–4420 (1982).
18. R. Coldea *et al.*, *Phys. Rev. Lett.* **86**, 5377–5380 (2001).
19. P. Dai, *Rev. Mod. Phys.* **87**, 855–896 (2015).
20. V. R. Fanelli *et al.*, *J. Phys. Condens. Matter* **26**, 225602 (2014).
21. A. D. Christianson *et al.*, *Phys. Rev. Lett.* **96**, 117206 (2006).
22. A. P. Murani, A. Severing, M. Enderle, P. Steffens, D. Richard, *Phys. Rev. Lett.* **101**, 206405 (2008).
23. O. Sakai, *J. Phys. Soc. Jpn.* **79**, 114701 (2010).
24. R. I. Bewley *et al.*, *Physica B* **385-386**, 1029–1031 (2006).
25. D. L. Abernathy *et al.*, *Rev. Sci. Instrum.* **83**, 015114 (2012).
26. R. A. Ewings *et al.*, *Nuclear Instr. Meth. Phys. Res. A* **834**, 132–142 (2016).
27. F. Weber *et al.*, *Phys. Rev. Lett.* **109**, 057001 (2012).
28. M. Jarrell, *Phys. Rev. Lett.* **69**, 168–171 (1992).
29. H. Park, K. Haule, G. Kotliar, *Phys. Rev. Lett.* **107**, 137007 (2011).
30. See further information in the supplementary materials.
31. P. Blaha *et al.*, *Wien2K* (Karlheinz Schwarz, Technische Universitat Wien, Austria, 2001).
32. C. Loong, J. Zarestky, C. Stassis, O. D. McMasters, R. M. Nicklow, *Phys. Rev. B Condens. Matter* **38**, 7365–7369 (1988).

ACKNOWLEDGMENTS

The research at the Joint Institute for Nuclear Research was supported by the Russian Foundation for Basic Research project 16-02-01086. The research at Argonne National Laboratory and Los Alamos National Laboratory was supported by the Materials Sciences and Engineering Division, Office of Basic Energy Sciences, U.S. Department of Energy. The research at Oak Ridge National Laboratory's Spallation Neutron Source was supported by the Scientific User Facilities Division, Office of Basic Energy Sciences, U.S. Department of Energy. Neutron experiments were performed at the Spallation Neutron Source, Oak Ridge National Laboratory, and the ISIS Pulsed Neutron Source, Rutherford Appleton Laboratory. We gratefully acknowledge the computing resources provided on Blues, a high-performance computing cluster operated by the Laboratory Computing Resource Center at Argonne National Laboratory. We are also grateful for useful discussions with P. Riseborough. Files containing the data sets used in this paper are available for download from http://dx.doi.org/doi:10.18126/M2C914. Four-dimensional inelastic neutron scattering data from ARCS are stored in HDF5 files, conforming to the NeXus standard (www.nexusformat.org), which can be viewed using the open-source application, NeXpy (http://nexpy.github.io/nexpy/). The data can be compared with DFT+DMFT calculations, produced using the Wien2K+DMFT package (*31*), which are also stored in NeXus files. Inelastic neutron scattering data from MERLIN are available as RAR archives, containing files produced by the Horace suite of MATLAB programs (http://horace.isis.rl.ac.uk/), which can be used to extract cuts and slices through the 4D data.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/359/6372/186/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S8
References (*33*–*40*)

27 February 2017; accepted 1 December 2017
10.1126/science.aan0593

MALARIAL GENOMICS

Mapping the malaria parasite druggable genome by using in vitro evolution and chemogenomics

**Annie N. Cowell,[1] Eva S. Istvan,[2] Amanda K. Lukens,[3,4] Maria G. Gomez-Lorenzo,[5] Manu Vanaerschot,[6] Tomoyo Sakata-Kato,[3] Erika L. Flannery,[1] Pamela Magistrado,[3] Edward Owen,[7] Matthew Abraham,[1] Gregory LaMonte,[1] Heather J. Painter,[7] Roy M. Williams,[1] Virginia Franco,[5] Maria Linares,[5] Ignacio Arriaga,[5] Selina Bopp,[3] Victoria C. Corey,[1] Nina F. Gnädig,[6] Olivia Coburn-Flynn,[6] Christin Reimer,[1] Purva Gupta,[1] James M. Murithi,[6] Pedro A. Moura,[6] Olivia Fuchs,[1] Erika Sasaki,[1] Sang W. Kim,[1] Christine H. Teng,[1] Lawrence T. Wang,[1] Aslı Akidil,[8] Sophie Adjalley,[8] Paul A. Willis,[9] Dionicio Siegel,[10] Olga Tanaseichuk,[11] Yang Zhong,[11] Yingyao Zhou,[11] Manuel Llinás,[7] Sabine Ottilie,[1] Francisco-Javier Gamo,[5] Marcus C. S. Lee,[6,8] Daniel E. Goldberg,[2] David A. Fidock,[6,12] Dyann F. Wirth,[3,4] Elizabeth A. Winzeler[1,10]*

Chemogenetic characterization through in vitro evolution combined with whole-genome analysis can identify antimalarial drug targets and drug-resistance genes. We performed a genome analysis of 262 *Plasmodium falciparum* parasites resistant to 37 diverse compounds. We found 159 gene amplifications and 148 nonsynonymous changes in 83 genes associated with drug-resistance acquisition, where gene amplifications contributed to one-third of resistance acquisition events. Beyond confirming previously identified multidrug-resistance mechanisms, we discovered hitherto unrecognized drug target–inhibitor pairs, including thymidylate synthase and a benzoquinazolinone, farnesyltransferase and a pyrimidinedione, and a dipeptidylpeptidase and an arylurea. This exploration of the *P. falciparum* resistome and druggable genome will likely guide drug discovery and structural biology efforts, while also advancing our understanding of resistance mechanisms available to the malaria parasite.

Malaria has a disproportionately negative impact on human health because its causal protozoan parasites are adept at changing their genomes to evade antimalarial drugs and the human immune system. A single human infection may result in upwards of 10^{12} asexual blood-stage parasites. Thus, even with a relatively slow random mutation rate (~10^{-9} per nucleotide site per mitotic division) in the parasite, within a few cycles of replication, each base in the *P. falciparum* genome can acquire a random genetic change that may render at least one parasite resistant to the activity of a drug or a human-encoded antibody. The recent evolution of artemisinin-resistant parasites in Southeast Asia now threatens both life-saving treatments and malaria-control efforts (*1*).

Although this rapid evolution impedes our ability to control the disease, in vitro evolution in the presence of known antimalarials, followed by whole-genome sequencing of resistant clones, can be used to discover mediators of drug resistance (*2*). Testing for the evolution of resistance can also reveal antimalarial drug targets (*3*). Compared with targets that are validated using genetic knockdown methods, chemically validated drug targets are more valuable because their activity can be inhibited in cultured parasites by a small molecule. Furthermore, the inhibitor provides a tool for crystallization and chemical genetic studies. Most studies using this method to date have focused on single gene mutations in response to single compounds, even though, in many cases, additional allelic changes have been noted in *P. falciparum* clones during the acquisition of compound resistance.

Next-generation sequencing shows both neutral and positive selection

We systematically studied patterns of *P. falciparum* genome evolution by analyzing the sequences of clones resistant to diverse compounds with antimalarial activity across the *P. falciparum* life cycle. To investigate the genomic evolutionary response to treatment with small molecules, we assembled a collection of isogenic *P. falciparum* clones that had acquired resistance to an array of chemically distinct small-molecule growth inhibitors. Of the 37 different small molecules used during in vitro resistance evolution (table S1), 26 compounds were identified as having potent in vitro antimalarial activity from previous *P. falciparum* phenotypic screens (*4*–*7*). Others were chosen on the basis of medicinal chemistry optimization: the spiroindolone NITD678 (*8*) and the imidazolopiperazines GNF179, GNF452, and GNF707 (*9*). Respectively, NITD678 and the three imidazolopiperazines are similar to KAE609 and KAF156, which are recently discovered antimalarials that are now in clinical trials (*10*, *11*). We also included the licensed antimalarials atovaquone and primaquine. Although three carbazoles (MMV019017, MMV009063, and MMV665882) and the three imidazolopiperazines (GNF707, GNF452, and GNF179) were structurally similar to one another (ChemAxon fingerprint score ≥ 0.7) (fig. S1), most compounds tested had a distinct variety of functional groups and heterocyclic substructures, with some compounds showing activity affecting different aspects of the parasite life cycle (tables S1 and S2 and fig. S1).

Although some *P. falciparum* clones included herein have been previously described with respect to their drug sensitivity (*12*), other clones, including those resistant to primaquine and those resistant to GNF179, were generated in this study over a period of 3 to 6 months (fig. S2) using a stepwise, high-pressure intermittent, or constant method of compound exposure (table S1). Parasite clones exhibited a gain of resistance relative to their isogenic parent clones, with a fold shift of 2.8 and 37.2 in the lower and upper quartile, respectively, of the half-maximal effective concentration (EC_{50}) for resistant clones (table S3). This resistance shift remained stable after compound pressure was removed for 30 to 115 days. Genomic DNA was available from an average of 6.45 (median = 5) independently derived clones for each compound.

To identify the genetic basis of drug resistance, 204 clones (including resistant clones and sensitive isogenic parent clones) were fully sequenced with the paired-end read method. We included 58 published genomic sequences of clones resistant to 12 compounds, including cladosporin (*13*), atovaquone (*14*), and GNF179 (*15*). Alignment to the *P. falciparum* reference genome showed 80.3-fold average coverage of the 23.3–million base pair (Mbp) genome, with 85.3% of bases covered by 20 or more reads (table S3). We discovered 1277 single nucleotide variants (SNVs) and 668 small

[1]School of Medicine, University of California San Diego (UCSD), 9500 Gilman Drive, La Jolla, CA 92093, USA. [2]Departments of Medicine and Molecular Microbiology, Washington University School of Medicine, St. Louis, MO 63110, USA. [3]Department of Immunology and Infectious Disease, Harvard T.H. Chan School of Public Health, 665 Huntington Avenue, Boston, MA 02115, USA. [4]Infectious Disease Program, The Broad Institute, 415 Main Street, Cambridge, MA 02142, USA. [5]Tres Cantos Medicines Development Campus, Malaria Discovery Performance Unit, GlaxoSmithKline, Severo Ochoa 2, Tres Cantos 28760, Madrid, Spain. [6]Department of Microbiology and Immunology, Columbia University College of Physicians and Surgeons, New York, NY 10032, USA. [7]Department of Biochemistry and Molecular Biology, Pennsylvania State University, University Park, PA 16802, USA. [8]Malaria Programme, Wellcome Sanger Institute, Wellcome Genome Campus, Hinxton, Cambridgeshire CB10 1SA, UK. [9]Medicines for Malaria Venture, Post Office Box 1826, 20 Route de Pre-Bois, 1215 Geneva 15, Switzerland. [10]Skaggs School of Pharmacy and Pharmaceutical Sciences, UCSD, 9500 Gilman Drive, La Jolla, CA 92093, USA. [11]Genomics Institute of the Novartis Research Foundation, 10675 John J Hopkins Drive, San Diego, CA 92121, USA. [12]Division of Infectious Diseases, Columbia University College of Physicians and Surgeons, New York, NY 10032, USA.
*Corresponding author. Email: ewinzeler@ucsd.edu

insertions or deletions (indels) that arose in the 262 whole-genome sequences during resistance acquisition (Fig. 1A, Table 1, and table S4).

Previous whole-genome analyses of *P. falciparum* with microarrays and sequencing (*14*, *16*) showed that during long-term in vitro growth, genes encoding proteins involved in antigenic variation (such as *var* and *rifin*), usually located in the subtelomeric regions, rapidly acquire mutations. Thus, we hypothesized that we would find a similar pattern for small variants in our data set. Indeed, upon verification, 1107 of these variants were located in the 6% of the genome that encodes proteins involved in antigenic variation, referred to as the noncore genome (Table 1). We further predicted that if mutations in the core genome conferred a selective advantage, then we would find an enrichment of nonsynonymous coding changes. As predicted, the ratio of nonsynonymous coding to synonymous coding alleles in the 23-Mbp core genome was 7:1 (148:21), indicating an uneven distribution (accumulative binomial $P = 7.7 \times 10^{-10}$) and detectable positive selection. In contrast, this ratio was ~2:1 (216:128) in the noncore region, indicating near-neutral selection and confirming the mutability of these genes in the absence of selection. In total, we observed 148 nonsynonymous changes (table S5) in 83 genes (table S6) in the core genome (containing 838 variants total), with less than one nonsynonymous change per resistant clone (table S7).

We also predicted that these 83 genes would be enriched for annotated roles in drug resistance, with mutations detected in the appropriate compound-specific groups of clones. Gene ontology enrichment testing revealed that the set of 83 genes was statistically enriched with a "response to drug" gene ontology biological process designation (GO:0042493; accumulative hypergeometric $P = 5.78 \times 10^{-7}$; table S8). To the best of our knowledge, there were no false negatives, and our results were in agreement with other studies (*8*, *17*).

Most drug-resistant clones acquire copy number variants in addition to SNVs

Copy number variants (CNVs) also contribute to drug resistance in *P. falciparum* (*18*–*20*). For example, amplification of the gene encoding multidrug-resistance protein 1 (*pfmdr1*; PF3D7_0523000) confers multidrug resistance, including resistance to mefloquine (*21*). However, traditional methods of CNV detection based on sequencing read depth are more problematic in *P. falciparum* owing to its AT-rich content (~81%), which leads to uneven sequencing coverage (*22*). Thus, we established an automated pipeline for CNV detection using >3.0-fold normalized coverage. Briefly, average read depth was computed for coding regions, because intergenic regions of *P. falciparum*

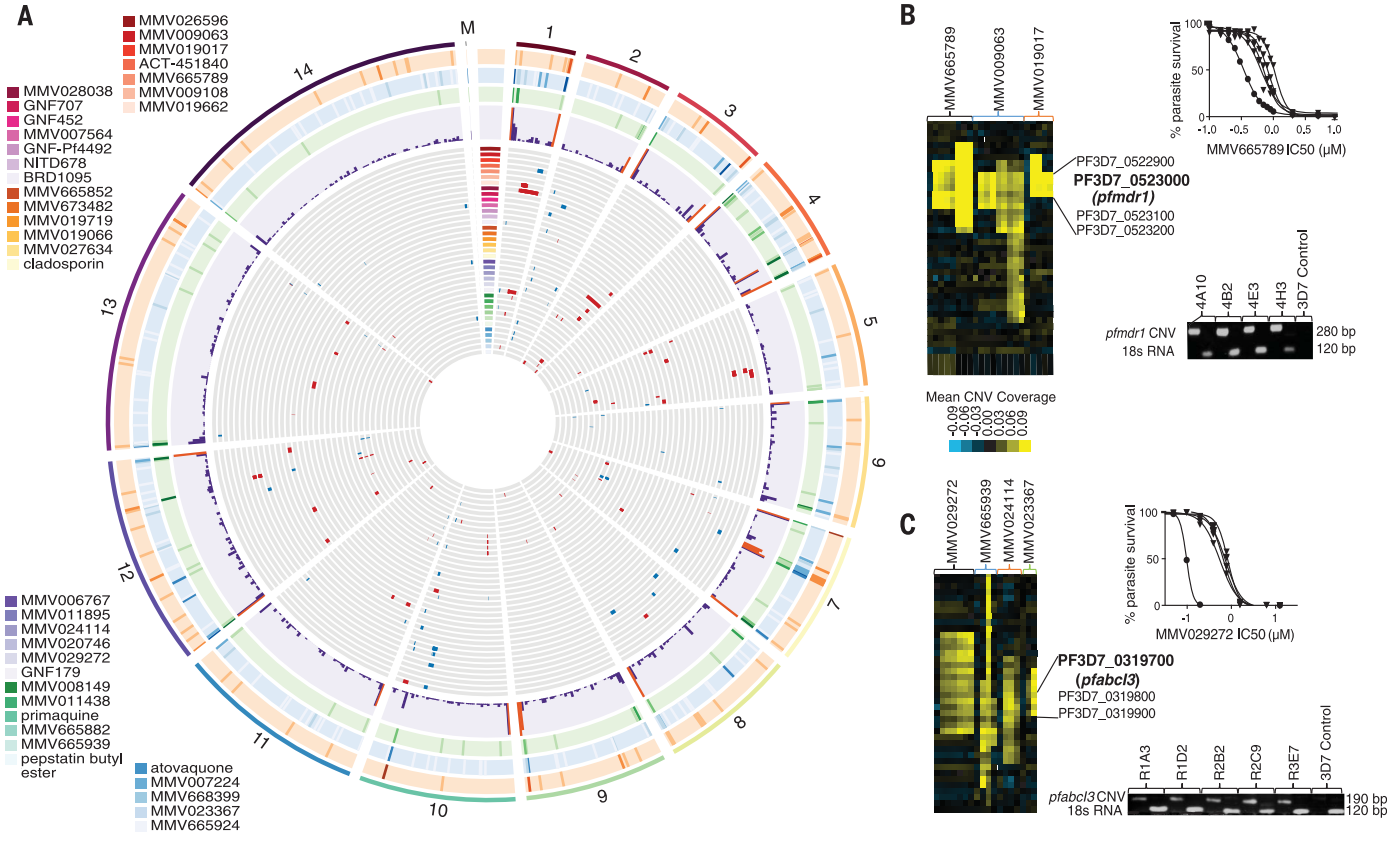


Fig. 1. Genetic changes acquired by 235 compound-resistant *P. falciparum* clones. (**A**) Circos plot (*68*) summarizing single nucleotide variants (SNVs), insertions and deletions (indels), and copy number variants (CNVs) acquired by the 235 *P. falciparum* clones resistant to 37 diverse compounds with antimalarial activity, grouped by chromosome. Each bar on the outer three rings represents 30,000 bp, and the darkness of the bar indicates a greater number of mutations. Orange ring, mutations lead to loss of function; blue ring, mutations lead to protein modification (nonsynonymous change or inframe deletion); green ring, no protein change (synonymous mutations or introns). The purple ring displays a histogram showing total counts of all three types of mutations, with orange bars where counts exceed 15. The gray rings display variants for resistant clones grouped by compound, with each ring representing one of 37 compounds. Red bars represent the location of CNVs, and blue bars represent the location of SNVs. (**B**) CNVs in the known drug-resistance gene *pfmdr1*. The heatmap (left) shows clones grouped by compound and is brighter yellow for genes with higher coverage (>2.5-fold above the standard deviation). Resistance shifts (top right) relative to the sensitive parent (circles) were demonstrated in MMV665789-resistant clones (triangles). PCR products (bottom right) shows that resistant clones contain the *pfmdr1* CNV junction sequence, whereas a 3D7 control lacks this sequence. (**C**) CNVs in *pfabcl3* are associated with resistance to multiple compounds. Resistance shifts relative to the sensitive parent (circles) were demonstrated in MMV029272-resistant clones (triangles). Again, the CNV junction sequence was not found in the 3D7 control DNA, whereas it was identified in clones resistant to MMV029272.

Table 1. Summary of effects for resistant clones in the core versus noncore regions of the *P. falciparum* genome. Noncore regions, making up 6% of the total genome, were located in subtelomeric regions and internal *var* gene clusters and encoded mostly PfEMP1s, RIFINs, and STEVORs. Effect changes are as classified using SnpEff (*67*).

Effect	Number		
	Total	Core	Noncore
Indels			
Codon change plus codon insertion	1	1	0
Disruptive inframe deletion	22	17	5
Disruptive inframe insertion	20	11	9
Frameshift variant	64	26	38
Frameshift variant plus stop-gained	1	0	1
Inframe deletion	43	37	6
Inframe insertion	30	20	10
Intergenic	412	239	145
Intron variant	69	64	5
Noncoding exon variant	1	1	0
Splice region variant plus intron variant	5	5	0
Total indels	668	449	219
SNVs			
Intergenic	700	178	522
Intron variant	49	30	19
Nonsynonymous coding	363	148	215
Splice region variant plus intron variant	3	2	1
Start lost	1	1	0
Stop-gained	10	9	1
Synonymous coding	1	0	1
Synonymous stop	1	0	1
Synonymous variant	149	21	128
Total SNVs	1277	389	888
Total indels and SNVs	**1945**	**838**	**1107**

exhibit reduced alignment confidence owing to 90 to 95% AT content and AT-repeat segments (*23*). The read coverage data were normalized in three groups representing the genetic backgrounds 3D7, Dd2, or 7G8, allowing for the identification of increased read coverage in background-specific amplified regions. Sets of two or more contiguous genes showing a ~2.0-fold change relative to the mean were identified. Potential amplified regions were then filtered to yield 159 high-confidence CNVs (core genome, mean average coverage per gene set > ~3.0-fold relative to mean coverage, and Benjamini-Hochberg–corrected $P < 0.001$) (tables S8 and S10). Altogether, 159 CNVs were observed in *P. falciparum* clones resistant to 27 of the 37 compounds (tables S8 and S11). CNVs were found primarily on chromosomes 1, 3, 5, and 12, with an average size of ~65 kilo–base pairs (kbp) (Fig. 1 and table S8). Of note, 76 core genome deletions were also identified (table S12), but most of these were small (<5 genes), located in subtelomeric regions, and apparently unrelated to the acquisition of drug resistance.

To validate our CNV-detection algorithm, we analyzed randomized, permuted average read coverage data and identified only about eight CNVs fulfilling these criteria. To assess our false negative rate, we confirmed that we could detect known CNVs previously detected by microarray analysis, including those that span the gene encoding lysyl tRNA synthetase (*pfkrs1*; PF3D7_1350100) in three cladosporin-resistant clones (*13*) (Benjamini-Hochberg–corrected *t* test; $P = 1.9 \times 10^{-40}$, 2.3×10^{-23}, and 1.3×10^{-12}, respectively) and the amplification event encompassing the multidrug-resistance protein 1 locus (*pfmrp1*; PF3D7_0112200) in the atovaquone-resistant clone R5a (*14*) (Benjamini-Hochberg–corrected *t* test; $P = 4.1 \times 10^{-33}$). The known amplification surrounding *pfatp4* in NITD678-resistant clones (*8*) was not detected, likely because this clone was sequenced with an older, 60-bp-read-length technology and was therefore not comparable to the rest of the set (although the amplification could be visually detected when compared with other samples sequenced with 60-bp reads). All 159 CNVs could be visually identified on a heatmap of normalized sequencing coverage. We found that CNV read coverage correlated with the shift in EC_{50} in resistant clones when no other resistance-conferring mutations were present, such as the *pfmdr1* amplification in clones resistant to compound MMV665789 (Fig. 1B). We used quantitative polymerase chain reaction (qPCR) to further validate CNVs that spanned genes encoding phosphatidylinositol 4-kinase (*pfpi4k*; PF3D7_0509800), ABC [adenosine triphosphate (ATP)–binding cassette] transporter I family member 1 (*pfabcI3*; PF3D7_0319700), and the aminophospholipid-transporting P-ATPase (*pfatp2*; PF3D7_1219600) (fig. S3). In addition, we identified paired-end reads from sequencing library fragments that spanned the CNV boundaries in the Integrative Genomics Viewer (*24*, *25*); this was followed by PCR over the boundaries (Fig. 1, B and C, and fig. S4) (*26*) for CNVs that covered *pfmdr1*, *pfabcI3*, and *pfpi4k*.

Resistance mechanisms are reproducible and structure-dependent

We predicted that treating parasites with compounds that have similar chemical structures would yield reproducible genomic changes. Clustering on the basis of compound similarity showed that parasites acquired similar resistance alleles when treated with structurally similar compounds (Fig. 2 and fig. S1). For example, independent clones resistant to imidazolopiperazines (GNF452, GNF707, and GNF179) and the closely related compound MMV007564 acquired mutations resulting in coding changes in the cyclic amine resistance locus (*pfcarl*; PF3D7_0321900). Furthermore, amplification of the *pfmdr1* region was observed in all 3D7 clones that acquired resistance to three closely related carbazoles: MMV009063, MMV019017, and MMV665882 (table S10 and Fig. 2).

The *P. falciparum* resistome

Our data set of CNVs and SNVs revealed 35 genes with two or more types of evidence suggesting a role as a drug-resistance determinant or actual target (Fig. 2). Evidence types include SNVs within an amplification event, the presence of two different alleles in the same gene for the same compound, or the same allele appearing in response to two different compounds. Mutations in *pfmdr1* were discovered for six different compounds, confirming its known role in drug resistance. Notably, a likely resistance mechanism or target gene was discovered for each compound examined. Below, we highlight genes with the strongest evidence of mediating drug resistance.

Discovery of alleles in known drug-resistance genes

Genetic detection of alleles associated with antimalarial resistance is the primary method of tracking resistance in clinical samples, because in vitro phenotypic susceptibility testing is difficult and costly in malaria parasites. Analysis of the mutations present in our set of resistant *P. falciparum* clones revealed both known and unknown genes implicated in drug resistance. We found several previously unidentified mutations in known *P. falciparum* resistance mediators. For example, we found four hitherto unrecognized alleles in the gene encoding the *P. falciparum* chloroquine resistance transporter (*pfcrt*; PF3D7_0709000), including S65R, A138V, K76Q, and S90N amino acid changes. Although one of the selection compounds, MMV006767, bears an aminoquinoline core, similar to chloroquine, neither MMV011895 nor MMV024114 do, providing evidence that PfCRT

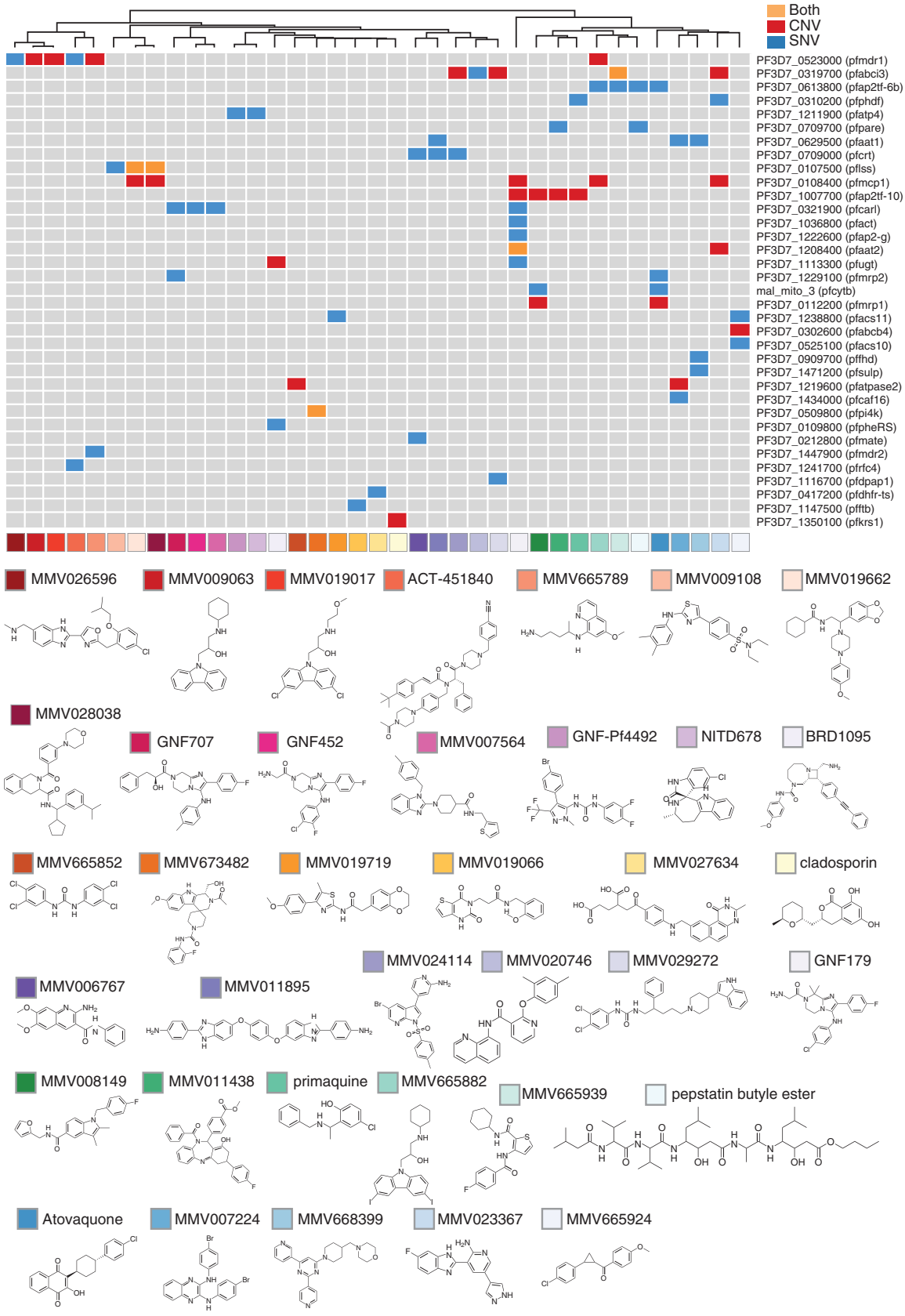


Fig. 2. The *P. falciparum* resistome. Compounds were clustered first on the basis of their target profile similarities, then by their chemical structural similarities. Each compound was assigned a color code, which is shared with Fig. 1. Genes that were detected as mutated (CNV, SNV, or indel) in independently created clones are listed.

is a pleotropic transporter regulating drug levels in the digestive vacuole. We also discovered a F806L allele in *pfmdr1* in MMV026596-resistant lines, which showed a 1.7-fold increase in the MMV026596 EC_{50} (table S3). An equivalent increase in the MMV026596 EC_{50} was also observed (table S13) upon assaying a separate NF54 line that had been selected using an unrelated compound (ACT-451840) and that had also acquired a F806L mutation in *pfmdr1* (*27*). In contrast, the same mutation increased parasite susceptibility of 3D7 and NF54 lines to known antimalarials such as mefloquine (13.8-fold and 7.3-fold, respectively), dihydroartemisin (4.2-fold and 1.4-fold, respectively), and lumefantrine (5.4-fold in NF54; the 3D7 line was not tested). In addition, multicopy amplification of the *pfmdr1* region was observed in all 3D7 clones that had acquired resistance to MMV009063 (eight of eight clones; average amplification, 10-fold), MMV019017 (four of four clones), MMV665882 (five of five clones), and MMV665789 (seven of seven clones) (table S10). A nonsynonymous SNV was also discovered in *pfmdr2* (K840N in response to compound MMV665789). In addition, mutations were detected in the gene encoding a plasma membrane–localized ABC transporter (*pfmrp2*; PF3D7_1229100) (*28*) in a clone that was resistant to imidazolopiperazines (selected with KAD707), as well as in two clones resistant to atovaquone. Changes in *pfmrp2* transcript levels have been associated with levels of *P. falciparum* resistance to quinoline drugs (*29*), including chloroquine and mefloquine. A SNV was found in the ATP-binding cassette of PfMRP2 at amino acid position D976N in the KAD707-resistant clone, whereas an A403P change in the transmembrane domain and a frameshift mutation (N1974fs) were detected in the atovaquone-resistant lines. Subsequent testing identified clones with cross-resistance to mefloquine, consistent with *pfmdr1* amplification (fig. S5).

Newly discovered drug-resistance genes

Knowing the identity of genes that impart multidrug resistance is important for the design of new drugs, understanding how existing therapeutics can lose their efficacy in clinical settings, and identifying likely causative alleles in genome-wide association studies. We observed that particular genes were mutated repeatedly in response to diverse compounds, indicating that they are more likely to be mediators of pleiotropic drug resistance. One such candidate is the putative ABC transporter encoded by *pfabcI3* (PF3D7_0319700) (*28*). Compounds that selected for mutations in *pfabcI3* (MMV020746, MMV023367, MMV024114, MMV029272, and MMV665939) are structurally diverse, and some are active across the parasite life cycle (*12*). The predicted gene product acquired point mutations [L690I, R2180G, R2180P (two times), and Y2079C] during resistance selections and was central to 12 different amplification events on chromosome 3 (table S7 and Fig. 1C).

A previously unrecognized type of potential resistance mediator in this study was a predicted amino acid transporter of the SLC32 family of solute carrier proteins (encoded by *pfaat1*; PF3D7_0629500). Orthologs of this conserved protein are responsible for the transport of amino acids in synaptic vesicles in humans (*30*). Mutations in the encoding gene were associated with resistance to three diverse compounds: MMV007224 (-135I inframe insertion), MMV668399 (P380S, K238N, and V185L), and MMV011895 (F230L). This gene was found to be associated with levels of chloroquine resistance in *P. falciparum* in a genome-wide association study (*31*), and we found that the product of *pfaat1* localizes to the *P. falciparum* digestive vacuole (fig. S6). Thus, the amino transporter encoded by *pfaat1* may play a role in the efflux of multiple drugs from the digestive vacuole, similar to PfCRT.

Nonessential genes and drug activators

Because genes with premature stop codons (stop-gained mutations) are unlikely to encode critical drug targets, genes with these mutations that were identified in our screen likely play a role in compound detoxification. The *P. falciparum* gene encoding the prodrug activation and resistance esterase (*pfpare*; PF3D7_0709700), which is annotated as a lysophospholipase, provides resistance to MMV011438 and pepstatin butyl ester (*32*). Another gene annotated as a lysophospholipase is PF3D7_0218600, which bears a FabD/lysophospholipase–like domain and harbors frameshift mutations in two independent clones (MALDA-Primaquine-PQG10 and MALDA-Primaquine-PQA11) that acquired blood-stage primaquine tolerance after 5 months of exposure (fig. S2). This gene was not mutated in any of the other 261 clones and contained two of the eight codon-changing core mutations (three frameshift, four missense, and one stop-gained) in seven different primaquine-resistant clones.

In addition, the gene encoding the amino acid transporter (*pfaat2*; PF3D7_1208400) was found to have a stop codon in a clone (NMicro-GNF179-S2-3D7-2C) that is resistant to GNF179, an imidazopiperazine that is closely related to the clinical candidate KAF156 (*10*). *pfaat2* is predicted to contain 10 transmembrane domains and a pfam01490 amino acid transporter domain. The position-903 stop mutation was found in concert with a splice acceptor intronic mutation in the gene encoding the acetyl coenzyme A transporter (*pfact*; PF3D7_1036800)—a gene whose disruption confers resistance to imidazolopiperazines (*15*). CRISPR-Cas9 introduction of the *pfaat2* L903 stop codon into GNF179-sensitive Dd2 parasites revealed that this mutation also confers resistance to GNF179 on its own (EC_{50} fold shift of 38 compared with the sensitive parent clone) (fig. S7).

Transcription factors

Our study unveiled several resistance mediators that are likely to play a role in the parasite's transcriptional response to drugs. Although there are no examples in *P. falciparum*, work in *S. cerevisiae* and other microbes has shown that mutations in transcription factors can often result in multidrug resistance (*33*). In our study, the most commonly mutated class of genes was that encoding the Apicomplexan AP2 transcription factor family (10 different variants observed in five different AP2 transcription factor–encoding genes, as well as several potential intergenic promoter mutations). In plants, AP2 transcription factors are involved in the cellular response to stress (*34*), and in *Plasmodium*, they regulate a variety of developmental transitions including commitment to sexual development and sporogony (*35*, *36*). The most prominent AP2 transcription factor in our set was encoded by PF3D7_0613800, which showed evidence of selection three times with three independent compounds (MMV665882, MMV665939, and MMV011438). In two cases, the same codon deletion (QMEGDNEMEGDNE197Q) was observed, and in one case, there was a nonsynonymous codon change (Q197E) at the same position. In addition, a region on chromosome 10 that encompasses a gene encoding another AP2 transcription factor [*pfap2tf-10*, also called *pfap2-i* (*37*); PF3D7_1007700] was amplified in nine clones (table S9). The amplification on chromosome 10 appeared in one primaquine-resistant clone (consisting of only seven genes with an estimated four to five copies; $P = 6.6 \times 10^{-5}$), as well in clones resistant to GNF179, MMV008148, or MMV011438. Variants in the gene encoding the chromosome 6 transcription factor (PF3D7_0613800) were associated with resistance to quinine in genome-wide association studies (*38*). However, we still need to determine whether these changes are associated with multidrug resistance or are related to long-term in vitro culturing.

The Forkhead-associated domain is a phosphopeptide recognition domain found in some kinases and transcription factors (*39*). Mutations in a gene encoding an uncharacterized protein bearing this domain (*pffhd*; PF3D7_0909700) were found in all samples resistant to MMV668399 ($P = 1 \times 10^{-18}$ using an accumulative binomial distribution). Mutations included a M1V start-lost mutation, a K392 stop-gained mutation, L95R and G720E missense mutations, and a gene deletion (MALDA-MMV668399-2F2). Given that three of the mutations were predicted to result in loss of function, we hypothesize that this is a gene that inhibits drug action, although additional work needs to be done to characterize this further. These same resistant clones also contained mutations in the genes encoding the amino acid transporter (*pfaat1*; PF3D7_0629500) described above and the inorganic anion transporter (*pfsulp*; PF3D7_1471200) discussed below.

New targets and compound target–inhibitor pairs

In addition to discovering genes involved in drug resistance, another goal of this study was to identify antimalarial drug targets to fill the need for alternative malaria treatments and advance the campaign to eliminate malarial infections in humans. In general, we considered enzyme-encoding genes with mutations as potential drug target candidates, which was further supported if the mutated gene was compound-specific and docking and homology modeling showed mutations

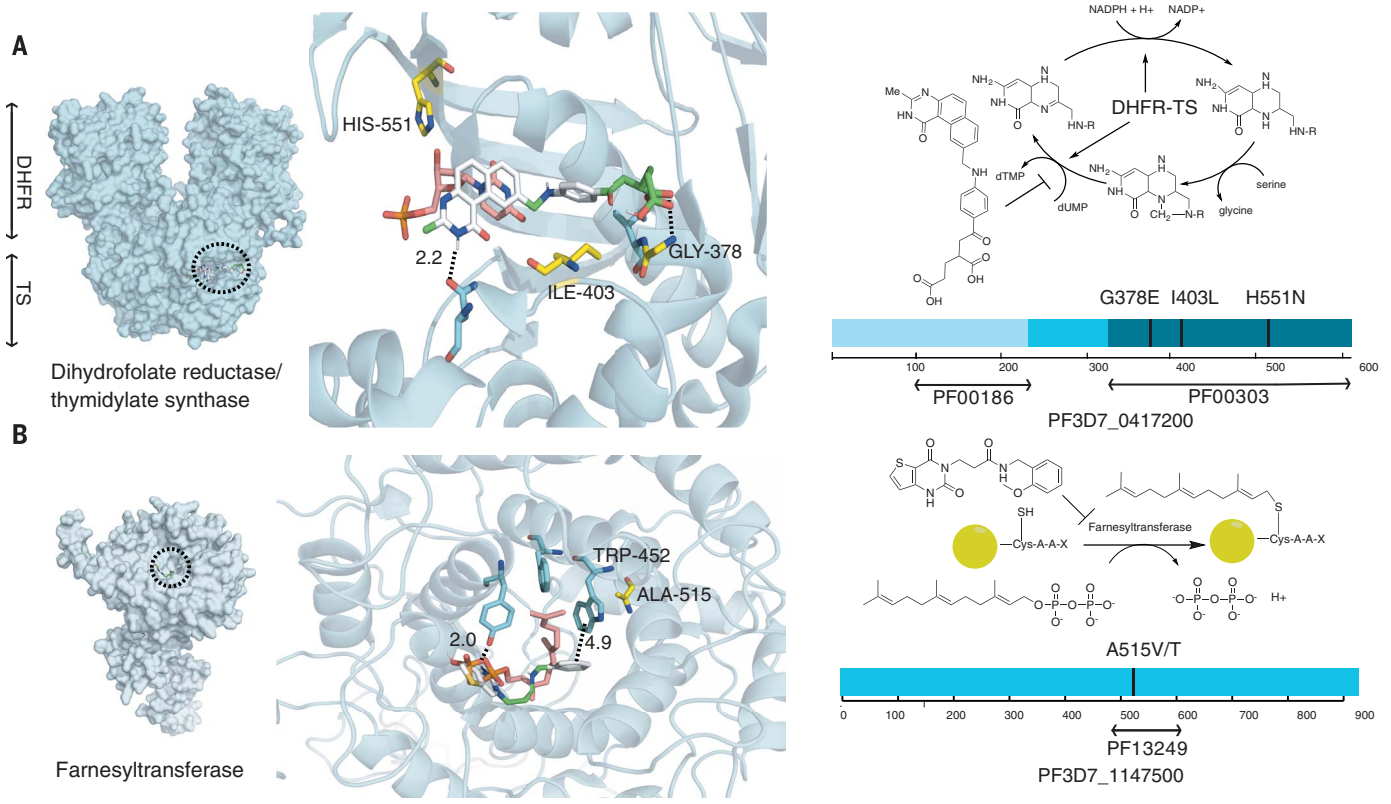


Fig. 3. A computational model of active antimalarial small molecules docked against their respective targets. Schematics of the biochemical pathway, the compound, and conserved pfam domains are shown in the right panel. (**A**) MMV027634 occupies the thymidylate synthase active site. The benzoquinazoline head group of MMV027634 is stabilized by hydrogen-bonding within the dUMP active site of dihydrofolate reductase–thymidylate synthase (DHFR-TS). Further stabilization occurs at the tail with multiple hydrogen bonds to Gly^{378}, mutation of which confers resistance to this compound. Mutations are all found in the thymidylate synthase portion of the molecule (http://pfam.xfam.org/family/pf00303). (**B**) MMV019066 and farnesyl pyrophosphate (FPP) are shown concurrently docked within the binding pocket of farnesyltransferase. A model of the *P. falciparum* farnesyltransferase beta subunit was constructed using the rat homolog (Protein Data Bank ID, 1ZIR) as a template. Despite substantial interspecies protein sequence variance, the FPP binding pocket is largely conserved (*69*). The preferential binding states of FPP and MMV019066 are shown competing for similar hydrophobic space. Resistance mutations are found in the squalene-hopene cyclase domain (http://pfam.xfam.org/family/pf13249). The yellow circle represents the rest of the donor protein to which the farnesyl group is attached. dTMP, deoxythymidine monophosphate; $NADP^+$, nicotinamide adenine dinucleotide phosphate; NADPH, reduced form of $NADP^+$; Me, methyl; R, benzoyl-L-glutamic acid. 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.

in the catalytic site. A previously reported example is the set of BRD1095-resistant clones, which bear four different amino acid position changes (*40*) located in the predicted active site of the alpha unit of the cytosolic phenylalanyl-tRNA synthetase.

Thymidylate synthase

One likely target-inhibitor pair is dihydrofolate reductase–thymidylate synthase (encoded by *pfdhfr-ts*; PF3D7_0417200) and the benzoquinazolinone MMV027634. Although dihydropyrimidine inhibitors of the bifunctional enzyme are well-known antimalarials (e.g., pyrimethamine), all target the dihydrofolate reductase portion of the dual-function protein. We identified three different nonsynonymous mutations mapping to the thymidylate synthase portion of the bifunctional molecule in the MMV027634-resistant lines (Fig. 3A). Mapping of the mutations onto a published thymidylate *P. falciparum* crystal structure (*41*) showed that each mutation flanks the 2'-deoxyuridylic acid (dUMP)–binding site of the enzyme (Fig. 3A). Metabolomics (discussed below) and independent docking of MMV027634 provided evidence that both substrates occupy the thymidylate synthase active site with a calculated affinity of −8.9 kcal/mol.

Farnesyltransferase

Farnesyltransferase is encoded by another target gene (*pfftb*; PF3D7_1147500), which acquired two different mutations in amino acid 515 (A515V and A515T) in three clones resistant to the pyrimidinedione MMV019066 (Fig. 3B). Modeling shows that the mutation at amino acid 515 likely disrupts interactions between the thienopyrimidine and the farnesylation active site, resulting in a resistant phenotype. Previous work has shown mutations in the lipid substrate–binding site of farnesyltransferase in tetrahydroquinoline-resistant parasites (*42*).

Dipeptidyl aminopeptidase 1

An additional potential target gene encodes dipeptidyl aminopeptidase 1 (DPAP1) (*pfdpap1*; PF3D7_1116700), an exopeptidase that localizes to the digestive vacuole and cleaves amino terminal dipeptides from proteins or oligopeptides (*43*). Each parasite clone resistant to MMV029272, an arylurea, carried a mutation (L437S, L415P, or N62H) in *pfdpap1*. This gene is considered to be essential because it cannot be disrupted (*44*). All resistant clones also contained amplifications that encompassed the ABC transporter gene (*pfabcI3*; PF3D7_0319700), highlighting how resistance mechanisms can be found along with targets.

Aminophospholipid-transporting P-type ATPase

As was previously reported for cladosporin and lysyl tRNA synthase (*13*), gene amplification events can occur around drug targets. We detected six CNVs in a genomic region on chromosome

12 that encompass a gene encoding a predicted aminophospholipid-transporting P-type ATPase (*pfatpase2*; PF3D7_1219600) (table S7 and fig. S3), previously named PfATP2. Tandem amplifications with approximate sizes of 8.7, 29, and 25 kbp were found in MMV007224-resistant clones, and amplifications of 102, 95, and 38 kbp were found in MMV665852-resistant lines. The 8.7-kbp amplification events encompassed only two genes—namely, the transporter gene and a truncated *var* pseudogene (PF3D7_1219500), indicating that the transporter is likely the target. This protein is closely related to PfATP4, an important antimalarial drug target. In the closely related rodent malaria parasite, *P. berghei*, this transporter was refractory to targeted gene deletion attempts, demonstrating its essentiality (*45*). MMV007224 and MMV665852 are structurally similar to one another and are both active against liver-stage parasites. *S. cerevisiae* orthologs of *pfatp2* encode the Dnf1/Dnf2 aminophospholipid translocases (flippases), which maintain membrane lipid asymmetry at the plasma membrane and contribute to endocytosis (*46*). In support of this activity, additional nonsynonymous variants in MMV007224-resistant lines were predicted to encode proteins involved in similar processes: Sec24, a component of the coat protein complex II (COPII) that promotes vesicle budding from the endoplasmic reticulum (ER), and Yip1, a protein required for fusion of these ER-derived vesicles with the Golgi. MMV665852 is a triclocarban, an antibacterial agent whose proposed mechanism is hypothesized to be inhibition of fatty acid synthesis, similar to triclosan, which was proposed to inhibit lipid and membrane function in *Plasmodium* parasites (*47*, *48*).

Last, in some cases, it can be difficult to determine whether a gene encodes a promising target or is a resistance gene. For example, PF3D7_0107500 is annotated as encoding a lipid-sterol symporter in the resistance-nodulation-division (RND) transporter family. RND transporters are well characterized in Gram-negative bacteria, where they are involved in extruding toxins, exporting virulence determinants, and maintaining overall homeostasis (*49*). Although specific pumps in this family have been associated with multidrug resistance in bacteria (*50*), it is unclear what specific function the transporter encoded by PF3D7_0107500 has in *P. falciparum*. We detected five different SNVs in this gene in clones resistant to MMV028038, MMV019662, and MMV009108, in addition to 12 amplification events surrounding this gene in clones resistant to MMV028038 and MMV019662. Additional heterozygous SNVs were found in this gene in the MMV019662-resistant clones with chromosome 1 amplification events (table S14). MMV028038 and MMV019662 both have an amide bond, whereas MMV009108 does not. In addition, three different SNVs (D520Y, K615N, and I596M) were detected in an annotated inorganic anion transporter gene (*pfsulp*; PF3D7_1471200) in clones resistant to the pyrimidine MMV668399. The probability of randomly observing the same gene three times in six selections is extremely low ($P = 4.0 \times 10^{-11}$). PF3D7_1471200 encodes PfSulP, a protein of the SLC26 membrane protein family (*51*) with 11 transmembrane domains. In *P. falciparum*, this transporter has been localized to the surface of the parasite and is thought to play a role in the transport of anions across the parasite plasma membrane (*52*), although its role in drug resistance has not been well defined. Homology models with a fumurate transporter (*53*) showed that all mutations were located in the STAS domain, a cytoplasmic extension that in some species plays a role in signal transduction.

Metabolomic profiling of compounds used to generate resistance

To functionally validate the modes of action associated with the targets predicted by resistance generation, metabolic perturbations associated with drug exposure (*54*) were assessed for 25 of the 37 compounds by using ultrahigh-performance liquid chromatography–mass spectrometry (UHPLC-MS). Overall, the detected metabolic perturbations predict modes of action in accordance with the generated resistance mutations observed (figs. S8 and S9 and table S16). MMV008149, which resulted in a mutation in cytochrome b (*pfcytb*; mal_mito_3) (*12*), gives a classical response in the pyrimidine biosynthetic pathway with accumulation of *N*-carbamoyl-L-aspartate and dihydroorotate, indicative of disruption of cytochrome b or dihydroorotate dehydrogenase function. Similarly, as predicted by the thymidylate synthase resistance mutations and computational docking (Fig. 3), metabolomic analysis of MMV027634 revealed a clear signature in folate biosynthesis (figs. S8 and S9). MMV024114 has an analogous metabolic response and likely targets folate metabolism, although resistance mutations for this compound arose in *pfcrt* and *pfabcI3*, suggesting an adaptive efflux-based resistance by these parasites. Cluster analysis of the metabolic responses showed that compounds that yielded resistance mutations in the transporter genes *pfmdr1*, *pfaat1*, and *pflss* (MMV009108, ACT-451840, MMV665789, MMV019662, MMV007224, and MMV009063) had similar profiles. For many of the compounds that showed minimal metabolic perturbation (GNF179, MMV008149, MMV011438, and primaquine), resistance selection resulted in CNVs in *pfap2tf-10* [*pfap2-i* (*37*)]. Although not all compounds resulted in metabolic responses, the general concordance between metabolic changes and resistance mutations supports the combined use of these approaches for compound target identification.

Discussion and conclusions

Our study represents a controlled examination of antimalarial drug resistance acquisition by *P. falciparum*. Prior studies examining the parasite's genetic response during drug resistance development have evaluated the response only to known antimalarials (*31*, *38*, *55*, *56*) or have focused on coding mutations in one target gene in response to a single compound class (*8*, *13*, *32*, *40*, *57*–*60*). It is likely that the genes that are identified here will prove important in both clinical studies and the process of drug development.

One noteworthy finding from this data set is the high enrichment of mutations under positive selection. We focused here on genes for which there were multiple lines of supporting evidence across our study; however, it is likely that other important genes were identified. Excluding genes contained within CNVs, the 57 singleton genes with nonsynonymous mutations encode potential druggable targets, such as guanosine triphosphatases, mRNA decapping enzymes, components of V-type ATPase complexes, kinases, and ubiquitin ligases (table S5). Although some of these nonsynonymous changes could have been maintained within the population by chance, some are plausible targets of selection. For example, although resistance to ACT-451840 is conferred by mutations in *pfmdr1* (*61*), a mutation was observed in the gene encoding a putative replication factor C subunit 4 (PF3D7_1241700) that results in a S42T amino acid change near the ATP-binding pocket in two independent ACT-451840–resistant clones (*61*). Replication factor C subunit 4 is the subunit ATPase in the clamp-loading DNA polymerase complex, is likely essential, and is a good drug target. Modeling shows that ACT-451840 could bind between subunits 1 and 4 of replication factor C, potentially disrupting ATPγ ingress (fig. S10). Our findings also highlight the large number of uncharacterized genes in the *P. falciparum* genome and the need for further functional annotation.

Not all mutations discovered in this study will confer resistance—they may compensate, specifically or nonspecifically, for fitness losses that result from a resistance-conferring mutation. Alternatively, they may be associated with adaptation to long-term in vitro culture. Genome editing or creating recombinant lines may be used to confirm that alleles provide resistance (*8*, *15*, *27*, *40*, *58*, *62*, *63*). On the other hand, given the multigenic nature of resistance, it may be difficult to recreate a resistance phenotype (*7*). For example, the nonsynonymous changes that emerged in phenylalanine tRNA ligase during development of resistance to BRD1095 were accompanied by copy number changes at unrelated sites, which potentially explains why attempts at validation through genome editing were unsuccessful.

Beyond changes in coding sequences, the low frequency at which noncoding variants were found in our data set suggests a possible role. We identified only 21 silent mutations within genes that are plausible drug targets, including those encoding protein kinase 4, Ark3 kinase, cytochrome c oxidase 3, a histone-lysine *N*-methyltransferase, a putative serine threonine protein kinase, a DNA helicase, and an 18*S* ribosomal RNA. Several of the genes with silent mutations were highly expressed, including the cytochrome c oxidase subunit 3 gene, in which the mutation substitutes a rare codon for a frequently used one [acA (23.8% usage for threonine) to acG (3.6%)]. Silent mutations in human *mdr1* can confer resistance to cancer drugs

owing to altered protein folding (*64*). Further work will be needed to establish whether these silent substitutions play a role in drug resistance in *P. falciparum*. Similarly, intergenic and intronic mutations were also common. Although in most cases, resistance was explained by the presence of nonsynonymous coding mutations in the target or resistance genes, intergenic mutations were also common. Likewise, intron variants were more likely to be found in the core genome than in subtelomeric regions (64:5), suggesting a possible functional role.

Last, it is likely that among the genes that we identified, several may contribute to clinical resistance at some level. Although pharmacokinetics are different in vivo than in vitro, we repeatedly rediscovered mutations in genes important for clinical resistance. It is therefore likely that previously unidentified genes are candidates to be under selection in clinical isolates. Mutations in 3247 clinical *P. falciparum* isolates reveal that four mediators of resistance (the ABC transporter gene *pfabcI3*, the putative amino acid transporter gene *pfaat2*, the AP2 transcription factor gene *pfap2tf-6b*, and the farnesyltransferase gene *pfftb*) have nonsynonymous-to-synonymous ratios greater than 2:1, suggesting that they are under positive selection (*65*) (table S15). However, the role of these genes in clinical drug resistance is unknown. In addition, our data highlight the importance of CNVs in conferring multidrug resistance and suggest that high-coverage genome sequencing of clinical isolates will provide information on selective pressures in the field (*66*).

It is notable that we were able to identify a likely target or resistance gene for every compound for which resistant parasites were generated. The haploid genome and the lack of transcriptional feedback loops suggest that *P. falciparum* is a particularly good model for both target identification and resistome studies. Our characterization of the chemogenetic landscape of *P. falciparum* will guide the design of small-molecule inhibitors against this deadly eukaryotic pathogen.

REFERENCES AND NOTES

1. E. A. Ashley *et al.*, *N. Engl. J. Med.* **371**, 411–423 (2014).
2. F. Ariey *et al.*, *Nature* **505**, 50–55 (2014).
3. E. L. Flannery, D. A. Fidock, E. A. Winzeler, *J. Med. Chem.* **56**, 7761–7771 (2013).
4. T. Spangenberg *et al.*, *PLOS ONE* **8**, e62906 (2013).
5. F.-J. Gamo *et al.*, *Nature* **465**, 305–310 (2010).
6. W. A. Guiguemde *et al.*, *Nature* **465**, 311–315 (2010).
7. D. Plouffe *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **105**, 9059–9064 (2008).
8. M. Rottmann *et al.*, *Science* **329**, 1175–1180 (2010).
9. S. Meister *et al.*, *Science* **334**, 1372–1377 (2011).
10. K. L. Kuhen *et al.*, *Antimicrob. Agents Chemother.* **58**, 5060–5067 (2014).
11. N. J. White *et al.*, *N. Engl. J. Med.* **371**, 403–410 (2014).
12. V. C. Corey *et al.*, *Nat. Commun.* **7**, 11901 (2016).
13. D. Hoepfner *et al.*, *Cell Host Microbe* **11**, 654–663 (2012).
14. S. E. Bopp *et al.*, *PLOS Genet.* **9**, e1003293 (2013).
15. M. Y. Lim *et al.*, *Nat. Microbiol.* **1**, 16166 (2016).
16. N. V. Dharia *et al.*, *Genome Res.* **20**, 1534–1544 (2010).
17. E. L. Flannery *et al.*, *ACS Chem. Biol.* **10**, 413–420 (2015).
18. S. Nair *et al.*, *Mol. Biol. Evol.* **24**, 562–573 (2007).
19. S. Nair *et al.*, *PLOS Genet.* **4**, e1000243 (2008).
20. J. L. Guler *et al.*, *PLOS Pathog.* **9**, e1003375 (2013).
21. B. Blasco, D. Leroy, D. A. Fidock, *Nat. Med.* **23**, 917–928 (2017).
22. S. M. Teo, Y. Pawitan, C. S. Ku, K. S. Chia, A. Salim, *Bioinformatics* **28**, 2711–2718 (2012).
23. M. J. Gardner *et al.*, *Nature* **419**, 498–511 (2002).
24. J. T. Robinson *et al.*, *Nat. Biotechnol.* **29**, 24–26 (2011).
25. H. Thorvaldsdóttir, J. T. Robinson, J. P. Mesirov, *Brief. Bioinform.* **14**, 178–192 (2013).
26. D. Menard *et al.*, *PLOS Negl. Trop. Dis.* **7**, e2489 (2013).
27. C. L. Ng *et al.*, *Mol. Microbiol.* **101**, 381–393 (2016).
28. R. A. Kavishe *et al.*, *Malar. J.* **8**, 205 (2009).
29. S. Mok *et al.*, *Mol. Microbiol.* **91**, 918–934 (2014).
30. B. Gasnier, *Pflugers Arch.* **447**, 756–759 (2004).
31. Z. Wang *et al.*, *Sci. Rep.* **6**, 33891 (2016).
32. E. S. Istvan *et al.*, *Nat. Commun.* **8**, 14240 (2017).
33. S. Paul, W. S. Moye-Rowley, *Front. Physiol.* **5**, 143 (2014).
34. J. L. Riechmann, E. M. Meyerowitz, *Biol. Chem.* **379**, 633–646 (1998).
35. T. L. Campbell, E. K. De Silva, K. L. Olszewski, O. Elemento, M. Llinás, *PLOS Pathog.* **6**, e1001165 (2010).
36. K. Modrzynska *et al.*, *Cell Host Microbe* **21**, 11–22 (2017).
37. J. M. Santos *et al.*, *Cell Host Microbe* **21**, 731–741.e10 (2017).
38. J. P. Wendler *et al.*, *PLOS ONE* **9**, e96486 (2014).
39. A. W. Almawi, L. A. Matthews, A. Guarne, *Prog. Biophys. Mol. Biol.* **127**, 105–110 (2017).
40. N. Kato *et al.*, *Nature* **538**, 344–349 (2016).
41. J. Yuvaniyama *et al.*, *Nat. Struct. Biol.* **10**, 357–365 (2003).
42. R. T. Eastman *et al.*, *Mol. Biochem. Parasitol.* **152**, 66–71 (2007).
43. F. Wang *et al.*, *Mol. Biochem. Parasitol.* **175**, 10–20 (2011).
44. M. Klemba, I. Gluzman, D. E. Goldberg, *J. Biol. Chem.* **279**, 43000–43007 (2004).
45. S. Kenthirapalan, A. P. Waters, K. Matuschewski, T. W. Kooij, *Nat. Commun.* **7**, 10519 (2016).
46. C. C. Paulusma, R. P. Oude Elferink, *Biochim. Biophys. Acta* **1741**, 11–24 (2005).
47. H. P. Schweizer, *FEMS Microbiol. Lett.* **202**, 1–7 (2001).
48. M. Yu *et al.*, *Cell Host Microbe* **4**, 567–578 (2008).
49. C. Alvarez-Ortega, J. Olivares, J. L. Martínez, *Front. Microbiol.* **4**, 7 (2013).
50. S. Coyne, N. Rosenfeld, T. Lambert, P. Courvalin, B. Périchon, *Antimicrob. Agents Chemother.* **54**, 4389–4393 (2010).
51. S. L. Alper, A. K. Sharma, *Mol. Aspects Med.* **34**, 494–515 (2013).
52. R. I. Henry, R. E. Martin, S. M. Howitt, K. Kirk, *Biochem. Biophys. Res. Commun.* **363**, 288–291 (2007).
53. E. R. Geertsma *et al.*, *Nat. Struct. Mol. Biol.* **22**, 803–808 (2015).
54. E. L. Allman, H. J. Painter, J. Samra, M. Carrasquilla, M. Llinás, *Antimicrob. Agents Chemother.* **60**, 6635–6649 (2016).
55. J. Mu *et al.*, *Nat. Genet.* **42**, 268–271 (2010).
56. S. Borrmann *et al.*, *Sci. Rep.* **3**, 3318 (2013).
57. M. B. Jiménez-Díaz *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **111**, E5455–E5462 (2014).
58. C. W. McNamara *et al.*, *Nature* **504**, 248–253 (2013).
59. E. S. Istvan *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **108**, 1627–1632 (2011).
60. B. Baragaña *et al.*, *Nature* **522**, 315–320 (2015).
61. C. Boss *et al.*, *ChemMedChem* **11**, 1995–2014 (2016).
62. E. Sonoiki *et al.*, *Nat. Commun.* **8**, 14574 (2017).
63. G. LaMonte *et al.*, *MBio* **7**, e00696-16 (2016).
64. C. Kimchi-Sarfaty *et al.*, *Science* **315**, 525–528 (2007).
65. MalariaGEN *Plasmodium falciparum* Community Project, *eLife* **5**, e08714 (2016).
66. B. Witkowski *et al.*, *Lancet Infect. Dis.* **17**, 174–183 (2017).
67. P. Cingolani *et al.*, *Fly* **6**, 80–92 (2012).
68. M. Krzywinski *et al.*, *Genome Res.* **19**, 1639–1645 (2009).
69. M. P. Glenn *et al.*, *J. Med. Chem.* **49**, 5710–5727 (2006).

ACKNOWLEDGMENTS

This work was supported by grants from the Bill and Melinda Gates Foundation (OPP1040406 and OPP1119049), the National Institutes of Health (P50GM085764), and the National Institute of Allergy and Infectious Diseases (NIAID; R01AI103058 and R01AI50234). Whole-genome sequencing was performed at the Institute for Genomic Medicine Genomic Center, UCSD. We thank the Penn State Metabolomics Core Facility (University Park, PA) for critical analytical expertise. A.C. received support from the UCSD Division of Infectious Diseases institutional NIAID training grant (T32AI007036). E.L.F. received support from a NIAID National Research Service Award fellowship (F32AI102567). V.C. received support from a UCSD Genetics Training Program through an institutional training grant from the National Institute of General Medical Sciences (T32GM008666). G.M.L. is supported by an A.P. Giannini Post-Doctoral Fellowship. The authors declare no competing financial interests. The red blood cells used in this work were sourced ethically, and their research use was in accord with the terms of the informed consents. All relevant sequence data that were not previously published have been deposited in the National Center for Biotechnology Information (NCBI) Sequence Read Archive (SRA) database with accession code SRP096299. Previously published sequences are deposited in the NCBI SRA with the accession codes listed in table S3. This work is licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0) license, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/. This license does not apply to figures/photos/artwork or other content included in the article that is credited to a third party; obtain authorization from the rights holder before using such material.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/359/6372/191/suppl/DC1
Materials and Methods
Figs. S1 to S10
Tables S1 to S16
References (*70*–*93*)

15 May 2017; accepted 2 November 2017
10.1126/science.aan4472

REPORT

MARTIAN GEOLOGY

Exposed subsurface ice sheets in the Martian mid-latitudes

Colin M. Dundas,[1]* Ali M. Bramson,[2] Lujendra Ojha,[3] James J. Wray,[4] Michael T. Mellon,[5] Shane Byrne,[2] Alfred S. McEwen,[2] Nathaniel E. Putzig,[6] Donna Viola,[2] Sarah Sutton,[2] Erin Clark,[2] John W. Holt[7]

Thick deposits cover broad regions of the Martian mid-latitudes with a smooth mantle; erosion in these regions creates scarps that expose the internal structure of the mantle. We investigated eight of these locations and found that they expose deposits of water ice that can be >100 meters thick, extending downward from depths as shallow as 1 to 2 meters below the surface. The scarps are actively retreating because of sublimation of the exposed water ice. The ice deposits likely originated as snowfall during Mars' high-obliquity periods and have now compacted into massive, fractured, and layered ice. We expect the vertical structure of Martian ice-rich deposits to preserve a record of ice deposition and past climate.

One-third of the Martian surface contains shallow ground ice. This ice is a critical target for science and exploration: it affects modern geomorphology, is expected to preserve a record of climate history, influences the planet's habitability, and may be a potential resource for future exploration. The extent of Martian ground ice and the depth to the ice table have been predicted in theory (*1*–*3*) and have been tested both in situ (*4*) and from orbital observations (*5*–*11*). However, the vertical structure of subsurface ice remains poorly known, including its layering, thickness, and purity, which record its emplacement and subsequent modification processes. Information about the structure, depth, and purity of shallow ice is also required to plan possible in situ resource utilization (ISRU) on future missions (*12*).

Early theoretical predictions suggested that Martian subsurface ice would be ice-cemented ground (*2*). Orbital neutron-spectrometer data have revealed ice contents greater than the likely pore space volume in the upper few centimeters of the ice table in many locations (*5*–*7*). Shallow ice (<1 to 2 m) exposed by fresh impacts remains distinct for months or years, also indicating a low rock or dust content, although possibly modified by the impact process (*9*, *10*, *13*). The Phoenix lander on the northern plains uncovered both ice-cemented regolith and deposits of pure (~99 volume %) ice (*4*, *14*) at a few centimeters depth. The ice at that site likely extends to 9 to 66 m depth on the basis of shallow radar reflections (*15*), but geomorphic interpretations suggest that ice-cemented ground is dominant there (*16*). However, ice-loss landforms indicate that several regions on Mars have high ice volume fractions extending through substantial subsurface depths (*17*, *18*), and radar echo power data have suggested high ice contents beneath areas of the northern plains (*8*). Subsurface radar reflections indicate the presence of debris-covered glaciers (*19*, *20*) as well as buried regional ice sheets in the Utopia and Arcadia Planitiae regions that are up to 170 m thick and nearly pure ice (*21*, *22*). Because the radar did not resolve the top of the ice, it is likely to be within ~20 m of the surface, the limit of the radar instrument's ability to identify shallow signals (*15*). The smaller-scale structure of the ice sheets and rocky cover are unresolvable with current radar data. It remains unknown whether the ice within a few meters of the surface has the same origin and age as the deeper ice because the upper ice is most readily modified and best coupled to the recent climate.

We describe observations of the vertical structure of ground ice using the Mars Reconnaissance Orbiter (MRO). The observations target eight locations that have steep, pole-facing scarps created by erosion (Figs. 1 and 2 and figs. S1 to S3) in images from the High-Resolution Imaging Science Experiment (HiRISE); seven are located in the southern hemisphere, and the eighth location is a cluster of scarps in Milankovič Crater in the north (table S1). Each of the scarps is relatively blue (compared with surrounding terrain) in enhanced-color HiRISE images, and three locations have water-ice signatures in mid-summer spectral data (Fig. 3 and figs. S4 and S5) taken by MRO's Compact Reconnaissance Imaging Spectrometer for Mars (CRISM) (*23*). CRISM spectra show an H_2O ice absorption feature at 1-μm wavelength (fig. S6), which can be masked by as little as 1% soil (*14*). Tens of micrometers of dust are needed to make an opaque cover (*24*), so the dust content is low and/or wind removes dust.

Several lines of evidence indicate that the scarps are exposures of subsurface ice rather than persistent seasonal frost. First, they remain distinct from the surrounding terrain in color in all HiRISE images, including many observations acquired long after seasonal frost has sublimed from steep pole-facing slopes at higher latitudes. Second, Mars Odyssey Thermal Emission Imaging System (THEMIS) observations indicate late-afternoon scarp temperatures above the likely atmospheric frost point (*23*). Last, nearby pits lack relatively blue material that would be expected for topographically controlled frost (fig. S7).

The units hosting the scarps drape the terrain, with some meter-scale boulders on the surface

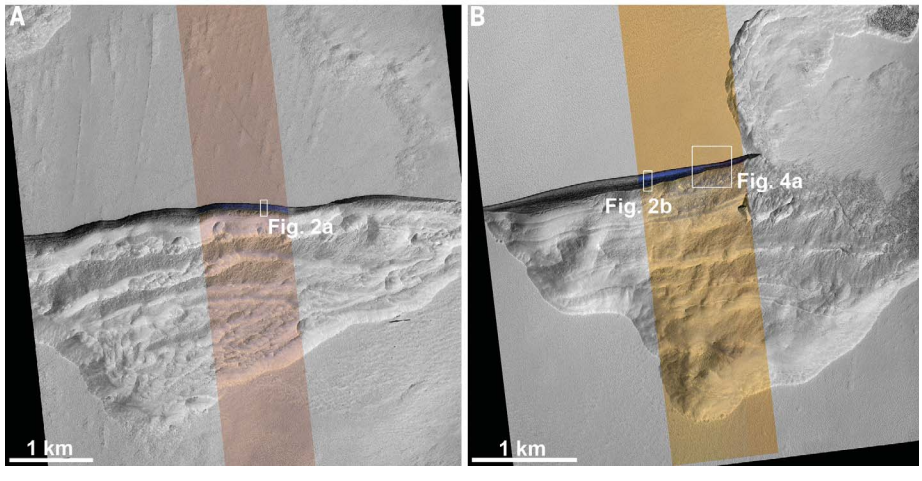


Fig. 1. Pits with scarps exposing ice. (**A** and **B**) Scarps 1 and 2. Both (A) and (B) show HiRISE red-filter data merged with the center color strip (*23*) in early-summer observations. Parallel ridges indicate retreat of scarps (fig. S1). North is up and light is from the left in all figures.

[1]Astrogeology Science Center, U.S. Geological Survey, 2255 N. Gemini Drive, Flagstaff, AZ 86001, USA. [2]Lunar and Planetary Laboratory, University of Arizona, Tucson, AZ, USA. [3]Department of Earth and Planetary Sciences, The Johns Hopkins University, Baltimore, MD, USA. [4]School of Earth and Atmospheric Sciences, Georgia Institute of Technology, Atlanta, GA, USA. [5]The Johns Hopkins University/Applied Physics Laboratory, Laurel, MD, USA. [6]Planetary Science Institute, 1546 Cole Boulevard, Suite 120, Lakewood, CO 80401, USA. [7]Institute for Geophysics, Jackson School of Geosciences, University of Texas at Austin, Austin, TX 78758, USA.
***Corresponding author. Email: cdundas@usgs.gov**

and few impact craters. The surface textures show some fracturing (fig. S8), probably due to thermal contraction. Ice is expected to be stable near 10 cm depth under level ground at these locations (*3*). The scarps are sharply defined and nearly straight, up to ~6 km long, and face slightly east of poleward. Each scarp constitutes the equatorward side of a few-kilometer-scale pit with a rugged and sometimes boulder-rich floor, except in Milankovič Crater, where they are part of the ragged edge of a smooth plateau (fig. S2). In several cases, ridge structures parallel the scarp within the pit and merge with cusps along the margin. The base of the exposed ice appears to be covered by colluvium (Fig. 2). In several cases, surface expressions of long fractures with surrounding depressions occur on the terrain above the scarp (fig. S8).

Topographic data (figs. S9 and S10) for Scarp 1 (numbering is provided in table S1) indicates that the ice exposure has a slope of ~45° (up to 55° in some locations), transitioning to a shallower slope for the regolith-covered lower scarp. We estimate the scarp-hosting material to be at least 130 m thick after correcting for the regional slope, although the exposed ice has ≤100 m of relief because the lower slope is buried. The former value provides a better estimate of the ice-rich thickness if all of the relief is due to ice loss.

These scarps reveal the vertical structure of the ice. At several locations, ice reaches close to the top of the scarp (Fig. 2), indicating that massive ice begins at shallow depths (≤1 m). Banded patterns or variations in color or slope at some sites suggest subunits within the ice. At Scarps 6 and 7, the banded patterns show unconformities where layers cut across each other (fig. S3). The uppermost section of the scarp at each site appears steeper and less blue than the remainder. Like the Martian polar deposits (*25*), the color and brightness of the surface may be partially controlled by a thin coating of dust (likely a sublimation lag) and thus not representative of the bulk properties. Because lags and contaminants will reduce the visibility of ice, it is likely that clean ice near the surface is more extensive than the limited locations where it is exposed (Fig. 2).

At Scarp 2, a lens-shaped section of the exposure lacks relatively blue coloration (Fig. 4A) and contains a concentration of boulders. Repeat images taken 3 Mars years apart show that meter-scale blocks fell from within or below this lens-shaped section (Fig. 4, B and C), providing evidence of active slope retreat. Because these blocks did not disappear after they fell [unlike ice blocks excavated by craters (*10*)] and originated as protrusions from the icy layer, they must be rocks that were embedded within the deposit, not chunks of ice broken off the scarp. Other sites have not shown similar blockfalls, but mottled shifts in tone and texture are common, and dark markings a few meters across appeared and disappeared at the base of some scarps (fig. S11).

The Shallow Radar (SHARAD) on MRO has detected radar contacts (reflectors) that have been interpreted as the base of ice layers in the northern plains of Mars (*15*, *19*–*22*). Similar features occur near several of the scarps (fig. S12) but are difficult to distinguish from radar clutter caused by local rugged topography. Candidate reflectors in the three SHARAD tracks closest to Scarp 1 do not appear in clutter simulations by using topography derived from either Mars Orbiter Laser Altimeter (MOLA) or High-Resolution Stereo Camera (HRSC) data (*23*), suggesting that they could be due to the bottom of the massive ice layer. However, candidate reflectors near Scarp 5 have possible counterparts in the HRSC-based clutter simulations, suggesting that they are not subsurface features. Only these two sites have topographic data available at suitable resolution, and given the rough terrain, we cannot currently rule out unresolved clutter as the cause of candidate reflectors. Nondetections may be due to surface roughness or stratigraphic contacts that are gradual or otherwise do not produce a sharp dielectric contrast on the scale of the radar footprint.

We interpret these scarps as exposures of subsurface ice originating as dusty snow or frost (*26*–*29*) and subsequently compacted and recrystallized. This interpretation is consistent with the high ice content and the mantling appearance of the host unit. Alternative explanations, such as growth of ice lenses (*30*) or enhanced vapor diffusion (*31*), are expected to be slow and operate at shallow depths. They should also produce layering parallel to the ground surface, unlike the layers at Scarps 6 and 7 (fig. S3). The latest ice deposition could have been geologically recent because there are few craters on the surface. The fractures and steep slopes indicate that the ice is cohesive and strong. The presence of banding and color variations suggest layers, possibly deposited with changes in the proportion of ice and dust under varying climate conditions, similar to the Martian polar-layered deposits (*25*). The lens-like body of lower ice content at Scarp 2 (Fig. 4A) may be a buried moraine transported by ice flow because subtle arcuate surface topography suggests glacial flow at this location (fig. S13). Alternatively, it could be a former sublimation lag; that would require some additional process to explain the presence of boulders, but some are found on the surface of the mantling deposit elsewhere (*23*), so this cannot be ruled out.

It is likely that the scarps are currently retreating owing to sublimation. Slope retreat caused by sublimation would explain rocks falling from Scarp 2, although the proximate cause could be seasonal frost processes or thermal cycles. Because the boulders that fell from the scarp are meter-scale, with a few percent of the boulders present having fallen over an interval of 3 Mars years, the likely retreat rate is on the order of a few millimeters each summer (unresolvable with remote imagery). This estimated retreat rate indicates that the pits may have formed over a time scale on an order of 10^6 years; however, the retreat rate is probably not constant. Ridges and bands paralleling the scarps may indicate former positions (Fig. 1).

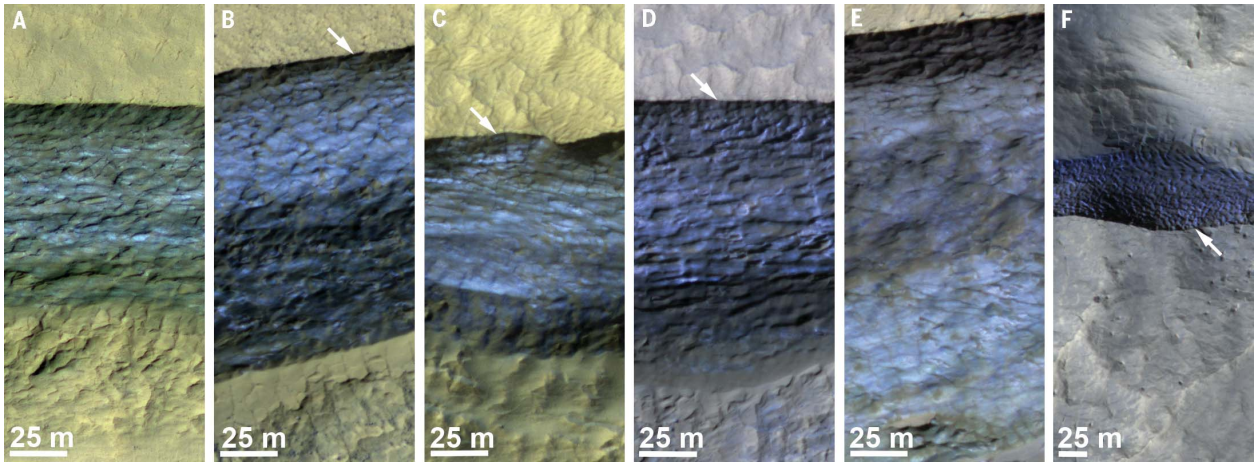


Fig. 2. Enhanced-color transverse sections of icy scarps in late spring/early summer. (**A** to **F**) Arrows indicate locations where relatively blue material is particularly close to the surface. Downhill is to the bottom in (A) to (E) (Scarps 1 to 5, respectively) and to the top in (F) (Milankovič Crater). Banding or layering is visible in several scarps, and (E) shows a distinct change in color as well as multiple fractures cutting the ice.

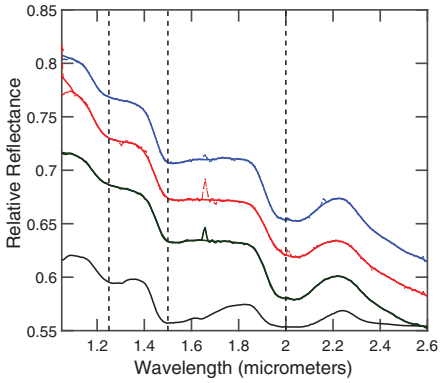


Fig. 3. CRISM spectra of three icy scarps. (Top to bottom lines) Scarps 2, 3, and 1 (vertically offset for clarity) compared with a laboratory water ice spectrum (*32*). Dashed lines indicate the centers of H_2O ice absorption features. The spike near 1.65 μm is an instrument artifact; differences in band shape are likely due to grain size (*33*) or minor effects of the neutral ratio region used in the CRISM spectra.

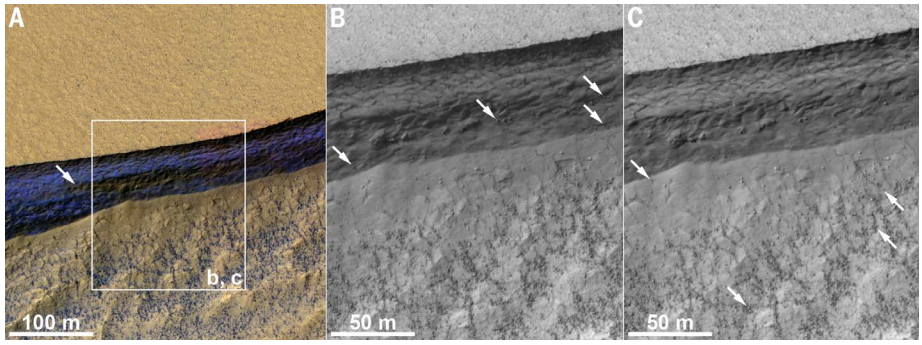


Fig. 4. Blocks falling from an ice-rich scarp. (**A**) Enhanced-color context with arrow pointing to a lens that is redder than the rest of the scarp. (**B** and **C**) Before and after images, respectively, with arrows indicating boulders that have fallen from the lens (B) to the boulder-rich pit floor (C). (B) and (C) are separated by ~3 Mars years (table S2).

The vertical structure of the ice deposits in these exposures appears simple (fig. S14). Ice with low rock and dust content occurs at shallow depths as small as ~1 to 2 m and extends to at least many tens of meters in depth. The ice may be capped by a cover of ice-cemented lithic material, but this veneer cannot be reliably distinguished from massive ice with a superficial lag. In some cases, the scarp color does not clearly indicate massive ice at the shallowest depths, but the textures are continuous across the ice unit. Variations in the apparent depth to massive ice could arise from both true lateral heterogeneity in ice content and from debris covering the uppermost ice, so the depth to the top of the ice sheet may be variable. Debris falling from an ice-free surface layer would be particularly effective at covering the uppermost massive ice. This simple structure matches radar interpretations of thick, regional-scale sheets of ice elsewhere on Mars, covered by lags thinner than the radar-constrained upper bound of 20 m (*21*, *22*). Furthermore, the observations demonstrate that the massive ice in many cases occurs at depths much shallower than that upper bound. The stratigraphy indicated by the scarps may be relevant to other locations with ice-rich deposits, such as Arcadia and Utopia Planitiae. We expect that the structure at lower latitudes will be more complex (for example, more layering) or have thicker covers of ice-free and ice-cemented soil because of reduced overall stability and more frequent unstable conditions (*28*, *29*).

Erosional scarps on Mars reveal the vertical structure of geologically young, ice-rich mantling deposits at mid-latitudes. The ice exposed by the scarps likely originated as snow that transformed into massive ice sheets, now preserved beneath less than 1 to 2 m of dry and ice-cemented dust or regolith near ±55° latitude. These shallow depths make the ice sheets potentially accessible to future exploration, and the scarps present cross-sections of these ices that record past episodes of ice deposition on Mars.

REFERENCES AND NOTES

1. R. B. Leighton, B. C. Murray, *Science* **153**, 136–144 (1966).
2. M. T. Mellon, B. M. Jakosky, *J. Geophys. Res.* **98**, 3345–3364 (1993).
3. M. T. Mellon, W. C. Feldman, T. H. Prettyman, *Icarus* **169**, 324–340 (2004).
4. M. T. Mellon *et al.*, *J. Geophys. Res.* **114**, E00E07 (2009).
5. W. V. Boynton *et al.*, *Science* **297**, 81–85 (2002).
6. W. C. Feldman *et al.* in *The Martian Surface*, J. F. Bell, Ed. (Cambridge Univ. Press, 2008).
7. B. Diez *et al.*, *Icarus* **196**, 409–421 (2008).
8. J. Mouginot *et al.*, *Icarus* **210**, 612–625 (2010).
9. S. Byrne *et al.*, *Science* **325**, 1674–1676 (2009).
10. C. M. Dundas *et al.*, *J. Geophys. Res.* **119**, (2014).
11. J. L. Bandfield, W. C. Feldman, *J. Geophys. Res.* **113**, E08001 (2008).
12. A. Abbud-Madrid *et al.*, Report of the Mars Water In-Situ Resource Utilization (ISRU) Planning (M-WIP) study (April 2016); https://mepag.jpl.nasa.gov/reports/Mars_Water_ISRU_Study.pdf.
13. C. M. Dundas, S. Byrne, *Icarus* **206**, 716–728 (2010).
14. S. Cull *et al.*, *Geophys. Res. Lett.* **37**, L24203 (2010).
15. N. E. Putzig *et al.*, *J. Geophys. Res.* **119**, (2014).
16. M. T. Mellon, R. E. Arvidson, J. J. Marlow, R. J. Phillips, E. Asphaug, *J. Geophys. Res.* **113**, E00A23 (2008).
17. C. M. Dundas, S. Byrne, A. S. McEwen, *Icarus* **262**, 154–169 (2015).
18. S. J. Conway, M. R. Balme, *Geophys. Res. Lett.* **41**, 5402–5409 (2014).
19. J. W. Holt *et al.*, *Science* **322**, 1235–1238 (2008).
20. J. J. Plaut *et al.*, *Geophys. Res. Lett.* **36**, L02203 (2009).
21. A. M. Bramson *et al.*, *Geophys. Res. Lett.* **42**, 6566–6574 (2015).
22. C. M. Stuurman *et al.*, *Geophys. Res. Lett.* **43**, 9484–9491 (2016).
23. Materials and methods are available as supplementary materials.
24. E. N. Wells, J. Veverka, P. Thomas, *Icarus* **58**, 331–338 (1984).
25. K. E. Herkenhoff, S. Byrne, P. S. Russell, K. E. Fishbaugh, A. S. McEwen, *Science* **317**, 1711–1715 (2007).
26. J. W. Head, J. F. Mustard, M. A. Kreslavsky, R. E. Milliken, D. R. Marchant, *Nature* **426**, 797–802 (2003).
27. J.-B. Madeleine *et al.*, *Icarus* **203**, 390–405 (2009).
28. N. Schorghofer, *Nature* **449**, 192–194 (2007).
29. N. Schorghofer, F. Forget, *Icarus* **220**, 1112–1120 (2012).
30. H. G. Sizemore, A. P. Zent, A. W. Rempel, *Icarus* **251**, 191–210 (2015).
31. D. A. Fisher, *Icarus* **179**, 387–397 (2005).
32. R. F. Kokaly *et al.*, *USGS Spectral Library Version 7: U.S. Geological Survey Data Series 1035* 10.3133/ds1035 (2017).
33. R. N. Clark, *J. Geophys. Res.* **86**, 3087–3096 (1999).

ACKNOWLEDGMENTS

Observation planning was funded by the MRO project, and analysis was funded by NASA grants NNH13AV85I (to C.M.D.) and NNX16AP09H (to A.M.B.). L.O. was supported by the Blaustein Postdoctoral Fellowship at Johns Hopkins University. We thank NASA/Jet Propulsion Laboratory and the instrument teams for their efforts collecting and processing data. The SHARAD instrument was provided to NASA's MRO mission by the Italian Space Agency (ASI). We thank M. Christoffersen for help with the HRSC-based clutter simulations, which are available at https://doi.org/10.6084/m9.figshare.5537893. The authors declare no competing financial interests. All of the primary spacecraft data used in this study are available via the Planetary Data System. HiRISE and CTX data are at https://pds-imaging.jpl.nasa.gov/volumes/mro.html, SHARAD data are at http://pds-geosciences.wustl.edu/missions/mro/sharad.htm, CRISM data are at http://pds-geosciences.wustl.edu/missions/mro/crism.htm, and THEMIS data and derived products are at http://viewer.mars.asu.edu/viewer/themis#T=0; observation ID numbers are listed in tables S2 to S5. Derived thermal inertia data are available via https://se.psi.edu/~than/inertia and dust cover index and albedo via www.mars.asu.edu/data.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/359/6372/199/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S17
Tables S1 to S5
References (*34*–*59*)

21 June 2017; accepted 4 December 2017
10.1126/science.aao1619

ASTROCHEMISTRY

Detection of the aromatic molecule benzonitrile (*c*-C_6H_5CN) in the interstellar medium

Brett A. McGuire,[1,2]* Andrew M. Burkhardt,[3] Sergei Kalenskii,[4] Christopher N. Shingledecker,[5] Anthony J. Remijan,[1] Eric Herbst,[3,5] Michael C. McCarthy[2,6]

Polycyclic aromatic hydrocarbons and polycyclic aromatic nitrogen heterocycles are thought to be widespread throughout the universe, because these classes of molecules are probably responsible for the unidentified infrared bands, a set of emission features seen in numerous Galactic and extragalactic sources. Despite their expected ubiquity, astronomical identification of specific aromatic molecules has proven elusive. We present the discovery of benzonitrile (c-C_6H_5CN), one of the simplest nitrogen-bearing aromatic molecules, in the interstellar medium. We observed hyperfine-resolved transitions of benzonitrile in emission from the molecular cloud TMC-1. Simple aromatic molecules such as benzonitrile may be precursors for polycyclic aromatic hydrocarbon formation, providing a chemical link to the carriers of the unidentified infrared bands.

The mid-infrared spectra, roughly from 3 to 20 μm, of the interstellar medium (ISM) and photodissociation regions in both our Galaxy (*1*) and external galaxies (*2*) are dominated by emission features commonly referred to as the unidentified infrared (UIR) bands. Because prominent UIR features closely agree with the characteristic vibrational frequencies of aromatic C-C and C-H bonds, it is now widely accepted that polycyclic aromatic hydrocarbons (PAHs), along with fullerenes such as the recently detected C_{60}^+ molecule (*3*) and their closely related derivatives, are probably the carriers responsible for most of these features (*4*). A substantial fraction of interstellar carbon is calculated to be in the form of PAHs [~10% (*5*)], yet the origin of these aromatic species is a topic of considerable debate. In the diffuse ISM and PDR regions, where 30 to 60% of the carbon is locked up in dust grains (*6*), top-down models of PAH formation—through the destruction of dust grains by the harsh radiation environment, shock waves, or both—may be viable pathways (*7*). In denser molecular clouds, which are not subject to the ultraviolet radiation and have not been subject to shocks, other pathways must exist to synthesize these species from smaller precursor molecules.

Despite the widespread acceptance of PAHs as a common class of interstellar molecules, no specific PAH has been identified in the ISM, either by rotational spectroscopy or by observations of its infrared features, despite sustained efforts (*8*). In the microwave and (sub-)millimeter regimes, some laboratory data do exist (*9*), but such studies are relatively uncommon. Many PAHs are poor candidates for detection through radio astronomy, both because of unfavorably large rotational partition functions and because they are either apolar or weakly polar, and thus lack sufficiently intense rotational lines (relative to linear molecules of similar composition and size). A notable exception is corrannulene ($C_{20}H_{10}$), a bowl-shaped molecule with a relatively large permanent dipole moment [2.07 debye (*10*)], but astronomical searches for that molecule have been unsuccessful as well (*11*). In the infrared, although a concerted effort has been undertaken to catalog both laboratory and theoretical vibrational and Raman spectra of PAHs (*12*), the structural similarities among individual species result in spectra that are often indistinguishable at the modest resolving power that can routinely be achieved by astronomical observations; as a result, aggregate spectra consisting of many PAHs are invoked to reproduce astronomical features (*4*).

For these reasons, many attempts to understand the chemistry of PAHs have focused on the possible formation pathways that proceed through more readily detectable molecules. Much effort has been centered on modeling the formation of small five- and six-membered aromatic rings and their subsequent reactions with smaller hydrocarbons and nitrogen species to produce PAHs and polycyclic aromatic nitrogen heterocycles [PANHs (*13*)]. To date, the only interstellar detection of a five- or six-membered aromatic ring is benzene (C_6H_6), through the observation of a single weak absorption feature arising from its ν_4 bending mode near 14.85 μm in a handful of sources (*14*–*17*). The lack of a permanent dipole moment, however, precludes the identification of benzene via its rotational transitions.

Here, we searched for a number of simple aromatic molecules, including several PA(N)Hs and nitriles (R–C≡N), a class of molecules believed to give rise to a common UIR feature at 6.2 μm (*18*). The astronomical source targeted in these observations was the cold-core Taurus Molecular Cloud 1 (TMC-1), which has long been known to possess a rich chemistry dominated by unsaturated carbon-chain molecules such as the cyanopolyynes (HC_nN; n = odd) [e.g., (*19*–*22*)]. The initial search was performed by construction of velocity-stacked composite-average spectra of 12 target molecules (Fig. 1) (*23*) using existing survey data taken with the Nobeyama 45-m telescope (*19*). This method enhances the signal-to-noise ratio (SNR) of a potential molecular detection by averaging the signal from multiple transitions of a molecule in velocity space. These composite averages are effective preliminary indicators of a molecule in a source such as TMC-1, where spectral features are narrow (0.3 to 0.5 km s^{-1}), the line density is relatively low (~1 line per 200 km s^{-1}), and the molecules occupy a narrow range in local standard of rest (LSR) velocity (v_{lsr} = 5.5 to 5.9 km s^{-1}) (*19*). As shown in Fig. 1, the composite spectra show highly suggestive evidence for benzonitrile in this source. Nonetheless, the observation of individual transitions is required to establish a firm detection and to enable the robust determination of the molecular abundance. The sensitivity and spectral resolution of the existing survey observations, however, were insufficient for that task.

We performed observations with the 100-m Robert C. Byrd Green Bank Telescope to confirm the detection of benzonitrile by observing nine of its individual rotational transitions, using deep integrations at high spectral resolution. Because the spectral features of other molecules in TMC-1 are so narrow, ^{14}N nuclear hyperfine structure is expected to be partially resolved in benzonitrile's lower rotational transitions. Existing spectral catalogs for benzonitrile in public databases did not contain hyperfine-splitting frequencies, and existing laboratory work at high resolution was limited to measurements below 11 GHz (*24*). For these reasons, additional transitions of benzonitrile were measured in the laboratory at high resolution between 7 and 29 GHz to ensure that the astronomical data could be interpreted (*25*).

Molecules in TMC-1 are typically well described by a single excitation temperature between 5 and 10 K (*21*, *22*). Under these conditions, the strongest benzonitrile transitions fall between 20 and 40 GHz. A total of 1.875 GHz of bandwidth was covered to high sensitivity (T_A^* = 2 to 5 mK) between 18 and 23 GHz. In this range, eight of the nine strongest predicted rotational transitions were observed, each with SNR ≥ 3 (Fig. 2). For six of these, characteristic ^{14}N nuclear hyperfine splitting is partially or fully resolved for one or more components (Table 1). The emission features are best described by a v_{lsr} value of 5.83 km s^{-1}, a typical velocity for molecules in this source (*19*). All other strong transitions between 18 and 23 GHz

[1]National Radio Astronomy Observatory, Charlottesville, VA 22903, USA. [2]Harvard-Smithsonian Center for Astrophysics, Cambridge, MA 02138, USA. [3]Department of Astronomy, University of Virginia, Charlottesville, VA 22904, USA. [4]Astro Space Center, Lebedev Physical Institute, Russian Academy of Sciences, Moscow, Russia. [5]Department of Chemistry, University of Virginia, Charlottesville, VA 22904, USA. [6]School of Engineering and Applied Sciences, Harvard University, Cambridge, MA 02138, USA.
*Corresponding author. Email: bmcguire@nrao.edu

fell into gaps in the spectral coverage, or in regions where insufficient noise levels were achieved. Taken together, these findings establish the presence of benzonitrile in TMC-1.

A joint analysis of all the lines yields a total column density $N_T = 4 \times 10^{11}$ cm^{-2} (*25*), about 5% that of HC_7N [1.1×10^{13} cm^{-2} (*21*)], an unsaturated linear cyanopolyyne with the same carbon and nitrogen composition as benzonitrile, in the same source. Because the upper state energies of the observed transitions span only a narrow energy range (3.6 to 5.7 K), the excitation temperature T_{ex} could not be constrained from the present observations. Our analysis therefore assumed T_{ex} = 7 K, in the middle of the range of 5 to 10 K, derived from other molecules in this source (*21*, *22*). We also constrained the linewidth to 0.4 km s^{-1}, based on the three fully resolved hyperfine components. Although these components are some of the lowest SNR features, and our data are limited by the resolution of the observations (0.08 km s^{-1}), this linewidth is consistent with that previously derived for other molecules in this source (*19*). Simulated spectra under these conditions (Fig. 2) are in agreement with the observations.

The pathways leading to the formation of benzonitrile at low temperature and in low-density environments have not been studied in detail. Perhaps the only promising astrochemically relevant formation pathway discussed in the literature is the neutral-neutral reaction

$$CN + c\text{-}C_6H_6 \rightarrow c\text{-}C_6H_5CN + H \qquad (1)$$

This barrierless, exothermic reaction has been considered previously (*26*, *27*). In an effort to determine the contribution of Reaction 1 to the observed abundance of benzonitrile in TMC-1, we have modified the Kinetic Database for Astrochemistry (KIDA) gas-phase reaction network (*28*) to include this reaction, as well as destruction pathways from photons, ions, and depletion onto grains (*25*). This network was then combined with the NAUTILUS-1.1 modeling code (*29*) assuming elemental abundances and physical conditions appropriate for TMC-1 (*30*). A number of additional gas-phase formation routes for the precursor benzene were also considered and included in the modified network (*25*). A column density of $H_2 = 10^{22}$ cm^{-2} (*22*) was used to convert from modeled abundances to column densities to compare with observations. Figure 3 shows the derived column densities and those predicted by the model for benzonitrile, benzene, CN, and the cyanopolyynes HC_3N, HC_5N, HC_7N, and HC_9N. Although the calculated column densities of most of the cyanopolyynes agree with observational results within a factor of 2, the predicted benzonitrile column density is smaller than the derived value by nearly a factor of 4.

This difference may be explained by other formation routes for benzene and/or benzonitrile that are not considered in our model. For instance, experiments have found that benzene can be formed in electron-irradiated acetylene ices (*31*). In astrochemical models, the addition of cosmic ray–driven irradiation chemistry in the solid phase has been found to improve agreement between observational and theoretical abundances for other large interstellar molecules (*32*), although for such grain-surface processes to contribute to gas-phase abundances, there must exist efficient nonthermal desorption mechanisms. Recent theoretical and experimental work suggests that interactions between cosmic rays and grain surfaces could result in the liberation of solid-phase species into the gas phase via processes that are viable in cold cores such as TMC-1 (*33*).

Benzene is also known to be produced in irradiated acetylene gas (*34*). Radiation chemistry differs from photochemistry in a number of ways (*35*) and may be a viable formation pathway for aromatic and PAH molecules as a result of their increased photostability relative to simpler organics (*36*). However, there is insufficient theoretical and experimental work to include such

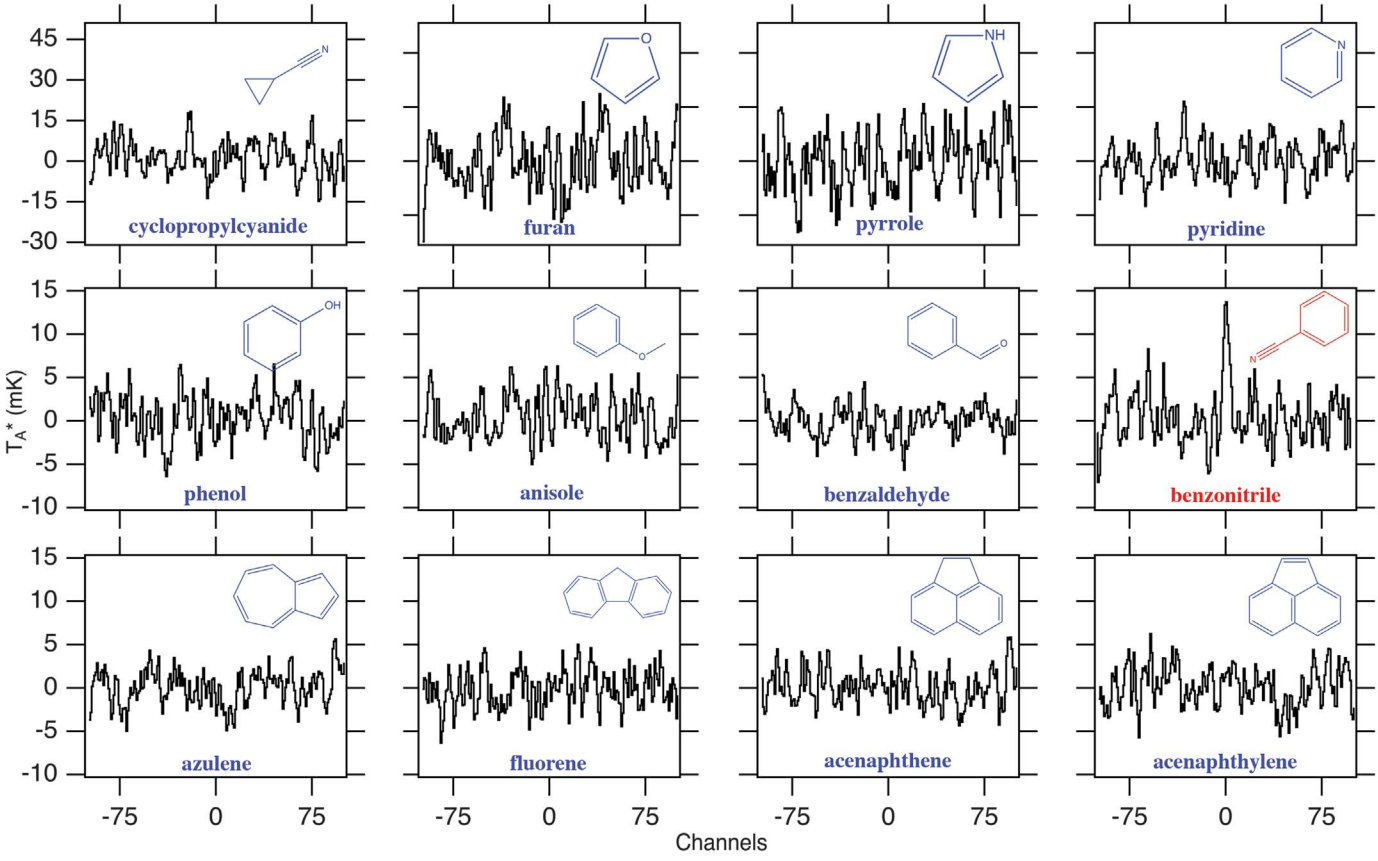


Fig. 1. Composite averages of molecules toward TMC-1. Velocity-stacked composite averages of all transitions of a given molecule with upper state energy (E_U) < 70 K constructed from the entire survey (8.8 to 50 GHz) of TMC-1 (*19*) are shown. The channel spacing is 20 kHz. If a molecule is present, signal in antenna temperature (T_A*) would be expected at channel 0. A detectable signal was present only for benzonitrile (red).

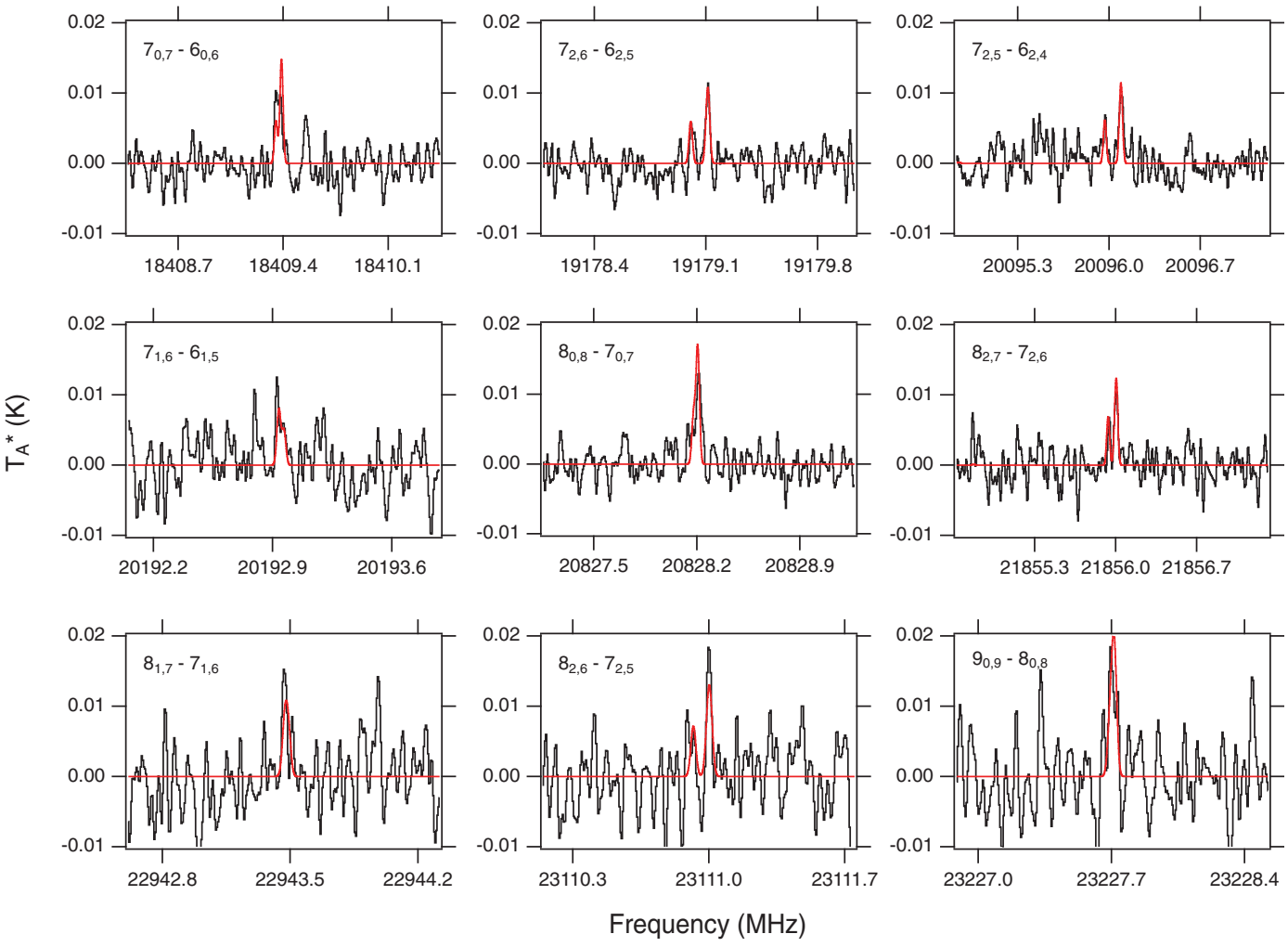


Fig. 2. Detected emission lines of benzonitrile in TMC-1. Observational spectra are shown in black smoothed to a resolution of 5.7 kHz (0.08 km s^{-1}) and shifted to a v_{lsr} value of 5.83 km s^{-1}. A simulated spectrum of benzonitrile (linewidth = 0.4 km s^{-1}, N_T = 4 × 10^{11} cm^{-2}, T_{ex} = 7 K) is overlaid in red (*25*). Rotational quantum numbers are displayed at the upper left of each panel. The four transitions with well-resolved hyperfine structure are shown on an expanded frequency axis in (*25*).

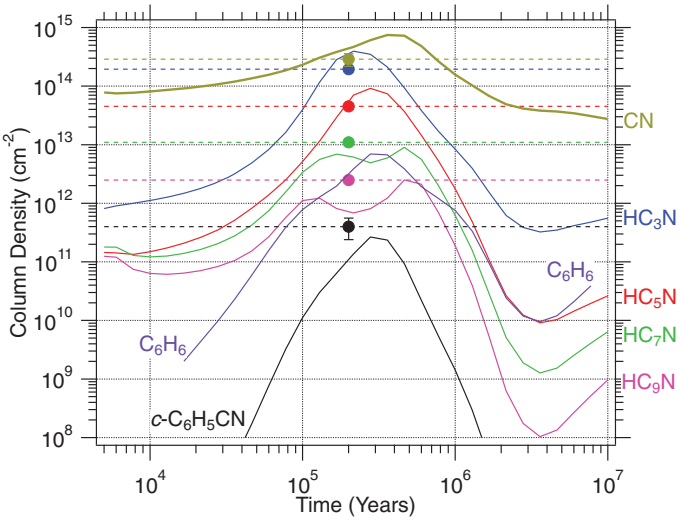


Fig. 3. Chemical model of TMC-1. Results of a three-phase astrochemical model are shown, updated to include the benzonitrile formation pathway given in Reaction 1. The model gas-phase column densities as a function of time are given as solid colored lines. Column densities of benzonitrile, cyanopolyynes, and CN derived from observations are shown as dots with dashed horizontal lines [from top to bottom: CN, HC_3N, HC_5N, HC_7N, HC_9N, and benzonitrile (*21*, *39*)]. The derived column densities are plotted at the chemical age of TMC-1 [~2 × 10^5 years (*30*)].

pathways in our model. Previous work also suggests a possible connection between benzonitrile and the cyanopolyynes (*21*). That work showed a sharp decrease in the calculated abundance of $HC_{11}N$ relative to the abundance trend of the n(odd) = 3 to 9 cyanopolyynes, possibly due to cyclization processes for HC_nN molecules that could eventually lead to functionalized aromatics such as benzonitrile.

We also considered the possibility that benzonitrile itself may be a contributor to the UIR bands. The vibrational spectrum of benzonitrile has been studied in the infrared, both experimentally (*37*) and theoretically (*38*), but it does not appear in spectral databases (*12*) and is not commonly considered as a potential UIR carrier. Nonetheless, the interstellar IR emission features at 3.3 μm (C–H aromatic stretch) and 4.48 μm (C≡N stretch) are both in agreement with very strong IR modes of benzonitrile (*38*), thus making it a potential carrier in its own right, as well as a likely precursor to polyaromatic species.

Table 1. Detected benzonitrile transitions. Shown are quantum numbers, frequencies, upper state energies (E_U), line strengths ($S_{ij}\mu^2$), observed intensities (ΔT_A*), and signal-to-noise ratio (SNR) of detected benzonitrile transitions. ΔT_A* values reflect uncertainty in the Gaussian fit; a conservative 30% uncertainty in the absolute flux calibrated value is assumed (*25*). Statistical uncertainties (1σ), derived from the best-fitting constants in (*25*), are given in parentheses in units of the last significant digit.

Transition		Frequency (MHz)	E_U (K)	$S_{ij}\mu^2$ (debye^2)	ΔT_A* (mK)	SNR
$J'_{K_a,K_c} - J''_{K_a,K_c}$	$F' - F''$					
$7_{0,7}$–$6_{0,6}$	6–5	18409.3490(2)	3.61	39.5		
	8–7	18409.3840(2)	3.61	52.9	10.3(8)[b]	4.4
	7–6	18409.3879(2)	3.61	45.7		
$7_{2,6}$–$6_{2,5}$	7–6	19179.0017(2)	4.52	42.4	5.4(5)	2.7
	6–5	19179.1027(2)	4.52	36.6	10.6(5)[b]	5.4
	8–7	19179.1128(2)	4.52	49.1		
$7_{2,5}$–$6_{2,4}$	7–6	20095.9645(2)	4.62	42.5	6.6(5)	3.3
	6–5	20096.0824(2)	4.62	36.7	10.0(2)[b]	4.9
	8–7	20096.0917(2)	4.62	49.2		
$7_{1,6}$–$6_{1,5}$	6–5	20192.9325(2)	4.11	39.0		
	7–6	20192.9342(2)	4.11	45.1	8(1)[b]	2.1
	8–7	20192.9632(2)	4.11	52.2		
$8_{0,8}$–$7_{0,7}$	7–6	20828.1746(2)	3.20	46.2		
	9–8	20828.2012(2)	3.20	59.6	11.7(9)[b]	5.9
	8–7	20828.2045(2)	3.20	52.5		
$8_{2,7}$–$7_{2,6}$	8–7	21855.9322(3)	5.57	49.7	7.1(4)	3.1
	7–6	21855.9944(3)	5.57	43.8	10.7(5)[b]	4.7
	9–8	21856.0064(3)	5.57	56.4		
$8_{1,7}$–$7_{1,6}$	7–6	22943.4640(3)	5.21	45.8		
	8–7	22943.4729(3)	5.21	52.0	16.2(8)[b]	3.2
	9–8	22943.4885(3)	5.21	59.1		
$8_{2,6}$–$7_{2,5}$	8–7	23110.9171(3)	5.73	49.9	6.2(9)	1.2
	7–6	23110.9923(3)	5.73	43.9	19.6(9)[b]	3.8
	9–8	23111.0042(3)	5.73	56.7		
$9_{0,9}$–$8_{0,8}$	8–7	23227.6903(3)	5.72	52.9		
	10–9	23227.7105(3)	5.72	66.3	16(2)[b]	3.3
	9–8	23227.7127(3)	5.72	59.3		

[b]Indicates blended hyperfine components; ΔT_A* is the peak value of the observed feature.

REFERENCES AND NOTES

1. F. J. Low *et al.*, *Astrophys. J.* **278**, L19 (1984).
2. M. W. Regan *et al.*, *Astrophys. J. Suppl. Ser.* **154**, 204–210 (2004).
3. E. K. Campbell, M. Holz, D. Gerlich, J. P. Maier, *Nature* **523**, 322–323 (2015).
4. A. G. G. M. Tielens, *Annu. Rev. Astron. Astrophys.* **46**, 289–337 (2008).
5. J. E. Chiar, A. G. G. M. Tielens, A. J. Adamson, A. Ricca, *Astrophys. J.* **770**, 78 (2013).
6. B. T. Draine, *Astrophys. J.* **831**, 109 (2016).
7. O. Berné, J. Montillaud, C. Joblin, *Astron. Astrophys.* **577**, A133 (2015).
8. V. Kofman, P. Sarre, R. Hibbins, I. ten Kate, H. Linnartz, *Mol. Astrophys.* **7**, 19–26 (2017).
9. S. Thorwirth, P. Theule, C. A. Gottlieb, M. C. McCarthy, P. Thaddeus, *Astrophys. J.* **662**, 1309–1314 (2007).
10. F. J. Lovas *et al.*, *J. Am. Chem. Soc.* **127**, 4345–4349 (2005).
11. P. Pilleri *et al.*, *Mon. Not. R. Astron. Soc.* **397**, 1053–1060 (2009).
12. C. Boersma *et al.*, *Astrophys. J. Suppl. Ser.* **211**, 8 (2014).
13. R. I. Kaiser, D. S. N. Parker, A. M. Mebel, *Annu. Rev. Phys. Chem.* **66**, 43–67 (2015).
14. J. Cernicharo *et al.*, *Astrophys. J.* **546**, L123–L126 (2001).
15. S. E. Malek, J. Cami, J. Bernard-Salas, *Astrophys. J.* **744**, 16 (2011).
16. K. E. Kraemer *et al.*, *Astrophys. J.* **652**, L25–L28 (2006).
17. D. A. García-Hernández, P. García-Lario, J. Cernicharo, D. Engels, J. V. Perea-Calderón, *J. Phys. Conf. Ser.* **728**, 052003 (2016).
18. D. M. Hudgins, C. W. Bauschlicher Jr., L. J. Allamandola, *Astrophys. J.* **632**, 316–332 (2005).
19. N. Kaifu *et al.*, *Publ. Astron. Soc. Jpn.* **56**, 69–173 (2004).
20. B. A. McGuire *et al.*, *Astrophys. J.* **843**, L28 (2017).
21. R. A. Loomis *et al.*, *Mon. Not. R. Astron. Soc.* **463**, 4175–4183 (2016).
22. P. Gratier *et al.*, *Astrophys. J. Suppl. Ser.* **225**, 25 (2016).
23. S. V. Kalenskii, in *Proceedings of the Russian-Indian Workshop on Radio Astronomy and Star Formation, October 10–12, 2016*, I. Zinchenko, P. Zemlyanukha, Eds. (Institute of Applied Physics RAS, 2017), p. 43.
24. K. Wohlfart, M. Schnell, J.-U. Grabow, J. Küpper, *J. Mol. Spectrosc.* **247**, 119–121 (2008).
25. See supplementary materials.
26. D. E. Woon, *Chem. Phys.* **331**, 67–76 (2006).
27. A. J. Trevitt, F. Goulay, C. A. Taatjes, D. L. Osborn, S. R. Leone, *J. Phys. Chem. A* **114**, 1749–1755 (2010).
28. V. Wakelam *et al.*, *Astrophys. J. Suppl. Ser.* **217**, 20 (2015).
29. M. Ruaud, V. Wakelam, F. Hersant, *Mon. Not. R. Astron. Soc.* **459**, 3756–3767 (2016).
30. U. Hincelin *et al.*, *Astron. Astrophys.* **530**, A61 (2011).
31. L. Zhou *et al.*, *Astrophys. J.* **718**, 1243–1251 (2010).
32. M. J. Abplanalp *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **113**, 7727–7732 (2016).
33. Fde. A. Ribeiro *et al.*, *Phys. Chem. Chem. Phys.* **17**, 27473–27480 (2015).
34. F. H. Field, *J. Phys. Chem.* **68**, 1039–1047 (1964).
35. J. F. Wishart, *Photochemistry and Radiation Chemistry: A Perspective* (ACS Publications, 1998).
36. H. W. Jochims, H. Baumgartel, S. Leach, *Astrophys. J.* **512**, 500–510 (1999).
37. J. Green, D. J. Harrison, *Spectrochim. Acta A* **32**, 1279–1286 (1976).
38. A. G. Császár, G. Fogarasi, *Spectrochim. Acta A* **45**, 845–854 (1989).
39. R. M. Crutcher, E. Churchwell, L. M. Ziurys, *Astrophys. J.* **283**, 668 (1984).

ACKNOWLEDGMENTS

Supported by NASA through Hubble Fellowship grant HST-HF2-51396 awarded by the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy Inc. for NASA under contract NAS5-26555 (B.A.M.); NSF through the Grote Reber Fellowship Program administered by Associated Universities Inc./National Radio Astronomy Observatory (A.M.B.); Basic Research Program P-7 of the Presidium of the Russian Academy of Sciences (S.V.K.); NSF grant AST 15-14844 (E.H. and C.N.S.); and NSF grant AST-1615847 (M.C.M.). We thank G. A. Blake for a critical reading of the manuscript and J. Mangum for helpful discussions. The National Radio Astronomy Observatory and the Green Bank Observatory are facilities of the National Science Foundation operated under cooperative agreement by Associated Universities Inc. The Nobeyama observations are archived at www.cv.nrao.edu/PRIMOS/datasummary.html, and the GBT observations at https://archive.nrao.edu/archive/advquery.jsp under project codes AGBT02C_012, AGBT17A_164, and AGBT17A_434. Laboratory data are tabulated in the supplementary materials. The modifications we applied to KIDA and the output of our chemical model are available at doi:10.18130/V3/4IFDBP.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/359/6372/202/suppl/DC1
Materials and Methods
Figs. S1 and S2
Tables S1 to S5
References (*40*–*58*)

31 July 2017; accepted 15 November 2017
10.1126/science.aao4890

POROUS MATERIALS

Ordered macro-microporous metal-organic framework single crystals

Kui Shen,[1] Lei Zhang,[1] Xiaodong Chen,[1] Lingmei Liu,[2] Daliang Zhang,[3] Yu Han,[2] Junying Chen,[1] Jilan Long,[1] Rafael Luque,[4] Yingwei Li,[1]* Banglin Chen[5]*

We constructed highly oriented and ordered macropores within metal-organic framework (MOF) single crystals, opening up the area of three-dimensional–ordered macro-microporous materials (that is, materials containing both macro- and micropores) in single-crystalline form. Our methodology relies on the strong shaping effects of a polystyrene nanosphere monolith template and a double-solvent–induced heterogeneous nucleation approach. This process synergistically enabled the in situ growth of MOFs within ordered voids, rendering a single crystal with oriented and ordered macro-microporous structure. The improved mass diffusion properties of such hierarchical frameworks, together with their robust single-crystalline nature, endow them with superior catalytic activity and recyclability for bulky-molecule reactions, as compared with conventional, polycrystalline hollow, and disordered macroporous ZIF-8.

Porous materials have attracted considerable scientific attention because of their wide industrial applications ranging from gas separation to catalysis (*1*). Topological features—particularly pore sizes at each dimension and the uniformity of these pores—in porous materials are key factors that determine their preferred function in a particular application (*2*, *3*). The construction of hierarchical porous structures that maintain their overall crystalline order is desirable because the high structural regularity may, in turn, afford markedly improved properties. A few advantageous examples include substantially improved conductivities and electron mobilities for mesoporous TiO_2 single crystals (*4*) as compared with nanocrystalline TiO_2 and much stronger framework acidities and stabilities for mesoporous crystalline zeolites with respect to amorphous molecular sieves (*5*, *6*).

The synthesis of highly ordered meso- or macroporous crystalline materials at a wide range of length scales with different structural coherencies remains an important goal (*7*, *8*), particularly with respect to enlarging the pores of metal-organic frameworks (MOFs). Most reported MOFs exhibit porosities restricted to the microporous regime, which limits their applications to diffusion-limited processes (*9*, *10*). Synthetic strategies to generate controllable porosities within MOFs mainly include templating, etching, and template-free and defect-induced methods (*11*–*14*). However, these approaches cannot generate ordered macro-microporous MOF single crystals. Alternatively, extendable ligands and three-dimensional (3D)–ordered templates have been used to design periodic meso- and macropores (*15*–*19*). The mesopores built from long ligands often suffer from rapid collapse after guest removal, whereas template-duplicated macropores are generally derived from the assembly of intergrown MOF polycrystals. Until now, MOFs featuring both macropore ordering (pores with diameter > 50 nm) and single crystallinity have not been achieved. On the basis of these premises, the development of reliable methods to design ordered meso- or macroporous MOF single crystals with robust framework structures is imperative.

We report a single-crystalline MOF with 3D ordering of macro-micropores, and we used the ZIF-8 structure as a proof of concept. This structure's 3D framework is built up from connecting Zn^{II} ions through 2-methylimidazole linkers, showing a periodically microporous topology featuring large cavities (11.6 Å) and small apertures (3.4 Å) (*20*, *21*). First, monodisperse polystyrene spheres (PSs) were assembled into highly ordered 3D PS monoliths (Fig. 1A and figs. S1 and S2). ZIF-8 precursors—that is, 2-methylimidazole and $Zn(NO_3)_2$—were then filled into the PS monolith interstices to form "precursor@PS" monoliths, which were subsequently soaked in CH_3OH/$NH_3{\cdot}H_2O$ mixed solutions. We chose the mixed solutions as the solvent, with the idea that $NH_3{\cdot}H_2O$ can induce rapid crystallization of the precursors, whereas CH_3OH can effectively stabilize such precursors and adjust the balance between nucleation versus growth of ZIF-8. As a result, the precursors smoothly turned into oriented and 3D-ordered ZIF-8 single crystals (fig. S3) (*22*). Obviously, the proposed double-solvent–assisted method could overcome the energetic barrier to homogeneous nucleation, which is usually unavoidable in conventional nanocasting processes (*23*, *24*). This new type of macro-microporous MOF was denoted as SOM-ZIF-8(*x*, *y*) (SOM stands for single-crystal ordered macropore), in which *x* and *y* represented the feeding molar ratios of 2-methylimidazole/$Zn(NO_3)_2$ and the volume fraction of $NH_3{\cdot}H_2O$ in the solvent used for synthesis, respectively.

Low-resolution scanning electron microscopy (SEM) images (Fig. 1B and fig. S4) revealed that SOM-ZIF-8(3, 0.5) displayed a tetrakaidecahedron morphology, although some crystals looked different as various directions were viewed. The oriented arrangement of 3D-ordered macropores could be identified in the representative SEM images of individual crystals taken from four different directions (Fig. 1C). On inspection, we identified that the ordered arrangement of macropores on the surface of six square planes and eight triangle planes inherited {100} and {111} planes of the colloidal PS template, respectively.

The crystal growth of ZIF-8 is known to start with the formation of cubes composed of 6 (100) facets, which can subsequently grow into truncated rhombic dodecahedra composed of 12 (110) and 6 (100) facets (*25*, *26*). In our study, ZIF-8 growing on the external surface of a PS template also displayed a truncated rhombic dodecahedra morphology with 6 (100) and 12 (110) well-developed facets (Fig. 1D and fig. S5). Given that the difference between the two crystal morphologies was the existence of the PS template, we assumed that the PS template was the controlling agent that allowed for the preferential formation of tetrakaidecahedron SOM-ZIF-8 with oriented macropore arrangement.

To validate this assumption, we conducted a controlled experiment in which the precursor@PS monolith was first broken down to very small particles before soaking with CH_3OH/$NH_3{\cdot}H_2O$ for further crystallization (under otherwise identical synthetic conditions). As expected, almost all of the obtained crystals also exhibited an obvious truncated rhombic dodecahedra morphology (fig. S6). A closer inspection revealed that the (100) facets of many poorly developed crystals also inherited {100} planes of the colloidal PS template. A selected crystal (Fig. 1E) provided direct insights into the effects of ordered PSs as template on the morphological features of SOM-ZIF-8(3, 0.5). In one part, the 3D PS template guided the crystal growth into a macroporous tetrakaidecahedron with its (100) and (111) facets matching well with the {100} and {111} planes of macropore arrangements, respectively. However, the crystal had grown into conventional macropore-free truncated rhombic dodecahedra in the remaining part, owing to the absence of PSs. This oriented growth model was further confirmed by a broken crystal composed of an internal oriented 3D-ordered macroporous system and an external compact faceted surface, consistent with conventional ZIF-8 morphology (Fig. 1F). Thus, the shaping and templating effects of the 3D-ordered PS template could directly guide the morphological evolution of SOM-ZIF-8 into a tetrakaidecahedron morphology and match its crystal face to the

[1]School of Chemistry and Chemical Engineering, South China University of Technology, Guangzhou 510640, China. [2]Advanced Membranes and Porous Materials Center, Physical Sciences and Engineering Division, King Abdullah University of Science and Technology (KAUST), Thuwal 23955-6900, Saudi Arabia. [3]Imaging and Characterization Core Lab, KAUST, Thuwal 23955-6900, Saudi Arabia. [4]Departamento de Química Orgánica, Universidad de Córdoba, Edificio Marie Curie, Carretera Nacional IV-A, Km 396, E14014, Córdoba, Spain. [5]Department of Chemistry, University of Texas at San Antonio, One UTSA Circle, San Antonio, TX 78249, USA.
*Corresponding author. Email: liyw@scut.edu.cn (Y.L.); banglin.chen@utsa.edu (B.C.)

oriented 3D-ordered macroporous arrangement (Fig. 1G).

X-ray diffraction (XRD) patterns of SOM-ZIF-8(3, 0.5) are compared with conventional (denoted C-ZIF-8, fig. S7) and simulated ZIF-8 in Fig. 2A. SOM-ZIF-8(3, 0.5) exhibited only XRD reflections ascribed to ZIF-8 (*20*, *21*), confirming the formation of phase-pure ZIF-8 with good crystallinity. Nitrogen-adsorption experiments were further performed to investigate the porosity in the material (Fig. 2B). Both SOM-ZIF-8(3, 0.5) and C-ZIF-8 showed similar type I isotherms with a high nitrogen adsorption capacity at very low relative pressures, indicative of their microporous structures (*20*, *21*), as also confirmed by the corresponding micropore size distribution curves (Fig. 2B, inset). The Brunauer-Emmett-Teller surface area and micropore volume of SOM-ZIF-8(3, 0.5) were 1540 m^2/g and 0.59 cm^3/g, respectively—values higher than those of C-ZIF-8 (1397 m^2/g and 0.55 cm^3/g). The formation of a 3D-ordered macroporous system did not influence its inherent microporous structure. The presence of macropores in SOM-ZIF-8(3, 0.5) was fully confirmed by means of mercury intrusion porosimetry. As shown in fig. S8, SOM-ZIF-8(3, 0.5) displays a regular macroporous distribution with an average size of ~80 nm, which could be assigned to the throat size of "ink-bottle" pores between neighboring macropores.

SOM-ZIF-8(3, 0.5) was subsequently characterized in detail by transmission electron microscopy (TEM) to further confirm its macroporous structure and composition. A representative high-angle annular dark-field–scanning transmission electron microscopy (HAADF-STEM) image revealed that closely packed ordered PSs were imprinted into entire SOM-ZIF-8(3, 0.5) crystals (Fig. 2C). The average diameter of the interconnected macropores was ~270 nm, similar to that of PS latex, because the ordered macroporous structure was a negative replica of the PS template. The highly ordered macropore arrangement in SOM-ZIF-8(3, 0.5) was clearly evidenced from low-magnification TEM images of different directions (fig. S9). Selected-area electron diffraction (SAED) patterns taken from an entire crystal (Fig. 2D, inset) revealed the single-crystalline nature of SOM-ZIF-8(3, 0.5). High-resolution TEM (HRTEM) images at different magnifications along the <011> zone axis showed uniform lattice fringes with consistent orientations over the entire image region, without domain boundaries or interfaces observed (Fig. 2, D to F). Fourier transform (FT) patterns derived from HRTEM further confirmed the single crystalline nature of this material (Fig. 2E, inset). The elemental mapping images (fig. S10) pointed to a homogeneous distribution of elements (including C, N, and Zn) in whole SOM-ZIF-8(3, 0.5) crystals.

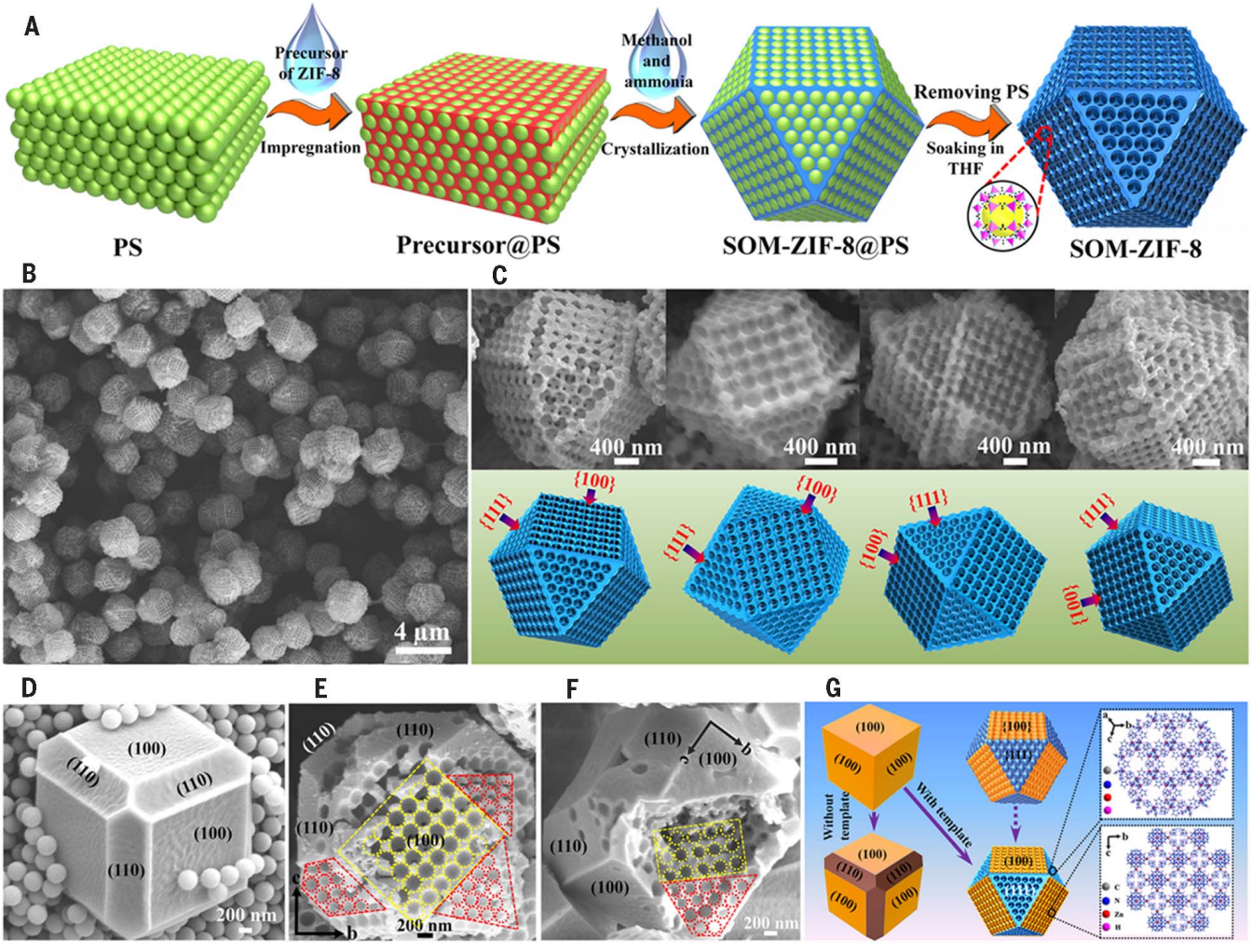


Fig. 1. In situ nanocasting synthesis of SOM-ZIF-8 and its structure confirmation. (**A**) Schematic diagram of our strategy to design SOM-ZIF-8. THF, tetrahydrofuran. (**B**) Representative SEM image of SOM-ZIF-8(3, 0.5). (**C**) SEM images of individual crystals taken from four different directions. (**D**) SEM image of the ZIF-8 grown on the external surface of the PS template. (**E** and **F**) SEM images of two selected semifinished SOM-ZIF-8(3, 0.5) crystals, confirming their oriented growth within ordered voids. Yellow areas, {100} planes of macropore arrangements; red areas, {111} planes of macropore arrangements. (**G**) Illustration of the shaping and templating effects of the 3D-ordered PS template on the morphologic evolution of SOM-ZIF-8.

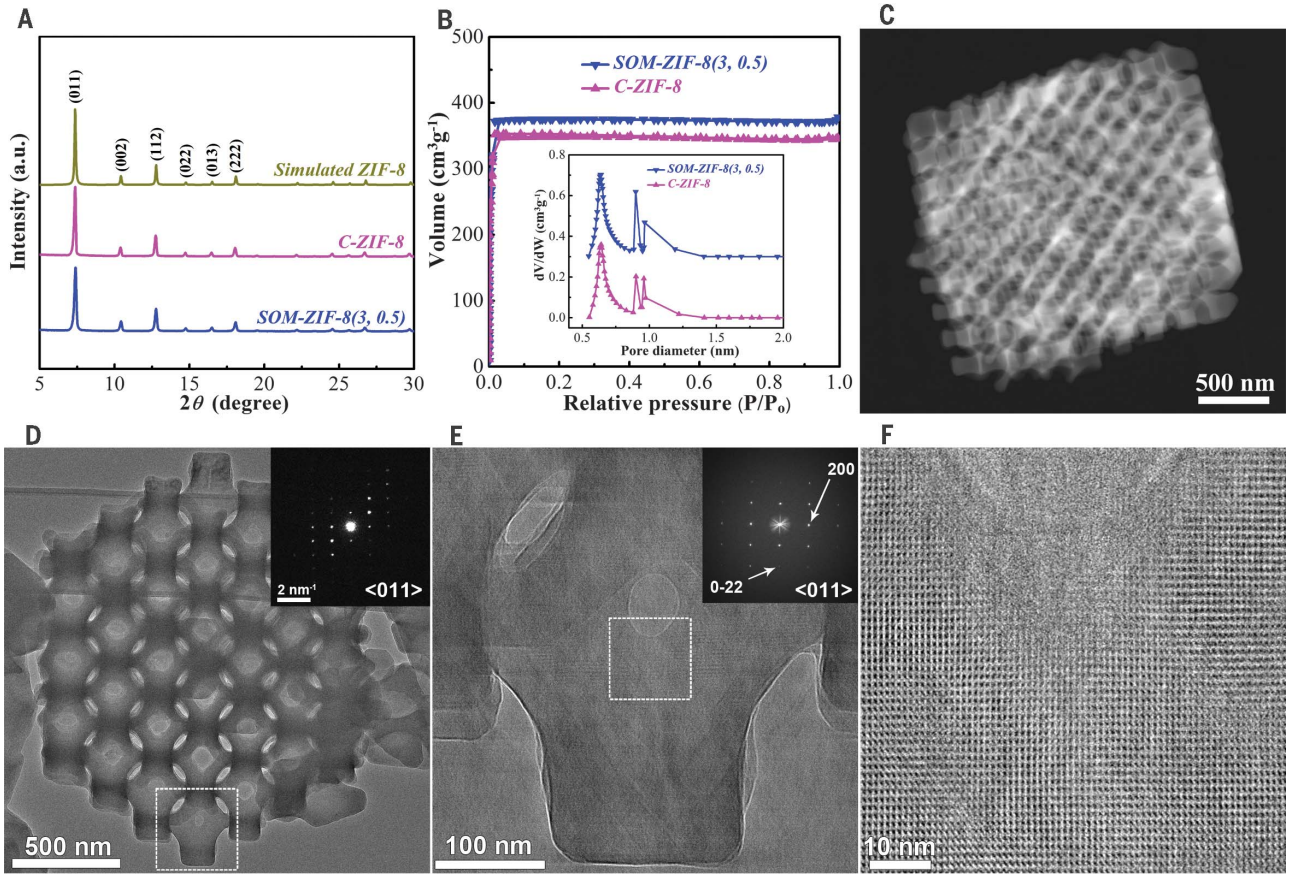


Fig. 2. Structure characterization of SOM-ZIF-8. (**A**) XRD patterns and (**B**) nitrogen adsorption-desorption isotherms of various samples. The inset of (B) shows the corresponding micropore size distribution from the Horvath-Kawazoe model, where SOM-ZIF-8(3, 0.5) has been shifted up by 0.3 $cm^3\ g^{-1}$ for clarity. a.u., arbitrary units. (**C**) HAADF-STEM image of an individual SOM-ZIF-8(3, 0.5) crystal. (**D** to **F**) TEM images of SOM-ZIF-8 at different magnifications taken along the <011> zone axis. The inset of (D) shows the corresponding SAED patterns, and the inset of (E) shows the indexed FT patterns. (E) and (F) are magnified views of the areas outlined in (D) and (E), respectively.

The prerequisite to achieve uniform SOM-ZIF-8 materials was the use of methanol/ammonia water mixtures as suitable solvents. Methanol is largely employed in the preparation of uniform ZIF-8 materials but generally does not provide high synthetic yields (fig. S11). Pure ammonia water could effectively increase the yield of ZIF-8, but with too high of a crystallization rate to prepare uniform ZIF-8 crystals (fig. S11). We successfully realized the controlled crystallization of SOM-ZIF-8(3, 0.5) within template-confined spaces by combining the advantages of the two types of solvents, as stated above. In the absence of $NH_3{\cdot}H_2O$, all ZIF-8 crystals nucleated and grew outside of the PS template, producing tiny solid ZIF-8 nanocrystals (Fig. 3A and figs. S12 to S14). Comparably, a homogeneous crystallization of the precursors occurred when using pure $NH_3{\cdot}H_2O$ as solvent, affording a monolithic SOM-ZIF-8 (figs. S12, S13, S15, and S16). Uniform SOM-ZIF-8 with single-crystalline nature could be efficiently obtained within a wide volume fraction of $NH_3{\cdot}H_2O$ in the solvent (Fig. 3, B and C, and figs. S17 and S18). The particle size grew from ~1.3 to ~4.3 μm when the volume fraction of $NH_3{\cdot}H_2O$ was increased from 0.33 to 0.67 (fig. S19), which demonstrates that crystal sizes can be flexibly tuned by simply changing the solvent's $NH_3{\cdot}H_2O$ content. Additionally, variations in the feeding molar ratio of 2-methylimidazole/$Zn(NO_3)_2$ from 2 to 4 indicated that uniform SOM-ZIF-8 with similar sizes could be prepared in a rather wide composition range (figs. S20 and S21). Lower temperatures appeared to favor the tetrakaidecahedron morphology; syntheses at a slightly higher temperature of 55°C produced near-sphere SOM-ZIF-8(3, 0.5) with larger particles (3.7 μm) (fig. S22). More importantly, a series of SOM-ZIF-8 with macropore sizes ranging from ~190 to ~470 nm could be prepared by carefully controlling the diameter of PS templates, indicating the general applicability of the proposed strategy (Fig. 3, D to K, figs. S23 to S27, and table S1).

The catalytic properties of SOM-ZIF-8 were investigated and compared to those of bulk C-ZIF-8, polycrystal hollow ZIF-8 (denoted as PH-ZIF-8, figs. S28 and S29), and macroporous ZIF-8 synthesized by disordered PSs as the template (denoted as M-ZIF-8, fig. S30) in the Knoevenagel reaction between benzaldehydes and malononitriles as a model (*27*). As expected, SOM-ZIF-8 exhibited substantially improved activities with respect to C-ZIF-8, as the conversion of benzaldehyde could be completed after 2 hours for the optimal SOM-ZIF-8(3, 0.33) versus 8 hours for C-ZIF-8 (Fig. 4A). SOM-ZIF-8(3, 0.33) also outperformed various state-of-the-art heterogeneous catalysts, including amino-functionalized molecular sieves and UiO-66-NH_2, under identical reaction conditions (figs. S31 to S37). Despite featuring similar macropore sizes, SOM-ZIF-8 exhibited much improved catalytic activities as compared with PH-ZIF-8 and M-ZIF-8. The interconnected macropores in SOM-ZIF-8 can be entered from the external crystalline surface, facilitating the accessibility to reactive sites. In contrast, macropores in PH-ZIF-8 and M-ZIF-8 were completely or partly occluded in the microporous matrix and thus were much less accessible to bulky reactants. To further support these hypotheses, the diffusion process of green fluorescent protein (GFP) from the outside to the interior of individual SOM-ZIF-8, M-ZIF-8, and C-ZIF-8 crystals was subsequently monitored in situ (fig. S38). Compared with M-ZIF-8 and C-ZIF-8, SOM-ZIF-8 clearly allowed a much faster GFP diffusion throughout the entire crystal, confirming the enhanced diffusion efficiency of

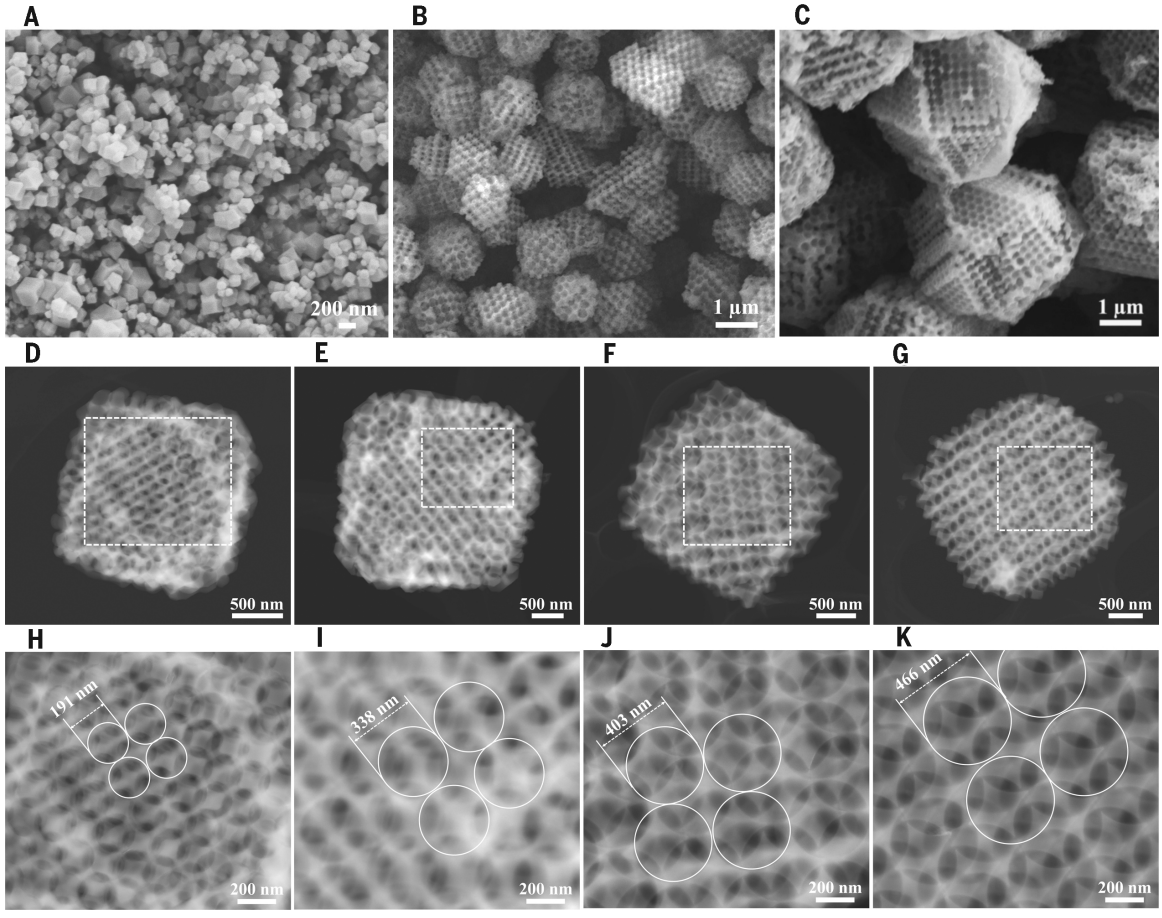


Fig. 3. Controlled growth of SOM-ZIF-8 by regulating solvents and PS sizes. (**A** to **C**) SEM images of various samples prepared with different volume fractions of $NH_3{\cdot}H_2O$: (A) 0, (B) 0.33, and (C) 0.67. (**D** to **K**) HAADF-STEM images of various SOM-ZIF-8 crystals with macropore sizes ranging from 190 to 470 nm, as determined by carefully controlling the diameter of the PS templates: (D and H) 190 nm, (E and I) 340 nm, (F and J) 400 nm, and (G and K) 470 nm. The images in (H) to (K) are magnified views of the areas outlined by the squares in (D) to (G).

3D-ordered macropores in SOM-ZIF-8 for bulky-molecule–involved applications. Furthermore, the reaction time for complete conversion of benzaldehyde was consistently 2 to 3 hours when the average particle size was decreased from 4.3 μm for SOM-ZIF-8(3, 0.67) to 1.3 μm for SOM-ZIF-8(3, 0.33).

Additionally, SOM-ZIF-8 exhibited superior structural stability and improved recyclability as compared with PH-ZIF-8 composed of intergrown polycrystals. SOM-ZIF-8(3, 0.33) could be reused at least seven times (Fig. 4B), with benzaldehyde conversion only decreasing from 94.6 to 87.0%. Comparably, a ≥5% decrease in catalytic activity was observed in every subsequent run for PH-ZIF-8. The TEM image of used PH-ZIF-8 indicates that the recycling operation may cause a severe collapse in its hollow structure, resulting in ultrasmall ZIF-8 nanocrystals that are lost in the separation step (fig. S39). In sharp contrast, SOM-ZIF-8 experienced much less structural damage, with its 3D-ordered macro-microporous structure being well preserved after seven reaction cycles. The durability of SOM-ZIF-8 can be attributed to its robust single-crystalline framework with much lower defect density in comparison with PH-ZIF-8 (fig. S40).

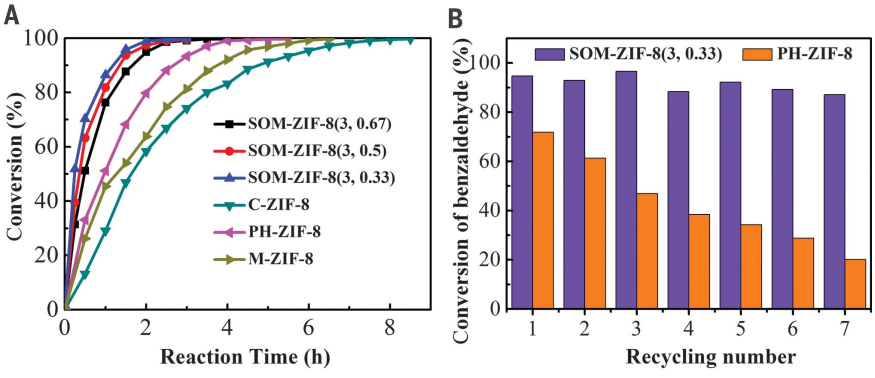


Fig. 4. Comparing the catalytic performance of SOM-ZIF-8 with C-ZIF-8, PH-ZIF-8, and M-ZIF-8. (**A**) Benzaldehyde conversions over various samples as a function of reaction time. (**B**) Recyclability tests of SOM-ZIF-8(3, 0.33) and PH-ZIF-8.

The reaction of various substituted benzaldehydes with malononitrile was also performed with SOM-ZIF-8(3, 0.33) and C-ZIF-8 (table S2). For all investigated substrates, SOM-ZIF-8(3, 0.33) always showed much higher catalytic activities than C-ZIF-8 under identical reaction conditions. The activity enhancement was even notable for benzaldehydes with bulky groups (table S2, entries

11 to 16). In view of the improved catalytic efficiencies in the Knoevenagel reaction, SOM-ZIF-8 could be expected to show additionally enhanced catalytic activities in various bulky-molecule reactions, which cannot be promoted or achieved using purely microporous C-ZIF-8.

REFERENCES AND NOTES

1. V. Valtchev, L. Tosheva, *Chem. Rev.* **113**, 6734–6760 (2013).
2. R. Krishna, *Chem. Soc. Rev.* **41**, 3099–3118 (2012).
3. M. E. Davis, *Nature* **417**, 813–821 (2002).
4. E. J. Crossland *et al.*, *Nature* **495**, 215–219 (2013).
5. K. Na *et al.*, *Science* **333**, 328–332 (2011).
6. M. Choi *et al.*, *Nat. Mater.* **5**, 718–723 (2006).
7. P. Yang, D. Zhao, D. I. Margolese, B. F. Chmelka, G. D. Stucky, *Nature* **396**, 152–155 (1998).
8. S. Inagaki, S. Guan, T. Ohsuna, O. Terasaki, *Nature* **416**, 304–307 (2002).
9. J.-R. Li, R. J. Kuppler, H.-C. Zhou, *Chem. Soc. Rev.* **38**, 1477–1504 (2009).
10. S. Furukawa, J. Reboul, S. Diring, K. Sumida, S. Kitagawa, *Chem. Soc. Rev.* **43**, 5700–5734 (2014).
11. L.-B. Sun, J.-R. Li, J. Park, H.-C. Zhou, *J. Am. Chem. Soc.* **134**, 126–129 (2012).
12. S. Han, M. S. Lah, *Cryst. Growth Des.* **15**, 5568–5572 (2015).
13. Y. Yue *et al.*, *J. Am. Chem. Soc.* **135**, 9572–9575 (2013).
14. G. Cai, H. L. Jiang, *Angew. Chem. Int. Ed.* **56**, 563–567 (2017).
15. M. Eddaoudi *et al.*, *Science* **295**, 469–472 (2002).
16. S. Lin *et al.*, *Science* **349**, 1208–1213 (2015).
17. L. Li, X. Jiao, D. Chen, C. Li, *Cryst. Growth Des.* **16**, 2700–2707 (2016).
18. Y. N. Wu *et al.*, *Angew. Chem. Int. Ed.* **50**, 12518–12522 (2011).
19. J. Reboul *et al.*, *Nat. Mater.* **11**, 717–723 (2012).
20. K. S. Park *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **103**, 10186–10191 (2006).
21. X. C. Huang, Y. Y. Lin, J. P. Zhang, X. M. Chen, *Angew. Chem. Int. Ed.* **45**, 1557–1559 (2006).
22. Supplementary materials are available online.
23. B. T. Holland, C. F. Blanford, A. Stein, *Science* **281**, 538–540 (1998).
24. P. Yang *et al.*, *Science* **282**, 2244–2246 (1998).
25. J. Cravillon *et al.*, *Chem. Mater.* **23**, 2130–2141 (2011).
26. Z. Li, H. C. Zeng, *Chem. Mater.* **25**, 1761–1768 (2013).
27. L. F. Tietze, *Chem. Rev.* **96**, 115–136 (1996).

ACKNOWLEDGMENTS

We thank the National Natural Science Foundation of China (grants 21606087, 21436005, 51621001, 21576095, and 21606088), the Fundamental Research Funds for the Central Universities (grants 2017PY004 and 2017MS069), the Guangdong Natural Science Foundation (grants 2014A030310445, 2016A050502004, and 2017A030312005), and the Welch Foundation (grant AX-1730) for financial support. K.S. thanks D. Tong for help with the schematic design and Y. Zhao for GFP preparation. K.S., Y.L., and B.C. conceived the idea and designed the experiments. K.S. synthesized the materials and carried out most of the structural characterization. Y.H., L.L., and D.Z. performed the HRTEM images and analyzed the results. K.S., L.Z., X.C., J.C., and J.L. performed the catalytic tests and some structural characterization. K.S., Y.L., B.C., and R.L. cowrote the paper. All authors discussed and commented on the manuscript. All data are reported in the main text and supplementary materials. K.S. and Y.L. are inventors on a patent/patent application (201710819873.9) held/submitted by South China University of Technology that covers the method for preparing ordered macroporous MOF single crystals.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/359/6372/206/suppl/DC1
Materials and Methods
Figs. S1 to S40
Tables S1 and S2
MS and NMR Spectra
References (*28*–*32*)

10 July 2017; accepted 4 December 2017
10.1126/science.aao3403

INFECTION

Antagonism toward the intestinal microbiota and its effect on *Vibrio cholerae* virulence

Wenjing Zhao, Florence Caro, William Robins, John J. Mekalanos*

The bacterial type VI secretion system (T6SS) is a nanomachine that delivers toxic effector proteins into target cells, killing them. In mice, we found that the *Vibrio cholerae* T6SS attacks members of the host commensal microbiota in vivo, facilitating the pathogen's colonization of the gut. This microbial antagonistic interaction drives measurable changes in the pathogenicity of *V. cholerae* through enhanced intestinal colonization, expression of bacterial virulence genes, and activation of host innate immune genes. Because ablation of mouse commensals by this enteric pathogen correlated with more severe diarrheal symptoms, we conclude that antagonism toward the gut microbiota could improve the fitness of *V. cholerae* as a pathogen by elevating its transmission to new susceptible hosts.

Bacterial pathogens deploy numerous virulence factors that directly influence their pathobiology. However, little is known about the impact of pathogen interactions with the commensal microbiome on virulence and pathogen fitness. *Vibrio cholerae*, the causative agent of cholera (*1*, *2*), assembles a dynamic organelle called the type VI secretion system (T6SS) that delivers toxic effector proteins to eukaryotic and prokaryotic cells (*3*, *4*). The T6SS of *V. cholerae* can kill heterologous bacterial species in vitro, such as *Escherichia coli* (*5*), but it is also activated in vivo (within infected experimental animals) to kill *V. cholerae* cells that lack immunity to its toxic effectors (*6*, *7*). Hence, we asked whether the antibacterial activity of the T6SS is directed against gut commensals that share a niche with *V. cholerae* (*6*, *8*). Here, we show that *V. cholerae* T6SS is used against commensal bacterial species and drives measurable changes in the pathogenicity of *V. cholerae*.

To test whether *V. cholerae* exerts antagonistic effects on the gut microbiota, we isolated several Gram-negative commensal bacteria from the small intestines of 5-day-old suckling (infant) mice. All isolates were susceptible to T6SS killing by the *V. cholerae* 2740-80 strain, which expresses an active T6SS in vitro (*9*) (table S1). Of the commensal strains isolated, we selected two *E. coli* strains that readily colonized the mouse small intestine—the same niche that *V. cholerae* also predominantly colonizes during infection. Although each strain was capable of colonizing the majority of animals within a mono-challenged group, two strains (WZ1-1 and WZ2-1), when used as a mixture, were found to reproducibly colonize all challenged animals (fig. S1). These two strains typically achieved a level of 10^4 to 10^6 colony-forming units (CFU) per mouse small intestine after 24 to 48 hours of infection and were therefore used as a defined commensal microbiota in subsequent studies. To establish whether the *V. cholerae* T6SS could affect the gut colonization of these *E. coli* commensal strains, we inoculated groups of neonatal mice with the two *E. coli* strains, returned them to their dams for 24 hours, and then arbitrarily selected three groups of pups for challenge with 10^5 CFU of wild-type *V. cholerae* [strain C6706 (*8*)], 10^5 CFU of an isogenic T6SS-defective *vipA* mutant, or buffer control. At 12 hours after wild-type *V. cholerae* challenge, the *E. coli* intestinal load was lower by a factor of ~300 than the load in mice challenged with the T6SS mutant at 24 hours (Fig. 1A; 5 ± 2 versus 1200 ± 400 CFU per mouse, $P < 0.05$, wild-type and T6SS-mutant treatment, respectively) or buffer (3840 ± 435 CFU per mouse). Surviving *E. coli* that were recovered from wild-type *V. cholerae*–treated mice did not show a reproducible skew toward either WZ1-1 or WZ2-1, indicating that both strains were equally targeted for elimination in vivo. Because the *V. cholerae* T6SS depends on the *vipA* gene, and because *V. cholerae* C6706 is known to express T6SS genes in vivo during mouse infections (*8*), we conclude that the antibacterial activity directed against the *E. coli* commensal isolates is dependent on expression of *V. cholerae* T6SS within the mouse intestine.

Surprisingly, we found that *V. cholerae* C6706, relative to its T6SS-defective *vipA* mutant, exhibited enhanced colonization in mice precolonized with *E. coli* commensals. At 8 hours after challenge, the wild-type *V. cholerae* bacterial load exceeded that of the T6SS mutant by a factor of ~10 (Fig. 1B; 6×10^4 versus 5×10^3 CFU per mouse, $P < 0.05$). By 24 hours after infection, this increased to a factor of ~15 difference (Fig. 1B; 1.2×10^7 versus 8.6×10^5 CFU per mouse, $P < 0.001$). To determine whether T6SS was allowing colonization of *V. cholerae* strains that could not express a functional T6SS, we simultaneously introduced the wild-type strain (C6706) and the *vipA* T6SS mutant into mice precolonized with *E. coli* commensal strains. We observed no significant difference between wild-type and T6SS-mutant colonization levels in these mice (fig. S2; competition index ~1.0), indicating that the wild-type strain could stimulate the colonization of the T6SS mutant, presumably by deploying its functional T6SS to eliminate the *E. coli* commensals that were antagonizing the *vipA* mutant.

T6SS-dependent killing of *E. coli* commensals also caused changes in the transcription of pro-inflammatory factors known to be induced in response to *V. cholerae* infection (*10*). Examination of intestinal extracts derived from mice 6 hours after infection with wild-type *V. cholerae* or with the T6SS-mutant strain showed that interleukin 6 (IL-6) transcriptional levels were higher by a factor of 3.3 (10.7 ± 1.8 versus 3.2 ± 0.7, $P < 0.05$), and that levels of KC [chemokine (C-X-C motif) ligand 1, also known as CXCL1] were higher by a factor of 5.1 (10.8 ± 2.4 versus 2.2 ± 0.3, $P < 0.05$), in the mice infected with wild-type *V. cholerae* (Fig. 2, A and B). These results

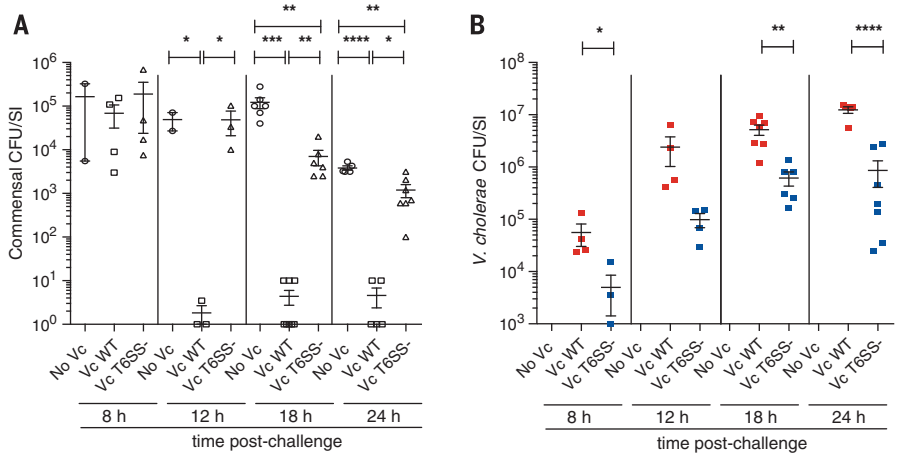


Fig. 1. T6SS facilitates *V. cholerae* colonization of the mouse gut by eliminating T6SS-sensitive commensals. (**A** and **B**) *E. coli* (A) and *V. cholerae* (B) CFU counts per small intestine (SI) homogenate collected from mice ($n = 2$ to 8) colonized first with commensal *E. coli* strains and later challenged with 10^5 CFU of wild-type *V. cholerae* C6706 (Vc WT), T6SS-defective mutant (Vc T6SS−), or buffer control (No Vc). Each data point represents one mouse; error bars represent mean ± SEM of recovered CFUs. *$P < 0.05$, **$P < 0.01$, ***$P < 0.005$, ****$P < 0.001$ (unpaired *t* test).

Department of Microbiology and Immunobiology, Harvard Medical School, Boston, MA 02115, USA.
*Corresponding author. Email: john_mekalanos@hms.harvard.edu

indicate that during the early stage of infection, the induction of innate immune factors IL-6 and KC is dependent on *V. cholerae* (C6706) possessing an active T6SS. If this early T6SS-dependent IL-6 and KC cytokine response depends on ablation of gut commensals, we reasoned that antibiotic treatment might suppress this response. We therefore eliminated much of the mouse gut microbiota by antibiotic treatment for 2 days, and then tested whether mice challenged with either wild-type *V. cholerae* (C6706) or the isogenic T6SS mutant elicited detectable differences in cytokine expression levels. In the absence of streptomycin-sensitive gut commensals, no significant differences in IL-6 or KC expression levels were observed between mice infected with wild-type *V. cholerae* and mice infected with the T6SS-mutant strain at 6 hours (Fig. 2C). Thus, T6SS-mediated killing of commensal organisms correlates with the up-regulation of IL-6 and KC during early stages of infection, and this cytokine response is suppressed by antibiotic treatment.

To gain a more comprehensive understanding of the T6SS-dependent modulation of the host immune response during early stages of infection, we harvested the small intestine of mice 6 hours after wild-type *V. cholerae* (C6706) or T6SS-mutant infection to extract host RNA for RNA sequencing (RNA-seq). We found 135 differentially expressed host genes; of the 19 genes involved in host immune function, 14 were up-regulated in the wild type–infected mice relative to the T6SS mutant–infected mice (table S2), including several known or predicted targets of the transcription factor NF-κB pathway (*cd14*, *nfkbia*, *fcgrt*, *egr1*, and *cebpd*) and regenerating islet-derived protein 3 gamma and beta (*reg3g* and *reg3b*), which were among the most highly induced genes (by factors of 16.8 and 13.1, respectively) (Fig. 2D and table S3). Thus, wild-type *V. cholerae* triggered a greater innate immune transcriptional response than did the T6SS mutant. We wondered whether the innate immune responses observed in these experiments were driven by release of microbe-associated molecular patterns (MAMPs), bacterial molecules recognized by the host innate immune system (*11*). Consistent with this hypothesis, we found that *V. cholerae* T6SS-dependent killing of *E. coli* releases MAMPs in vitro (fig. S3, A and B).

To test whether disease symptoms might be modulated by the T6SS-mediated killing of commensal species, we gavaged mice with 10^8 CFU of wild-type *V. cholerae* (C6706), its *vipA* T6SS mutant, or buffer and tested for fluid accumulation (FA ratio, indicative of diarrhea) after 22 hours. The FA ratio was highest for the wild type–challenged group (0.0875 ± 0.0025) and was significantly lower (0.0749 ± 0.0026) for the T6SS-mutant group (Fig. 3A, $P < 0.01$). We also monitored the bacterial load in the small intestine and colon of the infected mice. In both organs, the wild-type CFU exceeded that of the T6SS mutant by a factor of 3 to 9 (Fig. 3B, $P < 0.01$). Furthermore, elimination of host gut commensals by antibiotic treatment abolished the observed T6SS-dependent enhancement of the FA

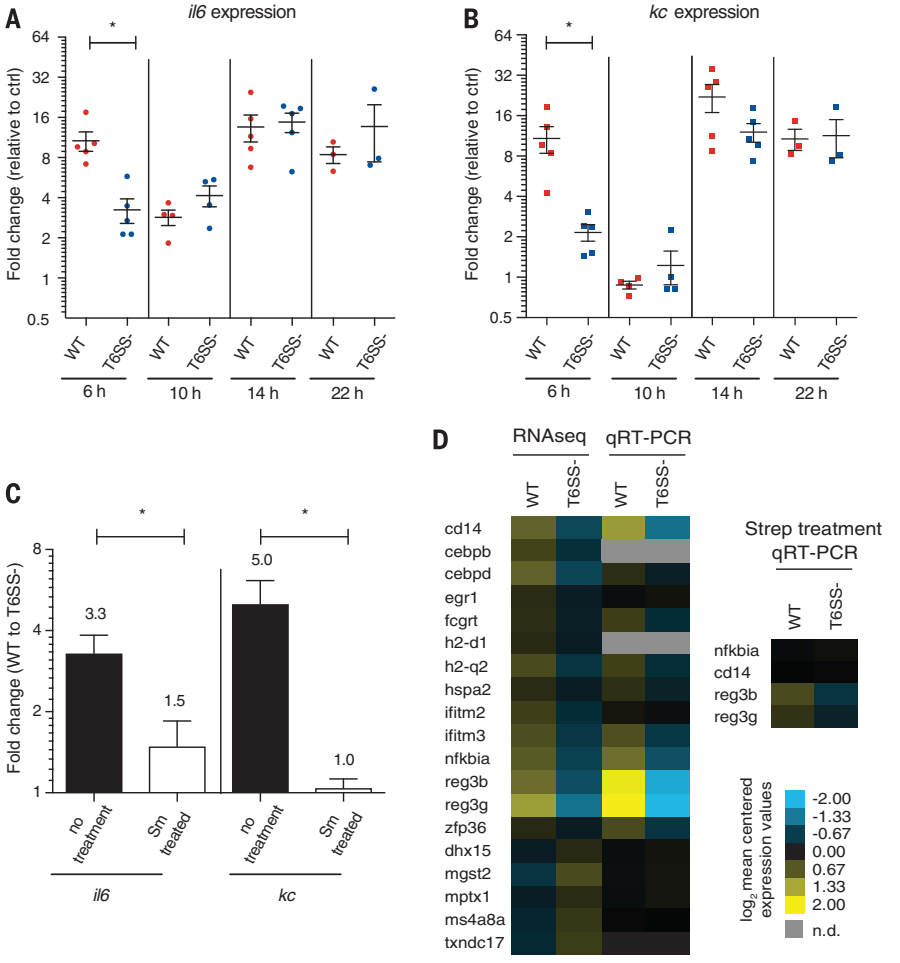


Fig. 2. Wild-type *V. cholerae* induces a more acute host innate immune response than the T6SS-defectivemutant during early stages of infection. (**A** and **B**) Transcript levels of inflammation markers IL-6 (A) and KC (B) were measured by quantitative reverse transcription polymerase chain reaction (qRT-PCR) from wild type–infected or T6SS mutant–infected mice (n = 3 to 5) and are expressed as relative (fold) change with respect to No Vc (buffer) control. (**C**) IL-6 and KC transcript levels were measured in streptomycin (Sm)–treated or untreated mice. Error bars represent mean ± SEM of fold change. *$P < 0.05$ (unpaired t test). (**D**) Heat map of host innate immune response transcript levels measured by RNA-seq and confirmed by qRT-PCR in wild type–infected or T6SS mutant–infected mice (each group n = 5).

ratio and bacterial load for animals infected with wild-type *V. cholerae* relative to the T6SS mutant (Fig. 3, C and D).

The *V. cholerae* ToxT regulatory cascade is known to enhance diarrheal responses and bacterial loads in vivo through its control of the *ctx* and *tcp* virulence operons (*12*–*15*). Thus, we asked whether T6SS-mediated killing of commensal organisms could release a commensal-derived signal that activates the ToxT regulatory cascade. We found that *tcpA* and *ctxA* transcript levels in bacteria recovered from infected mice were much higher in the wild type than in its T6SS mutant during the early stages of infection (fig. S4). This difference significantly dropped when mice where pretreated with streptomycin (Fig. 3, E and F), indicating that T6SS-dependent killing of streptomycin-sensitive commensal bacteria can activate *V. cholerae* virulence gene expression during early stages of infection. Because control experiments showed that T6SS-mediated killing of *E. coli* on laboratory media did activate *tcpA* or *ctxA* transcription in vitro (fig. S5), we conclude that the host is driving these T6SS-dependent changes in virulence gene expression.

Genome analysis has recently revealed that variant strains of the seventh pandemic El Tor *V. cholerae* clade are now responsible for the vast majority of cholera cases in South Asia, Africa, and Haiti (*12*). Because these variant strains cause more severe diarrhea in cholera victims, we questioned whether they showed higher levels of T6SS expression. Transcriptome analysis revealed that the variant strains H1 (isolated early in the Haitian cholera epidemic) and MDC126 [isolated in Bangladesh in 2008 (*12*)] showed greater T6SS gene transcript levels, by factors of 4 to 24, than two earlier seventh-pandemic strains [N16961 isolated in Bangladesh in 1971 (*13*) and C6706 isolated in Peru in 1991] (Fig. 4 and data S1). Variant strains also displayed more T6SS-dependent killing activity in vitro

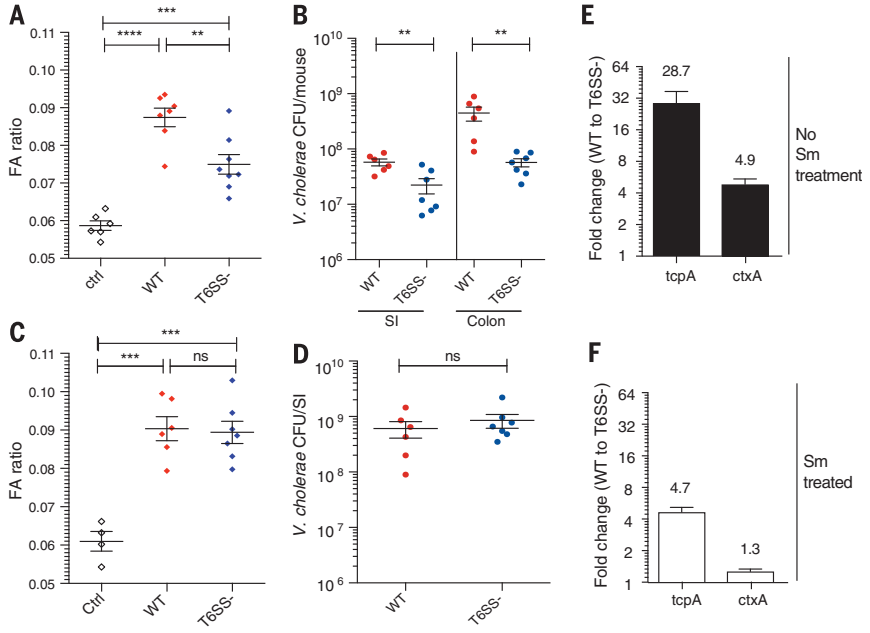


Fig. 3. *V. cholerae* T6SS enhances the development of diarrhea disease symptoms. (**A** and **B**) Fluid accumulation (FA) ratio (A) and *V. cholerae* CFU count per SI (B) of wild type–treated and T6SS mutant–treated mice. (**C** and **D**) FA ratio (C) and *V. cholerae* CFU (D) measured in mice pretreated with streptomycin. (**E** and **F**) *tcpA* and *ctxA* transcript levels were measured in mice not treated with streptomycin (No Sm) (E) or Sm-treated (F) at 6 hours after infection. $^{**}P < 0.01$, $^{***}P < 0.005$, $^{****}P < 0.001$ (unpaired *t* test); ns, not significant.

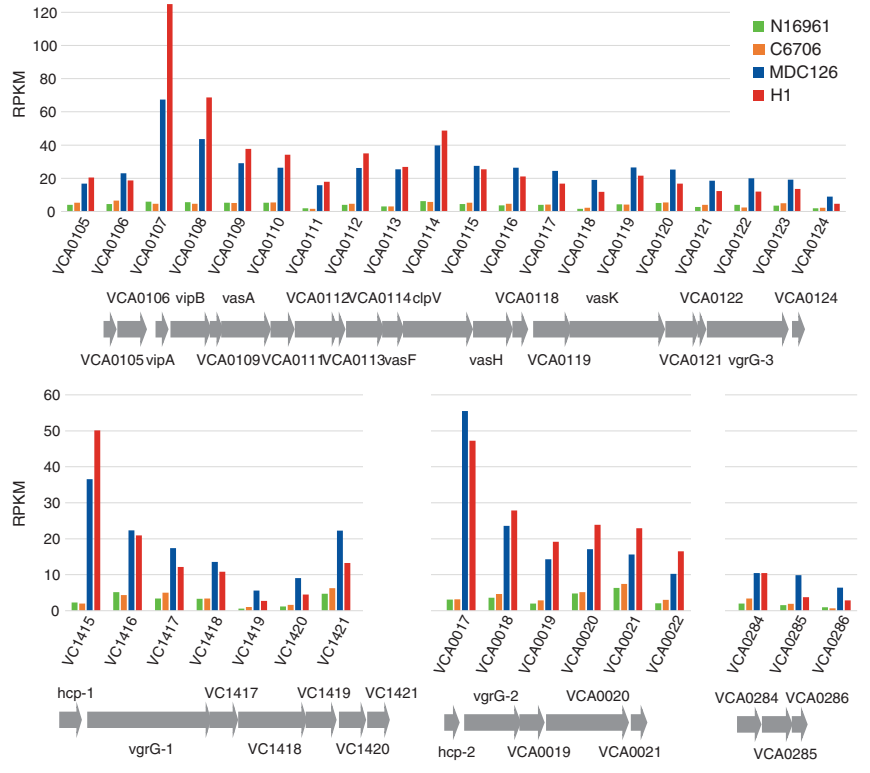


Fig. 4. Constitutive transcriptional up-regulation of T6SS genes in recent El Tor variant strains. The transcriptomes of past seventh pandemic El Tor strains N16961 (1971) and C6706 (1991) are compared to those of recent variant strains MDC126 (2008) and H1 (2010) within T6SS-related operons in both chromosomes I and II. The relative expression levels of T6SS core/accessory, hcp, vgrG, and effector/immunity genes are normalized as RPKM (reads per kilobase of transcript per million mapped reads) units.

(fig. S6). Thus, the enhanced T6SS expression in *V. cholerae* variant strains may improve their fitness through T6SS-mediated killing of commensal Gram-negative microbiota or even enteric pathogens, such as pathogenic *E. coli* that share an upper intestinal niche with *V. cholerae* and a similar epidemiological distribution (*14*).

T6SS-mediated antagonistic behavior was previously observed between enteric pathogens (such as *Salmonella typhimurium* and *Shigella sonnei*) and gut commensal organisms (*15*, *16*). Although secreted molecules such as autoinducers have been shown to activate or repress the expression of T6SS and virulence factors (*17*–*19*), it is unclear whether microbial antagonism affects interspecies communication in vivo and pathogen fitness. Our results suggest that microbial antagonism may also change the interaction of an enteric pathogen with its host through altering its expression of virulence determinants as well as driving host innate immune responses. In sum, our data support a working model (fig. S7) that provides a framework for how *V. cholerae* might use an antibacterial mechanism to clear its target niche of inhibitory competitors and simultaneously enhance disease symptom–associated transmission.

REFERENCES AND NOTES

1. A. Mutreja *et al.*, *Nature* **477**, 462–465 (2011).
2. J. B. Harris, R. C. LaRocque, F. Qadri, E. T. Ryan, S. B. Calderwood, *Lancet* **379**, 2466–2476 (2012).
3. S. Pukatzki *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **103**, 1528–1533 (2006).
4. M. Basler, M. Pilhofer, G. P. Henderson, G. J. Jensen, J. J. Mekalanos, *Nature* **483**, 182–186 (2012).
5. D. L. MacIntyre, S. T. Miyata, M. Kitaoka, S. Pukatzki, *Proc. Natl. Acad. Sci. U.S.A.* **107**, 19520–19524 (2010).
6. Y. Fu, M. K. Waldor, J. J. Mekalanos, *Cell Host Microbe* **14**, 652–663 (2013).
7. V. Bachmann *et al.*, *PLOS Negl. Trop. Dis.* **9**, e0004031 (2015).
8. A. Mandlik *et al.*, *Cell Host Microbe* **10**, 165–174 (2011).
9. M. Basler, B. T. Ho, J. J. Mekalanos, *Cell* **152**, 884–894 (2013).
10. A. L. Bishop, B. Patimalla, A. Camilli, *Infect. Immun.* **82**, 2434–2447 (2014).
11. M. A. Newman, T. Sundelin, J. T. Nielsen, G. Erbs, *Front. Plant Sci.* **4**, 139 (2013).
12. C. S. Chin *et al.*, *N. Engl. J. Med.* **364**, 33–42 (2011).
13. J. F. Heidelberg *et al.*, *Nature* **406**, 477–483 (2000).
14. A. M. Harris *et al.*, *Am. J. Trop. Med. Hyg.* **79**, 708–714 (2008).
15. T. G. Sana *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **113**, E5044–E5051 (2016).
16. M. C. Anderson, P. Vonaesch, A. Saffarian, B. S. Marteyn, P. J. Sansonetti, *Cell Host Microbe* **21**, 769–776.e3 (2017).
17. D. A. Higgins *et al.*, *Nature* **450**, 883–886 (2007).
18. A. Hsiao *et al.*, *Nature* **515**, 423–426 (2014).
19. S. Borgeaud, L. C. Metzger, T. Scrignari, M. Blokesch, *Science* **347**, 63–67 (2015).

ACKNOWLEDGMENTS

We thank M. Gack and J. Chiang for assistance with luciferase assays and for a kind gift of reporter. We thank all Mekalanos lab members for helpful comments. Supported by National Institute of Allergy and Infectious Diseases grant AI-01845 (J.J.M.). The authors declare no competing financial interests. All authors helped to design and analyze experiments; W.Z., F.C., and W.R. performed experiments; and W.Z., F.C., W.R., and J.J.M. wrote the paper. All data generated or analyzed during this study are included in this published article and its supplementary materials.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/359/6372/210/suppl/DC1
Materials and Methods
Figs. S1 to S7
Tables S1 to S5
References (*20*–*25*)
Data S1

5 September 2017; accepted 5 December 2017
10.1126/science.aap8775

NEUROSCIENCE

Spatial representations of self and other in the hippocampus

Teruko Danjo,[1] Taro Toyoizumi,[2] Shigeyoshi Fujisawa[1]*

An animal's awareness of its location in space depends on the activity of place cells in the hippocampus. How the brain encodes the spatial position of others has not yet been identified. We investigated neuronal representations of other animals' locations in the dorsal CA1 region of the hippocampus with an observational T-maze task in which one rat was required to observe another rat's trajectory to successfully retrieve a reward. Information reflecting the spatial location of both the self and the other was jointly and discretely encoded by CA1 pyramidal cells in the observer rat. A subset of CA1 pyramidal cells exhibited spatial receptive fields that were identical for the self and the other. These findings demonstrate that hippocampal spatial representations include dimensions for both self and nonself.

Spatial navigation requires the hippocampus (*1*, *2*). The cognitive map theory states that spatial recognition in the hippocampus is allocentric (*3*–*5*). Place cells in the hippocampus are the physiological correlate of this representation because they are highly sensitive to changes in landmarks and contexts (*6*–*13*). The characterization of additional types of navigational representations, including head-direction cells and grid cells, has promoted our understanding of the neural mechanisms of allocentric spatial representations in the hippocampal-entorhinal cortex network (*14*–*19*). The studies of these neural maps have focused on the animal's own position in space. It still remains unclear whether and how spatial information of nonself, such as the position of conspecifics, landmarks, and moving objects, is represented in the hippocampus.

We designed an observational T-maze task using a pair of rats (hereafter termed "self" and "other") and investigated how the other's position is represented in the self's hippocampus. The self was required to make a left/right choice to retrieve a reward based on the location of the other. We used two versions of the task, an "opposite-side rule" version in which the self rat had to choose the side opposite to the other's location (Fig. 1A, fig. S1, and movie S1) and a "same-side rule" in which the self rat had to choose the same side as the other rat (fig. S1 and movie S2). During the T-maze task, neuronal activity in the self's hippocampus (dorsal CA1) was recorded extracellularly (n = 3 pairs of rats; 88 ± 8.1% and 84 ± 11% correct performance with opposite-side and same-side rules, respectively). All analyses were performed on single units with pyramidal cell features and place fields in the task area (n = 1298 and 1205 units with opposite-side and same-side rules, respectively) (fig. S2) [see the supplementary materials (SM)].

We first examined how the location of the other rat was represented in the opposite-side rule observational task. Firing-rate maps of the self's positions ("self's rate map") and the other's positions ("other's rate map," which were obtained by replacing the self's positional data with the other's) revealed that in addition to the expected coding of the observer's own position in space, the majority of units also displayed obvious place fields for the other (Fig. 1, C to F, top). To understand the positional relationship of the rat pair at the time of firing, we constructed joint rate maps by linearizing the rats' trajectories and entering them into two-dimensional x-y axes (fig. S3). Here, the self's and other's place fields composed from identical spikes were combined into a single joint place field (Fig. 1, C to F, bottom, and fig. S4). We then examined whether these joint place fields were truly modulated by the other's positions and not a mere consequence of the constraints in the positional relationships of the two rats. The null hypothesis was that the firing only depended on the self's position and was independent of the other's. By computing the surrogate firing rates that followed this null hypothesis (see the SM) (*20*), we identified areas with significantly higher firing rates for the other (P < 0.05) (Fig. 1, C to F, third from top, and fig. S5). A significant firing dependency on the other's positions was detected for 85% of units (Fig. 1, G to I), indicating that the other's spatial information was generally encoded by the place cells in this T-maze task. Importantly, the significant areas were not dependent on the self's specific behavior or the other's positions (Fig. 1, C to I). Moreover, the firing activities of most units were not dependent on the time periods in the trial (fig. S6). Similar results were also observed when we analyzed the same-side rule version of the task (fig. S7).

In the hippocampus, spatial information is encoded not only by the average firing rates of the neurons but also by the timing of the spikes with respect to the phase of the underlying theta oscillations (7 to 11 Hz) (*21*–*24*). We next examined this temporal coding in the other's spatial representations. To control for confounds of the motion of the observer, we focused on units whose rate peaks were located in the self's starting positions in the opposite-side rule and analyzed the relationship between the other's positions and the theta phases of the spikes (n = 78 units). Although the self was not running during the time course of this analysis, the local field potential was dominated by theta oscillations (fig. S8 and SM). In contrast to the lack of theta-phase precession relative to the self's location, 75% of the units displayed significant theta-phase precession as a function of the other's location (P < 0.01) (Fig. 1, J to M). The observed theta-phase precession to the other's location was consistent with that to the self's location, as reported previously (*21*–*24*), demonstrating that the other's spatial information was also temporally encoded in the hippocampal pyramidal cells.

We next sought to dissociate two types of units—those preferentially modulated by the other's position ("other-preferred cells") and those preferentially modulated by the self's goal ["goal-preferred cells," also previously described as "prospective cells" (*25*, *26*)]. We took advantage of the two versions of our observational T-maze task (Fig. 2A). Other-preferred cells fire when the other is located on the same side of the side arms (left or right), irrespective of the rule during a given trial and, as a result, these neurons fire when the self's goal is on a different side across the two conditions (Fig. 2B). This should further support the allocentric representation of the other's place. In contrast, goal-preferred cells increase their firing rate only according to the future choice of the observer. By statistically comparing the other's rate maps between the two versions, we defined other-preferred cells and goal-preferred cells based on these criteria (see the SM). We identified 58 other-preferred cells (Fig. 2, C to G, and fig. S9) and 252 goal-preferred cells (Fig. 2, H and I). Although the number of other-preferred cells was smaller than that of goal-preferred cells, it represented 13% of the place-responsive units examined in this analysis (Fig. 2F and SM), and the firing-rate peaks of both unit groups distributed all along the side arms in the other's rate maps, constituting a full spatial map based on the position of the other rat (Fig. 2, G and I). Furthermore, the analysis of comparing the correct and the error trials in the opposite-side rule resulted in similar percentages of other-preferred and goal-preferred cells (fig. S10).

We next looked for another type of spatial representation, "a common place field" for the self and other, for which a unit fires when either the other or the self is in a specific spatial location. We found 45 units that had common place fields in both the other's and the self's rate maps under the opposite-side rule (Fig. 3, A to E, and SM). We further tested the trajectory specificity of common place fields by separating left- and right-side targeting trials. Notably, we identified 51 of 1244 units in the opposite-side rule task

[1]Laboratory for Systems Neurophysiology, RIKEN Brain Science Institute, 2-1 Hirosawa, Wako, Saitama, 351-0198, Japan. [2]Laboratory for Neural Computation and Adaptation, RIKEN Brain Science Institute, 2-1 Hirosawa, Wako, Saitama, 351-0198, Japan.
*Corresponding author. Email: fujisawa@brain.riken.jp

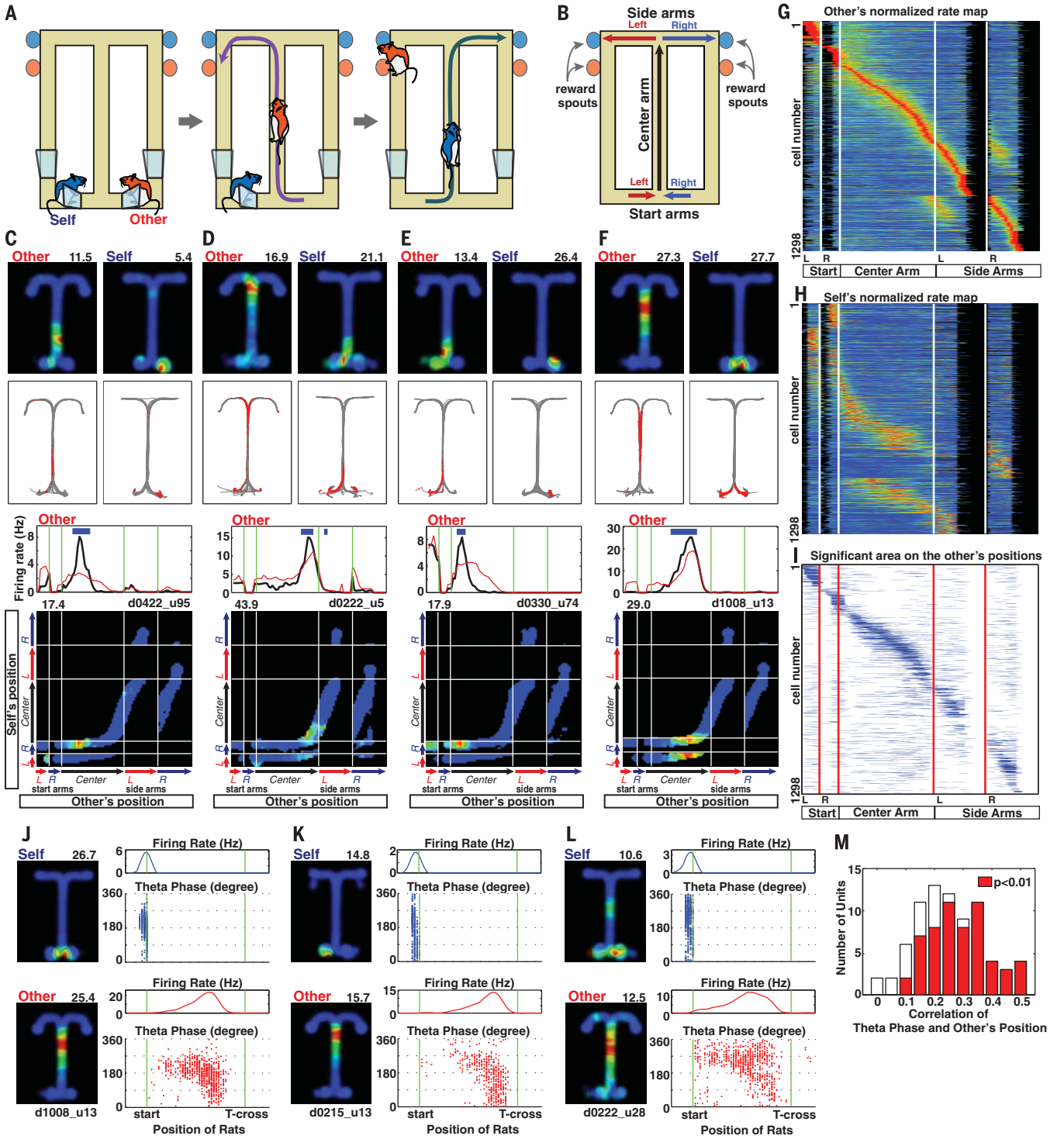


Fig. 1. Joint spatial representations of self and other in the observational task. (**A**) Schematic of the opposite-side rule. See fig. S1A. (**B**) Schematic of linearizing the trajectories on the T-maze. See figs. S3 and S4. (**C** to **F**) Units whose firing rates significantly depended on the other's position. (Top) Other's (left) and self's (right) rate maps. (Second from top) Trajectory plots of other (left) and self (right). Gray lines, rats' trajectories; red dots, spike positions. (Third from top) Linearized rate plots on other's position (black lines). Red lines indicate statistically significant bands ($P < 0.05$) (see fig. S5). Blue thick line indicates the significant area. (Bottom) Joint rate maps. Other's and self's linearized positions are coordinated in horizontal and vertical axes, respectively. The maximum firing rate is indicated at the top of each map. (**G** to **I**) Linearized and normalized rate maps based on the other's position (G), the self's position (H), and the significant area on the other's position (I), sorted by maximum rate positions on the other's map ($n = 1298$ units). (**J** to **L**) Units with theta-phase precession based on the other's position. (Top) Self's rate maps (left), linearized rate plots, and theta phases of the spikes as a function of self's position (right). (Bottom) Other's rate maps (left), linearized rate plots, and theta phases of the spikes as a function of other's position (right). (**M**) Correlation between theta phase and the other's position. Color bars indicate significant correlation ($P < 0.01$; $n = 59$ units).

that had trajectory-specific common place fields (Fig. 3, F and G; fig. S11; and SM), which was significantly higher than the number expected by chance ($P < 0.05$) (see SM).

Because the allocentric representation of self's place was also demonstrated by reliable reconstruction of self's positions only from the spikes of place-cell ensembles (*27*, *28*), we further tested whether the spikes of joint-place-cell ensembles could also reconstruct the other's positions. First, we performed Bayesian decoding analysis with a leave-one-out strategy in each rule (SM). The reconstructed other's trajectories revealed that the spikes of joint-place-cell ensembles contained sufficient information regarding the other's positions (Fig. 4A and fig. S12A). Although the error distances between the actual and decoded positions were larger for the other's decoding than those for the self, the differences in the error distances were as small as 5 cm per time window (200 ms) in both rules (other's versus self's error distances, 20.3 versus 16.1 cm in the opposite-side rule, 20.1 versus 15.1 cm in the same-side rule) (Fig. 4B). We then examined whether these spikes contained more information about the other's place than that obtained by distributions of time spent in the positional relationships of the self and the other. We computed a control estimation of the other's positions by reconstructing the self's positions and then accordingly referring to the probability distributions of the positional relationships (fig. S12B and SM). For the Bayesian reconstructions in this analysis, the prior templates were computed from trials including both rules, and results were examined based on error distances in a block-wise manner

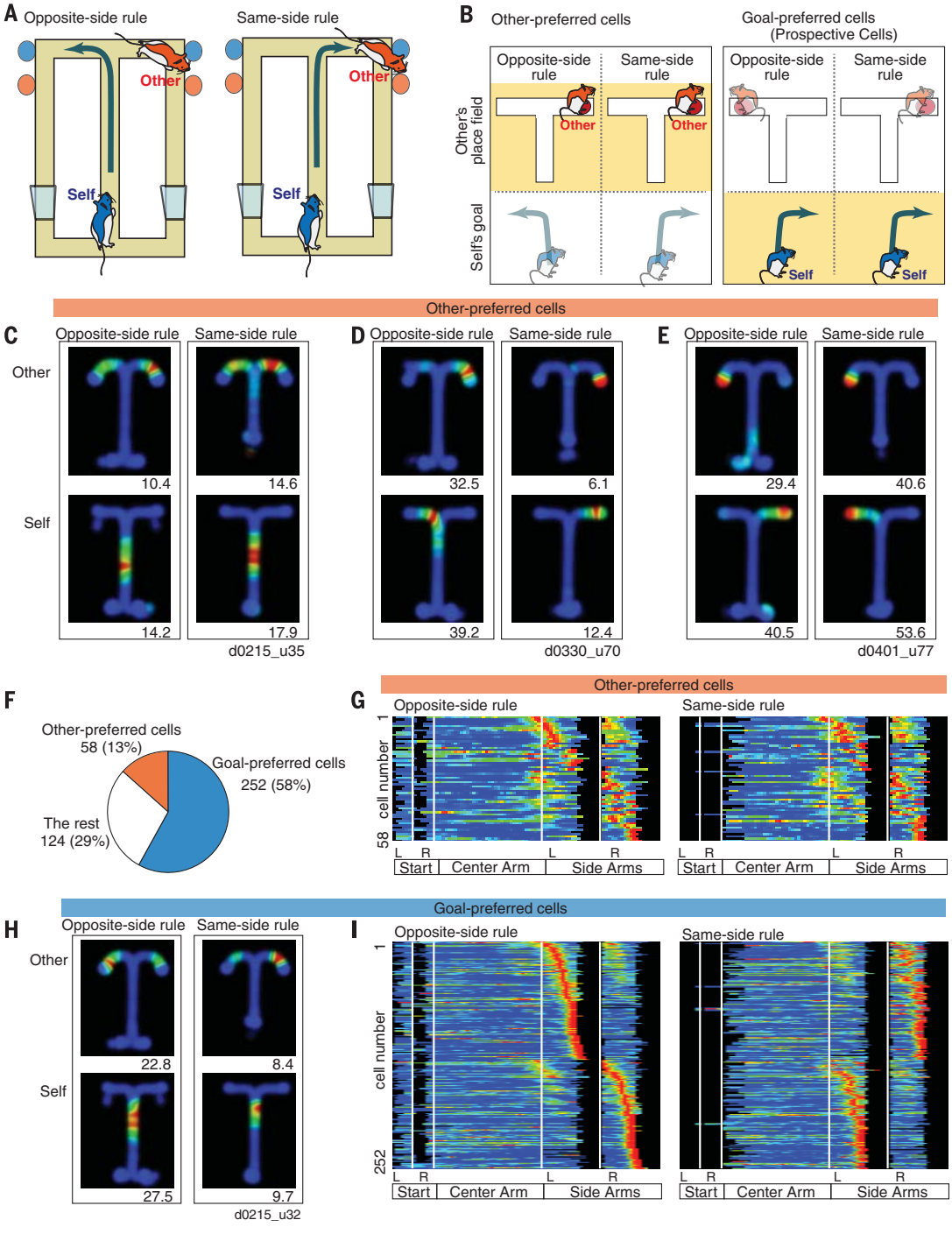


Fig. 2. Spatial representations of other-preferred and goal-preferred cells. (**A**) Schematic of the relationship of the other's position and the self's goal in opposite-side and same-side rules. (**B**) Schematic explaining other-preferred cells (left) and goal-preferred cells (right). Other-preferred cells fire preferentially when the other is on the same side, regardless of the rule. Similarly, goal-preferred cells fire when the self's goal is on the same side. (**C** to **E**) Rate maps of other-preferred cells. (Top) Other's rate maps in the opposite-side rule (left) and same-side rule (right). (Bottom) Self's rate maps in the opposite-side rule (left) and the same-side rule (right). See fig. S9 for statistical methods. (**F**) Summary of other-preferred cells and goal-preferred cells (n = total 434 units). (**G**) Linearized and normalized rate maps of other-referred cells, sorted by peak rate positions. (**H**) Other's (top) and self's (bottom) rate maps of goal-preferred cells in the opposite-side rule (left) and the same-side rule (right). (**I**) Linearized and normalized other's rate maps of goal-preferred cells in the opposite-side task (left) and same-side task (right), sorted by the maximum rate positions.

(Fig. 4, C and D, and SM). Reconstructing trajectories without rule information apparently made the error distances of the other's decoded positions larger, but they were significantly smaller than the control estimate of the other's positions (Fig. 4D). The overall averages of the error distances of the other's decoding were less than 40 cm and also were significantly smaller than those of the control estimation in both rules (Fig. 4E).

We propose an extended model of hippocampal spatial representations that can include dimensions for both self and other (fig. S13). Our model, encompassing various types of spatial representations, can categorize spatial representations into four types: own place fields, joint place fields, other's place fields, and common place fields (fig. S13, A to D). In particular, the common place field could be hypothesized to be a mirror representation of place (*29*, *30*). Combinatorial representations of spatial information of self and nonself would open the door to examining whether this allocentric spatial representation extends more generally to other nonliving moving objects (*31*–*33*) and how it is generated in the hippocampal-entorhinal cortex network. Our data indicate that the place cells in the hippocampus encode sufficient spatial information for organizing the recognition of other animals, which is essential for social behavior (*34*, *35*).

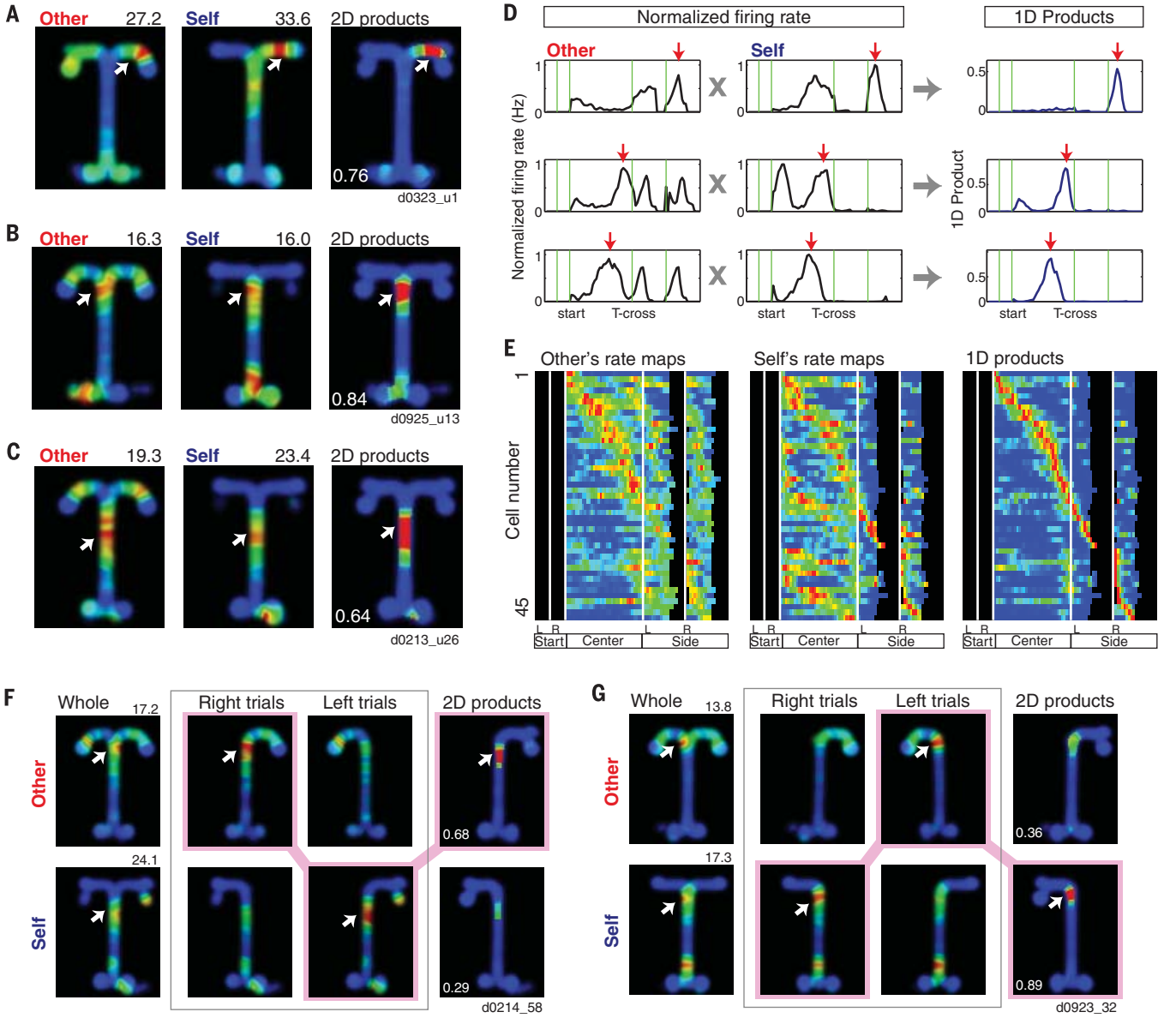


Fig. 3. Common place fields in self's and other's rate maps. (**A** to **C**) Common place fields of representative units in the other's (left) and the self's (middle) rate maps. (Right) The products (element-wise multiplications) of normalized other's and self's firing rates in each pixel. (**D**) Linearized other's (left) and self's (middle) firing rates and their products (right). The products are the element-wise multiplications of the normalized firing rates of the self and the other at each position, of the units shown in (A) to (C). This value was used to identify units with common place fields. (**E**) Linearized and normalized rate maps of units with common place fields aligned by the position with the maximum value of the products (n = 45 of 908 units). (**F** and **G**) Trajectory-specific common place fields of two representative units. (Left) Other's (top) and self's (bottom) rate maps of whole trials. (Middle) Rate maps with the left and right trials separated. (Right) The products of normalized other's and self's firing rates of right (top) and left (bottom) targeting trials in each pixel. Arrows indicate the centers of common place fields. The maximum color value of the product map was set as 0.5, and the values in the product map indicate the maximum of the product.

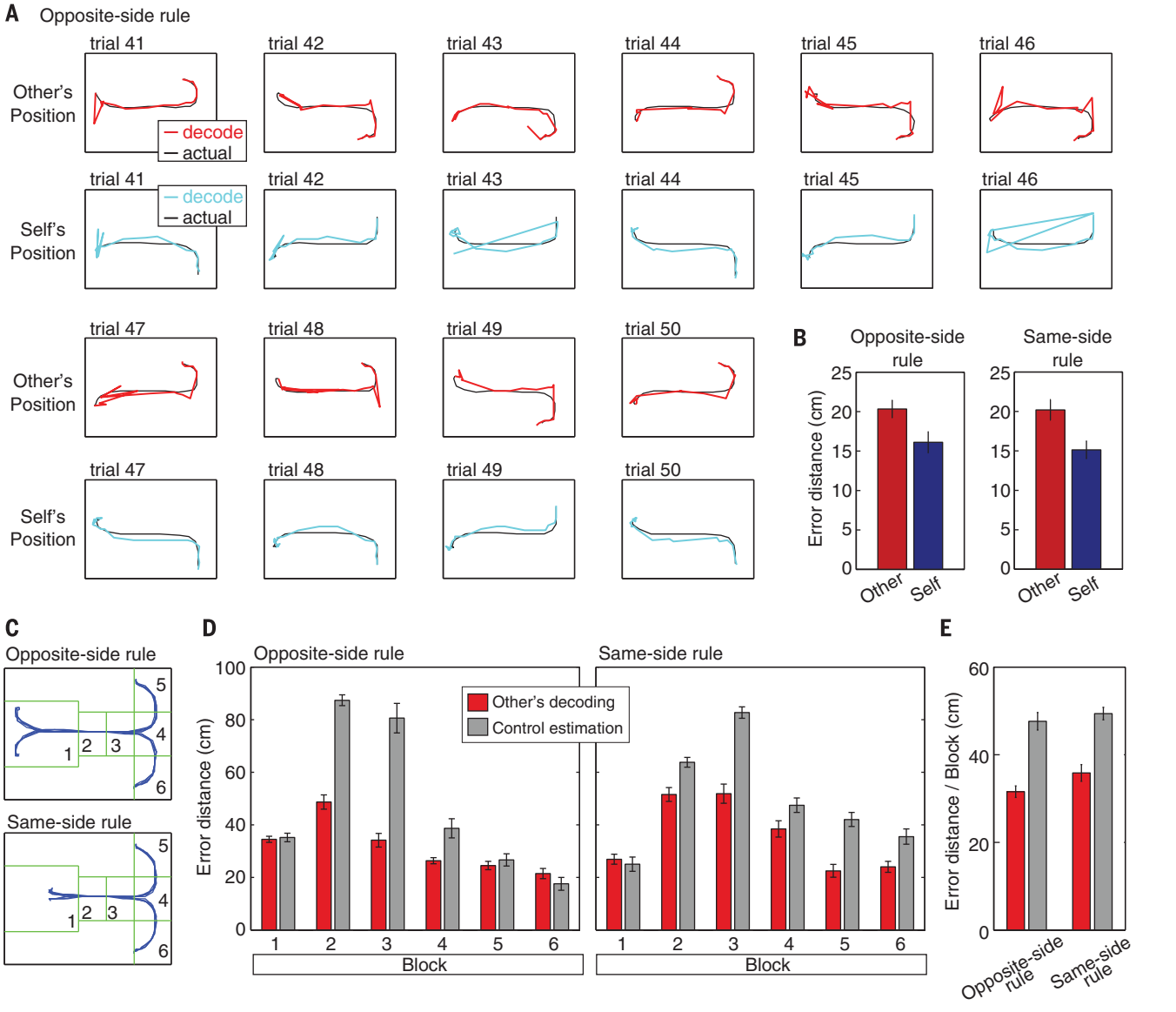


Fig. 4. Decoding other's and self's positions from spikes of hippocampal cells. (**A**) Reconstructed other's and self's positions in the opposite-side rule. Red and blue lines, reconstructed other's and self's positions, respectively. Black lines, actual trajectories of the other or self. (**B**) Mean error distances between actual and decoded positions (31 and 28 sessions from the opposite-side and same-side rules, respectively). (**C** to **E**) The reconstructions of other's positions by using prior templates calculated from trials including both rules. (C) Schematic of the division of other's trajectories into six blocks. (D) Error distances of reconstructed other's positions in each block. Red bars, distance between the actual and decoded other's positions. Gray bars, distance between the actual and estimated other's positions (n = 22 sessions) (two-way repeated measures analysis of variance: $P < 0.001$, $F_5 = 83.95$ for the opposite-side rule; $P < 0.001$, $F_5 = 72.6$ for the same-side rule). (E) Averaged error distances per block (two-tailed paired t test: $P < 0.001$, $t_{21} = -10.70$ for the opposite-side rule; $P < 0.001$, $t_{21} = -10.95$ for the same-side rule). Data are mean ± SEM.

REFERENCES AND NOTES

1. N. Burgess, E. A. Maguire, J. O'Keefe, *Neuron* **35**, 625–641 (2002).
2. A. D. Ekstrom *et al.*, *Nature* **425**, 184–188 (2003).
3. E. C. Tolman, *Psychol. Rev.* **55**, 189–208 (1948).
4. J. O'Keefe, L. Nadel, *The Hippocampus as a Cognitive Map* (Clarendon Press, Oxford, 1978).
5. G. Buzsáki, E. I. Moser, *Nat. Neurosci.* **16**, 130–138 (2013).
6. J. O'Keefe, J. Dostrovsky, *Brain Res.* **34**, 171–175 (1971).
7. E. R. Wood, P. A. Dudchenko, H. Eichenbaum, *Nature* **397**, 613–616 (1999).
8. I. Lee, D. Yoganarasimha, G. Rao, J. J. Knierim, *Nature* **430**, 456–459 (2004).
9. S. Leutgeb *et al.*, *Science* **309**, 619–623 (2005).
10. T. J. Wills, C. Lever, F. Cacucci, N. Burgess, J. O'Keefe, *Science* **308**, 873–876 (2005).
11. S. A. Ho *et al.*, *Neuroscience* **157**, 254–270 (2008).
12. S. S. Deshmukh, J. J. Knierim, *Hippocampus* **23**, 253–267 (2013).
13. X. Mou, D. Ji, *eLife* **5**, e18022 (2016).
14. J. S. Taube, R. U. Muller, J. B. Ranck Jr., *J. Neurosci.* **10**, 420–435 (1990).
15. T. Hafting, M. Fyhn, S. Molden, M. B. Moser, E. I. Moser, *Nature* **436**, 801–806 (2005).
16. D. J. Foster, M. A. Wilson, *Nature* **440**, 680–683 (2006).
17. P. Ravassard *et al.*, *Science* **340**, 1342–1346 (2013).
18. E. Kropff, J. E. Carmichael, M. B. Moser, E. I. Moser, *Nature* **523**, 419–424 (2015).
19. B. L. McNaughton, F. P. Battaglia, O. Jensen, E. I. Moser, M. B. Moser, *Nat. Rev. Neurosci.* **7**, 663–678 (2006).
20. S. Fujisawa, A. Amarasingham, M. T. Harrison, G. Buzsáki, *Nat. Neurosci.* **11**, 823–833 (2008).
21. J. O'Keefe, M. L. Recce, *Hippocampus* **3**, 317–330 (1993).
22. M. R. Mehta, A. K. Lee, M. A. Wilson, *Nature* **417**, 741–746 (2002).
23. G. Dragoi, G. Buzsáki, *Neuron* **50**, 145–157 (2006).
24. C. D. Harvey, F. Collman, D. A. Dombeck, D. W. Tank, *Nature* **461**, 941–946 (2009).

25. L. M. Frank, E. N. Brown, M. Wilson, *Neuron* **27**, 169–178 (2000).
26. E. R. Wood, P. A. Dudchenko, R. J. Robitsek, H. Eichenbaum, *Neuron* **27**, 623–633 (2000).
27. M. A. Wilson, B. L. McNaughton, *Science* **261**, 1055–1058 (1993).
28. K. Zhang, I. Ginzburg, B. L. McNaughton, T. J. Sejnowski, *J. Neurophysiol.* **79**, 1017–1044 (1998).
29. V. Gallese, L. Fadiga, L. Fogassi, G. Rizzolatti, *Brain* **119**, 593–609 (1996).
30. G. Rizzolatti, L. Craighero, *Annu. Rev. Neurosci.* **27**, 169–192 (2004).
31. S. Wirth *et al.*, *Science* **300**, 1578–1581 (2003).
32. R. Q. Quiroga, L. Reddy, G. Kreiman, C. Koch, I. Fried, *Nature* **435**, 1102–1107 (2005).
33. N. T. M. Robinson *et al.*, *Neuron* **94**, 677–688.e6 (2017).
34. F. L. Hitti, S. A. Siegelbaum, *Nature* **508**, 88–92 (2014).
35. T. Okuyama, T. Kitamura, D. S. Roy, S. Itohara, S. Tonegawa, *Science* **353**, 1536–1541 (2016).

ACKNOWLEDGMENTS

We thank A. Amarasingham for helpful suggestions on data analysis and T. J. McHugh and C. Yokoyama for comments on the manuscript. This work was supported by Japan Society for the Promotion of Science Grants-in-Aid for Scientific Research (KAKENHI) nos. 16K14561 and 16H01290 to T.D. and 15H05876 and 16H01519 to S.F. All data necessary to support the paper's conclusions are present in the paper and the supplementary materials.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/359/6372/213/suppl/DC1
Materials and Methods
Figs. S1 to S13
Movies S1 and S2
References (*36*–*41*)

14 July 2017; accepted 7 December 2017
10.1126/science.aao3898

NEUROSCIENCE

Social place-cells in the bat hippocampus

David B. Omer, Shir R. Maimon, Liora Las,*† Nachum Ulanovsky*†

Social animals have to know the spatial positions of conspecifics. However, it is unknown how the position of others is represented in the brain. We designed a spatial observational-learning task, in which an observer bat mimicked a demonstrator bat while we recorded hippocampal dorsal-CA1 neurons from the observer bat. A neuronal subpopulation represented the position of the other bat, in allocentric coordinates. About half of these "social place-cells" represented also the observer's own position—that is, were place cells. The representation of the demonstrator bat did not reflect self-movement or trajectory planning by the observer. Some neurons represented also the position of inanimate moving objects; however, their representation differed from the representation of the demonstrator bat. This suggests a role for hippocampal CA1 neurons in social-spatial cognition.

It is important for social animals to know the spatial position of conspecifics, for purposes of social interactions, observational learning, and group navigation. Decades of research on the mammalian hippocampal formation has revealed a set of spatial neurons that represent self-position and orientation, including place cells (*1*–*3*), grid cells (*4*–*6*), head-direction cells (*7*–*9*), and border/boundary cells (*10*–*12*). However, it remains unknown how the location of other animals is represented in the brain.

We designed an observational-learning task for Egyptian fruit bats (*Rousettus aegyptiacus*), which are highly social mammals that live in colonies with complex social structures (*13*). Bats were trained in pairs: In each trial, one bat ("observer") had to remain stationary on a "start ball" and to observe and remember the flight trajectory of the other bat ("demonstrator"), which was flying roughly randomly to one of two landing balls (Fig. 1A,"demonstrator flying" in trials *i* and *j*). After a delay, the observer bat had to imitate the demonstrator bat and fly to the same landing ball to receive a reward (Fig. 1A, "observer flying," and movies S1 and S2). This task had two key features: First, it required the observer to pay close attention to the demonstrator's position and to hold this position in memory during the delay period (the average delay between the demonstrator's return to the start ball and the observer's takeoff was rather long: 12.7 ± 8.6 s; mean ± SD). Second, because the observer was stationary during the demonstrator's flight, it allowed temporal dissociation between the effects of self-flights versus the flights of the other bat.

While the bats performed the task, we recorded the activity of 378 single neurons in the dorsal hippocampal area CA1 of four observer bats, using a wireless electrophysiology system (Fig. 1B) (*14*). For each neuron, we computed two firing-rate maps: a "classical" map, based on the self-movement flight trajectories of the observer—the standard depiction for place cells (Fig. 1C, "Self," left map for each neuron)—and a nonclassical map based on the spikes recorded from the observer's neuron together with the demonstrator's flight trajectories (Fig. 1C, "Demo," right maps) (*14*). We focused our analysis on the two-dimensional horizontal projections because the bats' flights were confined mostly to a narrow horizontal slab around the height of the landing balls (fig. S1). A subpopulation of hippocampal CA1 neurons encoded the position of the demonstrator-bat (Fig. 1C, cells 358, 254, 52, and 266—the right map in each example—and fig. S2). We termed these neurons "social place-cells."

We classified 68 of the 378 recorded CA1 neurons (18.0%) as significant social place-cells—significantly encoding the position of the other bat—based on spatial information (95th percentile in a shuffling analysis) (*14*). Using the same criteria, 261 of the 378 recorded neurons (69.0%) significantly encoded the self-position of the observer bat when it was flying and were thus classified as place cells (Fig. 1D). Of the 261 place cells, 14.9% were also social place-cells. Conversely, of the 68 social place-cells, 57.4% (39 neurons) were also place cells (Fig. 1, C—cells 358, 254, 52—and D), whereas the remaining 29 social place-cells (42.6%) were not place cells. Most of these neurons (16 of 29 cells; 55.2%) became completely inactive during self-flights, although they encoded the conspecific's position on interleaved demonstrator flights (examples are provided in Fig. 1C, cell 266, and fig. S2, cells 229 and 60).

This new type of social-spatial representation exhibited several features that were similar to the standard place cell representation: Both representations showed directional selectivity (Fig. 1E and fig. S3), and both place cells and social place-cells tiled space rather uniformly (Fig. 1F). However, we found also clear differences between the two representations: First, the firing rates of the social place-cells were significantly lower than for the classical place-cells (unpaired *t* test, $P < 0.01$) (Fig. 1G) [firing-rates of classical place cells were similar to our previous report from CA1 of flying bats (*15*)]. Second, in the 39 cells that encoded both self-position and conspecific position—were both place cells and social place-cells—we found a wide range of correlation values between the representations for self and other. Some neurons exhibited high similarity between their place field and social place field ("congruent cells," with positive correlations) (fig. S2, cells 68 and 45), whereas in other neurons, the place field and social place field were dissimilar ("noncongruent cells," with negative correlations) (Fig. 1C, cells 358 and 254, and fig. S2, cell 242). Overall, we found a continuum from noncongruent to congruent representations (Fig. 1H, top histograms), but we also found a slight overrepresentation of congruent cells among higher-firing neurons (Fig. 1H, bottom, gray bars, and top right histogram). These data suggest partial remapping between the hippocampal representations of self-position and conspecific-position, which can be interpreted as reflecting the contextual difference between observing a conspecific versus self-movement.

Next, we sought to rule out the possibility that social place fields might result from the observer's head movements during the demonstrator's flights. We therefore recorded head acceleration and head azimuth using a nine-axis motion-sensor that was placed on the observer's head (*14*). When the demonstrator bat was flying, the observer bat hardly moved its head: There was a lack of changes in head acceleration of the observer bat during the flights of the demonstrator bat (Fig. 2A, middle and bottom, gray areas). Consistent with this, in most of the demonstrator flights, the head azimuth of the observer changed by less than 20°, which is equivalent to a very small head movement of less than 6 mm (Fig. 2, B—black traces and rightmost *y* axis, in magenta—and C). Such small head movements did not modulate the firing of social place-cells outside the task (Fig. 2D). These bats have a wide visual field and no fovea (*13*) and hence did not need to move their head in order to track the demonstrator. However, in some of the demonstrator flights (35.4%), the observer bat did move its head more than 20° (the value of 20° corresponds roughly to ± 1 SD in azimuth) (Fig. 2, B, gray traces and C, gray vertical lines). These deviant flights might have potentially modulated the firing of the neurons. To rule out this possibility, we recomputed the social firing-rate maps after excluding the deviant flights and found that these maps were very similar to the original maps (Fig. 2, E, examples, and F, population analysis).

A second potential interpretation is that social place fields may reflect planning of the upcoming flight trajectory by the observer bat. To rule out this possibility, we conducted three analyses. (i) Trajectory planning by hippocampal cell assemblies has been linked to sharp-wave-ripples (SWRs) (*16*). We recorded the local field potential (LFP) in the observer bat, then detected SWRs (Fig. 2, G and H) and tested whether removing the observer flights that contained SWRs would affect social place fields (Fig. 2, I and J) (*14*). The removal of these flights hardly affected the social place field

Department of Neurobiology, Weizmann Institute of Science, Rehovot 76100, Israel.
*These authors contributed equally to this work.
†Corresponding author. Email: nachum.ulanovsky@weizmann.ac.il (N.U.); liora.las@weizmann.ac.il (L.L.)

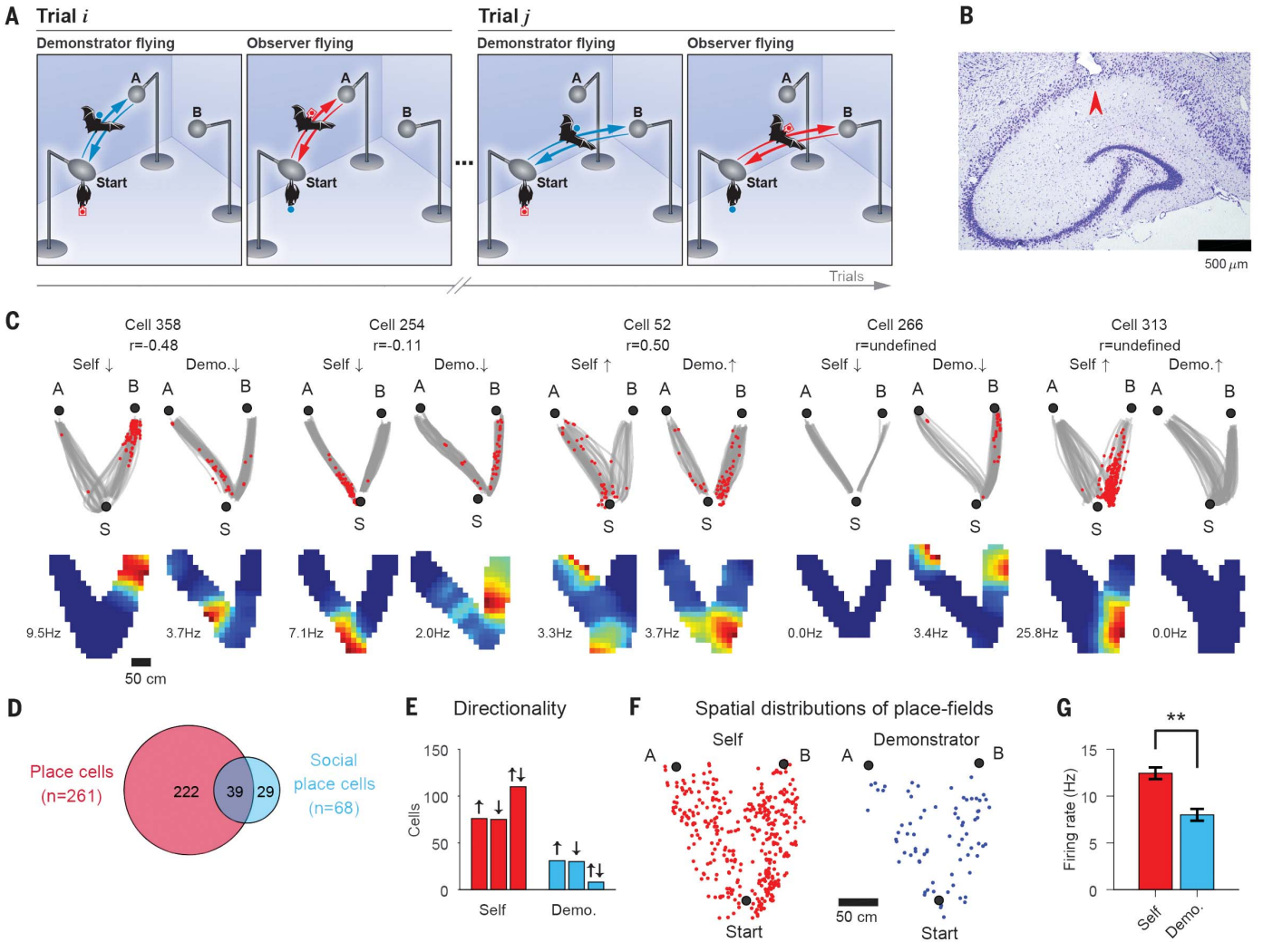


Fig. 1. Neurons in bat hippocampal area CA1 represent the position of conspecifics. (**A**) Behavioral setup inside a flight room (2.30 by 2.69 by 2.56 m). The demonstrator bat (blue) was trained to fly from the start ball, roughly randomly to either ball A or ball B, and back, and the observer bat had to imitate this flight. Two different trials are shown, one to ball A (trial *i*) and one to ball B (trial *j*). (**B**) Coronal section through dorsal hippocampus of one observer bat. Arrowhead, electrolytic lesion at end of tetrode track. (**C**) Five example cells (top view). For each cell, the left column shows the place-cell representation, based on spikes from the observer's neuron and the self-flight-trajectories of the observer (Self), and the right column shows the social place-cell representation, based on spikes from the same observer's neuron, plotted together with the demonstrator's trajectories (Demo.). (Top) Flight trajectories (gray) with spikes overlaid (red). A, B, and S are landing balls A, B, and start ball; arrows denote flight-direction (↑, flying away from start-ball; ↓, flying toward the start ball). (Bottom) Firing-rate maps. Color scale ranges from zero (blue) to maximal firing rate (red; value indicated). The four leftmost cells are social place-cells; some of these neurons are also place-cells (cells 358, 254, and 52), and some are not (cell 266). Cell 313 is a "pure" place cell. Correlations between the firing-rate maps for self and other are indicated for each cell (correlations are undefined for cells 266 and 313 because one of the maps is flat). Scale bar, 50 cm. (**D**) Total number of significant social place-cells versus significant classical place cells that we recorded. (**E**) Number of place cells and social place-cells that were significantly tuned to one flight-direction (↑), the other flight-direction (↓), or both directions (↑↓). Classical place cells are in red (n = 261), and social place-cells are in blue (n = 68). (**F**) Locations of peak firing for all the significant maps for place cells (red dots, n = 371 cells × directions), and social place-cells (blue dots, n = 76 cells × directions); cells that had significant tuning in both directions were depicted twice; hence, the counts here are larger than in (D). Dots were randomly jittered by up to ±5 cm (half bin) for display purposes. (**G**) Average peak firing rate for all the classical place cells (red, n = 371 cells × directions) and all the social place-cells (blue, n = 76 cells × directions). $^{**}P < 0.01$. (**H**) (Top) Distributions of correlation coefficients between classical place cell maps and social place-cell maps for all the neurons that encoded significantly either self-position or conspecific position and had >20 spikes per map (left histogram) or >300 spikes per map (right histogram). Gray, the data; black lines, cell-shuffling distributions (*14*). (Bottom) Map correlations increased with firing rate. Error bars, mean ± SEM; gray bars, the data; open bars, cell-shuffling; number of cells × directions included in the four bars: n = 334, 218, 137, and 91; $^{*}P < 0.05$; $^{**}P < 0.01$; n.s., nonsignificant.

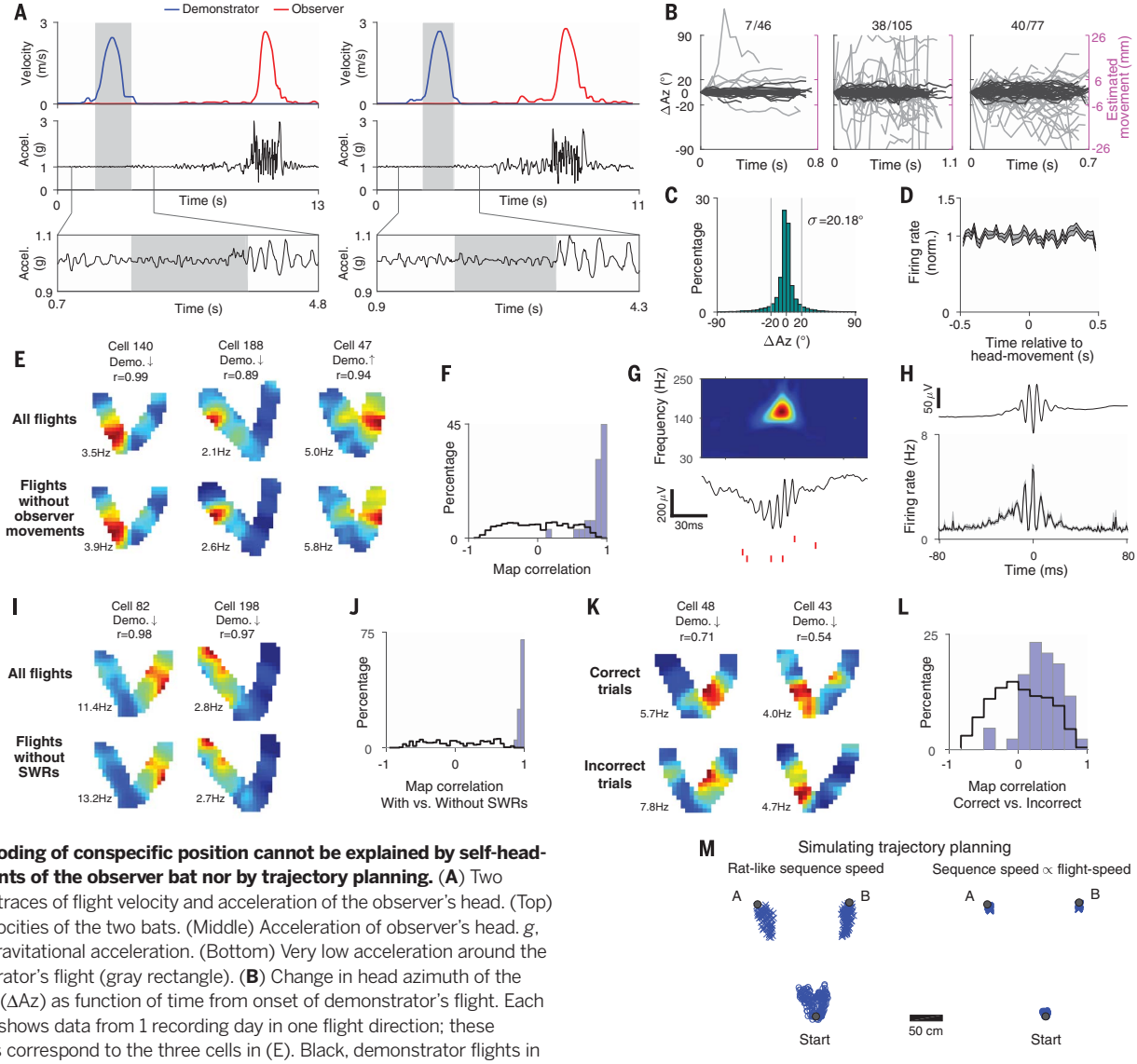


Fig. 2. Coding of conspecific position cannot be explained by self-head-movements of the observer bat nor by trajectory planning. (**A**) Two example traces of flight velocity and acceleration of the observer's head. (Top) Flight velocities of the two bats. (Middle) Acceleration of observer's head. *g*, Earth's gravitational acceleration. (Bottom) Very low acceleration around the demonstrator's flight (gray rectangle). (**B**) Change in head azimuth of the observer (ΔAz) as function of time from onset of demonstrator's flight. Each example shows data from 1 recording day in one flight direction; these examples correspond to the three cells in (E). Black, demonstrator flights in which the observer's head moved <1 SD (σ = 20.18°, which corresponds to <6 mm movement; right *y* axis) (*14*). Gray, demonstrator flights that included deviant head movements of the observer bat that exceeded ±σ. Numbers indicate proportion of deviant flights out of all the flights on this day. (**C**) Distribution of ΔAz of the observer's head, pooled over all days with significant social place-cells where motion-sensor data were recorded (n = 18 days, n = 35,284 samples). Gray lines mark 1 SD (σ = 20.18 °), which was the threshold used in (B) to define deviant flights. (**D**) Mean firing rate of social place-cells outside the task, triggered on the peak velocity of observer's head movements, for all the 1-s segments with small angular displacement <20° (n = 14,893 segments, pooled over all significant social place-cells with motion-sensor data; shaded area indicates mean ± SEM). (**E**) Three example cells, showing high correlation between social place field maps before (top) and after (bottom) removal of all the flights that included observer head-movements [At bottom, we removed all gray-colored flights in (B) and the corresponding spikes]. (**F**) Blue histogram, distribution of correlation coefficients between social place-cell maps with and without removal of flights with observer movements. Black line, cell-shuffling distribution. We included here all the significant social place-cells where motion-sensor data were recorded (n = 29 cells × directions). Shown are high correlations between maps with versus without removal of flights with observer movements (blue histogram); *t* test with unequal variances, compared with cell-shuffling control (black): $P < 10^{-26}$. (**G**) Example of a SWR. (Top) Spectrogram of the SWR. (Middle) Raw LFP trace (1 to 400 Hz bandpass). Scale bars, 30 ms and 200 μV. (Bottom) Spikes from four simultaneously recorded neurons (red ticks). Same time scale in all panels. (**H**) (Top) Mean SWR waveform, averaged across all recording days with social place-cells (n = 46 days; n = 9,092 SWRs). (Bottom) SWR-triggered firing rate, averaged over all neurons recorded during days with social place-cells (n = 276 neurons; shaded area, mean ± SEM). (**I**) Two social place-cells (columns), showing high stability with versus without flights that included SWRs (top versus bottom). (**J**) Distribution of correlation coefficients between social place-cell maps and the same maps after removal of flights with SWRs. Blue histogram, data for all cells with >20 spikes per map that had SWRs during observer flights (n = 20 cells × directions). Black line, cell-shuffling distribution. *t* test with unequal variances, data compared with cell-shuffling control: $P < 10^{-140}$. (**K**) Two social place-cells (columns), showing high stability in correct trials (top) versus incorrect trials (bottom). (**L**) Distribution of correlation coefficients between social place-cell maps computed by using correct trials versus incorrect trials. Blue histogram, data for all neurons with >20 spikes per map (n = 43 cells × directions). Black line, cell-shuffling distribution. *t* test with unequal variances, data compared with cell-shuffling control: $P < 10^{-8}$. We included in this analysis only cells with >15 correct flights and >15 incorrect flights; n = 43 cells × directions. (**M**) (Left) Simulated spatial distribution of social place fields, assuming that they are generated by place cell sequences with a ratlike sequence-speed of 8 m/s (*14*). (Right) Same, using a sequence speed of 43 m/s, which is scaled up to the flight speed of the demonstrator bat (corresponding to 20 times the bat's flight speed in our task). Blue circles and crosses denote cells with preferred direction ↑ and ↓, respectively.

maps, as indicated by very high map-correlations (Fig. 2, I, examples, and J, population), suggesting that social place fields are not created by SWR-associated trajectory planning. (ii) Next, we analyzed the neuronal activity during correct versus incorrect trials because studies in rats showed that hippocampal cell-assembly activity is strongly correlated to choice behavior on correct/incorrect trials (*17*). We reasoned that if the firing of the neurons reflects planning, then there would be a difference between social place-cell maps computed by using correct versus incorrect trials (where "incorrect" means that the demonstrator's flight was followed by an incorrect flight of the observer)—because before incorrect flights, the observer bat is likely planning to fly to the opposite landing ball from the demonstrator. However, we found high correlations between correct-trial maps and incorrect-trial maps (Fig. 2, K, examples, and L, population). (iii) Trajectory planning has been linked to hippocampal place cell sequences (*16*), and such sequences might potentially create the social place fields that we observed. However, this seems highly unlikely because place cell sequences play extremely rapidly—at a speed of 8 m/s in rats (*18*), which is ~20 times faster than the running speed of the animal (*18*)—and therefore, all the firing of the observer's neurons would be spatially compressed in one location, such as immediately after the takeoff of the demonstrator (*14*). Indeed, simulations of place-cell sequences confirmed this: All the place fields in this simulation were spatially compressed near the takeoff balls (Fig. 2M, blue crosses and circles), unlike the experimentally observed uniform distribution of social place fields (Fig. 1F, right). Together, this argues that social place fields cannot be explained via trajectory planning by the observer bat. Moreover, if trajectory planning in the observer's brain is somehow synchronized precisely to the timing and velocity of each of the demonstrator's flights—which seems rather unlikely—then it constitutes an explicit spatial representation of the position of the other bat.

Classical place cells in CA1 represent the animal's self-position in a world coordinate-frame: "allocentric coordinates" (*1*). To test whether social place-cells also form an allocentric representation, we exploited the fact that although the bats did not move their head much during the demonstrator flights (Fig. 2, A to C), the head did point in different azimuthal directions across different flights (we focused here on the azimuthal angle because the observer bats mainly moved their head in azimuth) (Fig. 3, A and B, and fig. S4) (*14*). For each of the social place-cells, we computed the median head azimuth of the observer (Fig. 3B, red line) and then used this median to divide all the demonstrator's flights into two halves, corresponding to the observer bat looking right versus looking left (Fig. 3C, top versus bottom, respectively). If social place-cells are allocentric, then we expect similar maps irrespective of the head azimuth of the observer. Indeed, maps computed during right-viewing and left-viewing were rather similar (Fig. 3, C, examples, and D, population), which is consistent with an allocentric representation. Further, there was no relation between the map correlation and the average head direction difference between looking right and looking left [correlation coefficient (r) = −0.12, P = 0.59; the head-direction differences spanned a broad range, from ΔAz = 30° to 102°] (Fig. 3E and fig. S4), which also is consistent with an allocentric representation. These neurons are thus fundamentally different from vectorial goal-direction cells in the bat hippocampus, which represent the direction to navigational goals in egocentric coordinates (*19*).

Last, we asked whether a flying conspecific is represented differently from inanimate moving objects. We conducted additional experiments in two of the four recorded bats. These experiments included three sessions (Fig. 4A). Session 1 was conducted as before (Fig. 1A). In session 2, we moved an object either to ball A or to ball B, and the observer bat had to imitate it; it was the same task as before, but with an object instead of a conspecific. We termed this object an "informative object" (Fig. 4A, session 2, and fig. S5B) (*14*). In session 3, the observer bat was trained to hang at a fixed position on the start ball and to do nothing, while we moved a different object, a "noninformative object" (in this session, the observer bat did not receive reward and hence did not fly) (Fig. 4A, session 3, and fig. S5C). Both objects were similar in size to a flying bat (fig. S5). Surprisingly, we found quite a few CA1 cells that encoded the position of inanimate moving objects (Fig. 4, B, four top examples, and C, population); to our knowledge, this is the first report that the position of moving objects is explicitly represented in the hippocampus [a previous study reported modulation of place cell firing by the movement of another object, but not an explicit spatial representation of that object (*20*)]. Some of the CA1 cells represented both the inanimate objects and the conspecific (Fig. 4B, cells 184, 169, and 361); some cells represented only the objects (Fig. 4B, cell 182); and some cells represented only the conspecific (Fig. 4B, cell 221 and C, population summary). There were some differences between the representations of the conspecific and the inanimate objects. First, there was a slight trend for a better encoding of space (higher spatial information) going from the demonstrator bat to the informative object and to the noninformative object (Wilcoxon rank-sum tests, informative object versus noninformative object, P < 0.05; demonstrator bat versus noninformative-object, P = 0.093; demonstrator bat versus informative-object, P = 0.824; Kruskal-Wallis test, P = 0.092) (Fig. 4D). Second, whereas the representation of

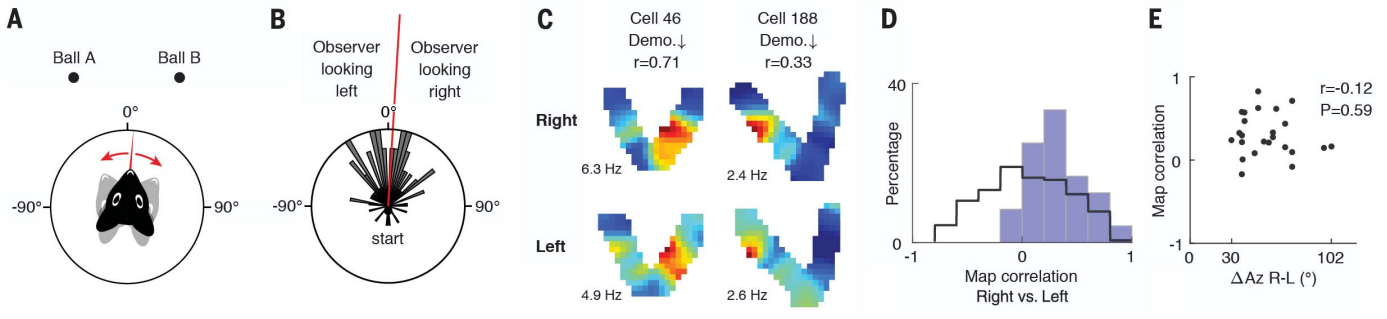


Fig. 3. The representation of conspecifics is allocentric, not egocentric. (**A** and **B**) Dividing the demonstrator's flight data based on the observer's head direction during demonstrator's flights. (A) Schematic drawing of directional notations of the bat's head relative to the two landing balls. (B) Distribution of the azimuthal head directions of the observer during demonstrator flights; data from 1 recording day. The median head direction (6.8°) is plotted in red. Direction 0° is parallel to the east-west wall of the room. (**C**) Two cells showing stability of their social place fields between right-pointing head directions (top) and left-pointing head directions (bottom). (**D**) Blue histogram, distribution of the correlation coefficients between right-looking maps and left-looking maps (blue), plotted for all the social place-cells for which we recorded motion-sensor data and had >20 spikes per map (n = 24 cells × directions); t test with unequal variances, compared with cell-shuffling control (black): $P < 10^{-4}$. Black line, cell-shuffling distribution, consisting of correlations between left-looking maps from cell i and right-looking maps from cell j across all the cell pairs where $i \neq j$. (**E**) Scatter plot of the similarity between right-looking and left-looking maps (y axis), versus the difference between the means of the right-looking and left-looking angles (x axis). No correlation was found (r = −0.12, P = 0.59; shown is a large span of azimuthal head-direction angles).

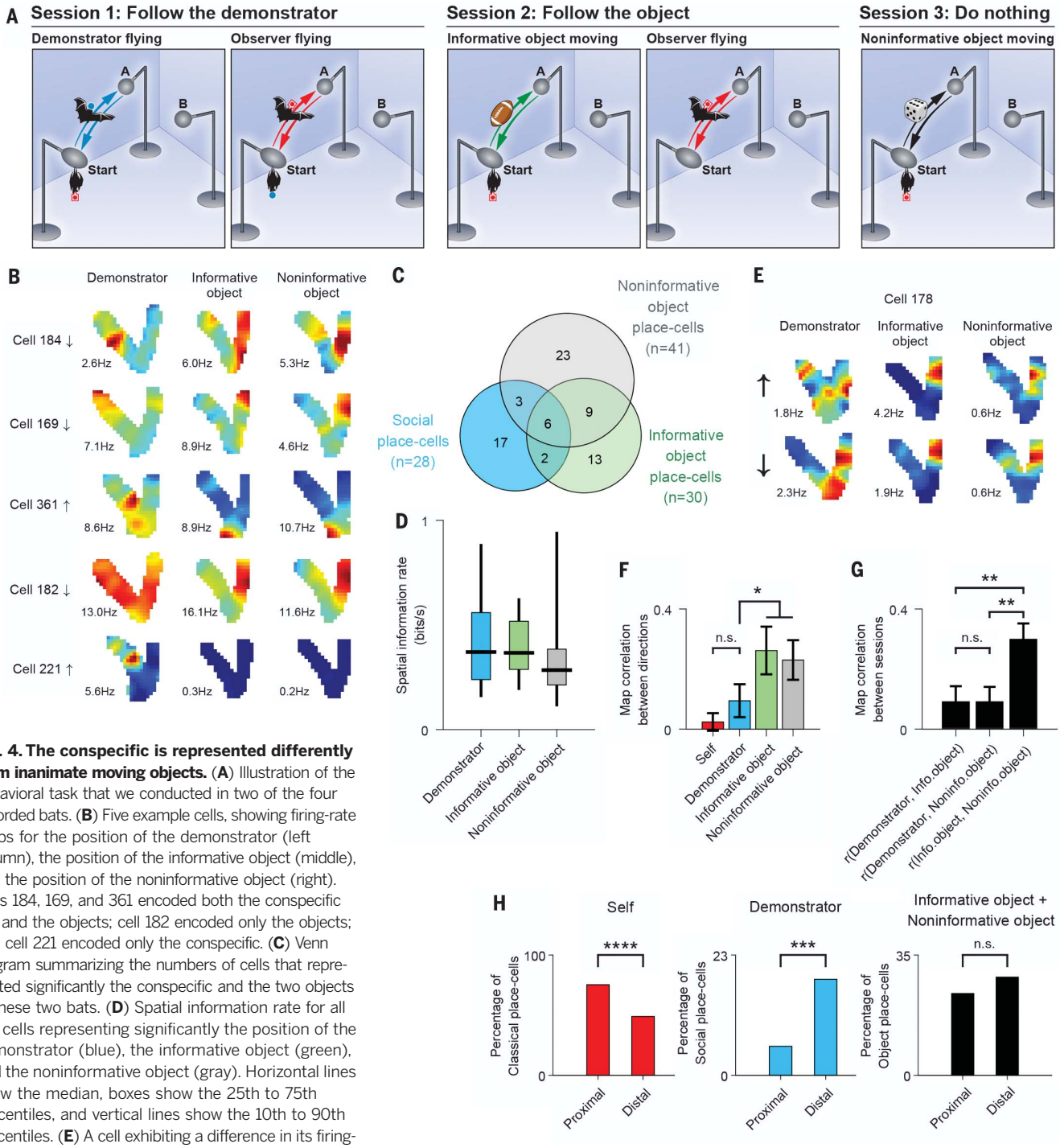


Fig. 4. The conspecific is represented differently from inanimate moving objects. (**A**) Illustration of the behavioral task that we conducted in two of the four recorded bats. (**B**) Five example cells, showing firing-rate maps for the position of the demonstrator (left column), the position of the informative object (middle), and the position of the noninformative object (right). Cells 184, 169, and 361 encoded both the conspecific bat and the objects; cell 182 encoded only the objects; and cell 221 encoded only the conspecific. (**C**) Venn diagram summarizing the numbers of cells that represented significantly the conspecific and the two objects in these two bats. (**D**) Spatial information rate for all the cells representing significantly the position of the demonstrator (blue), the informative object (green), and the noninformative object (gray). Horizontal lines show the median, boxes show the 25th to 75th percentiles, and vertical lines show the 10th to 90th percentiles. (**E**) A cell exhibiting a difference in its firing-rate maps between different flight directions of the demonstrator bat (left column), but showing no directionality for the two objects (middle and right columns); compare the top and bottom maps for the two objects (direction ↑ vs ↓). (**F**) Directionality: population summary. Shown are correlations of firing-rate maps between the two flight directions: for the self-representation, the demonstrator bat, and the informative and noninformative objects (data for all cells in which at least one flight direction exhibited a significant map, and both maps contained >50 spikes per map). The maps are much more directional (lower correlations) for the demonstrator than for the two objects; *t* test for the correlations between the two directions for demonstrator-bat versus the two pooled objects: $*P < 0.05$. (**G**) Correlations of firing-rate maps for demonstrator bat versus informative object (left), demonstrator bat versus noninformative object (middle), and informative object versus noninformative object (right). Correlations here were computed for all cells in which at least one of the two maps was significant, and only for maps with >50 spikes; *t* test of the object-object similarity versus the conspecific-object similarities: $**P < 0.01$ for both comparisons. To increase the robustness of comparisons between demonstrator and objects, (C), (D), (F), and (G) included only cells that met a strict criterion of >25 flights per map and >50 spikes per map. (**H**) Functional anatomy along the proximodistal axis of CA1, for one of the two bats tested with three sessions (*14*). Shown is the percentage of significant tuning, separately for proximal and distal tetrodes. (Left) Place cells (Self). (Middle) Social place-cells (Demonstrator). (Right) Object place cells (pooled over both objects). $***P < 10^{-3}$; $****P < 10^{-5}$.

the conspecific was directional—akin to the directionality exhibited by self place fields—the representation of both objects was nondirectional, that is, rather similar in both directions (Fig. 4, E, example, and F, population). Third, the representation of the demonstrator bat was significantly less similar to any of the object representations, as compared with the similarity between the two object representations (Fig. 4, B, cells 184,169, and 361; E, examples of similar firing for both objects; and G, population) (controls for spatial-coverage, velocity, and firing-rate are provided in fig. S6). Fourth, we found a significant difference in the functional-anatomical gradient between social place-cells and object place cells, along the proximodistal axis of CA1. Social place-cells were significantly more prevalent closer to the distal border of CA1 (log odds-ratio test: $P < 10^{-3}$) (Fig. 4H, middle, and fig. S7) (*14*), unlike object place-cells, which did not exhibit a significant proximodistal gradient (log odds-ratio test: $P = 0.17$) (Fig. 4H, right). Social place-cells exhibited the opposite pattern from classical place cells, which—consistent with previous reports in rats (*21*)—were significantly more prevalent near the proximal border of CA1 (log odds-ratio test: $P < 10^{-5}$) (Fig. 4H, left). However, this result (Fig. 4H) was obtained from a single animal (out of the two bats tested in all three sessions), in which we had a sufficient number of neurons and good proximodistal span of tetrodes (*14*); future studies will need to examine this in more detail. Together, these results suggest that the representation of the conspecific is rather different from the representation of inanimate objects, indicating that the spatial coding of the conspecific is not a simple sensory response driven by any sensory stimulus that moves through the social place field. Rather, these are context-dependent cognitive representations.

We found in this work a subpopulation of cells in bat dorsal CA1 that encode the position of conspecifics, in allocentric coordinates. This representation could not be explained by self-head-movements or by self-trajectory-planning. The responses to the conspecific were directional, which is in line with the directionality of classical place cells, but can also be interpreted through the social difference between an approaching and receding conspecific. Social place fields are unlikely to reflect either distance-coding (observer-demonstrator distance) or time-coding (time since demonstrator-takeoff) because in both cases, we would expect rather symmetric firing fields on flights to both ball A and ball B, whereas nearly all the social place-cells had a firing field on one side only. However, an alternative interpretation is that these neurons encode a position-by-time signal: Namely, they encode the spatial side to which the demonstrator bat is flying, together with its time from takeoff. We also found qualitative differences between the spatial representations of conspecifics versus inanimate moving objects. The different encoding of conspecifics versus objects may arise from (partially) different mechanisms. For example, spatial representation of moving objects in CA1 might arise from convergence of spatial inputs from grid cells in the medial entorhinal cortex (*4*, *6*) and object-related inputs from neurons in the lateral entorhinal cortex (*22*, *23*); by contrast, social place-cells may also involve socially modulated inputs from CA2 (*24*, *25*). Future studies are thus needed in order to search for social place-cells in the bat CA2, medial, and lateral entorhinal cortices, as well as in the ventral CA1, which was recently shown to be important for social memory (*26*).

It may seem surprising that social place-cells were not discovered previously in several studies of rat hippocampus that looked for a modulation of classical place fields by the presence of conspecifics (*27*–*29*). We believe that the key difference is in the task: In those previous studies, there was no incentive for the animal to pay attention to the position of the conspecific; our task, in contrast, required the bat to pay close attention to the position of the other bat and to hold this position in memory during a 12.7-s average delay, which revealed a spatial representation for the other. This interpretation is consistent with many studies that showed that hippocampal representations are highly task-dependent, plastic, and memory-dependent (*30*–*32*). Additionally, this task created a high level of social interactions between the two bats: When the bats were together at the start ball, they often approached and touched each other and emitted many social vocalizations (fig. S8), and this intensely social situation may have contributed to the representation of the conspecific.

There is an apparent similarity between the social place-cells, which encode the position of the other, and "mirror neurons" in monkeys, which encode the actions of the other (*33*). One difference, however, is that noncongruent social place-cells (Fig. 1C, cells 358 and 254) are still useful functionally because they encode meaningful information about the position of the other, whereas it is less clear how noncongruent mirror neurons in monkeys might be useful for the proposed functions of mirror neurons. Thus, social place-cells are conceptually different from mirror neurons, although both might possibly share a similar functional principle, whereby the same neuronal circuit can be used for self-representation as well as for representing conspecifics.

Last, we speculate that social place-cells may play a role in a wide range of social behaviors in many species—from group navigation and coordinated hunting to observational learning, social hierarchy, and courtship—and may be relevant also for the representation of nonconspecific animals—for example, for spatial encoding of predators and prey. These results open many questions for future studies: How are multiple animals represented in the brain? Is there a different representation for socially dominant versus subordinate animals, and for males versus females? These and many other questions await investigation in order to elucidate the neural basis of social-spatial cognition.

REFERENCES AND NOTES

1. J. O'Keefe, L. Nadel, *The Hippocampus as a Cognitive Map* (Oxford Univ. Press, 1978).
2. M. A. Wilson, B. L. McNaughton, *Science* **261**, 1055–1058 (1993).
3. N. Ulanovsky, C. F. Moss, *Nat. Neurosci.* **10**, 224–233 (2007).
4. T. Hafting, M. Fyhn, S. Molden, M.-B. Moser, E. I. Moser, *Nature* **436**, 801–806 (2005).
5. C. Barry, R. Hayman, N. Burgess, K. J. Jeffery, *Nat. Neurosci.* **10**, 682–684 (2007).
6. M. M. Yartsev, M. P. Witter, N. Ulanovsky, *Nature* **479**, 103–107 (2011).
7. J. S. Taube, R. U. Muller, J. B. Ranck Jr., *J. Neurosci.* **10**, 420–435 (1990).
8. A. Peyrache, M. M. Lacroix, P. C. Petersen, G. Buzsáki, *Nat. Neurosci.* **18**, 569–575 (2015).
9. A. Finkelstein *et al.*, *Nature* **517**, 159–164 (2015).
10. T. Solstad, C. N. Boccara, E. Kropff, M.-B. Moser, E. I. Moser, *Science* **322**, 1865–1868 (2008).
11. F. Savelli, D. Yoganarasimha, J. J. Knierim, *Hippocampus* **18**, 1270–1282 (2008).
12. C. Lever, S. Burton, A. Jeewajee, J. O'Keefe, N. Burgess, *J. Neurosci.* **29**, 9771–9777 (2009).
13. G. Neuweiler, *The Biology of Bats* (Oxford Univ. Press, 2000).
14. Materials and methods are available as supplementary materials.
15. M. M. Yartsev, N. Ulanovsky, *Science* **340**, 367–372 (2013).
16. B. E. Pfeiffer, D. J. Foster, *Science* **349**, 180–183 (2015).
17. J. Ferbinteanu, M. L. Shapiro, *Neuron* **40**, 1227–1239 (2003).
18. T. J. Davidson, F. Kloosterman, M. A. Wilson, *Neuron* **63**, 497–507 (2009).
19. A. Sarel, A. Finkelstein, L. Las, N. Ulanovsky, *Science* **355**, 176–180 (2017).
20. S. A. Ho *et al.*, *Neuroscience* **157**, 254–270 (2008).
21. E. J. Henriksen *et al.*, *Neuron* **68**, 127–137 (2010).
22. A. Tsao, M.-B. Moser, E. I. Moser, *Curr. Biol.* **23**, 399–405 (2013).
23. J. J. Knierim, J. P. Neunuebel, S. S. Deshmukh, *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **369**, 20130369 (2013).
24. F. L. Hitti, S. A. Siegelbaum, *Nature* **508**, 88–92 (2014).
25. S. M. Dudek, G. M. Alexander, S. Farris, *Nat. Rev. Neurosci.* **17**, 89–102 (2016).
26. T. Okuyama, T. Kitamura, D. S. Roy, S. Itohara, S. Tonegawa, *Science* **353**, 1536–1541 (2016).
27. L. Zynyuk, J. Huxter, R. U. Muller, S. E. Fox, *Hippocampus* **22**, 1405–1416 (2012).
28. M. von Heimendahl, R. P. Rao, M. Brecht, *J. Neurosci.* **32**, 2129–2141 (2012).
29. X. Mou, D. Ji, *eLife* **5**, e18022 (2016).
30. E. J. Markus *et al.*, *J. Neurosci.* **15**, 7079–7094 (1995).
31. M. A. Moita, S. Rosis, Y. Zhou, J. E. LeDoux, H. T. Blair, *Neuron* **37**, 485–497 (2003).
32. H. Eichenbaum, N. J. Cohen, *Neuron* **83**, 764–770 (2014).
33. G. Rizzolatti, C. Sinigaglia, *Nat. Rev. Neurosci.* **17**, 757–765 (2016).

ACKNOWLEDGMENTS

We thank K. Haroush, S. Romani, O. Forkosh, A. Rubin, M. Geva-Sagiv, A. Finkelstein, T. Eliav, G. Ginosar, A. Sarel, and D. Blum for comments on the manuscript; S. Kaufman, O. Gobi, and S. Futerman for bat training; A. Tuval for veterinary support; C. Ra'anan and R. Eilam for histology; B. Pasmantirer and G. Ankaoua for mechanical designs; and G. Brodsky for graphics. This study was supported by research grants to N.U. from the European Research Council (ERC-CoG–NATURAL_BAT_NAV), Israel Science Foundation (ISF 1319/13), and Minerva Foundation. The data are archived on the Weizmann Institute of Science servers and will be made available on request.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/359/6372/218/suppl/DC1
Materials and Methods
Figs. S1 to S8
References (*34*–*42*)
Movies S1 and S2

11 July 2017; accepted 7 December 2017
10.1126/science.aao3474

Cite as: S. Lamichhaney *et al.*, *Science* 10.1126/science.aao4593 (2017).

Rapid hybrid speciation in Darwin's finches

Sangeet Lamichhaney,[1]* Fan Han,[1] Matthew T. Webster,[1] Leif Andersson,[1,2,3]† B. Rosemary Grant,[4] Peter R. Grant[4]

[1]Department of Medical Biochemistry and Microbiology, Uppsala University, Uppsala, Sweden. [2]Department of Animal Breeding and Genetics, Swedish University of Agricultural Sciences, Uppsala, Sweden. [3]Department of Veterinary Integrative Biosciences, Texas A&M University, College Station, TX, USA. [4]Department of Ecology and Evolutionary Biology, Princeton University, Princeton, NJ, USA.

*Present address: Department of Organismic and Evolutionary Biology and Museum of Comparative Zoology, Harvard University, Cambridge, MA, USA.

†Corresponding author. Email: leif.andersson@imbim.uu.se

Homoploid hybrid speciation in animals has been inferred frequently from patterns of variation, but few examples have withstood critical scrutiny. Here we report a directly documented example from its origin to reproductive isolation. An immigrant Darwin's finch to Daphne Major in the Galápagos archipelago initiated a new genetic lineage by breeding with a resident finch (*Geospiza fortis*). Genome sequencing of the immigrant identified it as a *G. conirostris* male that originated on Española >100 km from Daphne. From the second generation onwards the lineage bred endogamously, and despite intense inbreeding, was ecologically successful and showed transgressive segregation of bill morphology. This example shows that reproductive isolation, which typically develops over hundreds of generations, can be established in only three.

Interbreeding of two species may result in the formation of a new species, reproductively isolated from the parental species (*1*). Hybrid speciation without chromosomal doubling, that is homoploid hybrid speciation, is rare (*1–4*). Possible examples have been reported in plants (*4*), butterflies (*5*), flies (*6*), fish (*7*), mammals (*8*) and birds (*9*). However, only one of these involving *Heliconius* butterflies (*5*), and three additional examples involving *Helianthus* sunflowers (*3*, *10*), meet stringent criteria that have been proposed for recognizing that hybridization was the cause of speciation (*2*). Here we report the results of a combined ecological and genomic study of Darwin's finches that documents hybrid speciation in the wild from its inception to the development of reproductive isolation.

An immature male finch immigrated to the small Galápagos Island of Daphne Major (0.34 km^2) in 1981 (*11–13*). It resembled the medium ground finch, *Geospiza fortis*, but was 70% larger and sang a unique song. Assignment tests with microsatellite markers from finches on neighboring islands indicated it was likely a *G. fortis* x *G. scandens* hybrid originating on the adjacent large island of Santa Cruz 8km from Daphne (*11*). We followed the survival and breeding of this individual and its descendants for six generations over the next 31 years.

The immigrant (generation 0) bred with a *G. fortis* female and one of its F_1 offspring bred with another *G. fortis* female, but all other matings occurred within this lineage, endogamously, and therefore from generation 2 onwards the lineage behaved as an independent species relative to other birds on the island (Fig. 1). Generations 4-6 were derived from a single brother-sister mating in generation 3. Despite close inbreeding, members of the lineage experienced high fitness as judged by their reproductive output and high survival (*12*). At maximum (2010) eight breeding pairs and 36 individuals were present on the island, and on our most recent visit (2012) there were eight breeding pairs and 23 individuals of generations 3 to 6. From observations and experiments with ground finches (*12*, *13*), bill morphology is likely to be a key factor in the success of these birds. The ability of finches to efficiently exploit the large woody fruits of *Tribulus cistoides* in dry seasons, and particularly during droughts and limited food supply, is a function of bill size, especially bill depth (*12*). Also, finch species imprint on features of their parents early in life, and later when choosing a mate they discriminate between members of their own and other species on the basis of bill size and shape, as well as body size and song (*13*).

We combined morphological measurements and whole genome sequencing of the great majority of individuals in the new lineage to establish the genetic basis of the founder population and characteristics associated with its success. We (a) assigned the founder male to species and source population, (b) confirmed pedigree assignments from observations and sequence data (*14*), (c) quantified patterns of gene transmission between generations, (d) assessed genetic diversity, and (e) searched for genetic clues to the success of the lineage. Since members of the new lineage are conspicuously large, we refer to it as the Big Bird lineage (*12*).

A phylogenetic tree analysis showed that the founder male (5110) was not a *G. fortis* x *G. scandens* hybrid as previously hypothesized (*12*), but a *G. conirostris* (Fig. 2A). This species

(large cactus finch) occurs on Española and its satellite Gardner (Fig. 2B) and nowhere else in the Galápagos archipelago; a population on Genovesa, formerly classified as *G. conirostris*, was recently reclassified as *G. propinqua* (*15*). Immigration from Española is remarkable and unexpected because it is located more than 100 km from Daphne Major, and a large island (Santa Cruz) lies between them (Fig. 2B). Rare long-distance movements of finches in the archipelago have been detected before, but until recently it was assumed these birds were vagrants that did not stay to breed (*16*–*18*).

The founder had an inbreeding coefficient (F) of 0.19 and appeared to be a typical member of the source population of *G. conirostris* from Española (F = -0.04-0.31) in terms of average genome-wide homozygosity, and ADMIXTURE (*19*) analysis classified it as a normal *G. conirostris* (Fig. 2C). The inbreeding coefficient was negative in the F_1 generation (Fig. 2D) as a result of the interspecies hybridization (*12*, *13*). A gradual increase in homozygosity was then observed over the next five generations (Fig. 2D), as expected from the small number of breeding pairs (1-8) causing genetic drift. Genome-wide average nucleotide diversity (π) showed a similar pattern, with a decline from 0.17 in generation 1 to ~0.13 in generations 4-6; values for *G. fortis* and *G. conirostris* were 0.15 and 0.16, respectively (fig. S1). Furthermore, the extensive linkage disequilibrium across the genome is consistent with a recent hybridization event (fig. S2). The Big Bird lineage also exhibited low quantitative variation. The population, all generations combined, varied less in bill length as measured by the coefficient of variation (3.82 ± 0.42, mean ± sem, n = 42) than *G. fortis* (7.55 ± 0.69, n = 60, $P < 0.005$) and *G. conirostris* (6.35 ± 0.56, n = 64, $P < 0.01$), and varied less in bill depth than *G. conirostris* (5.02 ± 0.55 vs. 7.71 ± 0.68, $P < 0.05$). The low values probably represent low additive genetic variation because the traits are highly heritable in *Geospiza* species (*13*, *20*).

The ecological success and reproductive isolation of the Big Bird lineage were most likely due to large bill and body size, and a distinctive song (*12*). We undertook a more detailed morphological analysis of the new lineage, together with both of the parental species *G. conirostris* and *G. fortis* [(*14*), table S1]. In body size, the members of the Big Bird lineage are intermediate on average between the means of the parental species (Fig. 3A), but closer to the *G. fortis* mean (table S2) as expected from their predominantly *G. fortis* genetic composition and polygenic inheritance. However, in contrast they are more similar to *G. conirostris* in bill size (Fig. 3A, fig. S3, and tables S1 and S2), despite the minority representation of *G. conirostris* genes in generation 3 and onwards. This represents a substantial allometric shift in the Big Bird lineage, possibly caused by natural selection. Selection is plausible because shifts in the elevations of static allometries have been produced relatively easily in a few generations of artificial selection in laboratory populations of animals (*21*). The pattern of change also has the characteristics of transgressive segregation (*10*, *22*). This is the production of progeny with extreme phenotypes beyond the range of parents that is likely caused by epistasis, which has been detected in other hybridizing finch species on Daphne (*23*), and/or by combining complementary alleles at different loci from different populations in F_2 and further generations (*22*, *24*, *25*).

In order to investigate the genetic basis for the relatively large bills of the Big Birds we examined the genotypes at *HMGA2* and *ALX1,* two closely linked loci (7 Mb apart) previously shown to be associated with variation in bill morphology in Darwin's finches (*15*, *26*). At the *HMGA2* locus, the allele frequency of the *L* allele associated with large bill size was 60.8% in generations 4-6 (fig. S4 and table S5). From generation 3 onwards all Big Birds were homozygous for *ALX1 B* alleles associated with blunt bills. A closer examination revealed that two variants of the *B* allele, denoted *B1* and *B2*, were segregating among the Big Birds (fig. S5 and table S5); the two alleles differ by 9 nucleotide substitutions within the 240 kb region showing a strong association with bill shape (*14*). The *B1* allele originated from the founder male that was genotyped as *P/B1*, where *P* refers to pointed, whereas the *B2* allele originated from *G. fortis* (table S5). Interestingly, *B2/B2* homozygotes had significantly shorter bills than the other two genotypes: $F_{2,32}$ = 10.5, P = 0.0003; Tukey post hoc tests, *B2/B2* < *B1/B1* (P = 0.005) and *B2/B2* < *B1/B2* (P = 0.0002). Although these associations should be confirmed with larger sample sizes, they are consistent with the hypothesis that the *ALX1* locus in Darwin's finches involves an allelic series with different effects on bill morphology (*15*).

ALX1 and *HMGA2* have large effects on bill dimensions, which are polygenic traits that are affected by other gene variants and these may have changed in frequency due to a combination of natural selection and random drift. A trend of increasing bill size across generations ($F_{1,42}$ = 6.0, P = 0.018, *adj* r^2 = 0.10) is more indicative of selection than of drift. In 2009, the only year with sufficient samples for an analysis of mortality, 19 adult survivors to the following year had a larger mean bill size than five adults that died ($F_{1,22}$ = 8.30, P = 0.009). The most important component of bill size is bill depth, and an increase in this dimension (Fig. 3B) is noteworthy for two reasons. First, the increase was independent of body size (Fig. 3C). The genetic correlation between bill size and body size that potentially constrains independent evolution of bill size is not known. However in *G. fortis* the genetic correlation between bill depth and body mass is strongly positive (0.67 ± 0.10 sem) (*23*). Second, bill length did not change in the population (fig. S6), hence bills became not only larger but progressively blunter on average across generations (fig.

S6). A possible scenario is that transgressive segregation produced genotype combinations that have been favored by natural selection causing the shift in beak morphology. The net result was morphologically based ecological segregation from the three sympatric competitor species *G. fortis*, *G. scandens* and *G. magnirostris* (Fig. 3D).

The final stage in speciation is the development of reproductive isolation from the parent population. In Darwin's finches a premating barrier to interbreeding is established by a difference in song and morphology (*12*, *13*). The test of reproductive isolation requires sympatry with the parental population(s) or a surrogate experiment, for example with finch models and/or playback of tape-recorded song (*27*). The new population on Daphne is reproductively isolated from one of the parental populations, *G. fortis*, but whether it is reproductively isolated from the other, *G. conirostris* on Española, is unknown because experiments have not been done there. Nevertheless, it is likely that the founder population has already become reproductively isolated from *G. conirostris* as bill size has changed in relation to body size (Fig. 3A). Together these traits are used as cues in the choice of mates arising from cultural, non-genetic, imprinting (*12*, *13*). Of particular relevance, experiments on Daphne with *G. scandens* showed that altering bill size in relation to body size of finch models significantly reduced responses from males (*28*). Additionally, males of the founder population sing a different song from *G. conirostris* on Española and Gardner, probably as a result of imperfect copying of a Daphne finch by the founder after it had first learned its father's song on Española (or Gardner) (*13*). Song and morphology are cues used in mate choice, and typically result in the avoidance of interspecific mating.

"—to understand the mechanism of speciation, the focus should be on cases of incipient speciation rather than on completed ones" (*29*). We have taken advantage of witnessing a rare colonization event to directly document the fate of a population founded by a single immigrant and his *G. fortis* mate. The newly founded population of Darwin's finches is an incipient hybrid species, reproductively isolated and ecologically segregated from coexisting finch species (Fig. 3D). The key features of success of the new lineage are reproductive isolation based on learned song and morphology, transgressive segregation producing novel phenotypes, and the availability of underexploited food resources. Homoploid hybrid speciation is believed to be a generally slow process extending over hundreds of generations (*29*), but as the present example shows it can be established in only three generations. Thus in small islands or island-like settings it may be easier to achieve than is currently believed (*1*, *30*–*32*).

Homoploid hybrid speciation of the Big Bird lineage exemplifies the potential evolutionary importance of rare and chance events. Expansion of the population from two individuals to three dozen was conditioned on the founder being a male with a distinctive song (*14*), and facilitated by the chance occurrence of strong selection against large bill size in a competitor species, *G. fortis*, in 2004-05 (*12*, *26*). The selection event, in turn, was mediated by *G. magnirostris*, a species that established a breeding population in 1983. Joint occurrence of rare and extreme events such as these may be especially potent in ecology and evolution (*33*, *34*).

REFERENCES AND NOTES

1. R. J. Abbott, N. H. Barton, J. M. Good, Genomics of hybridization and its evolutionary consequences. *Mol. Ecol.* **25**, 2325–2332 (2016). doi:10.1111/mec.13685 Medline
2. M. Schumer, G. G. Rosenthal, P. Andolfatto, How common is homoploid hybrid speciation? *Evolution* **68**, 1553–1560 (2014). doi:10.1111/evo.12399 Medline
3. B. L. Gross, L. H. Rieseberg, The ecological genetics of homoploid hybrid speciation. *J. Hered.* **96**, 241–252 (2005). doi:10.1093/jhered/esi026 Medline
4. S. B. Yakimowski, L. H. Rieseberg, The role of homoploid hybridization in evolution: A century of studies synthesizing genetics and ecology. *Am. J. Bot.* **101**, 1247–1258 (2014). doi:10.3732/ajb.1400201 Medline
5. J. Mavárez, C. A. Salazar, E. Bermingham, C. Salcedo, C. D. Jiggins, M. Linares, Speciation by hybridization in *Heliconius* butterflies. *Nature* **441**, 868–871 (2006). doi:10.1038/nature04738 Medline
6. D. Schwarz, B. M. Matta, N. L. Shakir-Botteri, B. A. McPheron, Host shift to an invasive plant triggers rapid animal hybrid speciation. *Nature* **436**, 546–549 (2005). doi:10.1038/nature03800 Medline
7. J. H. Kang, M. Schartl, R. B. Walter, A. Meyer, Comprehensive phylogenetic analysis of all species of swordtails and platies (Pisces: Genus *Xiphophorus*) uncovers a hybrid origin of a swordtail fish, *Xiphophorus monticolus*, and demonstrates that the sexually selected sword originated in the ancestral lineage of the genus, but was lost again secondarily. *BMC Evol. Biol.* **13**, 25 (2013). doi:10.1186/1471-2148-13-25 Medline
8. P. A. Larsen, M. R. Marchán-Rivadeneira, R. J. Baker, Natural hybridization generates mammalian lineage with species characteristics. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 11447–11452 (2010). doi:10.1073/pnas.1000133107 Medline
9. J. S. Hermansen, S. A. Saether, T. O. Elgvin, T. Borge, E. Hjelle, G.-P. Saetre, Hybrid speciation in sparrows I: Phenotypic intermediacy, genetic admixture and barriers to gene flow. *Mol. Ecol.* **20**, 3812–3822 (2011). doi:10.1111/j.1365-294X.2011.05183.x Medline
10. L. H. Rieseberg, C. Van Fossen, A. M. Desrochers, Hybrid speciation accompanied by genomic reorganization in wild sunflowers. *Nature* **375**, 313–316 (1995). doi:10.1038/375313a0
11. P. R. Grant, B. R. Grant, The secondary contact phase of allopatric speciation in Darwin's finches. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 20141–20148 (2009). doi:10.1073/pnas.0911761106 Medline
12. B. R. Grant, P. R. Grant, *40 Years of Evolution: Darwin's Finches on Daphne Major Island* (Princeton Univ. Press, 2014).
13. P. R. Grant, B. R. Grant, *How and Why Species Multiply* (Princeton Univ. Press, 2008).
14. See supplementary materials.
15. S. Lamichhaney, J. Berglund, M. S. Almén, K. Maqbool, M. Grabherr, A. Martinez-Barrio, M. Promerová, C.-J. Rubin, C. Wang, N. Zamani, B. R. Grant, P. R. Grant, M. T. Webster, L. Andersson, Evolution of Darwin's finches and their beaks revealed by genome sequencing. *Nature* **518**, 371–375 (2015). doi:10.1038/nature14181 Medline
16. P. R. Grant, B. R. Grant, Conspecific versus heterospecific gene exchange between populations of Darwin's finches. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **365**, 1065–1076 (2010). doi:10.1098/rstb.2009.0283 Medline
17. D. L. Lack, *Darwin's Finches* (Cambridge Univ. Press, 1947).
18. H. S. Swarth, The avifauna of the Galápagos islands. *Occas. Pap. Calif. Acad. Sci* **18**, 1–299 (1931).
19. D. H. Alexander, J. Novembre, K. Lange, Fast model-based estimation of ancestry

in unrelated individuals. *Genome Res.* **19**, 1655–1664 (2009). doi:10.1101/gr.094052.109
20. L. F. Keller, P. R. Grant, B. R. Grant, K. Petren, Heritability of morphological traits in Darwin's finches: Misidentified paternity and maternal effects. *Heredity* **87**, 325–336 (2001). doi:10.1046/j.1365-2540.2001.00900.x Medline
21. G. H. Bolstad, J. A. Cassara, E. Márquez, T. F. Hansen, K. van der Linde, D. Houle, C. Pélabon, Complex constraints on allometry revealed by artificial selection on the wing of *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 13284–13289 (2015). doi:10.1073/pnas.1505357112 Medline
22. L. H. Rieseberg, M. A. Archer, R. K. Wayne, Transgressive segregation, adaptation and speciation. *Heredity* **83**, 363–372 (1999). doi:10.1038/sj.hdy.6886170 Medline
23. P. R. Grant, B. R. Grant, Phenotypic and genetic effects of hybridization in Darwin's finches. *Evolution* **48**, 297–316 (1994). doi:10.1111/j.1558-5646.1994.tb01313.x Medline
24. K. J. Parsons, Y. H. Son, R. Craig Albertson, Hybridization promotes evolvability in African cichlids: Connections between transgressive segregation and phenotypic integration. *Evol. Biol.* **38**, 306–315 (2011). doi:10.1007/s11692-011-9126-7
25. R. Stelkens, O. Seehausen, Genetic distance between species predicts novel trait expression in their hybrids. *Evolution* **63**, 884–897 (2009). doi:10.1111/j.1558-5646.2008.00599.x Medline
26. S. Lamichhaney, F. Han, J. Berglund, C. Wang, M. S. Almén, M. T. Webster, B. R. Grant, P. R. Grant, L. Andersson, A beak size locus in Darwin's finches facilitated character displacement during a drought. *Science* **352**, 470–474 (2016). doi:10.1126/science.aad8786 Medline
27. B. R. Grant, P. R. Grant, Simulating secondary contact in allopatric speciation: An empirical test of premating isolation. *Biol. J. Linn. Soc. Lond.* **76**, 545–556 (2002). doi:10.1046/j.1095-8312.2002.00076.x
28. L. M. Ratcliffe, P. R. Grant, Species recognition in Darwin's finches (*Geospiza*, Gould) I. Discrimination by morphological cues. *Anim. Behav.* **31**, 1139–1153 (1983). doi:10.1016/S0003-3472(83)80021-9
29. A. W. Nolte, D. Tautz, Understanding the onset of hybrid speciation. *Trends Genet.* **26**, 54–58 (2010). doi:10.1016/j.tig.2009.12.001 Medline
30. J. Mavárez, M. Linares, Homoploid hybrid speciation in animals. *Mol. Ecol.* **17**, 4181–4185 (2008). doi:10.1111/j.1365-294X.2008.03898.x Medline
31. Marie Curie SPECIATION Network, What do we need to know about speciation? *Trends Ecol. Evol.* **27**, 27–39 (2012). doi:10.1016/j.tree.2011.09.002 Medline
32. R. J. Abbott, M. J. Hegarty, S. J. Hiscock, A. C. Brennan, Homoploid hybrid speciation in action. *Taxon* **59**, 1375–1386 (2010).
33. R. Paine, M. Tegner, E. Johnson, Compounded perturbations yield ecological surprises. *Ecosystems (N. Y.)* **1**, 535–545 (1998). doi:10.1007/s100219900049
34. P. R. Grant, B. R. Grant, R. B. Huey, M. T. J. Johnson, A. H. Knoll, J. Schmitt, Evolution caused by extreme events. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **372**, 20160146 (2017). doi:10.1098/rstb.2016.0146 Medline
35. B. Li, H. Li, P. Parker, J. Wang, The genome of Darwin's finch (*Geospiza fortis*). *GigaScience* 10.5524/100040 (2012).
36. S. Purcell, B. Neale, K. Todd-Brown, L. Thomas, M. A. R. Ferreira, D. Bender, J. Maller, P. Sklar, P. I. W. de Bakker, M. J. Daly, P. C. Sham, PLINK: A tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81**, 559–575 (2007). doi:10.1086/519795 Medline
37. J. Huisman, Pedigree reconstruction from SNP data: parentage assignment, sibship clustering and beyond. *Mol. Ecol. Resour.* **17**, 1775–0998 (2017). doi:10.1111/1755-0998.12665
38. M. N. Price, P. S. Dehal, A. P. Arkin, FastTree 2—Approximately maximum-likelihood trees for large alignments. *PLOS ONE* **5**, e9490 (2010). doi:10.1371/journal.pone.0009490 Medline
39. P. Danecek, A. Auton, G. Abecasis, C. A. Albers, E. Banks, M. A. DePristo, R. E. Handsaker, G. Lunter, G. T. Marth, S. T. Sherry, G. McVean, R. Durbin; 1000 Genomes Project Analysis Group, The variant call format and VCFtools. *Bioinformatics* **27**, 2156–2158 (2011). doi:10.1093/bioinformatics/btr330 Medline
40. P. R. Grant, Patterns of growth in Darwin's finches. *Proc. R. Soc. London* **212**, 403–432 (1981). doi:10.1098/rspb.1981.0046

ACKNOWLEDGMENTS

NSF (USA) funded the collection of material under permits from the Galápagos and Costa Rica National Parks Services and the Charles Darwin Research Station, and in accordance with protocols of Princeton University's Animal Welfare Committee. The project was supported by the Knut and Alice Wallenberg foundation and the Swedish Research Council. Sequencing was performed by the SNP&SEQ Technology Platform, supported by Uppsala University and Hospital, SciLifeLab and Swedish Research Council. Computer resources were supplied by UPPMAX. We would also like to thank two anonymous reviewers for valuable comments on the paper. The Illumina reads have been submitted to the short reads archive (http://www.ncbi.nlm.nih.gov/sra) with the accession number PRJNA392917. Raw tree files for constructing Fig. 2A and figs. S4, S5, and S7 have been submitted to the TreeBASE database with submission ID S21803 (http://purl.org/phylo/treebase/phylows/study/TB2:S21803?format=html). PRG and BRG collected the material. PRG, BRG and LA conceived the study. LA and MTW led the bioinformatic analysis of data. SL and FH performed the bioinformatic analysis. PRG, BRG, SL, and LA wrote the paper with input from the other authors. All authors approved the manuscript before submission.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/cgi/content/full/science.aao4593/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S7
Tables S1 to S5
References (*35–40*)

23 July 2017; accepted 9 November 2017
Published online 23 November 2017
10.1126/science.aao4593

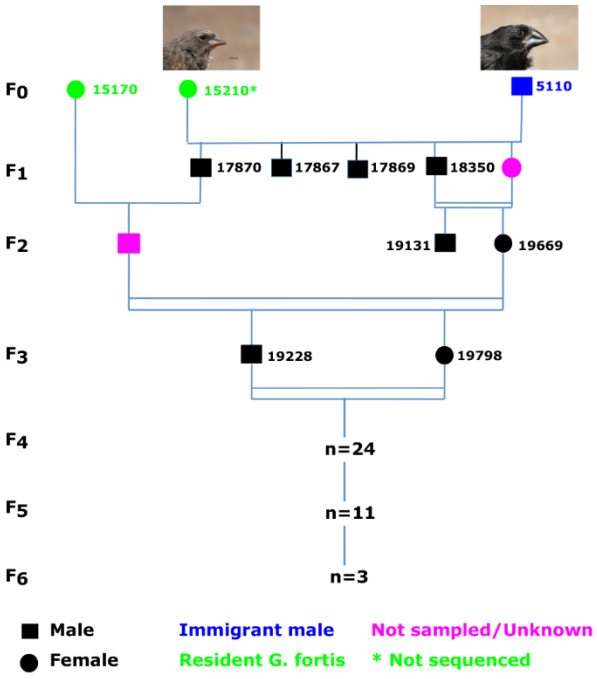


Fig. 1. The Big Bird lineage until the sixth generation. Interbreeding with two *G. fortis* females resulted in a reduction of the genetic contribution of the immigrant male from 0.50 in the first generation to 0.375 in the second and subsequent generations. Images © Peter & Rosemary Grant.

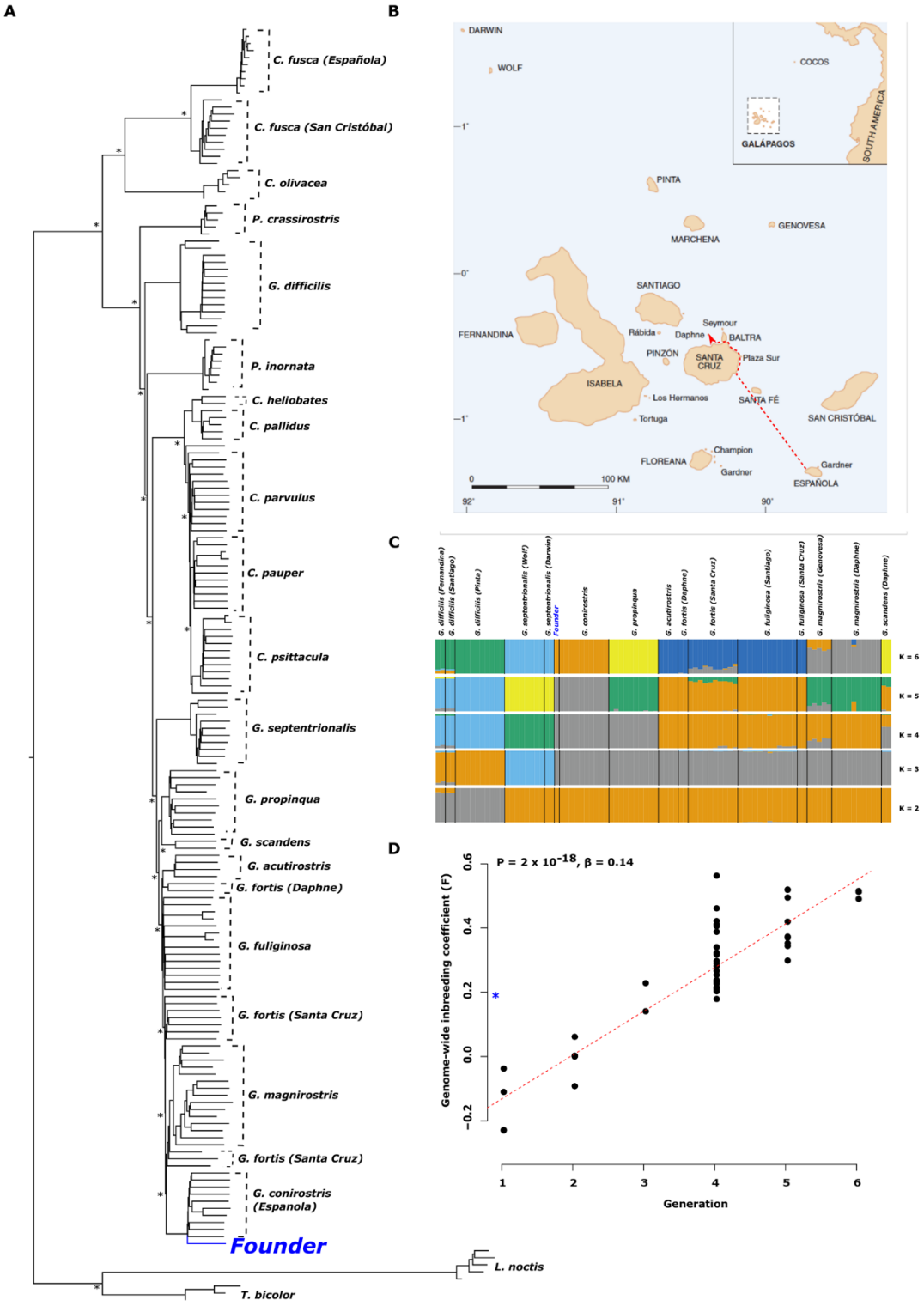


Fig. 2. Phylogeny, geography, and increase in homozygosity. (**A**) Maximum likelihood tree of Darwin's finches constructed from whole genomes (this study and *13*, *21*). The founder male of the Big Bird lineage is highlighted in blue. All nodes having full local support based on Shimodaira-Hasegawa test are marked with asterisks. (**B**) Map of Galápagos. The original colonist individual on Daphne Major originated on Española Island (or its satellite Gardner) in the southeast of the archipelago >100 km from Daphne. The hypothetical flight path is informed by observations (B.R.G. and P.R.G., 1973-75) of post-breeding movement of finches on Santa Cruz Island, northwards on the east coast and westwards on the north coast. (**C**) Maximum likelihood estimation of individual admixture proportions using genome-wide SNP data for a range of pre-assumed population (K=2-6). (**D**) Increase in homozygosity (genome-wide inbreeding coefficient, F) in the Big Bird lineage over generations. The estimate for the original colonist is shown by an asterisk.

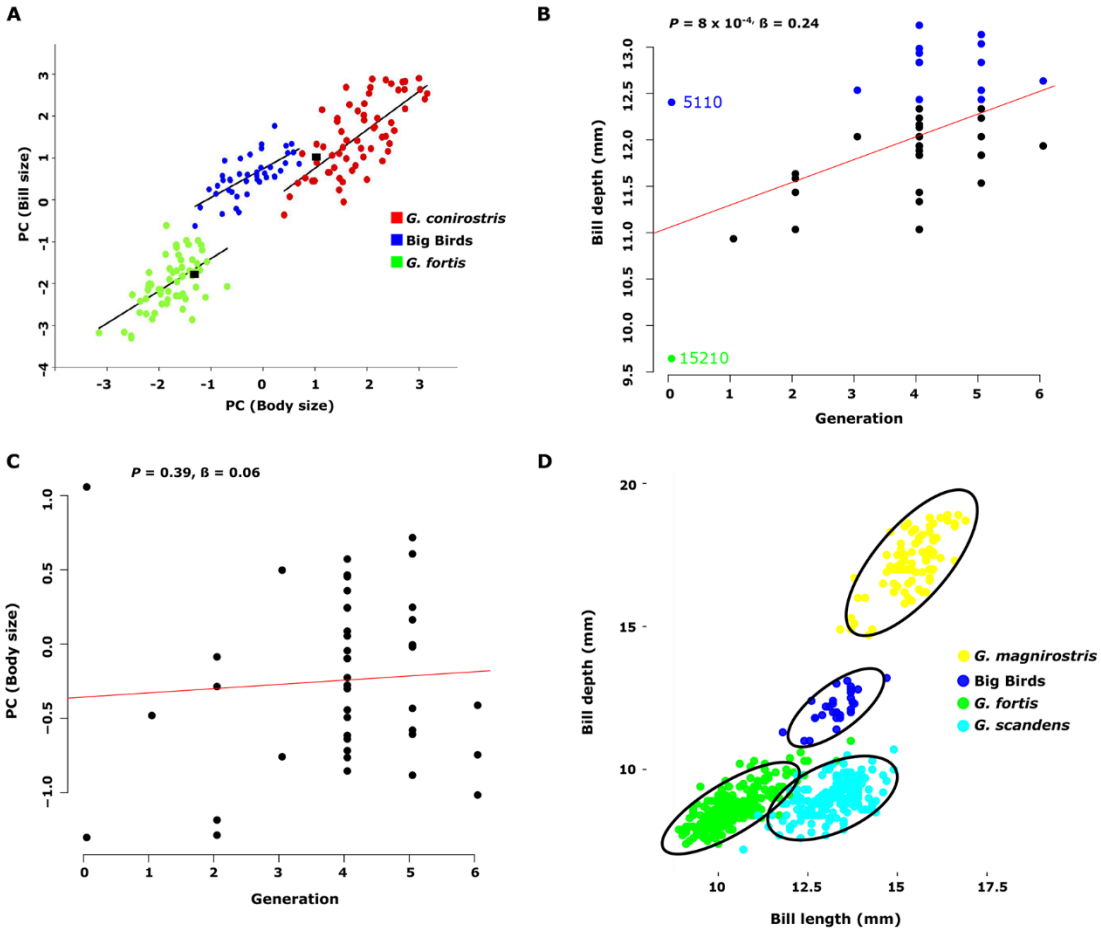


Fig. 3. Morphology. (A) Bill size variation in relation to body size among the Big Bird lineage (blue) and the two parental species *G. fortis* (green) and *G. conirostris* (red). All 42 Big Birds lie above a line connecting the two parents, indicated by black squares, and above a line connecting the means of the two populations. The ordinary least squares regression slopes of the three relationships (table S3) are homogeneous (ANCOVA, $F_{2,161} = 1.4$, $P = 0.26$), whereas the intercepts differ (ANCOVA with species*bill size interaction removed, $F_{2,163} = 140.9$, $P = 0.0001$). The 99% confidence limits on the Big Bird intercept estimate do not overlap those of the other two intercepts, whereas the 95% confidence limits of the *G. conirostris* and *G. fortis* slopes do overlap. This pattern is repeated in two of the components of bill size, depth and width [fig. S3, for bill width see (*14*)]. Images © Peter & Rosemary Grant. (B) Mean bill depth increased over generations. $F_{1,42} = 12.9$, $P = 0.0008$, adj $r^2 = 0.22$, slope = 0.25 ± 0.07. The relationship holds for generations 1-6 alone, i.e., without the founder and its mate ($F_{1,40} = 9.1$, $P = 0.004$, adj $r^2 = 0.16$, slope = 0.24 ± 0.08). Note the bill depth of the single member of generation 1 is 10.9 mm, which is close to the midpoint of the parental measurements (11.0mm). Transgressive segregation for bill depth in the Big Bird lineage is possibly indicated by the fraction of individuals that exceeded parental phenotypic values (highlighted in blue) (*10*), which was estimated at 0.5, 0.3, 0.4, and 0.3 in generations 3-6, respectively. (C) Mean body size remained unchanged across generations ($F_{1,42} = 0.76$, $P = 0.39$, adj $r^2 = 0.00$, slope = 0.06 ± 0.07). (D) The Big Bird lineage (blue) occupies unique morphological space among the coexisting ecological competitor species. Ellipses contain 95% of individuals. *G. magnirostris* (yellow), *G. fortis* (green) and *G. scandens* (aqua).

Cite as: H. E. Autzen *et al.*, *Science* 10.1126/science.aar4510 (2017).

Structure of the human TRPM4 ion channel in a lipid nanodisc

Henriette E. Autzen,[1,2] Alexander G. Myasnikov,[1] Melody G. Campbell,[1] Daniel Asarnow,[1] David Julius,[3]* Yifan Cheng[1,4]*

[1]Department of Biochemistry and Biophysics, University of California, San Francisco, CA 94143, USA. [2]Department of Molecular Biology and Genetics, University of Aarhus, 8000 Aarhus, Denmark. [3]Department of Physiology, University of California, San Francisco, CA 94143, USA. [4]Howard Hughes Medical Institute, University of California, San Francisco, CA 94143, USA.

*Corresponding author. Email: david.julius@ucsf.edu (D.J.); yifan.cheng@ucsf.edu (Y.C.)

Transient receptor potential (TRP) melastatin 4 (TRPM4) is a widely expressed cation channel associated with a variety of cardiovascular disorders. TRPM4 is activated by increased intracellular calcium in a voltage dependent manner, but unlike many other TRP channels is permeable to monovalent cations only. Here we present two structures of full-length human TRPM4 embedded in lipid nanodiscs at ~3Å resolution as determined by single particle electron cryo-microscopy. These structures, with and without calcium bound, reveal a general architecture for this major subfamily of TRP channels and a well-defined calcium binding site within the intracellular side of the S1-S4 domain. The structures correspond to two distinct closed states. Calcium binding induces conformational changes that likely prime the channel for voltage-dependent opening.

Transient receptor potential (TRP) channels comprise an extended superfamily of membrane ion channels that mediate diverse cellular and physiological functions. Permeability to both mono- and divalent cations is a defining feature for the majority of TRP channels (*1*). However, the TRP melastatin subfamily member 4 (TRPM4) channel is impermeable to Ca^{2+}, yet activated by intracellular Ca^{2+} (*2*). Activation of TRPM4 depolarizes the plasma membrane through Na^{+}-entry, which in turn enhances Ca^{2+}-influx through Ca^{2+}-permeable channels or otherwise modulates Ca^{2+} oscillations (*2*, *3*). Like many other TRP channel subtypes, TRPM4 displays voltage sensitivity as evidenced by pronounced outward rectification of Ca^{2+}-activated TRPM4 currents in response to voltage ramps (*4*). Binding of Ca^{2+} to TRPM4 is hypothesized to precede voltage-dependent opening, thereby giving rise to two closed states: Ca^{2+}-unbound and Ca^{2+}-bound (*5*). However, it is unclear where the Ca^{2+}-binding site is located in the channel and a structural basis for voltage sensitivity and ion selectivity remains elusive. Here, we present two single particle electron cryo-microscopy (cryo-EM) structures of the human TRPM4 channel in nanodiscs, in both Ca^{2+}-bound and unbound closed states. Comparison of the two structures reveals a Ca^{2+} binding site in the transmembrane domain, and conformational changes induced by Ca^{2+}. Lipid molecules, including cholesteryl hemisuccinate (CHS), form tight interactions with the channel, likely stabilizing it in the lipid bilayer.

Recombinant full-length human TRPM4 (hTRPM4b) was purified in detergent, and reconstituted into lipid nanodiscs as described for other TRP ion channels (*6*) (fig. S1). Three-dimensional (3D) reconstructions with a C4 symmetry were obtained in the presence of either 5 mM EDTA or $CaCl_2$ to overall resolutions of 3.2Å and 3.1Å (Fig. 1, A and B and figs. S2, A to F, and S3, A to F). In both structures, the transmembrane domain and part of the soluble domain are well resolved. However, the soluble regions distal to the membrane and the symmetry axis are of lower resolution, indicating conformational flexibility (figs. S2G and S3G). Focused refinement of the transmembrane domain with the central coiled-coil improved local resolution and enabled de novo atomic model building of these domains for both samples (figs. S2, F and H; S3, F and H; S4, B and D; and S5). For the rest of the cytoplasmic domains, further classification and focused refinement of individual monomers without symmetry improved the resolution to a level sufficient for de novo model building (Fig. 1, C to F and figs. S2, I to K; S3, I to K; S4, C and E; and S5). The complete data processing scheme is described in the supplementary methods and fig. S4.

The transmembrane core contains six helices (S1-S6) arranged in a domain-swapped architecture (Fig. 1, C and D) reminiscent of many tetrameric voltage-gated cation channels, and all other TRP channels to date (*7*–*13*). hTRPM4 has additional membrane embedded fragments that surround the S1-S4 domain from the inner leaflet of the membrane (Fig. 1, D and F), a feature that might be distinct to the TRPM subfamily. The C terminus of each subunit joins to form a parallel coiled-coil resembling that of TRPA1 channel (*11*). The coiled-coil is surrounded by a large, intertwined cytoplasmic domain, comprised of four N-terminal TRPM Homology Regions (MHR1-4), which are highly conserved within the TRPM subfamily (*14*). MHR1-2 contains an eight-stranded β-

sheet surrounded by eight α-helices. MHR3-4 is composed of stacked α-helices and linked to the transmembrane domain through MHR4, which clasps the TRP domain, thereby forming an interaction between the cytoplasmic domain and the transmembrane core (Fig. 1D). Between neighboring monomers, MHR1 of one subunit interacts with the MHR3 of the neighboring subunit forming a ring that does not interact with the central coiled-coil (Fig. 1E).

Comparison of Ca^{2+}-free and Ca^{2+}-bound structures reveals an extra density in a hydrophilic pocket within the cytoplasmic side of the S1-S4 domain (Fig. 2A and fig. S6). Several lines of analysis suggest that this density represents a *bona fide* bound Ca^{2+}-ion: 1) Direct comparison between both half maps and the combined maps of the $CaCl_2$ and EDTA samples confirms the existence and precise location of the additional density in the $CaCl_2$ map (fig. S6, A to C and E to G); 2) the omit densities calculated as the difference between the experimental maps and map calculated from the refined atomic models without Ca^{2+} show a strong density corresponding to a bound Ca^{2+}-ion (fig. S6D) in the $CaCl_2$ structure, but not in the EDTA structure; 3) the difference map between the $CaCl_2$ and EDTA maps reveals a clear density with a high signal-to-noise ratio (Fig. 2A and fig. S6H); 4) side chain densities of surrounding negatively charged residues are stronger in the $CaCl_2$ structure than in the EDTA structure (fig. S6, A and E), consistent with the observation that negatively charged side chains are generally weaker in a cryo-EM density map unless they are engaged in specific interactions (*15*).

The Ca^{2+} ion is directly coordinated by side chains of Glu828 and Gln831 from S2, and Asn865 from S3 (Fig. 2A). Asp868 is also near the ion and may participate in its coordination. Together, these side chains contribute at least four oxygens that can coordinate the Ca^{2+}. The Ca^{2+} binding site is located within a hydrophilic pocket that is large enough to accommodate additional water molecules for further coordination of the bound ion. This pocket is connected to the cytoplasmic space through a narrow path between the TRP domain and the S2-S3 linker (fig. S6, I and J). Within the pocket, Arg905 from a single 3_{10} helical turn in S4 and Tyr790 from the S1 are located above the Ca^{2+} binding site. In the absence of Ca^{2+}, Arg905 is coordinated by Glu828 and Asp868 (Fig. 2B). Upon Ca^{2+} binding, the S2-S3 linker translates ~1.5 Å such that Tyr859 supports a tighter conformation for coordination of Ca^{2+} by Glu828, Gln831, Asn865 and Asp868 (fig. S6L). Furthermore, the side chain of His908 in S4 transitions from a π-π stacking arrangement with Trp864 in S3 to form a new interaction with Cys867 (Fig. 2, B and C, and fig. S6, K to N). Without coordination of Glu828 and with rearrangement in S3 and S4, Arg905 moves slightly upward (fig. S6N). The configuration of Arg905 and Tyr790 is reminiscent of a positive gating charge (arginine) and charge transfer center (tyrosine) seen in voltage-gated potassium channels (*16*). We speculate that Ca^{2+} binding moves Arg905 up toward Tyr790 to prime the channel for voltage-dependent opening.

A previous mutagenesis study showed that the Glu1068Gln mutation significantly reduces TRPM4 Ca^{2+}-sensitivity (*17*). Glu1068 is located in the pathway leading to the Ca^{2+}-binding site from the cytoplasmic space (fig. S6, I and J), and this mutation likely reduces the Ca^{2+} accessibility to the site. Of the four residues that constitute the Ca^{2+} binding site, Asn865 and the negative charge of Asp868 are conserved throughout the TRPM subfamily, while Glu828 and Gln831 as well as Tyr859 and Glu1068 are conserved only in subfamily members that are shown to be Ca^{2+} dependent: TRPM2, TRPM4, TRPM5 and TRPM8 (*18–20*) (fig. S7). Therefore, it is likely that there is an analogous binding site in these channels. Electrophysiological and mutagenesis studies of TRPM8 suggest that Ca^{2+} binds directly to an analogous site in that channel to confer sensitivity to icilin, a synthetic cooling agent. Specifically, substituting the residues corresponding to Asn865 and Asp868 resulted in icilin-insensitive channels that retain robust Ca^{2+}-independent responses to cold and menthol (*21*).

The central pore of hTRPM4 is formed by S5 and S6, and the intervening re-entrant pore helix and pore loop. Together, these elements form an ion permeation pathway with two restriction sites, similar to other TRP channels (Fig. 3, A and B). As seen in TRPV1 (*10*), the outer pore has a funnel shape with a negatively charged residue facing the funnel, attracting cations (fig. S8, A and B). Two highly conserved residues, Phe975 and Gly976, located in the bottom of the pore loop, form the narrowest restriction point of the upper pore. The pore diameter at this site is very similar to that of the open gate in TRPV1 (*22*) (Fig. 3B), which is sufficient to accommodate partially dehydrated monovalent cations. However, no density corresponding to coordinated ions is found within the pore (fig. S8, C and D). The location of a previously identified putative selectivity filter (^{981}EDMDVA^{986}) (*23*) does not coincide with this restriction site, but is located further toward the extracellular face (Fig. 3C). At the bottom of the ion permeation pathway, side chains of opposing Ile1040 residues in S6 form a tight seal at the lower restriction site, signifying that both structures are in closed conformations (Fig. 3, A and B).

The pore helix of hTRPM4 has an additional turn compared to that in other TRP channels (Fig. 3C) (*9–13*). In the middle of the pore helix there is a single-turn π-helix (Arg964-Arg969) followed by a proline residue, Pro970, conserved in all human TRPM channels except TRPM2. Likewise, similar to what was first identified in TRPV1 (*10*) and subsequently observed in other TRP channels (*7*, *11*), there is a one-turn π-helix (Val1030-Leu1035) in the middle of the S6 helix (Fig.

3C). While the exact role of these single-turn π-helices is unclear, they could potentially facilitate helix bending under different functional states, leading to movement of the lower gate or modulation of the upper pore.

Outside of the pore, hTRPM4 has an extensive extracellular loop comprising more than 30 residues that connect the pore loop to the S6 helix. The loop is clearly resolved, likely stabilized by a disulfide bond between two cysteines (Cys993 and Cys1011) (Fig. 3C and fig. S8, F and G), a feature predicted to exist in the majority of TRPM channels. This disulfide bond is not seen in the $CaCl_2$ structure (fig. S7H), where disulfide bond breakage may have been caused by radiation damage (*24*). The loop also contains a glycosylation site (Asn992) with the attached glycan pointing toward the extracellular space (Fig. 3C and fig. S7, G and H).

In addition to the transmembrane S1-S6 domain, hTRPM4 has distinct membrane embedded α-helical segments (Fig. 4A). These segments surround the exterior of the S1-S4 domain within the inner leaflet of the membrane and mediate extensive interactions with the soluble domain. Preceding the S1 helix, there are two short helices shaped as an inverted "V" embedded within the inner leaflet of the membrane (Fig. 4B). We term it the "pre-S1 elbow," similar to that observed in the mechanotransduction channel NOMPC (*9*). A disordered loop unresolved in both structures, connects the pre-S1 elbow to an amphipathic helix termed as the "pre-S1 shoulder" positioned at the inner surface of the membrane (Fig. 4C). The pre-S1 shoulder helix contains large hydrophobic residues buried in the membrane and a number of charged residues facing the cytosol. Among these charged residues, Arg767 was identified as important for interactions with the phosphatidylinositol lipids, PIP2 and PIP3 (*25*).

An extended S2-S3 linker distinct to the TRPM subfamily forms a short amphipathic helix, positioned at the membrane surface, near the Ca^{2+}-binding site (Fig. 2B). The loop that connects S2 with the S2-S3 linker is ~12 residues longer in hTRPM4 compared to other human TRPM channels and extends beyond the membrane bilayer, interacting with MHR4. Ser839 in the linker is a predicted phosphorylation site important for trafficking TRPM4 to plasma membrane (*26*), however, this linker is only weakly resolved in the $CaCl_2$ structure. In addition, the S2-S3 linker also interacts loosely with the cytoplasmic domain MHR3 (Fig. 4D).

The last segment is positioned between the pre-S1 shoulder and S2-S3 linker helix (Fig. 4, A and F). Here, a loop fills the gap between S1 and S2 and connects the conserved TRP domain to the C-terminal Helix 1 (CH1), which is located at a steep angle relative to S1, extending downwards from the middle of the membrane bilayer to the cytoplasmic surface more than 20Å away from the end of the TRP domain (Fig. 4E). The C terminus of CH1 contains five arginine residues, similar to the pre-S1 shoulder mentioned above, which likely constitute interaction partners for negatively charged lipid headgroups at the membrane interface. CH1 connects to CH2 through a 19-residue linker which was only partially resolved in the structures and observed to associate with MHR4 and the C-terminal end of the TRP-domain.

A number of lipids were observed in our structures (fig. S9A). In addition to annular phospholipids, we identified three densities as CHS based on its characteristic shape (Figs. 3C and 4B and fig. S9B). CHS is often used as a cholesterol analog (*27*), thus a cholesterol molecule is likely to bind at the same sites. One CHS molecule fills a cleft at the backside of the pore, interacting with the S6 helix and the pore loop or the proposed selectivity filter of one subunit and the pore helix of the neighboring subunit. This arrangement, together with the relative rigidity of CHS over annular phospholipids, may stabilize the conformation of the pore (Fig. 3C and fig. S8E). A second CHS molecule is positioned at a location equivalent to the vanilloid binding pocket in TRPV1 (*6*) (fig. S9C). A third CHS is located at the pre-S1 elbow, which together with S1 and S4 creates a cavity (Fig. 4B).

While there is no noticeable change of the pore between the two structures (Fig. 3B and fig. S10, A and C), we observe subtle yet well-defined changes in the region surrounding the Ca^{2+} binding site (fig. S10). In addition, CH2 is horizontally displaced by ~2.5 Å around the central coiled-coil, measured from the C_α of Glu1116, in the distal end of CH2 (fig. S10D). The rotation results in a slight tightening of the central coiled-coil (fig. S10E), likely stabilizing it as evident by the observation that more residues in the C-terminal end of the coiled-coil were resolved in the $CaCl_2$ structure (figs. S2G and 3G). Given that the lower gate remains closed in both structures, such conformational changes are insufficient for channel opening. However, as the extensive soluble domains are implicated in interactions with various co-factors (*28*), such conformational changes may be important functional features that enable the channel to better detect or response to its co-factors (*28*).

REFERENCES AND NOTES

1. D. Julius, TRP channels and pain. *Annu. Rev. Cell Dev. Biol.* **29**, 355–384 (2013). doi:10.1146/annurev-cellbio-101011-155833 Medline
2. P. Launay, A. Fleig, A.-L. Perraud, A. M. Scharenberg, R. Penner, J.-P. Kinet, TRPM4 is a Ca^{2+}-activated nonselective cation channel mediating cell membrane depolarization. *Cell* **109**, 397–407 (2002). doi:10.1016/S0092-8674(02)00719-5 Medline
3. P. Launay, H. Cheng, S. Srivatsan, R. Penner, A. Fleig, J. P. Kinet, TRPM4 regulates calcium oscillations after T cell activation. *Science* **306**, 1374–1377 (2004). doi:10.1126/science.1098845 Medline
4. B. Nilius, J. Prenen, G. Droogmans, T. Voets, R. Vennekens, M. Freichel, U. Wissenbach, V. Flockerzi, Voltage dependence of the Ca^{2+}-activated cation channel TRPM4. *J. Biol. Chem.* **278**, 30813–30820 (2003). doi:10.1074/jbc.M305127200 Medline
5. B. Nilius, J. Prenen, A. Janssens, T. Voets, G. Droogmans, Decavanadate modulates gating of TRPM4 cation channels. *J. Physiol.* **560**, 753–765 (2004). doi:10.1113/jphysiol.2004.070839 Medline
6. Y. Gao, E. Cao, D. Julius, Y. Cheng, TRPV1 structures in nanodiscs reveal

mechanisms of ligand and lipid action. *Nature* **534**, 347–351 (2016). doi:10.1038/nature17964 Medline
7. Q. Chen, J. She, W. Zeng, J. Guo, H. Xu, X. C. Bai, Y. Jiang, Structure of mammalian endolysosomal TRPML1 channel in nanodiscs. *Nature* **550**, 415–418 (2017). doi:10.1038/nature24035 Medline
8. M. Hirschi, M. A. Herzik Jr., J. Wie, Y. Suo, W. F. Borschel, D. Ren, G. C. Lander, S.-Y. Lee, Cryo-electron microscopy structure of the lysosomal calcium-permeable channel TRPML3. *Nature* **550**, 411–414 (2017). doi:10.1038/nature24055 Medline
9. P. Jin, D. Bulkley, Y. Guo, W. Zhang, Z. Guo, W. Huynh, S. Wu, S. Meltzer, T. Cheng, L. Y. Jan, Y.-N. Jan, Y. Cheng, Electron cryo-microscopy structure of the mechanotransduction channel NOMPC. *Nature* **547**, 118–122 (2017). doi:10.1038/nature22981 Medline
10. M. Liao, E. Cao, D. Julius, Y. Cheng, Structure of the TRPV1 ion channel determined by electron cryo-microscopy. *Nature* **504**, 107–112 (2013). doi:10.1038/nature12822 Medline
11. C. E. Paulsen, J. P. Armache, Y. Gao, Y. Cheng, D. Julius, Structure of the TRPA1 ion channel suggests regulatory mechanisms. *Nature* **520**, 511–517 (2015). doi:10.1038/nature14367 Medline
12. P. Schmiege, M. Fine, G. Blobel, X. Li, Human TRPML1 channel structures in open and closed conformations. *Nature* **550**, 366–370 (2017). doi:10.1038/nature24036 Medline
13. P. S. Shen, X. Yang, P. G. DeCaen, X. Liu, D. Bulkley, D. E. Clapham, E. Cao, The structure of the polycystic kidney disease channel PKD2 in lipid nanodiscs. *Cell* **167**, 763–773.e11 (2016). doi:10.1016/j.cell.2016.09.048 Medline
14. C. B. Phelps, R. Gaudet, The role of the N terminus and transmembrane domain of TRPM8 in channel localization and tetramerization. *J. Biol. Chem.* **282**, 36474–36480 (2007). doi:10.1074/jbc.M707205200 Medline
15. A. Bartesaghi, D. Matthies, S. Banerjee, A. Merk, S. Subramaniam, Structure of β-galactosidase at 3.2-Å resolution obtained by cryo-electron microscopy. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 11709–11714 (2014). doi:10.1073/pnas.1402809111 Medline
16. X. Tao, A. Lee, W. Limapichat, D. A. Dougherty, R. MacKinnon, A gating charge transfer center in voltage sensors. *Science* **328**, 67–73 (2010). doi:10.1126/science.1185954 Medline
17. S. Yamaguchi, A. Tanimoto, K. Otsuguro, H. Hibino, S. Ito, Negatively charged amino acids near and in transient receptor potential (TRP) domain of TRPM4 channel are one determinant of its Ca^{2+} sensitivity. *J. Biol. Chem.* **289**, 35265–35282 (2014). doi:10.1074/jbc.M114.606087 Medline
18. M. Faouzi, R. Penner, TRPM2. *Handb. Exp. Pharmacol.* **222**, 403–426 (2014). doi:10.1007/978-3-642-54215-2_16 Medline
19. T. Hofmann, V. Chubanov, T. Gudermann, C. Montell, TRPM5 is a voltage-modulated and Ca^{2+}-activated monovalent selective cation channel. *Curr. Biol.* **13**, 1153–1158 (2003). doi:10.1016/S0960-9822(03)00431-7 Medline
20. D. D. McKemy, W. M. Neuhausser, D. Julius, Identification of a cold receptor reveals a general role for TRP channels in thermosensation. *Nature* **416**, 52–58 (2002). doi:10.1038/nature719 Medline
21. H. H. Chuang, W. M. Neuhausser, D. Julius, The super-cooling agent icilin reveals a mechanism of coincidence detection by a temperature-sensitive TRP channel. *Neuron* **43**, 859–869 (2004). doi:10.1016/j.neuron.2004.08.038 Medline
22. E. Cao, M. Liao, Y. Cheng, D. Julius, TRPV1 structures in distinct conformations reveal activation mechanisms. *Nature* **504**, 113–118 (2013). doi:10.1038/nature12823 Medline
23. B. Nilius, J. Prenen, A. Janssens, G. Owsianik, C. Wang, M. X. Zhu, T. Voets, The selectivity filter of the cation channel TRPM4. *J. Biol. Chem.* **280**, 22899–22906 (2005). doi:10.1074/jbc.M501686200 Medline
24. M. Weik, R. B. G. Ravelli, G. Kryger, S. McSweeney, M. L. Raves, M. Harel, P. Gros, I. Silman, J. Kroon, J. L. Sussman, Specific chemical and structural damage to proteins produced by synchrotron radiation. *Proc. Natl. Acad. Sci. U.S.A.* **97**, 623–628 (2000). doi:10.1073/pnas.97.2.623 Medline
25. K. Bousova, M. Jirku, L. Bumba, L. Bednarova, M. Sulc, M. Franek, L. Vyklicky, J. Vondrasek, J. Teisinger, PIP2 and PIP3 interact with N-terminus region of TRPM4 channel. *Biophys. Chem.* **205**, 24–32 (2015). doi:10.1016/j.bpc.2015.06.004 Medline
26. S. H. Shin, E. J. Lee, J. Chun, S. Hyun, S. S. Kang, Phosphorylation on TRPV4 serine 824 regulates interaction with STIM1. *Open Biochem. J.* **9**, 24–33 (2015). doi:10.2174/1874091X01509010024 Medline
27. M. Zocher, C. Zhang, S. G. Rasmussen, B. K. Kobilka, D. J. Müller, Cholesterol increases kinetic, energetic, and mechanical stability of the human β2-adrenergic receptor. *Proc. Natl. Acad. Sci. U.S.A.* **109**, E3463–E3472 (2012). doi:10.1073/pnas.1210373109 Medline
28. B. Nilius, J. Prenen, J. Tang, C. Wang, G. Owsianik, A. Janssens, T. Voets, M. X. Zhu, Regulation of the Ca^{2+} sensitivity of the nonselective cation channel TRPM4. *J. Biol. Chem.* **280**, 6423–6433 (2005). doi:10.1074/jbc.M411089200 Medline
29. A. Goehring, C.-H. Lee, K. H. Wang, J. C. Michel, D. P. Claxton, I. Baconguis, T. Althoff, S. Fischer, K. C. Garcia, E. Gouaux, Screening and large-scale expression of membrane proteins in mammalian cells for structural studies. *Nat. Protoc.* **9**, 2574–2585 (2014). doi:10.1038/nprot.2014.173 Medline
30. T. K. Ritchie, Y. V. Grinkova, T. H. Bayburt, I. G. Denisov, J. K. Zolnerciks, W. M. Atkins, S. G. Sligar, Reconstitution of membrane proteins in phospholipid bilayer nanodiscs. *Methods Enzymol.* **464**, 211–231 (2009). doi:10.1016/S0076-6879(09)64011-8 Medline
31. M. Ohi, Y. Li, Y. Cheng, T. Walz, Negative staining and image classification – Powerful tools in modern electron microscopy. *Biol. Proced. Online* **6**, 23–34 (2004). doi:10.1251/bpo70 Medline
32. S. Q. Zheng, E. Palovcak, J.-P. Armache, K. A. Verba, Y. Cheng, D. A. Agard, MotionCor2: Anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat. Methods* **14**, 331–332 (2017). doi:10.1038/nmeth.4193 Medline
33. K. Zhang, Gctf: Real-time CTF determination and correction. *J. Struct. Biol.* **193**, 1–12 (2016). doi:10.1016/j.jsb.2015.11.003 Medline
34. A. Punjani, M. A. Brubaker, D. J. Fleet, Building proteins in a day: Efficient 3D molecular structure estimation with electron cryomicroscopy. *IEEE Trans. Pattern Anal. Mach. Intell.* **39**, 706–718 (2017). doi:10.1109/TPAMI.2016.2627573 Medline
35. S. H. Scheres, RELION: Implementation of a Bayesian approach to cryo-EM structure determination. *J. Struct. Biol.* **180**, 519–530 (2012). doi:10.1016/j.jsb.2012.09.006 Medline
36. I. S. Fernández, X.-C. Bai, T. Hussain, A. C. Kelley, J. R. Lorsch, V. Ramakrishnan, S. H. W. Scheres, Molecular architecture of a eukaryotic translational initiation complex. *Science* **342**, 1240585 (2013). doi:10.1126/science.1240585 Medline
37. N. Grigorieff, Frealign: An exploratory tool for single-particle cryo-EM. *Methods Enzymol.* **579**, 191–226 (2016). doi:10.1016/bs.mie.2016.04.013 Medline
38. M. Zhou, Y. Li, Q. Hu, X. C. Bai, W. Huang, C. Yan, S. H. W. Scheres, Y. Shi, Atomic structure of the apoptosome: Mechanism of cytochrome c- and dATP-mediated activation of Apaf-1. *Genes Dev.* **29**, 2349–2361 (2015). doi:10.1101/gad.272278.115 Medline
39. S. H. Scheres, S. Chen, Prevention of overfitting in cryo-EM structure determination. *Nat. Methods* **9**, 853–854 (2012). doi:10.1038/nmeth.2115 Medline
40. P. B. Rosenthal, R. Henderson, Optimal determination of particle orientation, absolute hand, and contrast loss in single-particle electron cryomicroscopy. *J. Mol. Biol.* **333**, 721–745 (2003). doi:10.1016/j.jmb.2003.07.013 Medline
41. J. B. Heymann, Guidelines for using Bsoft for high resolution reconstruction and validation of biomolecular structures from electron micrographs. *Protein Sci.* 10.1002/pro.3293 (2017). doi:10.1002/pro.3293 Medline
42. P. Emsley, B. Lohkamp, W. G. Scott, K. Cowtan, Features and development of *Coot*. *Acta Crystallogr. D* **66**, 486–501 (2010). doi:10.1107/S0907444910007493 Medline
43. N. Stein, *CHAINSAW*: A program for mutating pdb files used as templates in molecular replacement. *J. Appl. Cryst.* **41**, 641–643 (2008). doi:10.1107/S0021889808006985
44. P. D. Adams, P. V. Afonine, G. Bunkóczi, V. B. Chen, I. W. Davis, N. Echols, J. J. Headd, L.-W. Hung, G. J. Kapral, R. W. Grosse-Kunstleve, A. J. McCoy, N. W. Moriarty, R. Oeffner, R. J. Read, D. C. Richardson, J. S. Richardson, T. C. Terwilliger, P. H. Zwart, *PHENIX*: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D* **66**, 213–221 (2010). doi:10.1107/S0907444909052925 Medline
45. M. Källberg, H. Wang, S. Wang, J. Peng, Z. Wang, H. Lu, J. Xu, Template-based protein structure modeling using the RaptorX web server. *Nat. Protoc.* **7**, 1511–

1522 (2012). doi:10.1038/nprot.2012.085 Medline
46. V. B. Chen, W. B. Arendall III, J. J. Headd, D. A. Keedy, R. M. Immormino, G. J. Kapral, L. W. Murray, J. S. Richardson, D. C. Richardson, *MolProbity*: All-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D* **66**, 12–21 (2010). doi:10.1107/S0907444909042073 Medline
47. B. A. Barad, N. Echols, R. Y.-R. Wang, Y. Cheng, F. DiMaio, P. D. Adams, J. S. Fraser, EMRinger: Side chain-directed model and map validation for 3D cryo-electron microscopy. *Nat. Methods* **12**, 943–946 (2015). doi:10.1038/nmeth.3541 Medline
48. O. S. Smart, J. G. Neduvelil, X. Wang, B. A. Wallace, M. S. Sansom, HOLE: A program for the analysis of the pore dimensions of ion channel structural models. *J. Mol. Graph.* **14**, 354–360 (1996). doi:10.1016/S0263-7855(97)00009-X Medline
49. E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng, T. E. Ferrin, UCSF Chimera—A visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004). doi:10.1002/jcc.20084 Medline
50. W. Humphrey, A. Dalke, K. Schulten, VMD: Visual molecular dynamics. *J. Mol. Graph.* **14**, 33–38 (1996). doi:10.1016/0263-7855(96)00018-5 Medline
51. X. Robert, P. Gouet, Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res.* **42**, W320–W324 (2014). doi:10.1093/nar/gku316 Medline

ACKNOWLEDGMENTS

We thank M. Diver, P. Dominik, A. Kintzer, R. Stroud for valuable discussions, M. Diver and E. Green for reading the manuscript. We thank Y. Jiang for coordinating co-submission of our studies. H.E.A. is supported by a postdoctoral fellowship from the Danish Council of Independent Research (Grant No. DFF-5051-00085). This work was supported by grants from National Institute of Health (R01NS047723 to D.J., R01GM098672, S10OD020054, S10OD021741 to Y.C.). Y.C. is an Investigator of the Howard Hughes Medical Institute. Accession numbers of the human TRPM4 structures in EDTA and $CaCl_2$ are: 6BQR and 6BQV (coordinates of atomic models), EMD-7132 and EMD-7133 (density maps), and EMPIAR-10126 and EMPIAR-10127 (raw particle stacks).

SUPPLEMENTARY MATERIALS

www.sciencemag.org/cgi/content/full/science.aar4510/DC1
Materials and Methods
Figs. S1 to S10
Tables S1 to S3
References (*29*–*51*)

9 November 2017; accepted 27 November 2017
Published online 7 December 2017
10.1126/science.aar4510

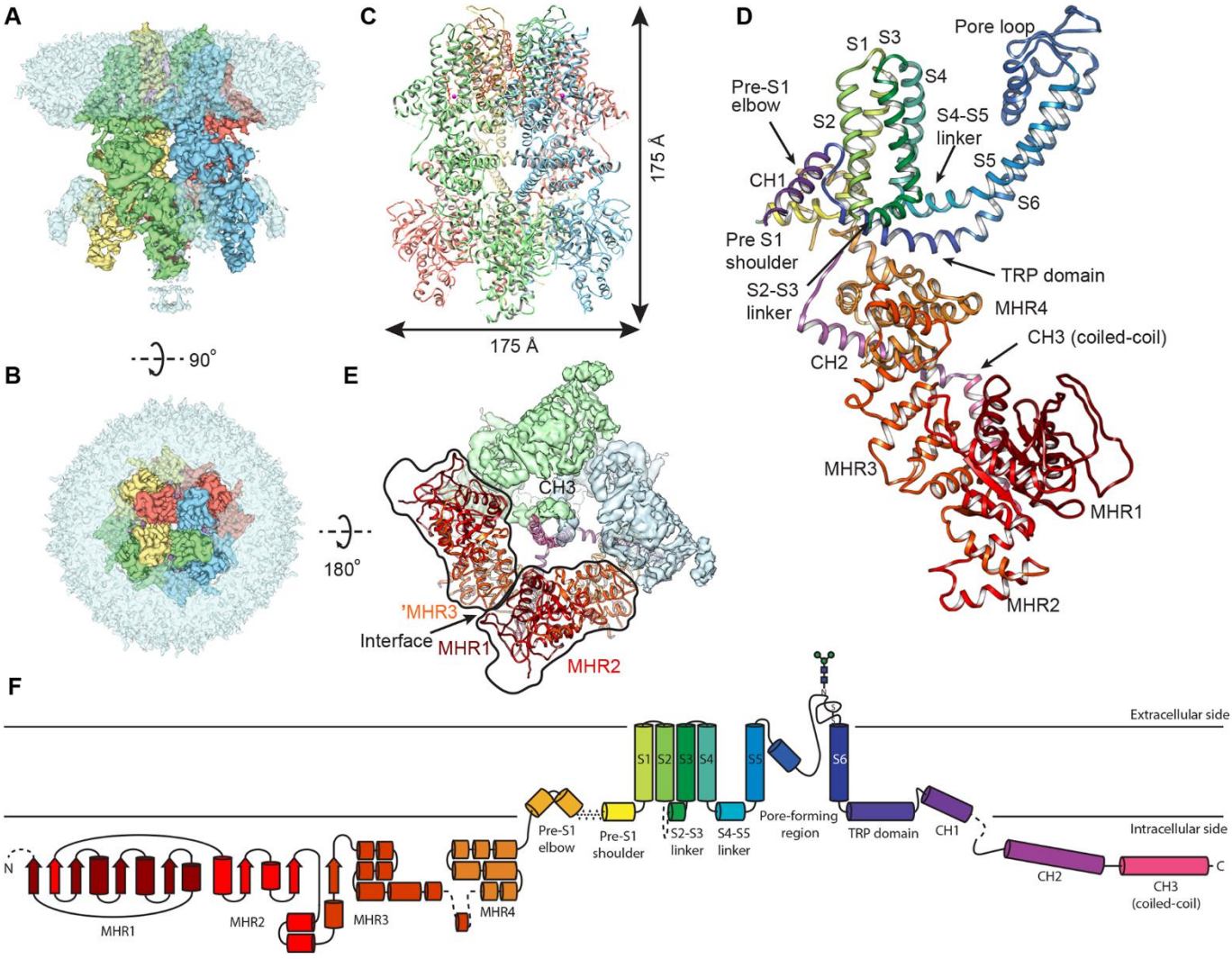


Fig. 1. Cryo-EM structure of hTRPM4. (A and B) Unsharpened (in transparent light blue) and sharpened cryo-EM density map of hTRPM4 in nanodiscs in side (A) and top (B) views with each subunit colored differently. (C) Atomic model of the tetrameric hTRPM4 in ribbons, in the same orientation and colors as the density map in (A). (D) Atomic model of the hTRPM4b monomer in ribbons with domains labeled. (E) bottom view of the hTRPM4 structure showing interactions between neighboring subunits. (F) A schematic representation of the major structural components in hTPPM4b. Dashed lines denote regions where density was insufficient for model building. Each domain is labeled and color-coded to match the domain representation in (D).

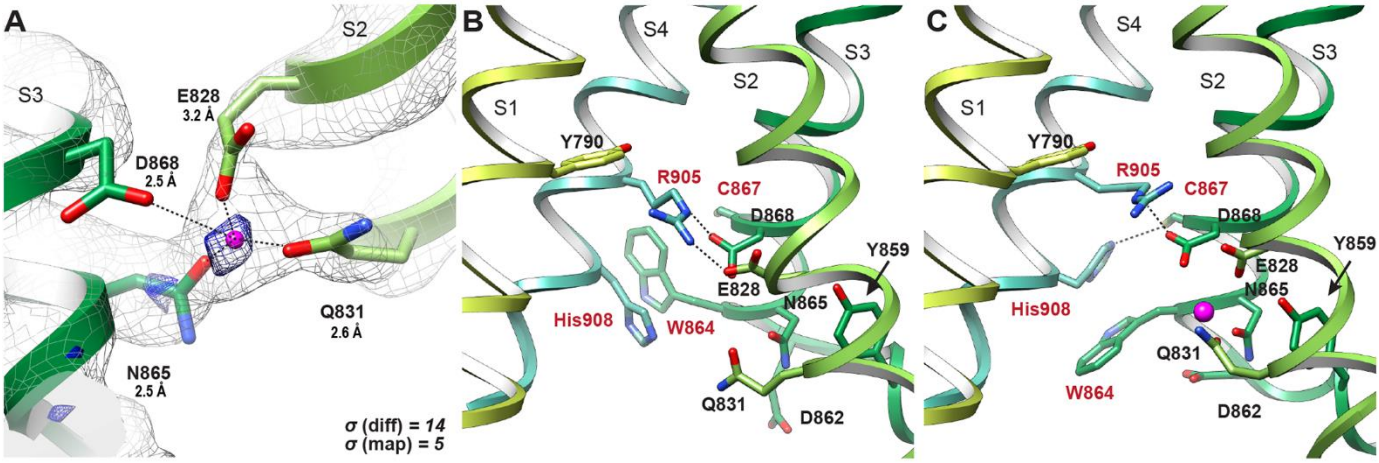


Fig. 2. The Ca^{2+} binding site. (A) Coordination of the bound Ca^{2+} ion by Glu828, Gln831, Asn865 and Asp868, with distances labeled. The density of the $CaCl_2$ structure (gray mesh) is contoured at $\sigma = 5$ and overlaid with the difference density (blue mesh) between the $CaCl_2$ structure and the EDTA structure, contoured at $\sigma = 14$. (B) Ca^{2+} binding site within the S1-S4 domain in the absence of bound ion. (C) The same site with a bound ion. Side chains are labeled. The bound Ca^{2+} ion is shown in magenta.

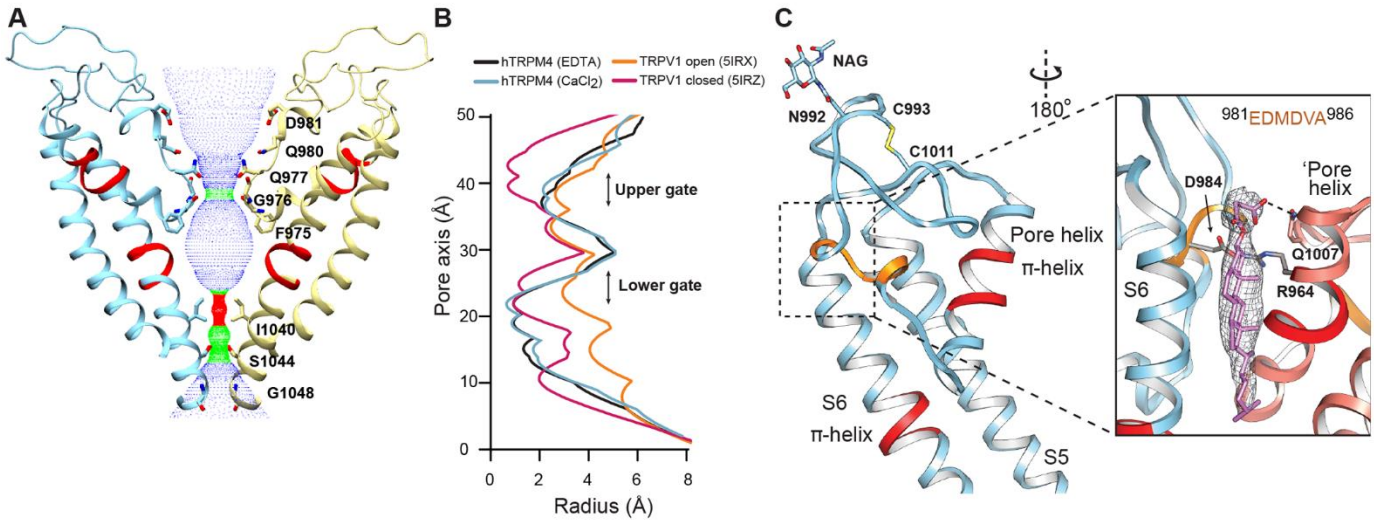


Fig. 3. Ion permeation pore. (A) The solvent-accessible pathway along the ion permeation pore represented by dots between two opposing monomers shown in ribbons colored yellow and blue. Residues aligning the pathway are shown in sticks and labeled. (B) The pore radius of the hTRM4 EDTA (black) and $CaCl_2$ (blue) structures overlaid with the pore radius of TRPV1 in its closed (purple) and open (orange) states. (C) A close-up view of the pore helix and pore loop. The two single-turn π-helices are marked with red in the middle of S6 and the pore helix. The putative selectivity filter is marked with orange. The insert represents the density of a bound lipid, fitted with the atomic model of CHS, which forms a tripartite complex with S1 from one monomer and the pore helix of the neighboring monomer ('Pore helix).

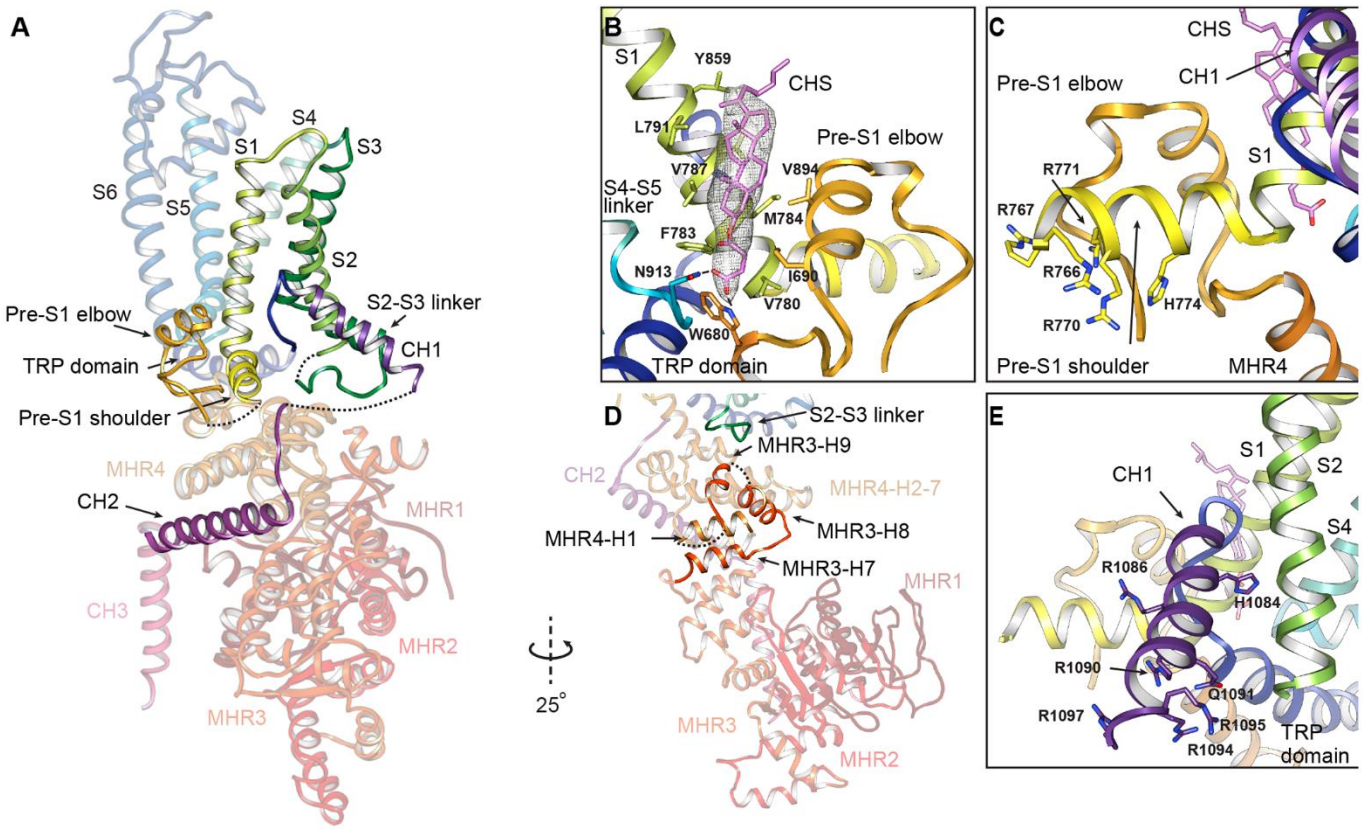


Fig. 4. Membrane embedded helical segments surround the S1-S4 domain. (**A**) Side view of the hTRPM4 monomer with membrane embedded helical segments labeled and in solid color. (**B**) The pre-S1 elbow (gold) and pre-S1 shoulder (yellow). Lipid density between the pre-S1 elbow and S1 is modeled with CHS. Side chains interacting with the CHS are shown as sticks. (**C**) Same as (B), viewed from the opposite direction. (**D**) Cytoplasmic domain MHR3 interacts with the S2-S3 linker. (**E**) The CH1 helix after the TRP domain. Charged residues in the pre-S1 shoulder (C) and CH1 (E) are shown in sticks. Missing links are shown in dashed lines in (A) and (D).

GUT IMMUNITY

CX3CR1$^+$ mononuclear phagocytes control immunity to intestinal fungi

Irina Leonardi,[1,2] Xin Li,[1,2] Alexa Semon,[1,2] Dalin Li,[3] Itai Doron,[1,2] Gregory Putzel,[2] Agnieszka Bar,[1,2] Daniel Prieto,[4] Maria Rescigno,[5] Dermot P. B. McGovern,[3] Jesus Pla,[4] Iliyan D. Iliev[1,2,6]*

Intestinal fungi are an important component of the microbiota, and recent studies have unveiled their potential in modulating host immune homeostasis and inflammatory disease. Nonetheless, the mechanisms governing immunity to gut fungal communities (mycobiota) remain unknown. We identified CX3CR1$^+$ mononuclear phagocytes (MNPs) as being essential for the initiation of innate and adaptive immune responses to intestinal fungi. CX3CR1$^+$ MNPs express antifungal receptors and activate antifungal responses in a Syk-dependent manner. Genetic ablation of CX3CR1$^+$ MNPs in mice led to changes in gut fungal communities and to severe colitis that was rescued by antifungal treatment. In Crohn's disease patients, a missense mutation in the gene encoding CX3CR1 was identified and found to be associated with impaired antifungal responses. These results unravel a role of CX3CR1$^+$ MNPs in mediating interactions between intestinal mycobiota and host immunity at steady state and during inflammatory disease.

Extensive studies on intestinal bacteria have demonstrated that alterations in the microbiome have a dramatic impact on host immunity and contribute to several diseases of inflammatory origin. Fungi are present in the mammalian intestine (*1*–*5*), yet little is known about their ability to influence immune homeostasis. Recent advances in deep sequencing technologies have redefined our understanding of fungal communities (mycobiota) colonizing mammalian barrier surfaces (*2*). Intestinal fungal dysbiosis has been shown to influence colitis, alcoholic liver disease, and allergic lung disease (*3*–*6*), providing evidence for its potential to influence both local and distal inflammation. Moreover, serum antibodies against *Saccharomyces cerevisiae* mannan (ASCA) are elevated in several inflammatory diseases, including Crohn's disease (CD) (*7*–*9*). Systemic ASCA can develop in response to intestinal fungi (*3*, *7*), providing a possible link between the gut mycobiota and host immunity. Despite the identification of receptors involved in the recognition and immunity to intestinal fungi (*3*, *10*), the cell subsets that initiate and regulate mucosal immune responses to the mycobiota remain unknown.

In the intestinal lamina propria (LP), several subsets of phagocytes respond to bacterial infections or to fluctuations in the commensal bacterial communities (*11*–*13*). Among those, mononuclear phagocytes (MNPs), expressing the fractalkine receptor CX3CR1 (CX3CR1$^+$ MNPs), and subsets of dendritic cells (DCs) marked by differential expression of the integrins CD11b and CD103 can initiate immunity and prime T helper 17 (T_H17) responses to both commensal and pathogenic bacteria in the gut (*11*, *12*, *14*). Despite their well-described ability to respond to gut bacteria, their role in mucosal immunity to gut fungi remains unknown.

To assess the in vivo ability of gut resident phagocytes to respond to fungi, we colonized mice with the opportunistic human fungal commensal *Candida albicans* and analyzed the changes in the surface expression of the costimulatory molecules. We found that colonization with *C. albicans* altered the surface expression of CD40 and CD86 among CX3CR1$^+$ MNPs but not among the other subsets (fig. S1, A and B). We thus assessed the ability of CX3CR1$^+$ MNPs to recognize intestinal fungi. We purified CX3CR1$^+$ MNPs from the intestinal LP and compared their RNA-sequencing (RNA-seq) expression profile to those of CD11b$^+$

[1]Gastroenterology and Hepatology Division, Joan and Sanford I. Weill Department of Medicine, Weill Cornell Medicine, New York, NY 10021, USA. [2]The Jill Roberts Institute for Research in Inflammatory Bowel Disease, Weill Cornell Medicine, New York, NY 10021, USA. [3]The F. Widjaja Inflammatory Bowel and Immunobiology Research Institute, Cedars-Sinai Medical Center, Los Angeles, CA 90048, USA. [4]Faculty of Pharmacy, Department of Microbiology II, Universidad Complutense de Madrid, 28040 Madrid, Spain. [5]Department of Experimental Oncology, European Institute of Oncology, I-20141 Milan, Italy. [6]Department of Microbiology and Immunology, Weill Cornell Medicine, New York, NY 10065, USA.
*Corresponding author. Email: iliev@med.cornell.edu

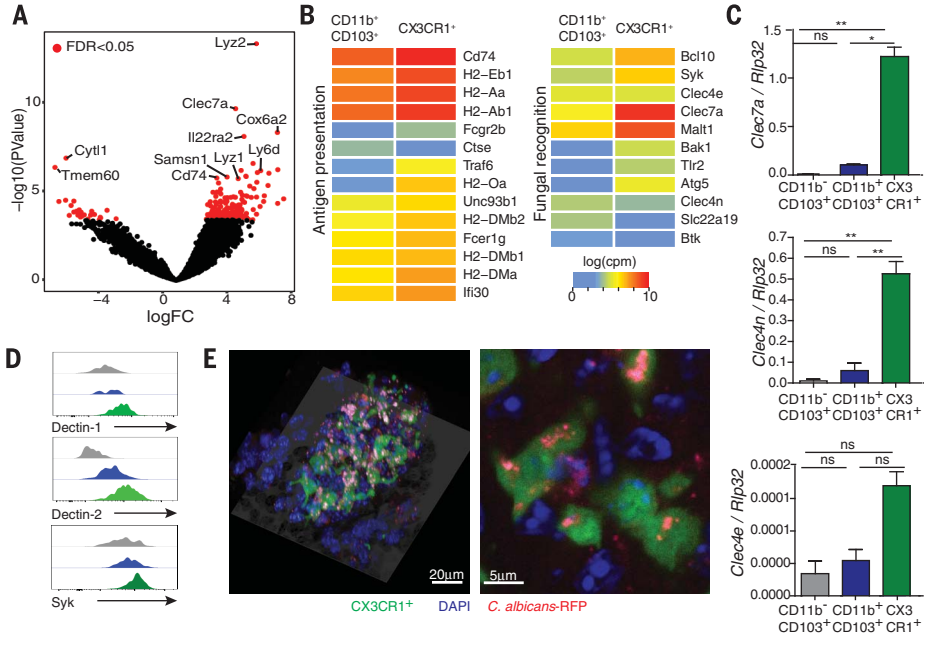


Fig. 1. CX3CR1 mononuclear cells express antifungal receptors and recognize fungi in the intestine. (**A**) RNA-seq analysis was performed on sorted CD11b$^+$ CD103$^+$ dendritic cells and CX3CR1$^+$ mononuclear phagocytes. Shown is a volcano plot of *P* value versus fold-change (FC) comparing gene expression in the two cell subsets; red dots indicate a false discovery rate (FDR) of <0.05. (**B**) Logarithmic count per million [log(cpm)] normalization of genes involved in antigen presentation (left) or fungal recognition (right). (**C**) The expression of antifungal CLRs was confirmed by means of quantitative reverse transcription–PCR. (**D**) Representative flow cytometry histogram of dectin-1, dectin-2, and Syk expression among CD11b$^-$ CD103$^+$, CD11b$^+$ CD103$^+$, and CD11b$^+$ CX3CR1$^+$ cells in colons of WT mice. (**E**) Representative confocal imaging of the intake of *C. albicans*–red fluorescent protein (RFP) (red) by CX3CR1$^+$ MNPs [green; CX3CR1$^+$, 4′,6-diamidino-2-phenylindole$^+$ (DAPI$^+$)] and other cell types (blue; CX3CR1$^-$, DAPI$^+$) in the intestine. Bar graphs represent mean ± SEM of individual mice (n = 4 to 7 mice), representative of at least two independent experiments. *P < 0.05, **P < 0.01, one-way analysis of variance (ANOVA).

CD103$^+$ DCs (Fig. 1, A and B, and fig. S2A) that have been shown to respond to lung fungal infection (*14*, *15*). Whereas both CD11b$^+$CD103$^+$ DCs and CX3CR1$^+$ MNPs expressed genes involved in antigen presentation, CX3CR1$^+$ MNPs showed a higher expression of genes involved in fungal recognition (Fig. 1B, and fig. S2B). Quantitative polymerase chain reaction (PCR) and flow cytometric analysis confirmed that transcripts encoding the fungal C-type lectin receptors (CLRs) dectin-1 (*Clec7a*), dectin-2 (*Clec6a*), and mincle (*Clec4e*) were highly present in CX3CR1$^+$ MNPs (Fig. 1, C and D, and fig. S3, A and B). Further, we examined the in vivo intake of *Candida* by phagocytes in the mouse intestine. Confocal microscopy

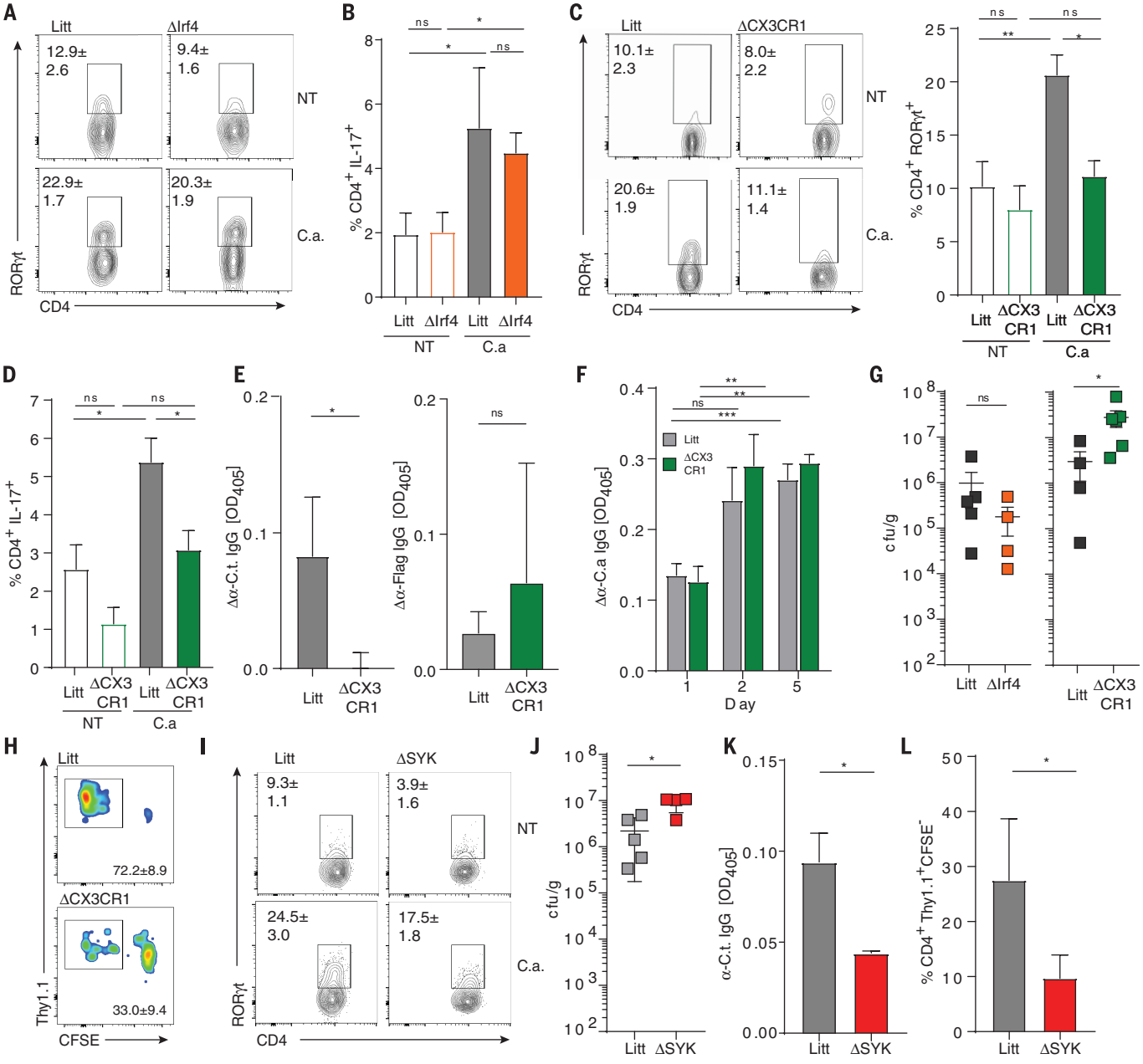


Fig. 2. CX3CR1$^+$ MNPs control gut antifungal immunity. Colonic LP cells were collected from ΔIrf4 mice (orange bars) or littermates (litt, gray bars) fed (C.a) or not (NT) with 5 × 10^7 *C. albicans* at day 10. (**A** and **B**) Expression of RAR-related orphan receptor–γt (RORγt) and interleukin-17 (IL-17) by CD4$^+$ T cells (pooled from two independent experiments). *Cd11c-Cre*$^{+/-}$ *CX3CR1*^{DTR} mice (ΔCX3CR1, green bars) or *Cd11c-Cre*$^{-/-}$ *CX3CR1*^{DTR} littermates (litt, gray bars) were treated with DT and fed with *C. albicans*. (**C** and **D**) RORγt and IL-17 expression by the CD4$^+$ T cells in the colon (pooled from three independent experiments). (**E**) IgG against the commensal *C. tropicalis* and flagellin. (**F**) Systemic IgG responses after intraperitoneal injection with *C. albicans* at day 1, 2, and 5 (pooled from two independent experiments). (**G**) *C. albicans* in the feces of control, ΔIrf4, and ΔCX3CR1 mice at day 10 (dots represent individual mice). ΔCX3CR1 mice and littermates were transferred with purified CD4$^+$ Thy1.1$^+$ OT-II cells, fed *C. albicans*–OVA, and euthanized after 10 days. (**H**) Representative plots of CD4$^+$ Thy1.1$^+$ OT-II cells proliferation in the colon (pooled from three independent experiments). *Cx3cr1-Cre-ERT*$^{+/-}$ *Syk*$^{fl/fl}$ mice (ΔSyk) or littermates (Litt) treated with tamoxifen and fed with *C. albicans*. (**I**) RORγt expression by CD4$^+$ T cells in the colon (pooled from two independent experiments). (**J**) *C. albicans* in the feces at day 10. Dots represent individual mice. (**K**) IgG responses against *C. tropicalis* at day 10 (n = 5 mice per group). (**L**) Quantification of proliferating CFSE$^-$ CD4$^+$ Thy1.1$^+$ OT-II cells in the colon (pooled from two independent experiments). Data are presented as mean ± SEM; *P < 0.05, **P < 0.01, ***P < 0.001; Mann-Whitney Test in (E), (G), (J), (K), and (L) and one-way ANOVA in (B), (C), (D), and (F).

examination revealed that *Candida* was efficiently recognized by intestinal phagocytes in vivo, with over 80% of all CX3CR1^{+} MNPs intaking *Candida* (Fig. 1E; fig. S3, C and D; and movie S1). These results indicate that gut-resident CX3CR1^{+} MNPs are equipped to efficiently recognize and respond to intestinal fungi in vivo.

Both CX3CR1^{+} MNPs and CD11b^{+}CD103^{+} DCs have been shown to play a role in the regulation of adaptive immunity to commensal and pathogenic bacteria (*12*, *14*, *16*, *17*). CX3CR1^{+} MNPs are involved in the induction of T_H17 responses to intestinal bacteria (*18*) and are essential for the killing of *Candida* in the kidneys during systemic infection (*19*). Conversely, several studies have suggested a central role for interferon regulatory factor–4 (IRF4)–dependent CD11b^{+}CD103^{+} DCs in intestinal T_H17 cell differentiation, as well as T_H17-induced bacterial and fungal clearance in the lung (*12*, *14*). Further, conventional migratory CD11b^{-} CD103^{+} DCs have been shown to be a cellular entry point to opportunistic pathogens and are absent in Batf3$^{-/-}$ mice (fig S10, A and B) (*20*). To directly assess the role of different phagocytic subsets in the induction of antifungal immune responses, we crossed inducible floxed *Irf4*$^{fl/fl}$ mice or inducible floxed *Cx3cr1*^{DTR} mice (*11*) with transgenic *Cd11c-Cre* mice. The first strategy allowed the specific ablation of *Irf4* in DCs, leading to the loss of intestinal CD11b^{+}CD103^{+} DCs (referred as ΔIrf4 mice) but intact CX3CR1^{+} MNPs (fig. S4, A and B) (*14*). The second strategy allowed for the selective depletion of intestinal CX3CR1^{+} MNPs upon administration of diphtheria toxin (DT; mice referred as ΔCX3CR1), without affecting CD11b^{+} CD103^{+} DCs (fig. S4, C to E) (*11*).

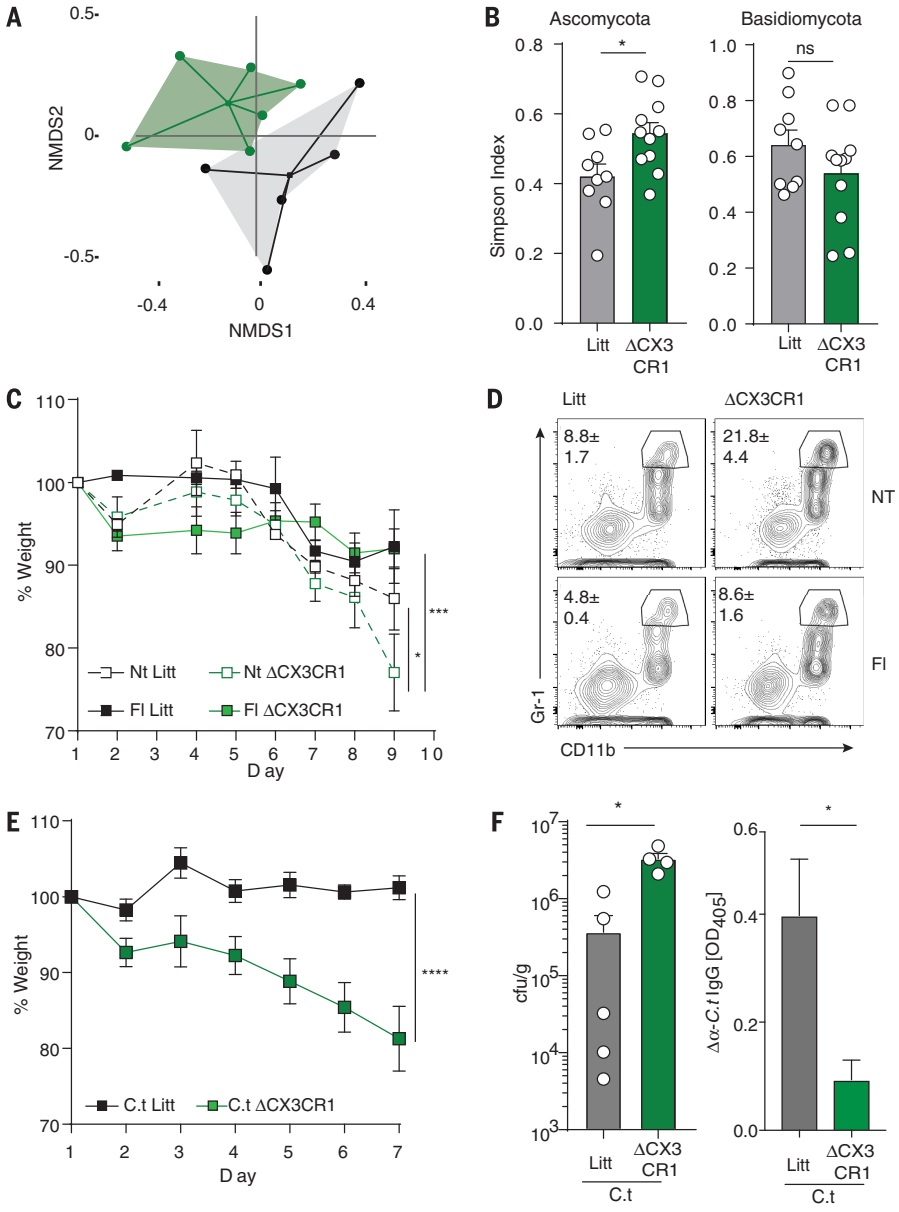


Fig. 3. Depletion of CX3CR1^{+} MNP affects gut mycobiota and results in exacerbated intestinal disease. Feces from ΔCX3CR1 or WT littermates (Litt) mice were collected 7 days after administration of the first DT dose. (**A**) Nonmetric multidimensional scaling (NMDS) plot of distance ordination based on Bray-Curtis dissimilarities in the colon for fungal OTUs (WT n = 5 mice; ΔCX3CR1 n = 6 mice). (**B**) α diversity (Simpson diversity index) among the Ascomycota (left) and Basidiomycota (right) phyla. Data are pooled from two independent experiments and show mean ± SEM. Each circle denotes one mouse. (**C**) Weight change during DSS colitis in ΔCX3CR1 mice or control littermates after short-term treatment with fluconazole (Fl) or no treatment (NT) (pooled from two independent experiments). (**D**) Representative plots of neutrophil infiltration (CD11b^{+} *Gr-1*^{high}) in the colon after DSS administration. (**E**) Weight change during DSS colitis in ΔCX3CR1 mice or control littermates fed with *C. tropicalis* (C.t). Data show mean ± SEM (litt n = 5 mice; ΔCX3CR1 n = 5 mice). (**F**) *C. tropicalis* colony-forming units (cfu) per gram in the feces of ΔCX3CR1 or control littermates at day 7. Dots represent individual mice, mean ± SEM. Systemic IgG responses against *C. tropicalis* were assessed by means of enzyme-linked immunosorbent assay (ELISA). Statistical analysis: *P < 0.05, **P < 0.01, ***P < 0.001; Mann-Whitney test in (B) and (F) and two-way ANOVA in (C) and (E).

Besides T_H17 responses, the development of systemic ASCA during intestinal inflammation is another hallmark of adaptive immunity activation in response to intestinal fungi (*3*, *7*). In addition to *S. cerevisiae*, *Candida* can act as an immunogen for ASCA production (*7*). Thus, we assessed whether lack of CX3CR1^{+} MNPs also alters the production of systemic immunoglobulin G (IgG) antibodies against the commensal fungus *C. tropicalis* (*3*) during steady state (fig. S8D). We found that CX3CR1^{+} MNPs depletion reduced IgG-antibody responses against *C. tropicalis* without affecting the response against the commensal bacterial antigen flagellin (Fig. 2E). To assess whether the induction of antifungal IgG is dependent on

recognition of luminal fungal antigens by CX3CR1$^+$ MNPs, we compared the induction of systemic IgG after intestinal colonization or systemic infection with *C. albicans*. Although both approaches led to the generation of systemic anti–*C. albicans* IgG in wild-type (WT) littermates, depletion of CX3CR1$^+$ MNPs led to a significant decrease of antibody production against intestinal *C. albicans* (fig. S9) without affecting the antibody production upon systemic infection (Fig. 2F). This suggests that the defect in antifungal antibody production in ΔCX3CR1 mice is specific to the gut. By contrast, neither the depletion of CD11b$^+$CD103$^+$ DCs nor the lack of CD11b$^-$CD103$^+$ DCs influenced systemic production of antibody to *Candida* (figs. S9 and S10K). Consistent with the decreased antifungal responses, *C. albicans* burdens are increased in ΔCX3CR1 mice but not in ΔIrf4 and Batf3$^{-/-}$ mice (Fig. 2G and fig. S10L).

To further assess the role of CX3CR1$^+$ MNPs in the induction of responses to antigens delivered by intestinal fungi, we fed ΔCX3CR1 mice and control littermates a recombinant *C. albicans* strain expressing a model major histocompatibility complex II–restricted OT-II peptide (*Ca*-OVA) (fig. S11) and adoptively transferred mice with carboxyfluorescein diacetate succinimidyl ester (CFSE)–labeled CD4$^+$ Thy1.1$^+$ OT-II T cells (fig. S12, A and B). CX3CR1$^+$ MNPs depletion lead to decreased antigen-specific proliferation of T_H17 cells in response to *C. albicans* in both the colon and the mLN (Fig. 2H and fig. S12, C to E). CX3CR1$^+$ MNPs might either directly transport fungal antigens to the mLNs for T cell priming or pass the antigens to other migratory phagocytes, although both scenarios are possible (*11*, *24*). Nevertheless, those data demonstrate that CX3CR1$^+$ MNPs play an important role in the induction of T_H17 and antibody responses to intestinal fungi, whereas CD11b$^+$ CD103$^+$ DCs and CD11b$^-$ CD103$^+$ DCs are not crucial.

Spleen tyrosine kinase (Syk) is crucial for the initiation of the signaling cascade downstream of several CLRs (fig. S2) and is highly expressed in CX3CR1$^+$ MNPs (Fig. 1, B and D, and fig S3B). We thus generated ΔSyk CX3CR1 mice (*Syk*$^{fl/fl}$ x *Cx3cr1-cre/ERT2*$^{+/-}$) to selectively delete Syk in CX3CR1$^+$ cells (fig. S13). Consistently, we observed fungal outgrowth but impaired T_H17 and antifungal antibody responses to intestinal *Candida* colonization (Fig. 2, I to L, and figs. S12F and S14), suggesting that intact CLR signaling in CX3CR1$^+$ cells is required to control fungal colonization and to induce effective adaptive immunity to fungi in the gut.

Having determined that CX3CR1$^+$ MNPs recognize intestinal fungi and control gut antifungal immunity, we explored whether these cells play a role in shaping the composition of gut fungal communities. We characterized the microbiota upon CX3CR1$^+$ MNPs depletion using high-throughput internal transcribed spacer 1 and 16*S* sequencing of fungal and bacterial ribosomal DNA. Despite their role in the control of commensal and pathogenic bacteria (*11*, *16*, *25*), ablation of CX3CR1$^+$ MNPs did not affect β and α diversity of the intestinal bacterial community (figs. S15, A to E, and S16A). By contrast, ablation of CX3CR1$^+$ MNPs, but not CD11b$^+$ CD103$^+$ DCs, altered the fungal community composition of these mice as compared with control littermates (Fig. 3A and figs. S15B; S16, B to D; S17; and S18). CX3CR1$^+$ MNPs depletion also led to an increase in fungal α-diversity that was mainly driven by an increased abundance and diversity among the Ascomycota phylum (Fig. 3B and figs. S15, B and E, and S16, C and D). This suggests that CX3CR1$^+$ MNPs play a role in immune surveillance and maintenance of a balanced gut fungal community.

Fungi are present at highest densities in the colon (*3*), where fungal colonization induced a strong T_H17 response (fig. S5A). Thus, we next tested whether ΔCX3CR1 mice are more susceptible to dextran sulfate sodium (DSS)–induced colitis. Consistent with their inability to mount an efficient antifungal response, we found that ΔCX3CR1 mice were more susceptible to DSS colitis as compared with their littermate controls (Fig. 3, C and D, and fig. S19, A and B). Fluconazole targets most *Candida* species and other dimorphic fungi and ameliorates colitis in mice with defects in antifungal immunity when used during the onset of intestinal disease (*2*, *3*, *10*). Because CX3CR1$^+$ MNPs depletion had a strong effect on Ascomycota, we treated ΔCX3CR1 mice with fluconazole. Fluconazole treatment ameliorated intestinal disease in ΔCX3CR1 mice (Fig. 3, C and D, and fig. S19, A and B). Transfer of feces from ΔCX3CR1 mice did not affect the outcome of colitis in ex-germ-free mice (fig. S19, C to H), indicating that ΔCX3CR1 microbiota is not per se pathogenic. Further supplementation with *C. tropicalis*, previously shown to affect intestinal conditions in dectin-1– and dectin-3–deficient mice (*3*, *10*, *26*), or with *C. albicans* led to severe wasting disease, colon shortening, and fungal overgrowth in the intestines of ΔCX3CR1 mice without worsening the disease in littermate controls (Fig. 3E and fig. S20, A and B). Despite the increased disease susceptibility and augmented *Candida* burden, ΔCX3CR1 mice failed to mount a systemic antifungal IgG antibody response (Fig. 3F and fig. S20, C and D), which is consistent with the defective antifungal immunity that we observed during the steady state. These results indicate that CX3CR1$^+$ MNPs play a crucial role in controlling fungal microbiota during intestinal disease.

Having established a role for CX3CR1$^+$ MNPs in the control of gut fungi during intestinal disease, and finding that CX3CR1$^+$ MNPs have the potential to intake other species of mouse and human commensal fungi (Fig. 4A), we explored whether genetic variations of the human *CX3CR1* gene

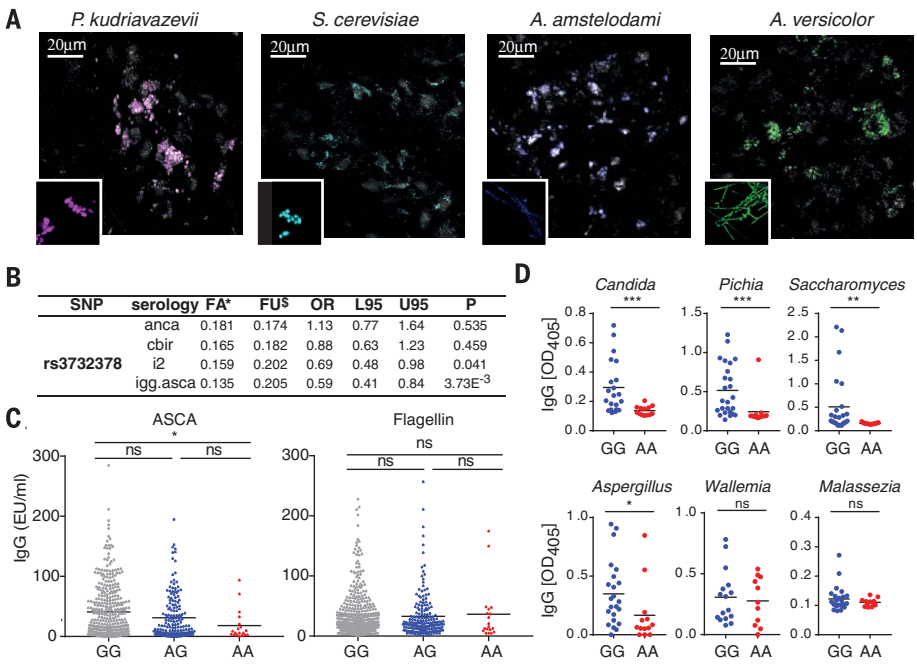


Fig. 4. Polymorphisms in the coding region of the *CX3CR1* gene are associated with decreased antifungal IgG responses in CD patients. (**A**) Representative images of the intake of fungal species (colored) by CX3CR1$^+$ MNPs (gray) in the colon. (**B**) Association between the missense mutation rs3732378 and the systemic serologic markers *anti-neutrophil cytoplasmic antibodies* (anca), flagellin (cbir), *Pseudomonas fluorescens*–associated sequence I-2 (i2), and antibodies to *S. cerevisiae* IgG (igg.asca) among 503 CD patients. FA, frequency affected; FU, frequency unaffected; L95 and U95, lower and upper 95th confidence interval. (**C**) IgG ASCA and antiflagellin (cbir) IgG responses were assessed in the sera from rs3732378 homozygous (AA), heterozygous (AG), and control (GG) CD patients by means of ELISA. Dots represent individual patients, and bars represent mean. (**D**) IgG responses against different commensal fungi. *C. albicans*, *Pichia kudrazevii*, *S. cerevisiae*, *Aspergillus amstellodamii*, *Wallemia sebi*, and *Malassezia restricta* were assessed. Dots represent individual patients, and bars represent mean. Statistical analysis: *P < 0.05, **P < 0.01, ***P < 0.001; Mann-Whitney test in (D) and one-way ANOVA in (C).

affect immunity to fungi in inflammatory bowel disease (IBD). Defects in CX3CR1 have been shown to increase susceptibility of mice and humans to systemic candidiasis (*19*) but not to vulvovaginal and oral candidiasis (*27*). By contrast, the role of CX3CR1 in the initiation of antifungal responses in the gut is unknown. We focused on polymorphisms in the coding region of the *CX3CR1* gene (fig. S21) that have been previously associated with human inflammatory diseases such as arteriosclerosis and coronary artery disease (*28*, *29*). Although none of these polymorphisms were associated with a predisposition to IBD (table S1), we identified a striking association of *CX3CR1* T280M (rs3732378) polymorphism specifically with IgG ASCA positivity in a cohort of 503 CD patients [odds ratio (OR) = 0.59, logistic regression $P = 3.73 \times 10^{-03}$] (Fig. 4B). By contrast, antibodies against bacterial and host antigens previously shown to increase in IBD (*7*) were not affected by this mutation (Fig. 4B and table S2). Because ASCA antibodies are directed against both *S. cerevisiae* and *C. albicans* (*7*), we next assessed whether antifungal antibody responses to common human commensals are also affected by *CX3CR1* T280M. Compared with nonaffected individuals, patients homozygous for *CX3CR1* T280M were severely impaired in their ability to generate systemic IgG recognizing a variety of gut-relevant fungi belonging to the phylum Ascomycota, while producing normal antibody responses against the bacterial antigen flagellin (Fig. 4, C and D, and fig. S22), which is consistent with the hypothesis that this mutation in *CX3CR1* leads to impaired responses to fungi in the gut.

We identified a specific subset of gut-resident phagocytes—namely, the CX3CR1$^+$ MNPs—that are able to recognize and respond to the gut mycobiota in a Syk-dependent manner. CX3CR1$^+$ MNPs can influence adaptive immune responses to gut fungi and control the mycobiota during experimental colitis. In humans, we found that a *CX3CR1* polymorphism is strongly associated with a decrease of antifungal antibody responses in CD patients. *CX3CR1* T280M is a common polymorphism [23.3% heterozygous and 4.4% homozygous (*30*)] and has been previously associated with extraintestinal inflammatory diseases (*28*, *29*). Conceivably, gut mycobiota and *CX3CR1*-dependent immune responses might further contribute to extraintestinal manifestations of inflammatory diseases. In support of this notion, antifungal antibodies are increased in patients with alcoholic liver disease, spondyloarthritis, Graves' disease, and systemic lupus erythematosus (*6*–*9*). Altogether, our findings provide evidence for the influence of gut fungal communities on both local and systemic immunity, which is mediated by the recognition of intestinal fungi by CX3CR1$^+$ MNPs.

REFERENCES AND NOTES

1. A. J. Scupham *et al.*, *Appl. Environ. Microbiol.* **72**, 793–801 (2006).
2. I. D. Iliev, I. Leonardi, *Nat. Rev. Immunol.* **17**, 635–646 (2017).
3. I. D. Iliev *et al.*, *Science* **336**, 1314–1317 (2012).
4. H. Sokol *et al.*, *Gut* **66**, 1039–1048 (2017).
5. M. L. Wheeler *et al.*, *Cell Host Microbe* **19**, 865–873 (2016).
6. A. M. Yang *et al.*, *J. Clin. Invest.* **127**, 2829–2841 (2017).
7. A. Standaert-Vitse *et al.*, *Gastroenterology* **130**, 1764–1775 (2006).
8. A. Mankaï *et al.*, *Endocr. Res.* **38**, 98–104 (2013).
9. L. Riente *et al.*, *J. Rheumatol.* **31**, 920–924 (2004).
10. T. Wang *et al.*, *PLOS Pathog.* **12**, e1005662 (2016).
11. G. E. Diehl *et al.*, *Nature* **494**, 116–120 (2013).
12. E. K. Persson *et al.*, *Immunity* **38**, 958–969 (2013).
13. M. Rescigno *et al.*, *Nat. Immunol.* **2**, 361–367 (2001).
14. A. Schlitzer *et al.*, *Immunity* **38**, 970–983 (2013).
15. M. A. Kinnebrew *et al.*, *Immunity* **36**, 276–287 (2012).
16. J. H. Niess *et al.*, *Science* **307**, 254–258 (2005).
17. T. Aychek *et al.*, *Nat. Commun.* **6**, 6525 (2015).
18. C. Panea *et al.*, *Cell Reports* **12**, 1314–1324 (2015).
19. M. S. Lionakis *et al.*, *J. Clin. Invest.* **123**, 5035–5051 (2013).
20. B. T. Edelson *et al.*, *Immunity* **35**, 236–248 (2011).
21. H. R. Conti *et al.*, *J. Exp. Med.* **206**, 299–311 (2009).
22. K. Trautwein-Weidner *et al.*, *PLOS Pathog.* **11**, e1005164 (2015).
23. K. Atarashi *et al.*, *Cell* **163**, 367–380 (2015).
24. E. Mazzini, L. Massimiliano, G. Penna, M. Rescigno, *Immunity* **40**, 248–261 (2014).
25. O. Medina-Contreras *et al.*, *J. Clin. Invest.* **121**, 4787–4795 (2011).
26. C. Tang *et al.*, *Cell Host Microbe* **18**, 183–197 (2015).
27. T. J. Break *et al.*, *Infect. Immun.* **83**, 958–965 (2015).
28. D. Moatti *et al.*, *Blood* **97**, 1925–1928 (2001).
29. D. H. McDermott *et al.*, *Circ. Res.* **89**, 401–407 (2001).
30. S. Brand *et al.*, *Am. J. Gastroenterol.* **101**, 99–106 (2006).

ACKNOWLEDGMENTS

This work was funded by the U.S. National Institutes of Health (grants DK113136, DK098310, and AI123819 to I.D.I; P01 DK046763 and U01 DK062413 to D.P.B.M.), Infect-ERA FUNCOMPATH (PCIN-2014-052 to J.P.), Ministerio de Economía y Competitividad (BIO2015-6477-P to J.P.), Swiss National Science Foundation (fellowship P2ZHP3_164850 to I.L), Kenneth Rainin Foundation (Innovator and Breakthrough awards to I.D.I), and support from the Jill Roberts Institute for Research in IBD. The data presented in this study are tabulated in the main text and supplementary materials. Microbiota sequencing data are deposited in National Center for Biotechnology Information (NCBI) Sequence Read Archive (SRA; www.ncbi.nlm.nih.gov/Traces/sra, run numbers SRP124782, SRP124783, SRP124736, and SRP124742). RNA-seq data are deposited in the NCBI Gene Expression Omnibus under accession no. GSE106594.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/359/6372/232/suppl/DC1
Materials and Methods
Figs. S1 to S22
Tables S1 and S2
References (*31*–*51*)
Movie S1

22 June 2017; resubmitted 15 November 2017
Accepted 9 December 2017
10.1126/science.aao1503

Cite as: Y. Yin *et al.*, *Science* 10.1126/science.aan4325 (2017).

Structure of the cold- and menthol-sensing ion channel TRPM8

Ying Yin,[1]* Mengyu Wu,[2]* Lejla Zubcevic,[1] William F. Borschel,[1] Gabriel C. Lander,[2]† Seok-Yong Lee[1]†

[1]Department of Biochemistry, Duke University School of Medicine, Durham, NC 27710, USA. [2]Department of Integrative Structural and Computational Biology, The Scripps Research Institute, La Jolla, CA 92037, USA.

*These authors contributed equally to this work. †Corresponding author. Email: seok-yong.lee@duke.edu (S.-Y.L.); glander@scripps.edu (G.C.L.)

Transient receptor potential melastatin (TRPM) cation channels are polymodal sensors that are involved in a variety of physiological processes. Within the TRPM family, member 8 (TRPM8) is the primary cold- and menthol-sensor in humans. We determined the cryo-electron microscopy structure of the full-length TRPM8 from the collared flycatcher at an overall resolution of ~4.1 Å. Our TRPM8 structure reveals a three-layered architecture. The amino-terminal domain with a fold distinct among known TRP structures, together with the carboxyl-terminal region, forms a large two-layered cytosolic ring that extensively interacts with the transmembrane channel layer. The structure suggests that the menthol binding site is located within the voltage-sensor-like domain and thus provides a structural glimpse of the design principle of the molecular transducer for cold and menthol sensation.

The transient receptor potential melastatin (TRPM) family, part of the TRP protein superfamily, is composed of eight members, TRPM1-TRPM8, and is involved in various processes including temperature sensing, ion homeostasis, and redox sensing (*1*, *2*). The TRPM8 channel cDNA was cloned and characterized as the long-sought-after cold- and menthol-sensor (*3*, *4*). Studies of TRPM8-deficient mice showed that the channel is required for environmental cold sensing (*5*–*7*), and that it is the principal mediator of menthol-induced analgesia of acute and inflammatory pain (*8*). Therefore, TRPM8 is a therapeutic target for treatments of cold-related pain, chronic pain, and migraine (*9*, *10*).

In addition to cold and menthol, TRPM8 is sensitive to voltage and phosphatidylinositol-4,5-biphosphate (PIP_2) (*11*, *12*). The molecular landscape of channel sensitization is shaped by the interplay between these four factors, suggesting that TRPM8 is a polymodal sensor that can integrate multiple chemical and physical stimuli into cellular signaling (*11*, *13*). Several thermodynamic models have been put forth to address the mechanism of polymodal sensing by TRPM8 (*14*–*16*), but our mechanistic understanding of polymodal sensing by TRPM8 remains limited.

All members of the TRPM family contain a large N-terminal region (~700 amino acids) that comprises four regions of high homology (Melastatin Homology Regions, MHRs) that do not contain any known structural motifs (*1*). These MHR repeat domains appear to be important for channel assembly and trafficking, but their functional roles are not known (*17*). TRPM8 was also predicted to contain a putative C-terminal coiled coil that is important for channel assembly, trafficking, and function (*17*, *18*).

Extensive electrophysiological and biochemical studies have identified residues involved in ligand binding, and homology models have attempted to provide a structural context for their locations (*19*–*22*). However, in the absence of a structure these predictions were speculative, and the mechanism of menthol-dependent TRPM8 gating remained unclear.

We conducted structural studies of full-length TRPM8 from the collared flycatcher *Ficedula albicollis* ($TRPM8_{FA}$) using cryo-electron microscopy (cryo-EM). $TRPM8_{FA}$ is highly homologous to human and chicken TRPM8 (83% and 94% sequence identity, respectively; fig. S1) which are cold- and menthol-sensitive (*14*, *20*). To prevent proteolysis, we introduced three mutations into $TRPM8_{FA}$ (Phe535Ala, Tyr538Asp, and Tyr539Asp). When expressed in HEK293 cells, both the wild type and the mutant channels exhibited similar menthol-evoked currents and calcium influx, indicating that the mutations do not appreciably affect $TRPM8_{FA}$ function (fig. S2). Notably, the EC_{50} value and the current-voltage relationship of $TRPM8_{FA}$ for menthol are comparable to those of human and chicken TRPM8 (*14*, *20*).

$TRPM8_{FA}$ was frozen in vitreous ice and imaged using a Titan Krios transmission electron microscope with a K2 Summit direct electron detector (See Methods and fig. S3). The final 3D reconstruction was resolved to an overall resolution of ~4.1 Å, with local resolutions ranging from ~3.8 Å at the core to ~8 Å at the periphery (Fig. 1 and fig. S3). The quality of the map allowed for de novo model building (Methods) of 75% of the $TRPM8_{FA}$ polypeptide (fig. S4 and table S1). The final model for $TRPM8_{FA}$ contains amino acids 122-1100 with multiple loops missing. Several regions (three

β-strands in pre-MHR, C-terminal domain helix 1, and the C-terminal coiled coil) are built as polyalanine (Methods).

The $TRPM8_{FA}$ homotetramer is shaped like a three-layered stack of tetragonal bricks with dimensions of approximately 110 Å by 110 Å by 125 Å (Fig. 1, A and B). The top layer comprises the transmembrane channel domain (TMD) and the lower two layers comprise the cytosolic domain (CD). Each protomer contains a large N-terminal region consisting of MHR1-4, a transmembrane channel region, and a C-terminal region (Fig. 1, C and D).

The TMs of $TRPM8_{FA}$ assume a fold similar to that of TRPV1/TRPV2, with a voltage-sensor-like domain (VSLD) made up of transmembrane helical segments S1-S4, a pore domain formed by S5 and S6 helices, and one pore helix (*23*, *24*) (Figs. 1 and 2). Similar to previously determined TRPV structures, the $TRPM8_{FA}$ TMD exhibits a domain-swapped arrangement, with the VSLD of one protomer interacting with the pore domain of the neighboring protomer. However, the TMD of $TRPM8_{FA}$ has three features that are distinct from other TRP ion channel architectures. First, relative to the structure of apo TRPV1, the pore helix of $TRPM8_{FA}$ is positioned ~12 Å further away from the central axis, tilted by ~8°, and translated toward the extracellular side by ~5 Å (Fig. 2, A and B). The putative selectivity filter is poorly resolved in the cryo-EM density map, which prevented accurate model building in this region (fig. S5). Sequence comparison with TRPV1 shows that TRPM8 lacks the turret connecting S5 and the pore helix, instead containing a much longer pore loop (fig. S5). Given the differences in the pore helix position and the sequence surrounding the selectivity filter, we speculate that the TRPM8 selectivity filter adopts an organization that is distinct from TRPV1. The second distinguishing feature of $TRPM8_{FA}$ is the absence of non-α-helical elements (e.g., 3_{10} or π helices) in its TMs, which in other TRP channels were proposed to provide helical bending points important for channel gating (*23*, *25*). In $TRPM8_{FA}$, a straight α-helical S4 is connected to α-helical S5 via a sharp turn induced by a conserved proline residue (Fig. 2, C to E). The lack of a bending point in S5 calls to question whether $TRPM8_{FA}$ possesses an S4-S5 linker, which is the structural element critical for vanilloid-dependent TRPV1 gating (*26*); however, it is possible that a transition from α to π helix in the $TRPM8_{FA}$ S5 may occur during gating, as was suggested for TRPV2 (*23*). Despite the absence of an obvious S4-S5 linker, $TRPM8_{FA}$ nonetheless forms a domain-swapped tetramer, which is in stark contrast to the calcium-activated K^+ channel Slo1, where a similarly short S4-S5 linker prevents formation of a domain-swapped tetramer (*27*). The overlay of $TRPM8_{FA}$ and TRPV1 protomers reveals that the C-terminal part of $TRPM8_{FA}$ S4 is longer and straight such that it can connect with S5 to achieve a domain-swapped configuration. Furthermore, the sharp turn that connects the α helical S4 and S5 results in a tilt of S5, giving rise to the distinct position of the pore helix as compared to TRPV1 (Fig. 2, B and C). Thirdly, while TRPV channels have a cytosolic pre-S1 helix, $TRPM8_{FA}$ contains three additional helices between S1 and the cytosolic pre-S1 helix in the putative membrane region: a helix-turn-helix (HTH) followed by an interfacial helix that connects to S1. We term this structural motif, including the cytosolic pre-S1 helix, the "pre-S1 domain".

The C-terminal portions of the four α helical S6s are in proximity of each other, providing a constriction point termed the S6 gate (Fig. 2G and fig. S5). The Leu^{973} residues on each protomer form a tight constriction, with diagonally opposing residues distanced 5.3 Å apart, suggesting that our structure represents a non-conductive conformation (Fig. 2, G, H, and I).

The cytoplasmic domain (CD) is composed of the N-terminal domain (NTD) and the C-terminal domain (CTD) (Fig. 3A). The NTD is composed of four MHRs (MHR1-MHR4). MHR1 and MHR2, along with a part of the pre-MHR region, form a single domain (MHR1/2) with an α/β fold (Fig. 3D). Unlike a typical α/β fold of five β strands surrounded by α helices, MHR1/2 contains a large parallel β sheet composed of about ten β strands sandwiched by about four α helices on each side. The pre-MHR and the tip of MHR1/2 are not well resolved in the reconstruction, precluding model building around this region (Fig. 3 and fig. S4). The MHR3 and 4 are mostly made up of helix-turn-helix (HTH) motifs, but unlike the ankyrin repeats in TRPV channels, the HTHs in MHR3 and MHR4 do not contain hairpin-like protruding loops. The HTHs within MHR3 stack in a sequential fashion, giving rise to the angled position of MHR1/2 relative to MHR4, and resulting in the L-shape of the protomer (Figs. 1D and 3, A and C). The MHR4 domain, located below the VSLD, consists of five (HTH4-HTH8) non-sequentially stacked HTH motifs (Fig. 3B). The CTD is composed of three helices that are located C-terminal to the TRP domain (Fig. 3A). The first CTD helix (CTDH1) extends from the TRP domain. It then connects to the second CTD helix (CTDH2) via a long loop, which is positioned below HTH6 and HTH7 of MHR4, and runs almost parallel with the TRP domain (Fig. 3, A and B). The C-terminal region of the CTDH2 points toward the cytoplasmic cavity and connects to a tetrameric coiled coil (CC) at the central axis (Fig. 3, A and B). The limited resolution of the cryo-EM density around the CC prevents accurate modeling of this region, but the relatively short (~25 amino acids) and parallel architecture of the $TRPM8_{FA}$ CC contrasts the long (>50 amino acids), anti-parallel architecture observed in the CC fragment structure of TRPM7 (*28*).

In contrast with TRPV1, $TRPM8_{FA}$ makes extensive intra- and inter-subunit interactions (Fig. 3E). First, while the in-

teraction between the TMD and CD is primarily mediated by the interfacial TRP domain in TRPV1, in $TRPM8_{FA}$ the pre-S1 domain establishes additional TMD-CD interactions through contacts with the tip of HTH7b in MHR4 from the adjacent subunit (Fig. 3F and fig. S6, A and B). Second, the CTDH2 runs parallel to the TRP domain, but is translated ~29 Å toward the cytosolic side (Fig. 3G), positioned beneath MHR4 and contacting the MHR1/2 of the adjacent subunit (Fig. 3G and fig. S6, C and D). The CTDH2 is therefore in contact with both the top and the bottom layers of the cytoplasmic ring. Based on its position at this inter-layer and inter-subunit nexus, the CTDH2 might have a role in channel gating as well as cytoplasmic ring assembly. Third, in the bottom layer of the cytoplasmic ring, the tips of two HTH motifs (HTH2a and HTH2b) from MHR3 and the loop of MHR1/2 from the neighboring protomer establish an inter-subunit network of interactions (Fig. 3H and fig. S6D). These extensive interfacial interactions are the distinguishing features of TRPM8, which may be important for either cold- or menthol-dependent channel gating.

Unlike TRPV channels, TRPM8 contains arginine residues in S4 of the VSLD (Fig. 4A). Arg^{842} in S4 and Lys^{856} in S5 (Arg^{841} and Lys^{855} in $TRPM8_{FA}$, respectively), contribute to the voltage-dependence in human TRPM8 (*13*). In $TRPM8_{FA}$, many aromatic and aliphatic residues within the VSLD form a large hydrophobic "seal" (Fig. 4A) between the extracellular side and the middle of the membrane. Between this hydrophobic seal and the TRP domain, there is a large cavity we term the VSLD cavity, which is also present in TRPV channels (*23*, *26*). In the $TRPM8_{FA}$ structure, Arg^{841} points toward the center of the VSLD cavity, where three acidic residues form a negatively charged cluster. Although we are unable to unambiguously assign the interaction pairs between the charged residues, this arrangement is more reminiscent of a canonical voltage sensor than the VSLD of TRPV1 (Fig. 4A) (*24*, *29*). We also observed that Lys^{855} is located at the beginning of S5 outside the VSLD, but it is unclear how this residue contributes to voltage sensing. Despite the similarities in the positions of charged residues in the $TRPM8_{FA}$ VSLD and the canonical voltage sensor, we postulate that the degree and the type of voltage-sensing motion in TRPM8 is different from canonical voltage-gated ion channels due to the large size of the hydrophobic seal and the fact that a small gating charge (~1 e_o) is associated with voltage gating of TRPM8 (*13*).

Many studies have been conducted to elucidate the mechanism of TRPM8 activation by ligand binding. It has been reported that Arg^{842} (Arg^{841} in $TRPM8_{FA}$) also affects menthol- and cold-dependent activation of human TRPM8. Based on [^{3}H]menthol binding studies, this residue appears to interact with menthol (*13*). Residues Tyr^{745} and Tyr^{1005} (Tyr^{1004} in $TRPM8_{FA}$) were shown to be involved in menthol binding, but not cold sensing (*19*). Tyr^{745} was also found to be critical for binding of the inhibitor SKF96365 (*21*), further indicating that this residue is central to ligand-dependent gating in TRPM8. In addition, studies have identified Asn^{799}, Asp^{802}, and Gly^{805} as important for icilin binding (*20*). It was predicted that these ligand-binding sites were all located at the membrane-facing region of S2-S3 (*20*–*22*). We can now place these functional studies in the proper structural context. Notably, residue Tyr^{745} is located at the middle of S1, directed toward the center of the VSLD cavity (Fig. 4B). This contrasts with its previously predicted location in the membrane-facing side of S2. Furthermore, all residues implicated in menthol binding (Tyr^{745}, Arg^{841}, and Tyr^{1004}) are located in the VSLD cavity (Fig. 4B). We propose that the VSLD cavity is the binding site for menthol and menthol-like molecules in TRPM8. Interestingly, the corresponding cavity in TRPV channels has been implicated in lipid binding (*23*, *26*) and modulation of channel gating (*23*). We propose that menthol binding in the VSLD cavity in TRPM8 might modulate the S6 gate through interactions with the TRP domain.

PIP_2 is necessary for TRPM8 activation, as depletion of PIP_2 has been shown to desensitize the channel (*11*–*13*). Furthermore, recent studies have suggested that PIP_2 alone can activate TRPM8 (*11*–*13*). Mutagenesis studies identified Lys^{995}, Arg^{998}, and Arg^{1008} (Lys^{994}, Arg^{997}, and Arg^{1007} in $TRPM8_{FA}$) as important for PIP_2-dependent channel gating (*11*). All of these residues are located in the TRP domain: Arg^{1007} is located on the side of the TRP domain facing the VSLD cavity, and Lys^{994} and Arg^{997} are located on the side of the TRP domain facing the pre-S1 domain and HTH6 in MHR4 from the neighboring subunit (fig. S7). Our structure suggests that the VSLD cavity is unlikely to bind PIP_2, since the net charge of its interior is negative and this cavity is not large enough to accommodate both menthol and PIP_2. The reported effects of mutations to Arg^{1008} may reflect downstream conformational changes following PIP_2 binding. Notably, we found that the interface between the TRP domain, the pre-S1 domain, and MHR4 contains many basic amino acid residues, including the previously identified Lys^{994} and Arg^{997} (fig. S7). We speculate that PIP_2 binds at this interfacial site, where it could modulate the position of the TRP domain to enable a non-desensitized state. The distinct location of this putative PIP_2 binding site compared to TRPV1 and TRPML reflects the diverse effects of PIP_2 on TRP channels (*25*, *26*).

In the context of previous mutagenesis data, our structural analyses suggest that the ligand-dependent gating mechanism of TRPM8 differs substantially from TRPV1. In TRPV1, the binding of vanilloid to the S4b and the S4-S5 linker is thought to induce a "swivel" motion in the TRP domain, which would pull on S6 to open the S6 gate (*23*,

30). We suggest that menthol binds in the VSLD cavity, which is distinct from the vanilloid binding site in TRPV1. Whereas vanilloid-mediated gating involves non-α-helical elements (3_{10} and π helices) and an S4-S5 linker, neither of these structural features appear to be present in the conformational state of our $TRPM8_{FA}$ structure. Thus, we speculate that ligand-induced repositioning of the TRP domain in TRPM8 may directly lead to the opening of the S6 gate.

Our observation of the extensive inter-subunit interactions between the TMD and the top layer of the CD ring leads us to speculate that the gating-related TRP domain motion may also involve the top layer of the CD ring (Fig. 3F and fig. S6B). Furthermore, the CTDH2, with its central position within the CD and its link with the TRP domain, may be important for coupling the movements of the TRP domain to those of the MHR elements and especially MHR4 (Fig. 3G and fig. S6D). In addition, the tetrameric coiled coil located in the bottom layer ring may play a role in regulating the position of CTDH2.

REFERENCES AND NOTES

1. A. Fleig, R. Penner, The TRPM ion channel subfamily: Molecular, biophysical and functional features. *Trends Pharmacol. Sci.* **25**, 633–639 (2004). doi:10.1016/j.tips.2004.10.004 Medline
2. A. A. Farooqi, M. K. Javeed, Z. Javed, A. M. Riaz, S. Mukhtar, S. Minhaj, S. Abbas, S. Bhatti, TRPM channels: Same ballpark, different players, and different rules in immunogenetics. *Immunogenetics* **63**, 773–787 (2011). doi:10.1007/s00251-011-0570-4 Medline
3. D. D. McKemy, W. M. Neuhausser, D. Julius, Identification of a cold receptor reveals a general role for TRP channels in thermosensation. *Nature* **416**, 52–58 (2002). doi:10.1038/nature719 Medline
4. A. M. Peier, A. Moqrich, A. C. Hergarden, A. J. Reeve, D. A. Andersson, G. M. Story, T. J. Earley, I. Dragoni, P. McIntyre, S. Bevan, A. Patapoutian, A TRP channel that senses cold stimuli and menthol. *Cell* **108**, 705–715 (2002). doi:10.1016/S0092-8674(02)00652-9 Medline
5. D. M. Bautista, J. Siemens, J. M. Glazer, P. R. Tsuruda, A. I. Basbaum, C. L. Stucky, S.-E. Jordt, D. Julius, The menthol receptor TRPM8 is the principal detector of environmental cold. *Nature* **448**, 204–208 (2007). doi:10.1038/nature05910 Medline
6. A. Dhaka, A. N. Murray, J. Mathur, T. J. Earley, M. J. Petrus, A. Patapoutian, TRPM8 is required for cold sensation in mice. *Neuron* **54**, 371–378 (2007). doi:10.1016/j.neuron.2007.02.024 Medline
7. R. W. Colburn, M. L. Lubin, D. J. Stone Jr., Y. Wang, D. Lawrence, M. R. D'Andrea, M. R. Brandt, Y. Liu, C. M. Flores, N. Qin, Attenuated cold sensitivity in TRPM8 null mice. *Neuron* **54**, 379–386 (2007). doi:10.1016/j.neuron.2007.04.017 Medline
8. B. Liu, L. Fan, S. Balakrishna, A. Sui, J. B. Morris, S.-E. Jordt, TRPM8 is the principal mediator of menthol-induced analgesia of acute and inflammatory pain. *Pain* **154**, 2169–2177 (2013). doi:10.1016/j.pain.2013.06.043 Medline
9. M. D. Andrews, K. Af Forselles, K. Beaumont, S. R. G. Galan, P. A. Glossop, M. Grenie, A. Jessiman, A. S. Kenyon, G. Lunn, G. Maw, R. M. Owen, D. C. Pryde, D. Roberts, T. D. Tran, Discovery of a Selective TRPM8 Antagonist with Clinical Efficacy in Cold-Related Pain. *ACS Med. Chem. Lett.* **6**, 419–424 (2015). doi:10.1021/ml500479v Medline
10. A. D. Weyer, S. G. Lehto, Development of TRPM8 Antagonists to Treat Chronic Pain and Migraine. *Pharmaceuticals (Basel)* **10**, 37 (2017). doi:10.3390/ph10020037 Medline
11. T. Rohács, C. M. Lopes, I. Michailidis, D. E. Logothetis, PI(4,5)P2 regulates the activation and desensitization of TRPM8 channels through the TRP domain. *Nat. Neurosci.* **8**, 626–634 (2005). doi:10.1038/nn1451 Medline
12. B. Liu, F. Qin, Functional control of cold- and menthol-sensitive TRPM8 ion channels by phosphatidylinositol 4,5-bisphosphate. *J. Neurosci.* **25**, 1674–1681 (2005). doi:10.1523/JNEUROSCI.3632-04.2005 Medline
13. T. Voets, G. Owsianik, A. Janssens, K. Talavera, B. Nilius, TRPM8 voltage sensor mutants reveal a mechanism for integrating thermal and chemical stimuli. *Nat. Chem. Biol.* **3**, 174–182 (2007). doi:10.1038/nchembio862 Medline
14. T. Voets, G. Droogmans, U. Wissenbach, A. Janssens, V. Flockerzi, B. Nilius, The principle of temperature-dependent gating in cold- and heat-sensitive TRP channels. *Nature* **430**, 748–754 (2004). doi:10.1038/nature02732 Medline
15. D. Baez, N. Raddatz, G. Ferreira, C. Gonzalez, R. Latorre, Gating of thermally activated channels. *Curr. Top. Membr.* **74**, 51–87 (2014). doi:10.1016/B978-0-12-800181-3.00003-8 Medline
16. D. E. Clapham, C. Miller, A thermodynamic framework for understanding temperature sensing by transient receptor potential (TRP) channels. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 19492–19497 (2011). doi:10.1073/pnas.1117485108 Medline
17. C. B. Phelps, R. Gaudet, The role of the N terminus and transmembrane domain of TRPM8 in channel localization and tetramerization. *J. Biol. Chem.* **282**, 36474–36480 (2007). doi:10.1074/jbc.M707205200 Medline
18. P. R. Tsuruda, D. Julius, D. L. Minor Jr., Coiled coils direct assembly of a cold-activated TRP channel. *Neuron* **51**, 201–212 (2006). doi:10.1016/j.neuron.2006.06.023 Medline
19. M. Bandell, A. E. Dubin, M. J. Petrus, A. Orth, J. Mathur, S. W. Hwang, A. Patapoutian, High-throughput random mutagenesis screen reveals TRPM8 residues specifically required for activation by menthol. *Nat. Neurosci.* **9**, 493–500 (2006). doi:10.1038/nn1665 Medline
20. H. H. Chuang, W. M. Neuhausser, D. Julius, The super-cooling agent icilin reveals a mechanism of coincidence detection by a temperature-sensitive TRP channel. *Neuron* **43**, 859–869 (2004). doi:10.1016/j.neuron.2004.08.038 Medline
21. A. Malkia, M. Pertusa, G. Fernández-Ballester, A. Ferrer-Montiel, F. Viana, Differential role of the menthol-binding residue Y745 in the antagonism of thermally gated TRPM8 channels. *Mol. Pain* **5**, 62 (2009). doi:10.1186/1744-8069-5-62 Medline
22. A. Pedretti, C. Marconi, I. Bettinelli, G. Vistoli, Comparative modeling of the quaternary structure for the human TRPM8 channel and analysis of its binding features. *Biochim. Biophys. Acta* **1788**, 973–982 (2009). doi:10.1016/j.bbamem.2009.02.007 Medline
23. L. Zubcevic, M. A. Herzik Jr., B. C. Chung, Z. Liu, G. C. Lander, S.-Y. Lee, Cryo-electron microscopy structure of the TRPV2 ion channel. *Nat. Struct. Mol. Biol.* **23**, 180–186 (2016). doi:10.1038/nsmb.3159 Medline
24. M. Liao, E. Cao, D. Julius, Y. Cheng, Structure of the TRPV1 ion channel determined by electron cryo-microscopy. *Nature* **504**, 107–112 (2013). doi:10.1038/nature12822 Medline
25. M. Hirschi, M. A. Herzik Jr., J. Wie, Y. Suo, W. F. Borschel, D. Ren, G. C. Lander, S.-Y. Lee, Cryo-electron microscopy structure of the lysosomal calcium-permeable channel TRPML3. *Nature* **550**, 411–414 (2017). doi:10.1038/nature24055 Medline
26. Y. Gao, E. Cao, D. Julius, Y. Cheng, TRPV1 structures in nanodiscs reveal mechanisms of ligand and lipid action. *Nature* **534**, 347–351 (2016). doi:10.1038/nature17964 Medline
27. X. Tao, R. K. Hite, R. MacKinnon, Cryo-EM structure of the open high-conductance Ca2+-activated K+ channel. *Nature* **541**, 46–51 (2017). doi:10.1038/nature20608 Medline
28. Y. Fujiwara, D. L. Minor Jr., X-ray crystal structure of a TRPM assembly domain reveals an antiparallel four-stranded coiled-coil. *J. Mol. Biol.* **383**, 854–870 (2008). doi:10.1016/j.jmb.2008.08.059 Medline
29. S. B. Long, X. Tao, E. B. Campbell, R. MacKinnon, Atomic structure of a voltage-dependent K+ channel in a lipid membrane-like environment. *Nature* **450**, 376–382 (2007). doi:10.1038/nature06265 Medline
30. E. Cao, M. Liao, Y. Cheng, D. Julius, TRPV1 structures in distinct conformations reveal activation mechanisms. *Nature* **504**, 113–118 (2013). doi:10.1038/nature12823 Medline
31. A. Goehring, C.-H. Lee, K. H. Wang, J. C. Michel, D. P. Claxton, I. Baconguis, T. Althoff, S. Fischer, K. C. Garcia, E. Gouaux, Screening and large-scale expression of membrane proteins in mammalian cells for structural studies. *Nat. Protoc.* **9**, 2574–2585 (2014). doi:10.1038/nprot.2014.173 Medline

32. J. Dubochet, M. Adrian, J.-J. Chang, J.-C. Homo, J. Lepault, A. W. McDowall, P. Schultz, Cryo-electron microscopy of vitrified specimens. *Q. Rev. Biophys.* **21**, 129–228 (1988). doi:10.1017/S0033583500004297 Medline
33. C. Suloway, J. Pulokas, D. Fellmann, A. Cheng, F. Guerra, J. Quispe, S. Stagg, C. S. Potter, B. Carragher, Automated molecular microscopy: The new Leginon system. *J. Struct. Biol.* **151**, 41–60 (2005). doi:10.1016/j.jsb.2005.03.010 Medline
34. G. C. Lander, S. M. Stagg, N. R. Voss, A. Cheng, D. Fellmann, J. Pulokas, C. Yoshioka, C. Irving, A. Mulder, P.-W. Lau, D. Lyumkis, C. S. Potter, B. Carragher, Appion: An integrated, database-driven pipeline to facilitate EM image processing. *J. Struct. Biol.* **166**, 95–102 (2009). doi:10.1016/j.jsb.2009.01.002 Medline
35. N. R. Voss, C. K. Yoshioka, M. Radermacher, C. S. Potter, B. Carragher, DoG Picker and TiltPicker: Software tools to facilitate particle selection in single particle electron microscopy. *J. Struct. Biol.* **166**, 205–213 (2009). doi:10.1016/j.jsb.2009.01.004 Medline
36. A. Rohou, N. Grigorieff, CTFFIND4: Fast and accurate defocus estimation from electron micrographs. *J. Struct. Biol.* **192**, 216–221 (2015). doi:10.1016/j.jsb.2015.08.008 Medline
37. T. Ogura, K. Iwasaki, C. Sato, Topology representing network enables highly accurate classification of protein images taken by cryo electron-microscope without masking. *J. Struct. Biol.* **143**, 185–200 (2003). doi:10.1016/j.jsb.2003.08.005 Medline
38. A. M. Roseman, FindEM—a fast, efficient program for automatic selection of particles from electron micrographs. *J. Struct. Biol.* **145**, 91–99 (2004). doi:10.1016/j.jsb.2003.11.007 Medline
39. D. Kimanius, B. O. Forsberg, S. H. Scheres, E. Lindahl, Accelerated cryo-EM structure determination with parallelisation using GPUs in RELION-2. *eLife* **5**, e18722 (2016). doi:10.7554/eLife.18722 Medline
40. S. Q. Zheng, E. Palovcak, J.-P. Armache, K. A. Verba, Y. Cheng, D. A. Agard, MotionCor2: Anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat. Methods* **14**, 331–332 (2017). doi:10.1038/nmeth.4193 Medline
41. S. H. Scheres, RELION: Implementation of a Bayesian approach to cryo-EM structure determination. *J. Struct. Biol.* **180**, 519–530 (2012). doi:10.1016/j.jsb.2012.09.006 Medline
42. S. Chen, G. McMullan, A. R. Faruqi, G. N. Murshudov, J. M. Short, S. H. W. Scheres, R. Henderson, High-resolution noise substitution to measure overfitting and validate resolution in 3D structure determination by single particle electron cryomicroscopy. *Ultramicroscopy* **135**, 24–35 (2013). doi:10.1016/j.ultramic.2013.06.004 Medline
43. P. Emsley, K. Cowtan, Coot: Model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126–2132 (2004). doi:10.1107/S0907444904019158 Medline
44. L. Holm, L. M. Laakso, Dali server update. *Nucleic Acids Res.* **44**, W351–W355 (2016). doi:10.1093/nar/gkw357 Medline
45. P. Jin, D. Bulkley, Y. Guo, W. Zhang, Z. Guo, W. Huynh, S. Wu, S. Meltzer, T. Cheng, L. Y. Jan, Y.-N. Jan, Y. Cheng, Electron cryo-microscopy structure of the mechanotransduction channel NOMPC. *Nature* **547**, 118–122 (2017). doi:10.1038/nature22981 Medline
46. P. D. Adams, P. V. Afonine, G. Bunkóczi, V. B. Chen, I. W. Davis, N. Echols, J. J. Headd, L.-W. Hung, G. J. Kapral, R. W. Grosse-Kunstleve, A. J. McCoy, N. W. Moriarty, R. Oeffner, R. J. Read, D. C. Richardson, J. S. Richardson, T. C. Terwilliger, P. H. Zwart, PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 213–221 (2010). doi:10.1107/S0907444909052925 Medline
47. V. B. Chen, W. B. Arendall 3rd, J. J. Headd, D. A. Keedy, R. M. Immormino, G. J. Kapral, L. W. Murray, J. S. Richardson, D. C. Richardson, MolProbity: All-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 12–21 (2010). doi:10.1107/S0907444909042073 Medline

ACKNOWLEDGMENTS

Cryo-EM data were collected at The Scripps Research Institute (TSRI) electron microscopy facility. We thank Yang Zhang and Huanghe Yang at Duke University who provided access to their calcium imaging apparatus and guidance to calcium imaging experiments. We thank Jean-Christophe Ducom at TSRI High Performance Computing facility for computational support, Bill Anderson for microscope support, and Mark Herzik and Saikat Chowdhury for helpful discussion and training. We thank Zach Johnson for advice on sample freezing. This work was supported by the National Institutes of Health (R35NS097241 to S.-Y. L., DP2EB020402 to G.C.L.). G.C.L is supported as a Searle Scholar and a Pew Scholar. Computational analyses of EM data were performed using shared instrumentation funded by NIH S10OD021634. The coordinates are deposited in the Protein Data Bank with the PDB ID 6BPQ and the electron density maps have been deposited in EMDB with the ID EMD-7127.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/cgi/content/full/science.aan4325/DC1
Materials and Methods
Figs. S1 to S7
Tables S1
References (*31–47*)

3 November 2017; accepted 27 November 2017
Published online 7 December 2017
10.1126/science.aan4325

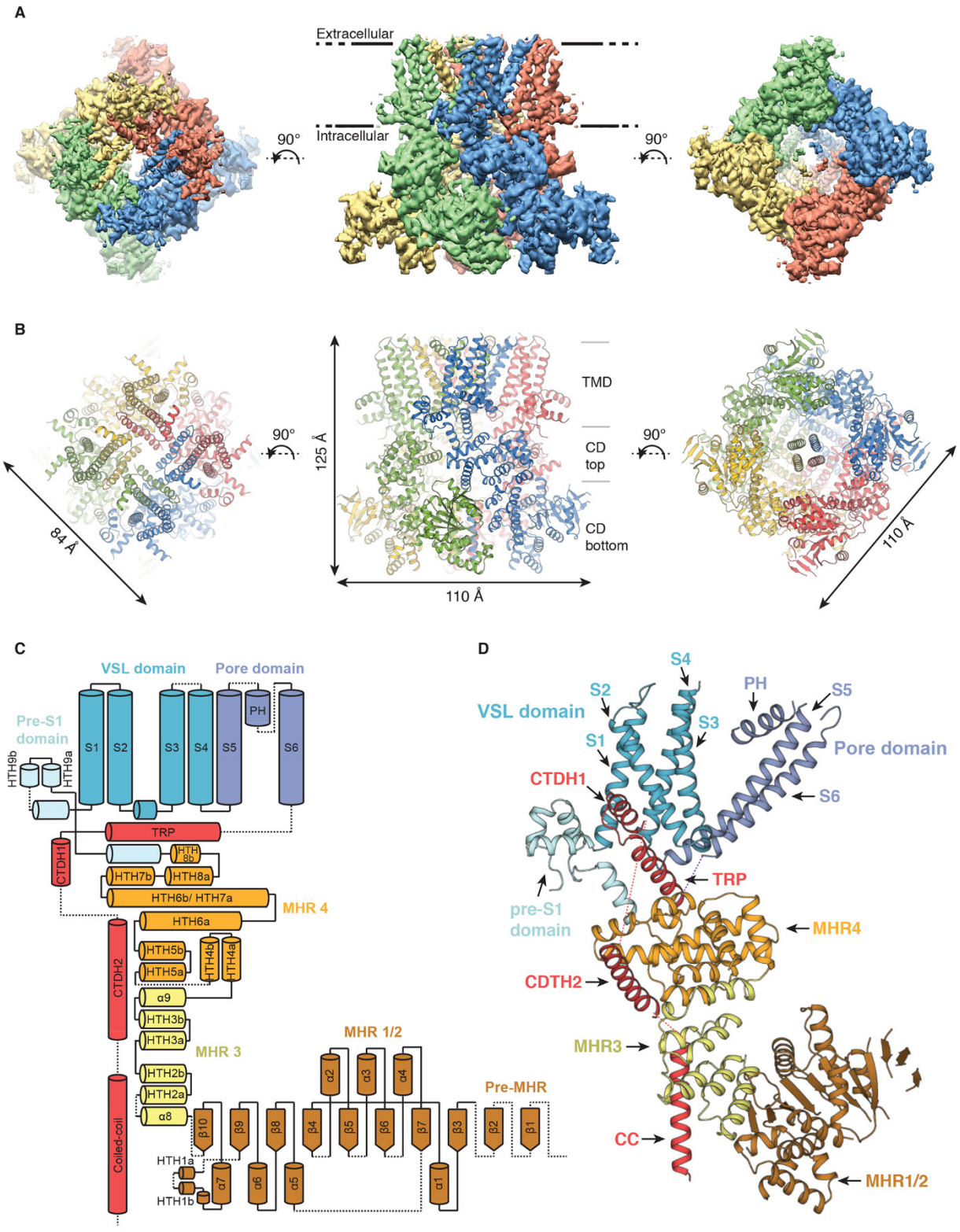


Fig. 1. Overall architecture of TRPM8. (**A**) Cryo-EM reconstruction and (**B**) model of TRPM8 viewed from the extracellular side (left), from the membrane plane (middle), and from the cytosolic side (right). (**C**) Topology diagram delineating the protein domains with secondary structure elements. (**D**) Detailed view of the atomic model of the TRPM8 protomer.

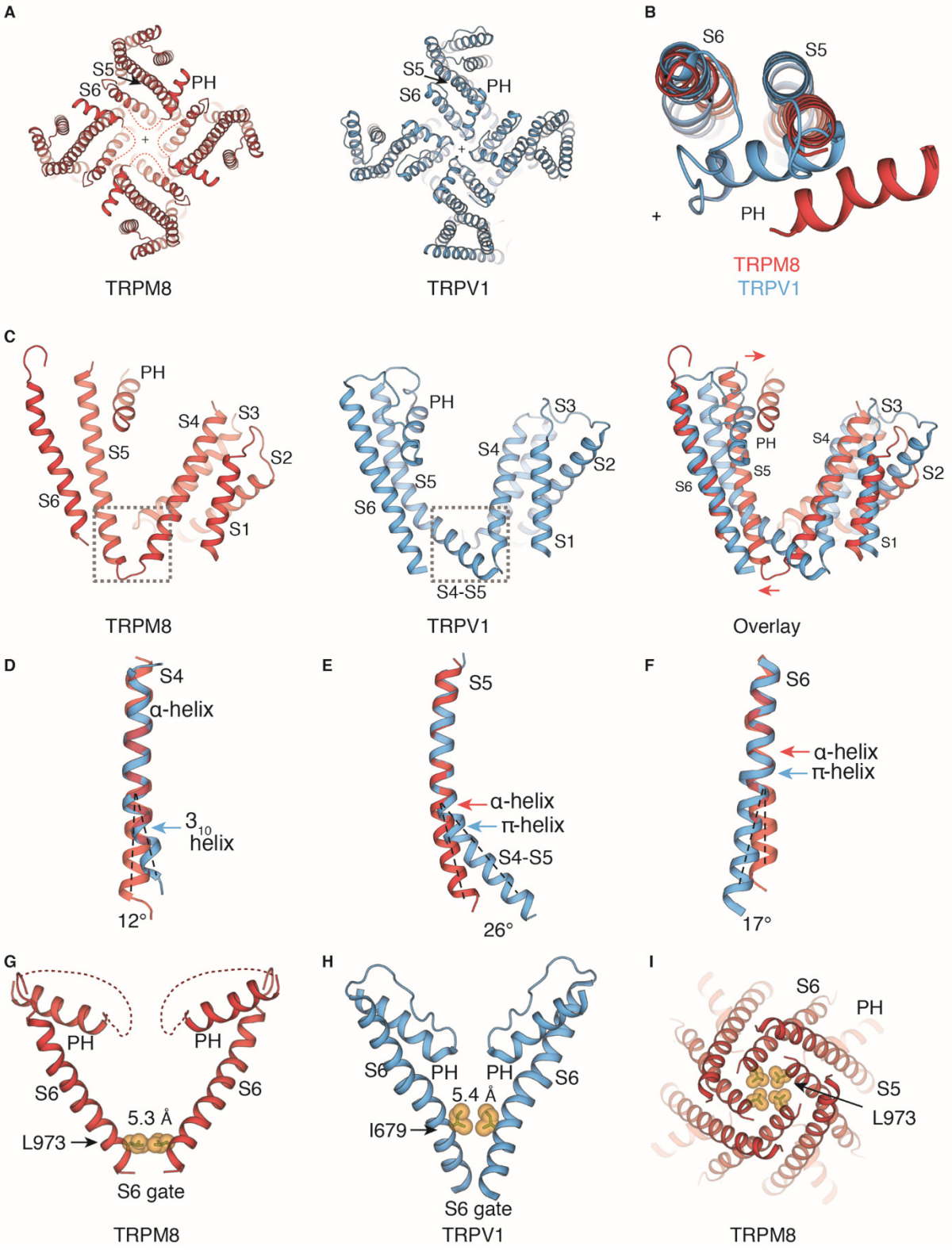


Fig. 2. Comparison of the transmembrane domain in TRPM8 and TRPV1. (**A**) View down the channel pore from the extracellular side. Dashed red lines indicate the selectivity filter and the pore loop linking the pore helix (PH) and S6 in TRPM8, which were not built in the structure due to the lack of well-resolved cryo-EM density in this region (see fig. S5B). In all panels, the "+" indicates the center of the four-fold symmetry axis. (**B**) Overlay of the pore domains of TRPM8 (red) and the apo TRPV1 (blue). The pore helix in TRPM8 is located farther away from the ion permeation pathway. (**C**) Comparison of the junction between S4 and S5. In TRPM8, S5 is straight and fully α-helical, and connected to S4 via a sharp turn, while an S4-S5 linker is identified in TRPV1. The overlay illustrates the differing arrangement of the domain swap between TRPM8 and TRPV1. (**D** to **F**) The non-α-helical features in the S4 (**D**), S5 (**E**) and S6 (**F**) of TRPV1 (blue) are absent in those of TRPM8 (red). The differences in tilt of helices originating at helical bending points are indicated with dashed lines. (**G** to **H**) Comparison of the pores of the TRPM8 and the apo TRPV1 channels, showing that the S6 gate in TRPM8 is formed by Leu^{973}. (**I**) The S6 gate of the TRPM8 structure viewed from the intracellular side. Leu^{973} is shown in sphere representation.

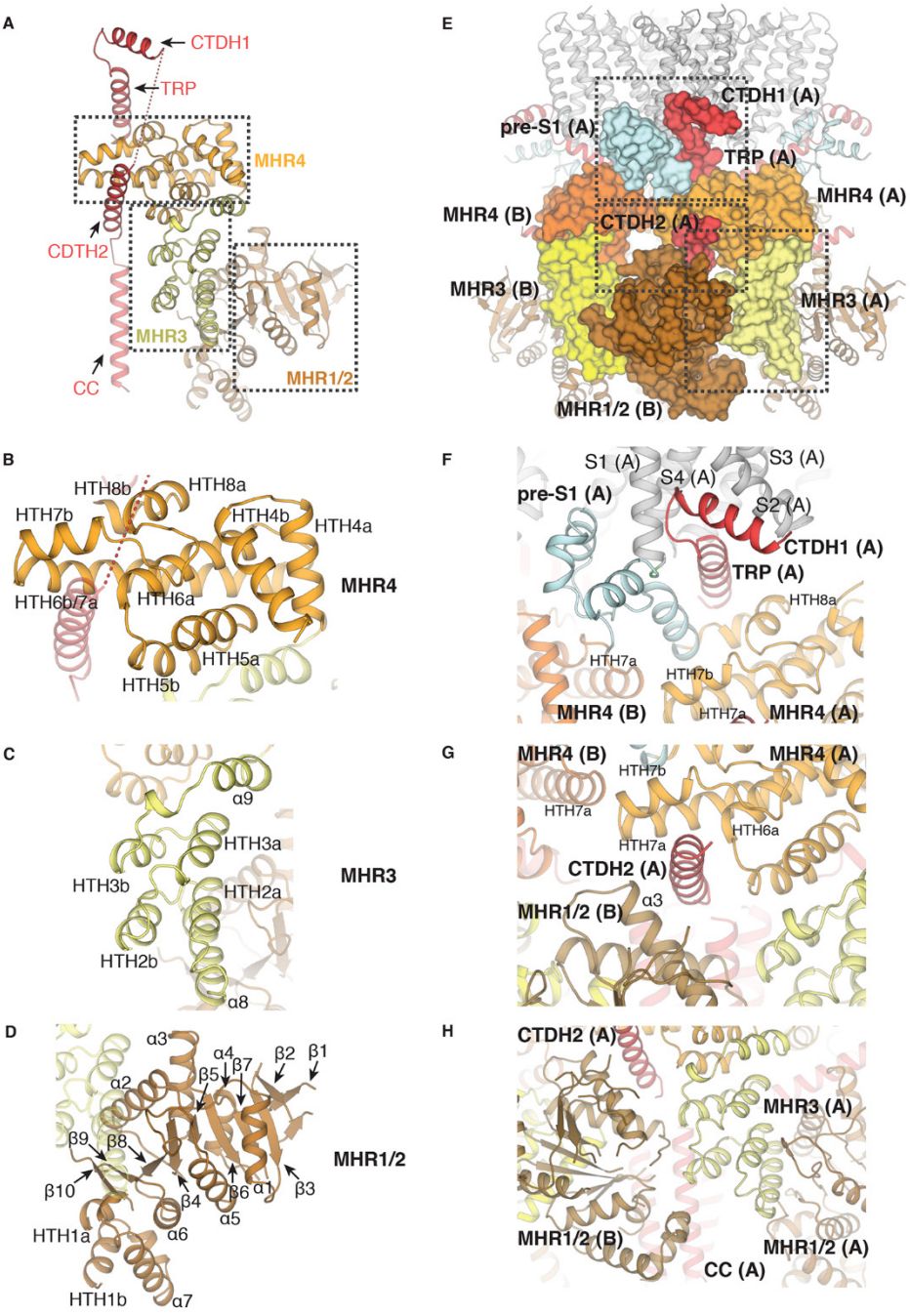


Fig. 3. The cytoplasmic domain of TRPM8. (**A**) The organization of the N- and C-terminal regions of the channel within the cytoplasmic domain (CD). (**B** to **D**) Close-up views of MHR4 (**B**), MHR3 (**C**), and the single domain MHR1/2 (**D**) that comprise the N-terminal domain. (**E**) Surface representation of the interactions between different domains from the neighboring subunits (A) and (B). (**F** to **H**) Close-up views of the domain interfaces that contribute to communication between the TMD and CTD (**F**), between the top and bottom layers of the cytoplasmic ring (**G**), and between protomers in the bottom layer of the cytoplasmic ring (**H**).

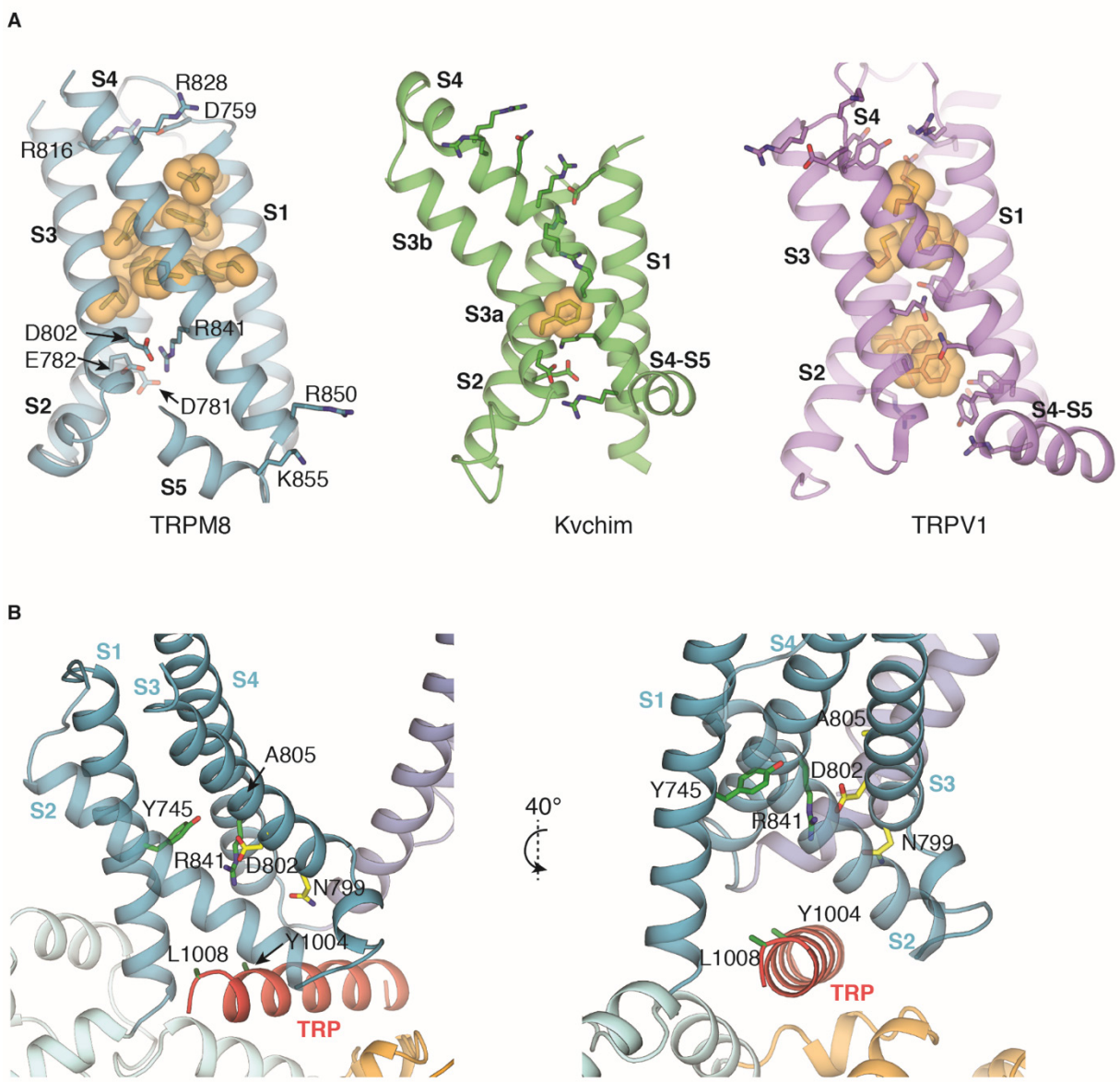


Fig. 4. The voltage-sensor like domain (VSLD) and the putative menthol binding site in the VSLD cavity. (**A**) Comparison of the VSLD in TRPM8 (marine) with the canonical voltage-sensor domain in the Kvchim channel (green) (PDB ID: 2R9R) and the VSLD of TRPV1 (purple) (PDB ID: 3J5P). A gating charge Arg^{841} in the VSLD of TRPM8 is near three negatively charged amino acids below a large hydrophobic seal (spheres in bright orange) in TRPM8 (left). Many gating charge arginines in S4 of Kvchim are located above and below a small hydrophobic seal (spheres in bright orange) and interact with negatively charged amino acids (middle). The interior of the VSLD of TRPV1 is lined with hydrophobic and polar amino acids. (**B**) Residues critical for the sensitivity of TPRM8 to menthol (shown in green stick representation) were mapped to S1 (Tyr^{745}), S4 (Arg^{841}), and the TRP domain (Tyr^{1004}). Residues implicated in icilin sensitivity of TRPM8 (yellow stick representation) were mapped to S3 (Asn^{799} and Asp^{802}). All of these residues point toward the VSLD cavity.

Science Robotics | AAAS

The First Journal of Interdisciplinary Robotics Research.

NEW AND UPCOMING SPECIAL ISSUES

NEW MARINE AND SPACE ROBOTICS

HUMANOIDS

NEW MATERIALS AND ACTUATORS

Submit Your Article for Publication in *Science Robotics*:

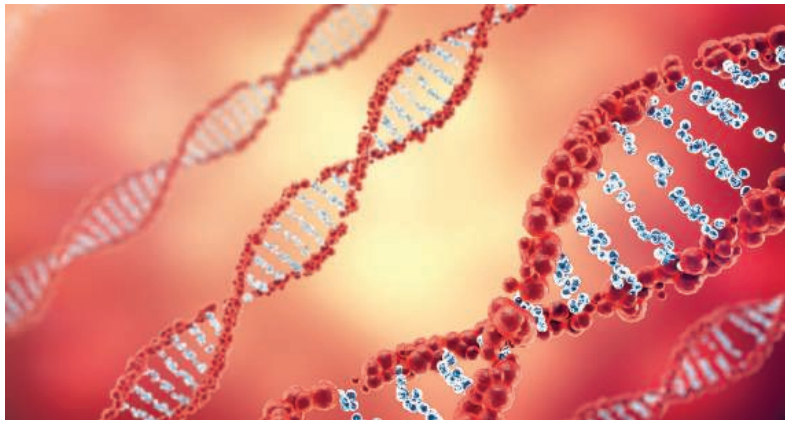
- Rapid review, scoop protection and no article page limits.
- Exposure to an international audience.
- Supportive and knowledgeable editorial staff.
- Promotion on the *Science Robotics* website and the potential for accompanying commentary or podcasts.
- AAAS social media and media relations support.

Learn more at **robotics.sciencemag.org**.

Send pre-submission inquiries to **sciroboteditors@aaas.org**.

MICHELSON MEDICAL RESEARCH FOUNDATION

HUMAN VACCINES PROJECT



CALL FOR ENTRIES

NEW! Michelson Prizes for Human Immunology and Vaccine Research

The Michelson Prizes are international awards that support young investigators applying innovative research concepts and disruptive technologies to significantly advance the development of future vaccines and immunotherapies.

Proposals will be reviewed by a distinguished committee of internationally recognized scientists.

Researchers under the age of 35 are invited to apply.

Learn more about the Prizes and how to apply at **humanvaccinesproject.org/michelsonprizes**

PRE-APPLICATION DEADLINE FEBRUARY 12, 2018

Be the first to win this prize and receive:

- Prize money of US $150,000 each will be awarded to two investigators
- Present your research at a conference on the Convergence of Human Immunology and Vaccine Research at the Michelson Center for Convergent Bioscience at the University of Southern California

APPLY TODAY!

Western Blot Imaging System

The C-DiGit Blot Scanner gives you a digital replacement for film—keeping the advantages of film and eliminating many of the drawbacks (e.g., multiple exposures, waiting in line). Keep it on your lab bench, at your desk, or anywhere you choose. As the cost of film rises, the C-DiGit Blot Scanner gives you an affordable replacement option. The scanner's Image Studio imaging software is designed explicitly for chemiluminescent Western blots. You can acquire data using a standard or high-sensitivity setting to collect a single digital film image; this one data file contains a wide range of different "exposures" representing all options that would be available with film. Image Studio offers powerful, easy-to-use image analysis, and helps to resolve issues caused by film's limited ability to quantify. The C-DiGit is compatible with both Mac and PC systems as well as convenient image acquisition using iPad, iPhone, or iPod.

LI-COR
For info: 888-645-7242
www.licor.com/bio/products/imaging_systems/cdigit

TQ-ICP-MS System

Thermo Scientific's easy-to-use iCAP TQ ICP-MS (triple quadrupole–inductively coupled plasma–mass spectrometry) system is designed to support ultralow detection limits and to overcome interferences common in challenging matrices. It provides reliable and reproducible data for a range of applications in fields such as metallurgy, clinical research, pharmaceutical compliance, environmental science, food safety, and geoscience. The system allows users to switch between single- and triple-quadrupole modes in a single multielement experiment, so that laboratories can keep their existing workflow and also add new capabilities over time. It requires minimal user maintenance and includes QCell flatapole technology, a small-volume collision/reaction cell with flatapole rods, and a dedicated gas-distribution unit for lab safety and maximum flexibility. The iCAP TQ ICP-MS also features self-aligning sample introduction components for reproducibility, and an ergonomic benchtop design with a compact footprint that saves valuable lab space.

Thermo Fisher Scientific
For info: 800-955-6288
www.thermofisher.com

Fluorescence Illuminators

Whether you are observing fixed or live cells, X-Cite carries a complete range of lamps, LED microscope light sources, and fluorescence illumination systems to optimize imaging and ensure greater data reliability. These products are known for maximum stability and offer superior illumination uniformity for fluorescence microscopy applications, with lamp-based solutions that provide the convenience of prealigned, long-life lamps and easy installation. Delivering high-intensity LED illumination, the X-Cite XLED1, X-Cite 120LED*Boost*, X-Cite 110LED, and X-Cite TURBO feature intuitive controls and no bulbs to install, making setup and operation simpler than ever. From our standard X-Cite 120Q model used for routine assays to our most advanced models, we have the ideal system for your application.

Excelitas Technologies
For info: 800-668-8752
www.excelitas.com/pages/product/x-cite.aspx

Imaging Flow Cytometer

The FlowSight Imaging Flow Cytometer stands apart from other flow cytometers by producing up to 12 images of each and every cell. It simultaneously produces darkfield (side scatter), brightfield, and up to 10 fluorescence images. The FlowSight operates at ~20X magnification, allowing visualization of fluorescence from the membrane, cytoplasm, or nucleus. Identifying cell conjugates or distinguishing single cells from doublets and debris is effortless—flow cytometry has never been so intuitive. A proprietary CCD camera, coupled with patented detection technology, yields fluorescence sensitivity at least 10 times that of photomultiplier tube (PMT)-based flow cytometers across all channels. The FlowSight is powerful enough for the core lab, but sized and priced for any lab. The system can be factory-configured or field-upgraded with up to four excitation lasers (405 nm, 488 nm, 561 nm, 642 nm), a 96-well plate AutoSampler, and powerful quantitative image processing.

EMD Millipore
For info: 800-645-5476
www.emdmillipore.com

Biosimilar and R&D Support Service

AffinityImmuno research products are made for scientists by scientists. We offer custom and off-the-shelf R&D solutions and support services for biological drug discovery, biosimilars, and analytical methodologies. Our team has extensive experience with many aspects of immunogenicity testing, epitope and idiotype analysis, carbohydrate analysis, pharmacokinetic testing, and in vitro and in vivo drug testing. We specialize in antibodies, antibody lead discovery, protein engineering, and assay development for novel targets for cancer, autoimmunity, and endocrine disorders.

AffinityImmuno
For info: support@affinityimmuno.com
affinityimmuno.com

qPCR Assays and Arrays

With our PrimePCR Assays and Arrays, we've designed, tested, and validated quantitative PCR (qPCR) primers and probes for your gene expression experiments. Our expertly designed primer/probe sets cover 12 genomes, while our PCR array plates have been curated for most biological pathways and diseases of interest. We've even bench-tested the primers and probes for every single human, mouse, and rat assay. See how you can eliminate at least four steps from your gene expression workflow using these assays.

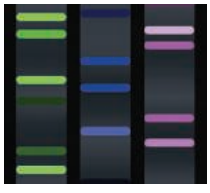
Bio-Rad
For info: 800-424-6723
www.bio-rad.com/en-us/product/primepcr-pcr-primers-assays-arrays

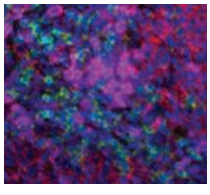
Electronically submit your new product description or product literature information! Go to www.sciencemag.org/about/new-products-section for more information.

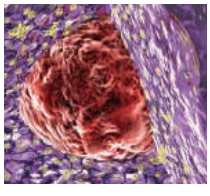
Newly offered instrumentation, apparatus, and laboratory materials of interest to researchers in all disciplines in academic, industrial, and governmental organizations are featured in this space. Emphasis is given to purpose, chief characteristics, and availability of products and materials. Endorsement by *Science* or AAAS of any products or materials mentioned is not implied. Additional information may be obtained from the manufacturer or supplier.

want new technologies?

antibodies
apoptosis
biomarkers
cancer
cytometry
data
diseases
DNA
epigenetics
genomics
immunotherapies
medicine
microbiomics
microfluidics
microscopy
neuroscience
proteomics
sequencing
toxicology
transcriptomics







watch our webinars

Learn about the latest breakthroughs, new technologies, and ground-breaking research in a variety of fields. Our expert speakers explain their quality research to you and answer questions submitted by live viewers.

VIEW NOW!
webinar.sciencemag.org

Science
AAAS

Brought to you by the *Science*/AAAS Custom Publishing Office

@SciMagWebinars

AAAS | ANNUAL MEETING
Feb 15–19 2018 ★ Austin, TX

ADVANCING SCIENCE | DISCOVERY TO APPLICATION

Browse the Online Program and Register Now

Register and book your hotel room now. Reduced rates are available until January 24, 2018.

aaas.org/meetings

AAAS, publisher of *Science*, thanks the sponsors and supporters of the 2018 Annual Meeting

European Commission

TEXAS

GREAT BRITAIN & NORTHERN IRELAND

Bristol-Myers Squibb

SANOFI Empowering Life

Johnson & Johnson INNOVATION

ARCONIC FOUNDATION

Symantec

As of December 26, 2017

Chair, Dept. of Cell & Regenerative Biology

The University of Wisconsin School of Medicine and Public Health (SMPH) invites applications and nominations for the position of Chair of the Department of Cell and Regenerative Biology (CRB).

The Department of Cell and Regenerative Biology is committed to understanding the fundamental mechanisms by which living systems operate at cellular and molecular levels of organization. By embracing a wide range of contemporary and emerging approaches and experimental systems, the department seeks to define signaling and regulatory pathways that provide the basis for understanding, diagnosis and treatment of human disease. Basic research is the centerpiece of the Department and serves as the driving force behind teaching and training efforts. The overarching research interests of the Department are highly interdisciplinary, emphasizing molecular, cellular and systems approaches to describe biological processes in molecular terms. To maintain its excellence and stature, the department is currently focusing on existing strengths in four research areas: Cell and Molecular Biology, Developmental Biology, Stem Cell and Regenerative Biology, and Cardiovascular Biology.

We seek a recognized leader with an outstanding academic background, strong research credentials, demonstrated commitment to education, experience in mentoring junior faculty, and proven leadership and management skills. The chair will provide professional and administrative leadership of the highest quality to this distinguished department in its teaching and research.

The successful candidate will have a compelling vision for the future of CRB in a leading academic medical center. Candidates must have a Ph.D., M.D., MD/Ph.D. or equivalent degree(s). They should possess substantial background and experience in administrative leadership, research, and teaching, and a strong academic background that would qualify for appointment as a tenured professor at the University of Wisconsin-Madison.

Please send nominations to: **Terri Young, MD, MBA, and David Gamm, MD, PhD, Co-Chairs of the CRB Chair Search Committee, c/o Staci Andersen, 4150Q HSLC, 750 Highland Avenue, Madison, WI, 53705- 2111, or use: slandersen2@wisc.edu**

To apply for this position, please use the University of Wisconsin applicant tracking system which is found at the link below. Click the "Apply Now" button. Applicants will be asked to upload a current cover letter, CV, and list of three references.

Application Link: jobs.hr.wisc.edu/crbchair

Applications from minorities and women are strongly encouraged. To receive full consideration, applications should arrive by **January 31, 2018**

Unless confidentiality is requested in writing, information regarding applicants must be released upon request. Finalists cannot be guaranteed confidentiality. Wisconsin Caregiver Law applies. The University of Wisconsin is an equal opportunity, affirmative action employer. For more information: med.wisc.edu, crb.wisc.edu.

School of Medicine and Public Health
UNIVERSITY OF WISCONSIN-MADISON

Open Level Professor in Cell Biology at University of California, Santa Barbara
Molecular Cellular and Developmental Biology

The Molecular, Cellular, and Developmental Biology Department at the University of California Santa Barbara invites applications for an open level faculty position at the Assistant, Associate, or Full Professor level in Molecular Cell Biology with an estimated start date of July 1, 2018. Candidates must hold a PhD (or equivalent) and postdoctoral experience in Molecular Cell Biology or a related field. Applicants should have established a record of creative, cutting edge research as well as a strong record of publication and external funding commensurate with experience. A commitment to teaching at the undergraduate and graduate levels is also important. The Department is especially interested in candidates who can contribute to the diversity and excellence of the academic community through research, teaching, and service as appropriate to position.

Applicants should submit a cover letter, statement of past and future research interests, statement of teaching, and a curriculum vitae directly to our website. Materials should be submitted electronically via **https://recruit.ap.ucsb.edu/apply/JPF01181**. Applications are due **January 20, 2018** for primary consideration. However, applications will continue to be accepted until the position is filled. Questions can be emailed to cpl@mcdb.ucsb.edu.

The University of California is an Equal Opportunity/Affirmative Action Employer. All qualified applicants will receive consideration for employment without regard to race, color, religion, sex, sexual orientation, gender identity, national origin, disability status, protected veteran status, or any other characteristic protected by law.

UCSB

UNIVERSITY OF CALIFORNIA
SANTA BARBARA

COLLEGE OF MEDICINE AND LIFE SCIENCES
THE UNIVERSITY OF TOLEDO

Chair
Department of Cancer Biology

The University of Toledo College of Medicine and Life Sciences is seeking nominations or applications for the position of Professor and Chair of the Department of Cancer Biology. The successful candidate will provide the vision and leadership to expand the portfolio of funded research in the department though strategic recruitment and support of the current faculty. The department is located on the University of Toledo Health Science Campus, in proximity to the University of Toledo Medical Center and the Eleanor N. Dana Cancer Center. The latter is slated for significant expansion, with emphasis on early-stage clinical trials and personalized approaches to cancer treatment. Candidates with research programs in all areas of cancer biology will be considered. Current research interests represented within the department include development of new therapeutic approaches based on knowledge of cancer signaling pathways, gene regulation, tumor microenvironment, and mechanisms of cancer invasion and metastasis. Specific information about the research interests of the faculty is available at **http://www.utoledo.edu/med/depts/cancer-biology/profiles.html**. In addition to its research mission, the department is strongly committed to graduate and medical education. The Cancer Biology track of the Biomedical Sciences graduate program is closely aligned with the department.

The next Chair is expected to have excellent communication and leadership skills, a strong commitment to faculty mentorship, experience in medical or graduate education, and a clear vision for advancing the research mission of the department. A substantial record of extramurally funded research in an area related to cancer biology and the ability to foster intra- and interdepartmental collaborations are important qualifications. The College of Medicine and Life Sciences is committed to growth of its cancer research programs, and the new Chair of Cancer Biology will be provided with renovated lab space and resources to accomplish this goal. Applications should include a letter of interest summarizing research, educational and leadership experience and a curriculum vitae. Application materials may be sent via e-mail as a single pdf file to the attention of the **Cancer Biology Search Committee, c/o Megan Newcomer, COMLS Dean's Office, 3000 Arlington Ave., University of Toledo Health Science Campus, Toledo, OH 43614**. **megan.newcomer@utoledo.edu**

The University of Toledo is committed to diversity and equal opportunity. Applications from women and minority candidates are strongly encouraged.

BIODISCOVERY INSTITUTE
UNT

BIODISCOVERY INSTITUTE, UNIVERSITY OF NORTH TEXAS
TWO ASSISTANT/ASSOCIATE PROFESSOR POSITIONS

The University of North Texas (UNT), Denton, Texas, invites applications from outstanding scientists for two academic-year positions at the assistant or associate professor levels in the BioDiscovery Institute (BDI), one of four recently established Institutes of Research Excellence. Candidates addressing important fundamental and applied questions in the synthesis and development of bio-based products for agriculture, materials engineering, bioenergy or health benefits are especially encouraged to apply. We are seeking individuals working in two general areas: (1) microbial metabolism and synthetic biology and (2) structure-based enzyme design. These new hires will complement and interact with existing researchers in systems modeling of metabolism, plant biochemistry, and biotechnology. Dedicated, newly renovated research space, competitive start-up packages, and state-of-the-art core facilities in genomics and metabolomics will support these new members and their colleagues in BDI.

Successful candidates will bring an internationally recognized research program, interact with diverse faculty, staff and students within BDI and across UNT, and contribute to graduate and/or undergraduate training. Applicants should have a Ph.D. in a biological, computational, physical or engineering sciences discipline, an established record of funding and publications in one of the target areas described above, and a commitment to excellence in research within an interdisciplinary environment. Applications should be made through the UNT Academic Resources website (**https://facultyjobs.unt.edu**) and should include a cover letter, curriculum vitae, two-page summary of research interests and accomplishments, one-page statement of teaching philosophy, and the names and contact information for three references. Questions may be directed to Brian G. Ayre (**bgayre@unt. edu**), Search Committee Chair, or Kent D. Chapman (**chapman@unt.edu**), BDI Director. Screening of applications will begin immediately and will continue until the positions are closed. For best consideration, apply by **February 15, 2018**.

The University of North Texas is an Equal Opportunity/Access/Affirmative Action/Pro Disabled and Veteran Institution committed to diversity in its employment and educational programs, thereby creating a welcoming environment for everyone.

西南交通大学
Southwest Jiaotong University

Southwest Jiaotong University, Chengdu, China Invites Applications for Academic Positions

Southwest Jiaotong University (SWJTU), founded in 1896, is one of the oldest institutions of high learning in China. In its proud legacy of 120 years, the University has been dedicated to Chinese higher education and has proudly nurtured generations of engineering and scientific leaders. As the most comprehensive leading research university in transportation, SWJTU is world-renown for pioneering the Chinese railway transportation engineering and industry, and for its leading contributions to the development of Chinese high-speed rail system. For its sustained excellence and prominence, the University is placed among the key, elite multidisciplinary "211" and "985" Tier-1 universities directly administered by the Chinese Ministry of Education. We offer comprehensive education and research programs in 19 faculties and institutes/centers, covering diverse disciplines in engineering, sciences, arts, and management leading from undergraduate to doctoral degrees. The University boasts 2,600 outstanding academic staffs, 15 doctoral/ 43 master/ 75 undergraduate programs and 11 post-doctoral stations, supported by more than 30 cutting edge key laboratories at the national and provincial levels.

Located strategically in Chengdu, the capital of Sichuan province—the China' s dynamically growing West, SWJTU is blessed with rich heritage, unparallel vibrancy, and a beautiful campus. "Prosperous and plentiful ever now and then, the City flourishes in hibiscus blossoms in no end," as so known, Chengdu has been long renowned for its historical and cultural heritage, and for its natural beauty and abundance. As a major cultural and economical center and a transportation hub, the City offers first-class cultural experience, education, employment, cuisine and living environment. Leveraging on these unique advantages and the University's strengths, SWJTU is vigorously implementing its strategic plan "Developing and Strengthening SWJTU: Attracting and Cultivating Talents" . We earnestly look forward to your joining our legacy and contributing to the University's continuing excellence.

More information is available at http://www.swjtu.edu.cn/

Openings in

Civil Engineering/Surveying Science and Engineering/Mechanical Engineering

Science of Transportation and Logics/ Science of Information and Communication

Electrical Engineering/Computer Science and Technology

Materials Science and Technology/Mechanics/Management Science and Technology

Managing Technology and Innovation/Environmental Science/Architecture

Physics Science/ Mathematics Science/Life Science/Medical Science

Chemical Science/Humanities and Social Science

Salary and Fringe Benefits

Salary will be highly competitive, commensurate with qualifications and experiences. The University offers a comprehensive fringe benefit package for eligible appointees, including relocation allowances, subsidy of rental residence, start-up funds for research. The University is committed to proving assistance in establishing scientific platform and research group as well as international-level training and promotion.

The University also assists the eligible appointees in child education. Special arrangements are open for discussion for exceptional appointees.

How to Apply

Interested candidates should send a full resume, copies of academic credentials, a publication list with abstracts of selected publications, a research plan, a teaching statement, together with names of three references to Human Resources Department Southwest Jiaotong University Western Park of High-Tech Zone Chengdu, Sichuan Province, China 611756

Telephone number: +86-28-66366202

Email: talent@swjtu.edu.cn

For inquiries, please contact Ms. Ye Zeng or Mr. Jian Wu at the above addresses.

High-level Global Talents Recruitment

Welcome back to hometown.
Thousands of academic job vacancies are in fast-developing China.

On-the-spot Recruitment in Britain:
March 2, 2018 University of Cambridge
March 3, 2018 University of Warwick
March 4, 2018 University of Oxford

Online Job Fair:
March 10, 2018 (GMT+8) www.edu.cn/cv

Qualification for Applicants:
Overseas scholars, Doctor and Post-doctor

Participating Approach:
Please send your CV to consultant@acabridge.edu.cn
for On-the-spot Recruitment and Online Job Fair

Pre-registration Now Open

Job Vacancies in China's Universities and Institutes

Please visit http://www.acabridge.cn/
Contact consultant@acabridge.edu.cn

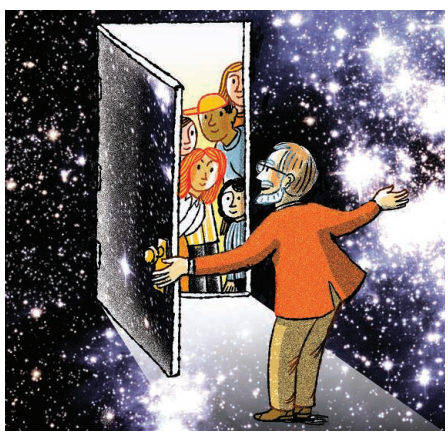
By **William H. Waller**

My second life as a teacher

For most of my educational and professional life, I pursued a fairly standard trajectory. A bachelor's degree in physics and astronomy, a master's in optical physics, and a Ph.D. in astronomy prepared me for a postdoctoral fellowship and subsequent work as a scientist at NASA's Goddard Space Flight Center. I moved on to a visiting professorship and then a research professorship at Tufts University. I thought I was well on my way to a stable career as an astronomer. Then it stalled, and my second life beckoned.

The turning point came in 2008, 18 years after I got my Ph.D., when my luck securing grants dried up. It can happen to any of us, and that was my year. Over the next 2 years, I applied for several more grants, hoping that I could sustain my career in academia, but none got funded. In the meantime, I began considering options beyond the hallowed halls.

During those years, I spent a summer teaching a high school astronomy course, worked on a book about the Milky Way, and served as a subject matter expert for an online textbook project. Yet, for the most part, I was either unemployed or severely underemployed. It was hard to not feel disappointed and frustrated. I wondered whether my hard-earned scientific background would ever be put to good use again.



"I was happy to dive into teaching high school students."

That's when my community service paid off. For several years, starting before I left my university post, I had been volunteering in the local public school system, serving as a member of a council dedicated to making sure that the educational environment was safe, welcoming, and enriching. So, when a physics teaching opportunity opened up in the high school, the principal—who knew me through my volunteer work—asked whether I would like to apply. I had truly relished teaching at the undergraduate and graduate levels, so I was happy to dive into teaching high school students. Besides, the salary was decent enough—certainly more than an adjunct college professor makes.

Now, 7 years later, I am glad my career path took this unexpected turn. Making the adjustment felt like a new and exciting challenge, not a downgrade in my prospects. In some ways, teaching high school students has been even more rewarding than my college-level teaching. I teach physics and occasional astronomy-related courses at levels that are typically more rigorous than the introductory courses I was teaching at Tufts. I also get to interact with my high school students in a more personal way, as the class sizes are smaller and I see the students more often over the course of the year. Yes, they are still teenagers who are prone to lapses in their executive functioning and who can manifest a fair degree of silliness and drama. But they also can be delightful, especially if given a chance to express themselves. I particularly enjoy mentoring students as they conceive and carry out research projects that they can then present in competition. For me, the experience is surprisingly similar to mentoring undergraduate and graduate students, and it is just as rewarding.

Teaching also leaves enough room in my life for my own intellectual interests, including reading scientific journals and magazines; attending professional meetings; and co-editing *The Galactic Inquirer*, an online journal for students and astronomy enthusiasts. During the summers, I can focus on writing and editing, hosting occasional astronomical workshops for educators and enthusiasts, and attending more far-flung scientific meetings.

It's not perfect, of course. Although I feel fulfilled in my second life as a science teacher and part-time science popularizer, I miss carrying out cutting-edge research. Over the years, my skills in data acquisition and analysis have waned, while my skills in classroom management and instruction have grown. But I've enjoyed building my new arsenal. I've also come to appreciate that my background in science has true and lasting value, regardless of where or how I apply it. By being flexible and open to new possibilities, I have become one of the many scientists who find joyful and rewarding work in unexpected places. ■

William H. Waller is a high school physical sciences teacher in Rockport, Massachusetts. Send your career story to SciCareerEditor@aaas.org.