France's bold bid to wean farmers from chemicals *p. 144*

More anti-CRISPRs discovered *pp. 156, 236, & 240*

Discrete brain circuits underlie parenting behavior *p. 168*

Science

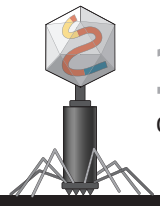
$15
12 OCTOBER 2018
sciencemag.org

AAAS

SPECIAL ISSUE

BRAIN DEVELOPMENT

CONTENTS



152 & 207
Gene swapping in bacteria

12 OCTOBER 2018 • VOLUME 362 • ISSUE 6411

SPECIAL SECTION

BRAIN DEVELOPMENT

ON THE COVER

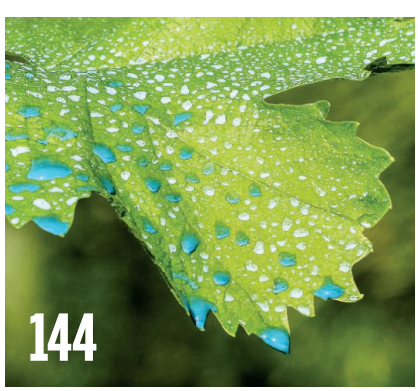
Neurons of the brain find their places during development guided by internal programs and external interactions. This confocal microscopy image from a *Drosophila* larva shows the characteristic placement of neurons in the optic lobe. The neurons are organized in layers according to their birth order, with the color of each layer corresponding to its neurons' unique expression of transcription factors, as detected by immunofluorescence. This special issue highlights the behind-the-scenes forces and constraints that guide brain development. See page 170. *Image: Isabel Holguera, Claude Desplan laboratory*

NEWS

IN BRIEF

IN DEPTH

FEATURES



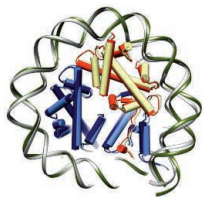
INSIGHTS

PERSPECTIVES

POLICY FORUM

Published by AAAS

CONTENTS



199
Remodeling DNA

12 OCTOBER 2018 • VOLUME 362 • ISSUE 6411

A virus threatens Mongolia's saiga population

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. **Authorization to reproduce** material for internal or personal use under circumstances not falling within the fair use provisions of the Copyright Act is granted by AAAS to libraries and others who use Copyright Clearance Center (CCC) Pay-Per-Use services provided that $35.00 per article is paid directly to CCC, 222 Rosewood Drive, Danvers, MA 01923. The identification code for *Science* is 0036-8075. *Science* is indexed in the *Reader's Guide to Periodical Literature* and in several specialized indexes.

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 www.sciencemag.org/authors/science-information-authors

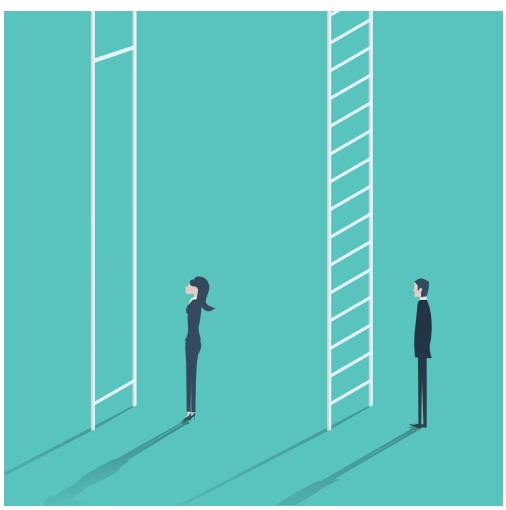
Published by AAAS

EDITORIAL

Japan needs gender equality

Last month, Tokyo Medical University (TMU) announced Yukiko Hayashi as its first female president. This comes on the heels of discovering that the insitution had manipulated entrance exam scores for many years to curb female enrollment. Hayashi may be an attempt by TMU to restore its reputation, but the scandal should be a wake-up call for Japanese society to ensure that men and women have equal opportunities to succeed.

TMU admitted to engaging in gender discrimination because affiliated hospitals wanted more male graduates. Apparently, the university believed that female physicians are more likely to leave the medical profession because they cannot cope with demanding hospital schedules. Whereas misconduct of this scale can make the headlines, more subtle barriers discourage women from attaining various roles in Japan. Lack of gender parity is particularly severe in the science, technology, engineering, and mathematics (STEM) and medical fields. For example, the Japanese female-to-male ratio of physicians is the lowest among member countries of the Organization for Economic Co-operation and Development (OECD), hovering around 21% compared with the OECD average of 47%.



"Gender equality will be the key for better lives for all Japanese."

Yet, Japan boasts one of the most sophisticated educational systems in the world. Japanese 15-year-old students are among the highest performers in mathematics and science according to the OECD's Programme for International Student Assessment (PISA). However, a gender gap is noticeable, especially among top-tier students. The latest PISA assessment in 2015 indicates that Japanese boys outperform girls in mathematics by ~15% of a standard deviation on the achievement scale. Among the top 10% of students in Japan, there is an even larger gender gap in both mathematics and science. But by international standards, Japanese girls perform at very high levels. For example, the highest-achieving girls in Japan performed significantly above the highest-achieving boys from most of the other 70 education systems assessed by PISA in both mathematics and science.

That girls' test scores at TMU were manipulated is all the more troubling because the health profession is the only STEM area in which Japanese female students express a stronger desire to work compared with their male counterparts. Japanese boys generally contemplate a wider range of STEM careers as compared with what girls consider.

PISA assessments suggest that the relative underachievement of girls in mathematics and science compared with boys may have to do with a lack of confidence of girls. When students are self-confident, they give themselves the freedom to fail and to engage in a trial-and-error process that is fundamental in acquiring knowledge in mathematics and science. Indeed, OECD surveys show that when Japanese boys and girls have similar levels of confidence in their abilities, there are no performance gaps, not even among top-performing students.

Fostering confidence in female students, trainees, and professionals would be a highly effective way to close the gender gap in Japan. Gender mainstreaming in various segments of Japanese society is crucial to address unconscious biases. A better gender balance in professional occupations, especially in STEM fields, would boost self-efficacy in female students, and TMU's decision to appoint a female president is a major step toward promoting more female role models in medicine. Medical schools should work with hospitals to improve working conditions so that female physicians are not hampered in their careers by life events such as pregnancies and child-rearing.

Halving Japan's gender gap in the labor force by 2025 could add almost 4 percentage points to projected growth in gross domestic product over the period from 2013 to 2025. Japan clearly needs to embrace women to bolster its economy. And having more highly educated women in medical professions is particularly beneficial for a country with a rapidly aging population. Gender equality will be the key for better lives for all Japanese.

–Yumiko Murakami and Francesca Borgonovi

Yumiko Murakami *is the head of the Organization for Economic Co-operation and Development/Tokyo Centre, Tokyo, Japan. yumiko.murakami@oecd.org*

Francesca Borgonovi *is a senior analyst at the Organization for Economic Co-operation and Development, Paris, France. francesca.borgonovi@oecd.org*

10.1126/science.aav6236

NEWS

> “I am sincerely and profoundly sorry for the image conveyed by this video.”
>
> **New Nobel laureate Gérard Mourou**, apologizing for a 2013 video, meant to promote a laser facility, that included dancing women in lab coats and other images critics found sexist.

IN BRIEF

Edited by **Jeffrey Brainard**

EARTH SCIENCE

Climate panel urges emissions cuts now

The report predicts a drastic melting of Arctic sea ice, like these floes in Norway’s Hinlopen Strait.

The U.N. climate panel warned this week that allowing the planet to warm by more than 1.5°C could have dire consequences. The Intergovernmental Panel on Climate Change (IPCC) said in a new report that a speedy transformation of the world’s energy systems is needed to avoid breaching that limit, which is notably tighter than the target of 2°C cited in the Paris agreement of 2015. Exceeding 1.5°C could result in greater sea level rise; loss of biodiversity; and increasingly severe storms, flooding, and drought, the panel said. Among other measures, coal needs to be all but eliminated as a source of electricity and renewable power must be greatly expanded. The scientists concluded that the global average annual temperature has already increased by 1°C since preindustrial times and is likely to reach 1.5°C between 2030 and 2052, meaning there is no time for further delay. “We have to alter course immediately; no longer can we say the window for action will close soon—we’re here now,” says Drew Shindell, an atmospheric scientist at Duke University in Durham, North Carolina, who took part in the IPCC assessment.

Critics fault economics Nobel

SCIENTIFIC PRIZES | This year’s Nobel Prize in economics has sparked controversy. Half of the $1 million prize honors Paul Romer, 63, an economist at New York University in New York City, for his theory of how economies spawn wealth-producing ideas. The other half goes to William Nordhaus, 77, an economist at Yale University, for integrating climate change into economic theory and developing models to assign a price to carbon emissions. Some economists object to Nordhaus’s emphasis on economic growth and his suggestion that near-term growth may make it cheaper to pay for longer-term environmental damage. “His kind of analysis has been used to delay, delay, delay” taking action on climate change, says Julia Steinberger, an ecological economist at the University of Leeds in the United Kingdom. Others say Nordhaus’s work was just a first step toward confronting climate change. “Environmental economists put a big emphasis on a carbon tax, but they don’t say, ‘You don’t need anything else,’” says Katheline Schubert, an environmental economist at the Paris School of Economics. Nevertheless, the climate change problem clearly remains unsolved.

Japan allows embryo editing

GENOMICS | Japan will allow genome editing of fertilized human embryos for basic research, following the release of draft guidelines on 28 September by a joint advisory committee of Japan’s health and science ministries. Under the guidelines, the embryos must be leftover from infertility treatments and procured with informed consent and without any compensation, and must not be subsequently implanted into human or animal wombs. Institutional review boards and both ministries must vet proposed experiments. Such research is banned in many European countries and permitted but regulated in the United Kingdom. There is no U.S. law forbidding the research, but federal funds cannot be used.

Father of ‘the God particle’ dies

PHYSICS | Leon Lederman, a Nobel Prize–winning physicist and passionate advocate

PHOTO: KEENPRESS/GETTY IMAGES

Published by AAAS

for science education, died last week at age 96. He and two colleagues shared the physics Nobel in 1988 for their discovery 26 years earlier that elusive particles called neutrinos come in more than one type. Lederman became identified with the neologism in the title of his 1993 book, *The God Particle: If the Universe Is the Answer, What Is the Question?* He was referring to the Higgs boson, the last missing piece of physicists' standard model of fundamental particles and forces, which was finally discovered in 2012. Some physicists scorned the title. Lederman puckishly claimed his publisher balked at *Goddamn Particle*, which would have conveyed how hard physicists were struggling to detect the Higgs.

Third probe explores asteroid

SPACE SCIENCE | Japan's Hayabusa2 spacecraft dropped a third probe onto the surface of the asteroid Ryugu last week. The Mobile Asteroid Surface Scout (MASCOT), about the size of a large lunchbox, carries a camera and instruments to measure day-to-night temperature changes and check for magnetism. It also has an instrument to study rock composition and look for evidence that the asteroid has or once hosted water or organic molecules. MASCOT survived for more than 17 hours, hopping between three locations to collect data. Later this month, Hayabusa2 is set to make the first of three touchdowns on the asteroid to collect samples for eventual return to Earth.

Bears helicoptered to Pyrenees

WILDLIFE CONSERVATION | After years of delay and despite local protests, France last week introduced two female brown bears (*Ursus arctos*) in the Pyrenees, a mountain range that forms the border with Spain. The females, from Slovenia, were released in the region's western part, where only two males had been spotted. Scientists expect breeding will increase the region's bear population, which numbered 43 animals last year, after France made similar introductions in 1996 and 2006. Even so, a French court ruled in March that the government's failure to pursue an adequate bear conservation strategy breached European rules. The airlift was carried out by helicopter, evading roadblocks by groups fighting the operation, including shepherds who say bears threaten their flocks and livelihood.

PHOTO: FRANCK FOUQUET/GETTY IMAGES

France imported two females to help increase its population of brown bears.

Database aids immunity studies

IMMUNOLOGY | A new, free repository of data from healthy adults provides comparison groups in studies of immune system dysfunctions such as chronic fatigue. To build the 10,000 Immunomes Project, a team at the University of California, San Francisco, pulled together data from 83 studies funded by the U.S. National Institute of Allergy and Infectious Diseases that had been deposited in a preexisting data portal called ImmPort. The team says in the 8 October issue of *Cell Reports* that the new, searchable tool is easier for researchers to use than the few existing immunological databases and contains a greater number of immune traits for each person. The authors hope it will help research on disease and also sharpen the definition of a healthy immune system.

Biotech group targets sexism

GENDER | The Biotechnology Innovation Organization (BIO), based in Washington, D.C., last week urged its members to fill their executive ranks and boardrooms with substantially more women; people of color; and lesbian, gay, bisexual, transgender, and questioning (LGBTQ) employees. In a 2 October letter, BIO challenged its roughly 1000 members to aim for 50% women at senior executive levels and 30% women on company boards by 2025. (The current levels are about 25% and 10%, respectively.) BIO's president, CEO, and workforce diversity committee chair

THREE QS

U.K. scientists to be 'all right'

In January, Sam Gyimah was named the United Kingdom's science and universities minister. Last month, he visited the United States, touring pharmaceutical companies in Boston and NASA's Johnson Space Center in Houston, Texas. During his visit, Gyimah spoke with *Science* about Brexit, the United Kingdom's impending withdrawal from the European Union, which is causing anxiety among U.K. scientists. (For a longer version of this interview, see https://scim.ag/UKscienceQA.)

Q: There is evidence that non-U.K. scientists are seeing the United Kingdom as a less appealing destination. What can you do to staunch brain drain?
A: I understand that mobility of scientists is essential to our success. We've launched a £1 billion "future leaders" program under the United Kingdom Research and Innovation [UKRI], which is open to the brightest, the best, and the talented from all around the world. And we're looking at our visa regimes. The UKRI visa program, it's going to make it easier for researchers to come to the U.K. and do their work.

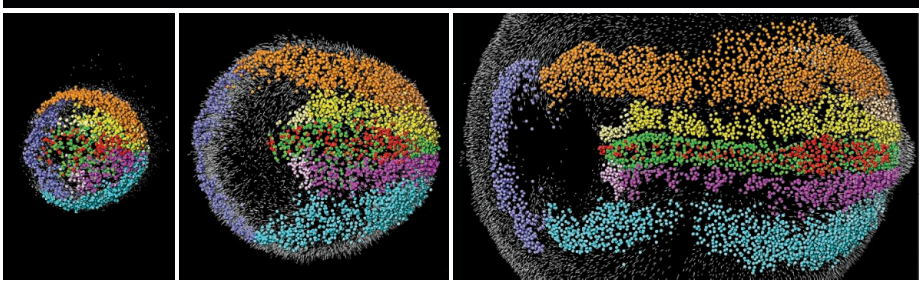
Q: The European Union says the United Kingdom will be shut out of future contracts for Galileo, the European Union's GPS system, including perhaps its military-grade signals. What's next?
A: It doesn't look like the EU is going to change its mind. So, we will do what any sovereign nation would do which has military interests to bear in mind and which needs access to this technology—which is to look to produce our own version of it. It'll probably be less complicated than one that's built to the spec of 28 different countries. What we want to be post-Brexit is nimble, agile—and this is one area where we can prove that.

Q: During recent Brexit negotiations, European leaders made it clear that withdrawal would not be easy. What do you say to U.K. scientists? Are they going to be all right?
A: They're going to be all right, and we're going to do everything to make sure that post-Brexit, the U.K. is a go-to place for science and innovation. We're proving that by increasing investment in science to record levels. ... It's neither in our interest nor the EU's for there not to be a deal. I think cool heads will prevail.

DEVELOPMENTAL BIOLOGY

Microscope tracks developing mouse embryo

A powerful new computer-assisted microscope has let researchers observe how a mouse embryo develops primitive organs, following the fate of individual cells along the way. The microscope, described this week in *Cell*, shines a sheet of laser light through a growing mouse embryo, allowing researchers to see cells deep inside. The microscope uses algorithms to sort through millions of images to track cells as they move and grow by an order of magnitude from day 6 to day 8 of development. Here, dark blue cells will become part of the heart, red cells will form the notochord—the precursor of the spine—and green cells will turn into the brain and spinal cord. The scientists, at the Howard Hughes Medical Institute's Janelia Research Campus in Ashton, Virginia, used similar techniques to construct "digital embryos" of fruit flies and zebrafish. They have made their "digital mouse embryo" publicly available along with the software and directions for building the microscope.



IMAGES: K. MCDOLE ET AL., *CELL*, 10.1016/J.CELL.2018.09.031 (2018)

also cautioned members that they could be ejected from the trade association for sexist, racist, or homophobic "actions, activities, or events." Critics have cited the topless dancers with BIO member company logos painted on their abdomens and thighs who performed at an unofficial party in June organized by several firms during BIO's annual convention in Boston, an incident first reported by BioCentury.com. BIO is also working to craft goals for racial diversity and LGBTQ representation in the biotech industry, a spokesperson said.

Pact bans Arctic fishing

MARINE CONSERVATION | Nine nations and the European Union have formally signed an agreement to bar commercial fishing for at least 16 years in a huge swath of the central Arctic Ocean. The pact, signed on 4 October in Greenland, marks the end of a lengthy effort by nations with a direct territorial stake in the Arctic, and others interested in fishing there, to develop a management plan before global warming makes the waters easily accessible to commercial fleets (*Science*, 8 December 2017, p. 1235). The deal covers some 2.8 million square kilometers surrounding the pole, making it the world's largest to preemptively block the exploitation of marine resources. The ban, which can be extended, is aimed at giving scientists and policymakers time to develop a better understanding of the region's resources.

Industry funders control trials

BIOMEDICINE | In most clinical trials funded solely by industry, the sponsors help design the studies, analyze results, and co-author publications, some of which downplay or omit industry involvement, a study has found. Researchers at the Nordic Cochrane Centre in Copenhagen examined 200 such trials of drugs, vaccines, and devices published from 2014 to 2017 in seven leading medical journals. In only 40% of the trials did participating academic scientists analyze trial results, and in only 4% did they conduct all aspects of the studies. The Cochrane authors say the findings, published in the 4 October issue of *The BMJ*, indicate a need for more accurate disclosure by authors and funders of their involvement in key aspects of these studies.

U.S. suit filed over copyrights

PUBLISHING | The American Chemical Society (ACS) in Washington, D.C., and Elsevier have extended their copyright fight with ResearchGate to the United States. The two publishers filed suit last week against the popular academic networking site in U.S. District Court in Maryland, alleging that ResearchGate was infringing their copyrights by hosting published, paywalled journal articles uploaded by scientists. Both publishers had sued ResearchGate in Germany in October 2017 on similar grounds; that case's outcome, not yet decided, would not bind ResearchGate's U.S. activities. "ResearchGate continues to reject the viable long-term solutions we have proposed to address the copyright infringement on its site," says ACS Senior Vice President James Milne. ResearchGate declined to comment.

Hoax papers spur uproar

PEER REVIEW | Three scholars caused an uproar last week by revealing a hoax in which they published fabricated papers with outlandish themes in journals about gender and sexual orientation. Writing in the online magazine *Areo*, the authors said reviewers overlooked obvious holes in methods and statistics because of their left-wing, ideological bias in favor of what the scholars called "grievance" studies that examine forms of oppression rooted in identity. Seven of 20 submitted hoax papers were accepted or published online. One, for example, studied "rape culture" of dogs at dog parks. Another, titled, "Our Struggle Is My Struggle," lifted passages from Hitler's *Mein Kampf* in an analysis of contemporary feminism. The episode recalled an infamous 1996 hoax in which the physicist Alan Sokal published a paper in a cultural studies journal that mixed postmodern philosophy and quantum gravity. Critics of the new hoax called it unethical and unfairly selective.

A mathematical maze in Norway

MATH | When Hans Munthe-Kaas was asked to help design a new botanical garden at his institution, the University of Bergen in Norway, he drew inspiration from his own discipline, math. The result debuted on 30 September: an 800-square-meter labyrinth whose interior walls of yew trees stand in repeating patterns that embody mathematical principles. Called the Archimedes Labyrinth, the maze's patterns include the Archimedes spiral—also seen in nature, in fiddlehead ferns. The maze also features symmetrical, repeating patterns called "wallpaper groups," which are common in the mosaics of ancient and medieval buildings, including Spain's Alhambra. (Although the patterns in principle can extend infinitely, the Bergen maze has an exit!) Fittingly, the garden will host festivities connected to next year's celebration of the Abel Prize, often referred to as the Nobel Prize of math.

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

PHOTO: ALASTAIR PHILIP WIPER/SCIENCE SOURCE

PARTICLE PHYSICS

Leaks put Italy's underground lab in jeopardy

Prosecutors charge Gran Sasso chiefs after chemical spills put spotlight on water safety

By **Edwin Cartlidge**

Scientists fear for the future of Gran Sasso National Laboratory, a world-leading underground physics lab in central Italy, after prosecutors charged four lab leaders with endangering drinking water supplies. Sparked by a number of accidental spills that released small amounts of toxic chemicals into groundwater feeding a local aqueduct, the 28 September legal action could lead to at least two major Gran Sasso experiments being shut down.

Gianpaolo Bellini, a particle physicist at the University of Milan in Italy and a former spokesperson for Borexino, one of the lab experiments in jeopardy, says fears of contamination are "groundless." But he says the lab itself is in a "very delicate situation." He worries that research groups, particularly from abroad, might be put off by the possibility of legal action and delays to their work. "This [investigation] damages the reputation of the lab," he says. "People will be more cautious about coming and therefore more cautious about investing their money."

The largest facility of its kind, Gran Sasso consists of three huge experimental halls carved out of a mountain next to a motorway tunnel that connects the cities of L'Aquila and Teramo in Italy's Abruzzo region. Sheltered from cosmic rays by 1400 meters of rock, the lab draws physicists from around the world to probe neutrinos, search for dark matter, and study other rare subatomic phenomena. But it has also attracted the ire of environmentalists. Some of the experiments rely on large tanks of organic compounds to detect subatomic particles, and critics worry that leaks could contaminate the surrounding mountain aquifer, which provides drinking water to hundreds of thousands of people.

Tensions came to a head in 2002 when researchers working on Borexino, which measures neutrinos from the sun, accidentally released some 50 liters of the hydrocarbon pseudocumene, which ended up in a local river. At the time, a judge in Teramo sealed off the hall containing Borexino, putting that detector out of action for 3 years and forcing another experiment to shut down prematurely. The leak also led the government to appoint a commissioner to improve safety at the lab.

But now the lab, run by Italy's National Institute for Nuclear Physics (INFN), again finds itself in trouble. Prosecutors from Teramo, backed by a 1000-page investigative report, accuse Gran Sasso Director Stefano Ragazzi, INFN President Fernando Ferroni, and the heads of the lab's environmental and technical operations of "negligence and imprudence" for having failed to correct safety flaws dating back to the 2002 spill.

The new investigation was triggered after researchers working on the CUPID neutrino experiment in August 2016 accidentally released a dichloromethane solvent, used to clean their detector's crystals. Small amounts of the solvent somehow ended up in Teramo's drinking water. The incident only came to light several months later, when the regional government indirectly revealed that the lab's water had been diverted away from the aqueduct. Augusto De Sanctis, president of the nonprofit environmental group Abruzzo Ornithological Station in Pescara, filed a complaint soon thereafter.

De Sanctis points out that other minor accidents have occurred, including a small air conditioner fire in June 2016 that shut down one experiment for several months, and a minor spill of chloroform in November 2016. But he is most concerned about the potential for a larger spill involving the 1300 tons of pseudocumene in Borexino and some 1000 tons of mineral spirits in the Large Volume Detector (LVD), which also studies neutrinos. De Sanctis argues that the use of the chemicals is illegal under a 2006 law forbidding the presence of dangerous substances within 200 meters of drinking water sources.

The prosecutors agree. They accuse the Gran Sasso management of failing to adopt "measures needed to remove" the lab's dangerous substances, particularly those in Borexino and the LVD. They also say the safety improvements ordered after the 2002

Gran Sasso National Laboratory sits within a mountain aquifer that feeds water to a local aqueduct.

spill were never completed. Costing €84 million altogether (including improvements to the road tunnels and aqueduct), this work was supposed to include resealing the lab floors and overhauling the drainage system. Eugenio Coccia, who was director of the lab at the time, says he can't say how much of the work was completed. "I was not responsible for the safety work," he says.

A spokesperson for INFN says Ferroni and Ragazzi don't want to speak about the "delicate situation." They say the lab's research is carrying on as normal, and the legal action "does not have a direct impact" on research. Spokespeople for Borexino and the LVD declined to comment.

In a 2017 statement, INFN argued that the 200-meter rule doesn't apply to Gran Sasso because the "underground laboratory infrastructure" predates the 2006 law. But earlier laws imposed the same minimum distance, De Sanctis says, and the 2006 law requires the removal of preexisting substances. In 2013, Italy's National Institute of Health told the lab that the 2006 law requires a "drastic reduction" in the lab's activities as long as nearby cities rely on groundwater from the mountain. The lab appears to be moving toward this concession: In January, Ferroni wrote to Ragazzi saying it is "indispensable, although painful" to remove most of the pseudocumene and mineral spirits by the end of 2020.

If lawyers for Ferroni and colleagues cannot persuade the prosecution to drop the case, a judge will then decide whether to put the scientists on trial. To try to avoid that, the lab could decide to disband Borexino and the LVD before 2020, De Sanctis says. "It would be strange if the lab didn't do something in response," he says.

Andrew Sonnenschein, a physicist at Fermi National Accelerator Laboratory in Batavia, Illinois, who worked on Borexino as a postdoctoral researcher just before the 2002 spill, says that accident was caused by "chaos in the management" of the plant used to purify Borexino's pseudocumene. But he insists the incident was "quite overblown" and says critics of Gran Sasso "don't think quantitatively" about the tiny risk involved. "Gran Sasso is a wonderful facility and it would be a shame if these experiments had to shut down," he says. "There is no public benefit to doing that."

De Sanctis sees things differently. "I am a supporter of scientific research," he says. "But research has to have its limits, particularly when it comes to basic rights such as the right to clean drinking water." ■

Edwin Cartlidge is a journalist in Rome.

CREDITS: (GRAPHIC) A. CUADRA/*SCIENCE*; (DATA) NASA

PLANETARY SCIENCE

NASA's next Mars rover aims to explore two promising sites

2020 rover could start with a river delta in Jezero crater

By **Paul Voosen**

Sometimes, a problem really can be solved by meeting halfway. For the past 4 years, planetary scientists have wrestled over where to send NASA's next Mars rover, a $2.5 billion machine to be launched in 2020 that will collect rock samples for eventual return to Earth. Next week, nearly 200 Mars scientists will gather for a final landing site workshop in Glendale, California, where they will debate the merits of the three candidate sites that rose to the top of previous discussions. Two, Jezero and Northeast Syrtis, hold evidence of a fossilized river delta and mineral springs, both promising environments for ancient life. Scientists yearn to visit both, but they are 37 kilometers apart—much farther than any rover has traveled.

Now, the Mars 2020 science team is injecting a compromise site, called Midway, into the mix. John Grant, a planetary scientist at the Smithsonian Institution's Center for Earth and Planetary Studies at the National Air and Space Museum in Washington, D.C., who co-leads the landing site workshops, says the team wanted to know whether a rover might be able to study the terrains found at Jezero and Northeast Syrtis by landing somewhere in the middle.

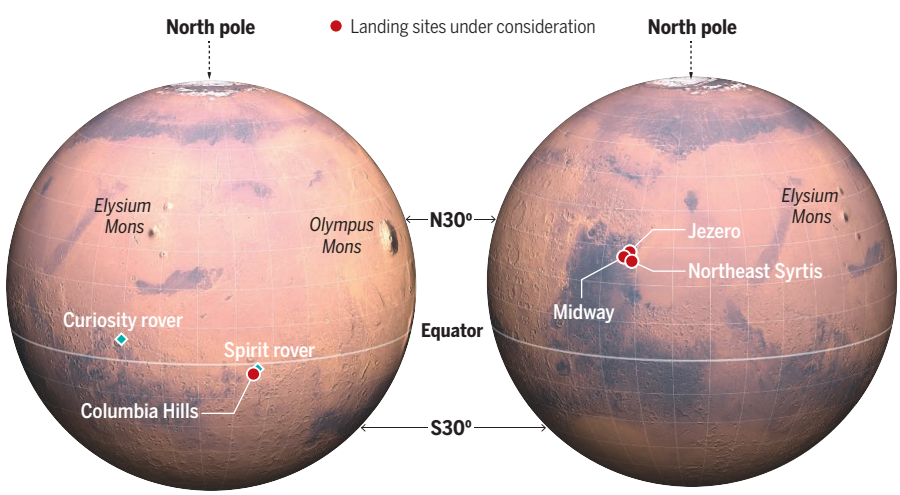
So far, the answer appears to be yes. The Mars 2020 rover borrows much from the design of the Curiosity rover that has been exploring another Mars site for 6 years. But it includes advances such as a belly-mounted camera that will help it avoid landing hazards during its harrowing descent to the surface. This capability allowed scientists to consider Midway, just 25 kilometers from Jezero and close enough to drive there. At the same time, Midway's rocks resemble those of Northeast Syrtis, says Bethany Ehlmann, a planetary scientist at the California Institute of Technology (Caltech) in Pasadena and member of the Mars 2020 science team.

Midway and Northeast Syrtis both hail from a time, some 4 billion years ago, when Mars was warmer and wetter—an era never explored by a Mars rover. Surveys from orbit suggest the sites harbor rocks that formed underground in the presence of water and iron, a potential food for microbes. The rocks, exposed on the flanks of mesas, include a layer of carbonate deposits that many scientists believe were formed by underground mineral springs. Sheltered from a harsh surface environment, these springs would have been hospitable to life, Ehlmann says. "We should go where the action was."

Nearby Jezero crater has its own allure, etched on the surface: a fossilized river delta.

A happy medium

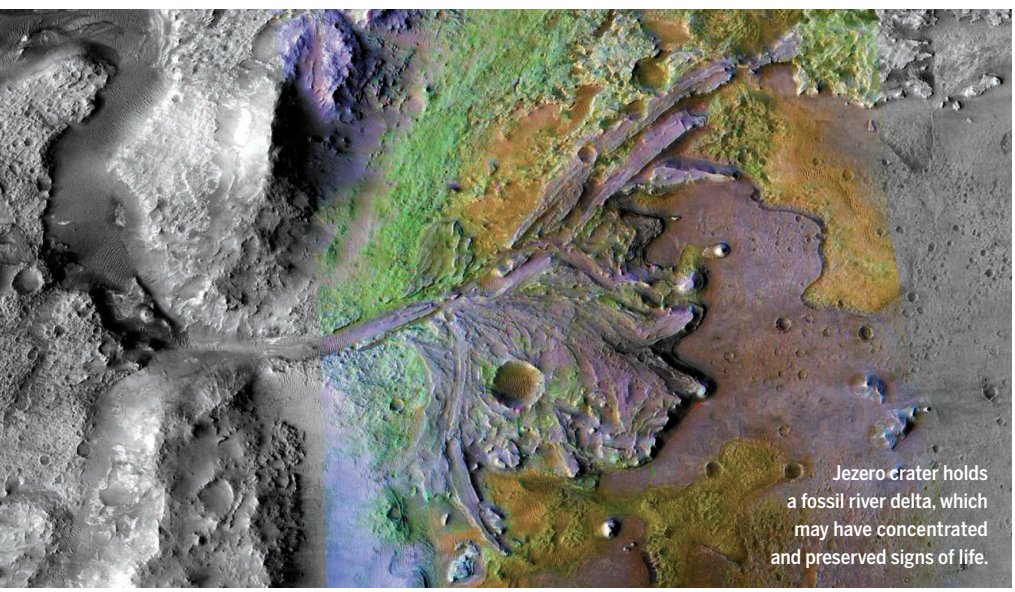
Jezero and Northeast Syrtis, two attractive landing sites for NASA's Mars 2020 rover, are close to each other. A new landing site, Midway, might allow the rover to study rocks from both terrains.



Nearly 4 billion years ago, water spilled into the crater, creating the delta. "It's right there," says Ray Arvidson, a planetary geologist at Washington University in St. Louis, Missouri. "It's beautiful." Geologists know deltas concentrate and preserve the remnants of life; they can see that on Earth in offshore deposits of oil—itself preserved organic matter—fed by deltas like the Mississippi's. New work to be presented at the workshop by Briony Horgan, a planetary scientist at Purdue University in West Lafayette, Indiana, will show that Jezero crater has a bathtub ring of carbonate—a strong sign that it once contained a lake. On Earth, such layers are often home to stromatolites, cauliflowerlike minerals created by ancient microbial life.

ing candidate site: Columbia Hills. "I have a sense there's a hill to climb," says the site's chief advocate, Steven Ruff, a planetary scientist at Arizona State University in Tempe. "I'll go in with a lot of questions of whether they can make that drive between Midway and Jezero." Columbia Hills sits within the large Gusev crater that the Spirit rover explored from 2004 to 2010. Driving backward while dragging a bad front wheel, Spirit gouged a trench that revealed opaline silica, a mineral that on Earth is a sure sign of life-supporting hot springs. Ruff has even proposed that the martian silica deposits are stromatolites.

The engineers building Mars 2020 will be glad to settle on a destination, says Matt Wallace, the rover's deputy project manager



Jezero crater holds a fossil river delta, which may have concentrated and preserved signs of life.

Right now, the Mars 2020 team favors landing at Jezero and driving uphill to Midway, says Matt Golombek, a planetary scientist at NASA's Jet Propulsion Lab (JPL) in Pasadena, and the other workshop co-leader. For the past year, the team has scoured potential routes between the two. "We haven't identified any deal-breakers," says Ken Farley, the mission's project scientist and a Caltech geologist. The rover's advanced autonomous driving should allow it to cover more ground than Curiosity, which often stops to plan routes. Even so, the path from Jezero to Midway would take nearly 2 years, Farley says. That means the rover could explore only one site during its primary 2-year mission, when it must drill and store 20 rock cores, to be picked up by future sample return missions. Exploration of the second site would have to come during an extended mission, after the rover's warranty expires. "The further away you land from your gold mine, the higher the risk you might not get there," Arvidson says.

Left out of those plans is the last lead-

at JPL. The lab's clean room is starting to fill up. The "sky crane" that will lower the rover to the surface is done. The spacecraft that will shepherd the rover to Mars is nearly complete—it just needs a heat shield, which is being rebuilt after testing revealed a crack. Several weeks ago, the chassis of the rover arrived and is now being filled with computers, batteries, and other electronics. Assembly of its complex drilling and sample storage system is underway, with other scientific instruments due by the end of February. "This is the mad scramble," Farley adds. "It is full bore get it done, get it done now."

At the workshop's end, scientists will vote on the candidates, followed by a closed-door meeting of the rover team to make a final choice. Engineers have deemed the sites safe for landing, Golombek adds, so it will come down to the science. The team's recommendation won't be the final word—the choice is ultimately up to NASA science chief Thomas Zurbuchen. But expect a decision within the next few months, if not sooner. ■

BIOMEDICINE

Replumbing the lymphatic system

Clinical trials for lymphedema begin to test approaches to restore lymph flow

By **Mitch Leslie**

Every night for the past 5 years, Patricia Egan has prepared for bed by pulling on a pair of chipped foam leggings that encase her limbs from her toes to the tops of her thighs. The custommade garments, which look like something a hockey goalie might wear, are supposed to prevent her legs from swelling. But they also make her feel claustrophobic and keep her awake, says Egan, 67, a retired writer from Chico, California. In the morning, she switches to less cumbersome toe gloves and knee-high compression socks that can still be uncomfortable.

Egan needs this bothersome and costly clothing because she suffers from lymphedema, a condition in which the fluid known as lymph doesn't drain normally from a limb or other part of the body. Many cases of lymphedema result from surgery for breast cancer and other tumors that severs the lymphatic vessels, which transport lymph and immune cells, but Egan's is a rare inherited form. "It's like being 8 months pregnant all the time," she says. "Your legs feel swollen and heavy, and they just won't move." As the lymph pools, it can also cause a host of other health problems.

Although Egan's compression clothing curtails swelling, it can't reestablish her lymphatic circulation—no U.S. Food and Drug Administration–approved therapy can. That's why she recently jumped at the chance to participate in a clinical trial at Stanford University School of Medicine in Palo Alto, California. Researchers there and at several other institutions were testing whether the drug ubenimex, a leukemia treatment used in Japan, can spur the growth of new lymphatic vessels.

Results of this phase II trial may be announced later this year, and researchers have started or are planning trials of other strategies for regenerating lymphatic vessels. If the treatments work, they could benefit many of the 5 million to 6 million

people in the United States—and more than 100 million around the world—with lymphedema. Improved lymph flow could also help in conditions other than lymphedema. Recent studies have implicated poor lymphatic drainage in glaucoma, rheumatoid arthritis, atherosclerosis, Alzheimer's disease, and other illnesses.

That lymphedema might be curable "was unthinkable a few years ago," says developmental biologist Guillermo Oliver of Northwestern University Feinberg School of Medicine in Chicago, Illinois. The new hope reflects growing knowledge about the lymphatic system. "We are starting to understand much better the molecular principles of how lymphatic vessels work," says vascular biologist Tatiana Petrova of the University of Lausanne in Switzerland. Researchers are also uncovering some of the molecules and cells that prevent lymphatic vessels from repairing themselves after they are damaged by cancer surgery, radiation treatment, or, in the developing world, parasite infections.

Lymphedema "starts out as a plumbing problem," says Stanford cardiologist Stanley Rockson, who first diagnosed Egan's lymphedema and is one of the researchers who ran the recent ubenimex trial. But over time, the uncleared lymph leads to "pretty profound changes in the structure of the affected part of the body," he says. One consequence is fibrosis—the proliferation of connective tissue—that can stiffen and thicken the skin. Fat cells proliferate, too, adding to the swelling.

The lymphatic system starts to grow before birth and is fully formed by soon afterward. However, adults can produce new lymphatic vessels, a process termed lymphangiogenesis, after injuries. The drug Rockson's group is studying blocks a lipid molecule, known as leukotriene B4, that curbs this regrowth. As Rockson and his colleagues reported last year in *Science Translational Medicine*, lymphedema patients have high levels of the molecule. The scientists also found that low concentrations of the molecule induce new lymphatic vessels to sprout. Ubenimex knocks leukotriene B4's concentration down to the progrowth range. "It unlocks the potential of the lymphatic vessels to repair themselves," says Rockson, who has no financial connection to the drug's manufacturer.

The biotech company Herantis Pharma in Espoo, Finland, is testing a more radical approach, gene therapy, to stimulate lymph vessel regeneration in women who developed arm lymphedema after lymph nodes in the armpit were removed during breast cancer surgery. During this treatment, patients undergo a second operation to replace the excised lymph nodes with nodes taken from elsewhere in their body. Although this transfer alone can combat lymphedema, it's unsuitable for many patients and others don't want to risk surgery. The Herantis approach aims to boost the effectiveness of the transfer and involves injecting the lymph nodes with a cold virus that carries the gene for the growth factor VEGF-C, which stimulates lymphangiogenesis. The immune system quickly eliminates the virus after the nodes are in place, notes the company's CEO, Pekka Simula, but its cargo will spur production of VEGF-C for about 2 weeks.

This woman's breast cancer treatment caused an extreme case of arm lymphedema.

PHOTO: VOISIN/SCIENCE SOURCE

Babak Mehrara, a reconstructive plastic surgeon at Memorial Sloan Kettering Cancer Center in New York City who treats lymphedema patients, questions the therapy's rationale. Sluggish lymphangiogenesis does not result from a shortage of the growth factor, he says. "Patients with lymphedema have tons of VEGF-C, but it's not generating functional lymphatic vessels." And because VEGF-C can promote tumor growth, he's wary of boosting its levels in patients recovering from cancer.

Simula counters that the treatment stimulates a pulse of VEGF-C that is unlikely to spark tumor growth. VEGF-C production is "local, transient, and highly specific to lymphangiogenesis," he says. Earlier this year, Herantis announced that the treatment appeared to be safe in a small trial that involved 15 patients, but it has yet to publish those findings. In June, the company started a phase II trial that will assess the approach's effectiveness in 40 women, who will receive either injections of the virus or a placebo shot.

Mehrara and his colleagues have followed a different strategy that capitalizes on recent findings about the immune system's role in lymphedema. "We have a very good understanding of the pathways of lymphedema, much better than we did 10 to 15 years ago," he says. One pathway involves the immune cells known as Th2 cells. As Mehrara's team has discovered over the past 5 years, these cells accrue after damage to lymphatic vessels. They swarm into the skin and pump out chemical signals called cytokines that curtail lymphangiogenesis and promote fibrosis. The researchers started a phase II trial of two antibodies that neutralize these cytokines but had to halt it when the drugs' manufacturer discontinued them. Now, they are planning a new trial using a different drug.

All three lymphedema approaches face significant hurdles. Even if a treatment does improve a patient's lymph flow, for instance, "the skin will still be fibrous and elephant-like, and the fat is still there" says physiologist Joseph Rutkowski of the Texas A&M College of Medicine in College Station. Flushing the excess lymph may eventually reverse fibrosis and fat accumulation, he says, but "we haven't had the means" to test it. Another issue is whether any successful treatment will be affordable for the millions of lymphedema patients in developing countries.

But in the Stanford trial, Egan, who has become a patient activist, says she has seen "the most remarkable improvement" in her lymphedema. She is certain she received the active compound rather than the placebo. "One morning I woke up and I could see an ankle bone, I could see a shin bone. I could see definition in my legs I hadn't seen in decades." The swelling in her legs has gone down so much, she says, that the leggings she wore at night were no longer snug. She is swapping them for a smaller, cheaper version that only covers her lower legs. ■

NOBEL PRIZE

Protein evolution earns chemistry Nobel

Darwinian principles applied to biomolecules led to new drugs, fuels, and detergents

By **Robert F. Service**

Breeders of everything from horses to tomatoes long ago learned to speed evolution by selecting and propagating offspring with desired traits. The Nobel Prize in Chemistry went on 3 October to researchers who developed tools to do the same with biological molecules. The findings have led to many commercial products, notably treatments for autoimmune diseases and cancer.

CHEMISTRY NOBEL

"for the directed evolution of enzymes" and "for the phage display of peptides and antibodies"

Frances Arnold

George Smith

Gregory Winter

Frances Arnold, a chemical engineer at the California Institute of Technology in Pasadena, received half of the prize for "directed evolution" of enzymes, proteins that carry out specific chemical reactions. Arnold is only the fifth woman to win the chemistry Nobel. But she takes heart in the rising number of women in the field: "You are going to see a steady stream of prizes, I predict, in chemistry given to women."

George Smith, a biologist at the University of Missouri in Columbia, and Gregory Winter, a biochemist at the Medical Research Council's Laboratory of Molecular Biology in Cambridge, U.K., shared the other half of the prize for using viruses that infect bacteria to direct the selection and evolution of new proteins, including antibodies.

Arnold says she turned to evolution in "desperation." In the early 1990s, she was struggling to re-engineer an enzyme called subtilisin so that it would work in organic solvents, rather than just in water. She tried the conventional approach of deliberately altering spots in the enzyme's makeup, but her changes failed repeatedly. So instead, she introduced random changes into the gene for subtilisin, then inserted copies of the mutated gene into bacteria, which produced thousands of variants of the protein. Initially she found one version that worked slightly better in the solvent than the original protein had. After two more rounds of introducing random mutations to the better-performing enzymes, she isolated a version that carried 10 mutations and worked 256 times better in the organic solvent than the original enzyme.

When Arnold started to use the random-mutation technique, she says, "we found beneficial mutations immediately. When we went and mapped those to the protein structures … we couldn't even explain them, much less predict them."

Arnold adds that many scientists resisted the idea of trying to improve proteins by trial and error. "The people who wanted to understand proteins were aghast. They said, 'That's not science.' I said, 'Well, I'm an engineer!'" Arnold recalls. "The industrial people, the people who had to make better proteins, said, 'Yeah, this makes total sense.'"

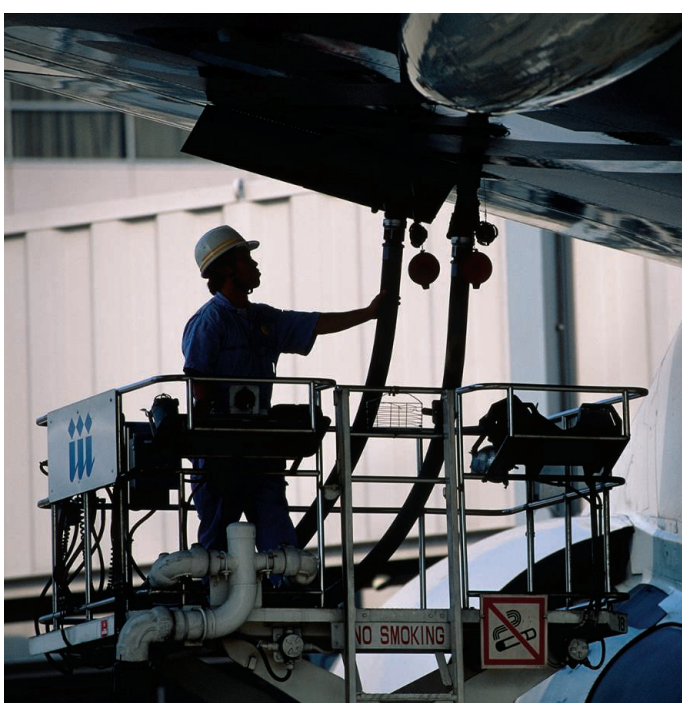
Plenty of scientists now agree. Researchers use directed evolution to improve enzymes that synthesize ethanol fuel and perform environmental cleanup. The method has also opened new vistas in science. Arnold's group has used it to give enzymes new functions, enabling microbes to incorporate new elements, such as silicon, into biomolecules. Others have evolved enzymes that can construct proteins from amino acid building blocks that natural life forms can't use. These unnatural proteins have yielded new drugs for treating cancer, heart failure, and infectious diseases in cattle. "It remains surprising to me how powerful evolution can be in the test tube," says John Dueber, a biochemist at the University of California, Berkeley.

For Smith, artificial selection was a tool for basic research. Like many others in the 1980s, he was struggling to identify unknown genes in humans and other organisms. Smith's idea was to use viruses that infect bacteria, called bacteriophages (phages), to trace a known protein to an unknown gene. He inserted a fragment of the mystery gene into the phage's own genetic material, so that when infected bacteria made new copies of the phage, the viruses would display protein fragments encoded by the gene on their surface. Smith then exposed the viruses to an antibody that binds to a known protein. When an antibody fished the protein out of the mix, the phage came along, which enabled Smith to identify the matching gene fragment.

Winter then turned Smith's idea on its head by engineering phages to display key pieces of antibodies on their surface, and then testing the phages against wide arrays of known proteins. This allowed him to identify (and later improve) genes that make antibodies to almost any protein. Winter's technique has allowed researchers to produce, among others, an antibody to a protein called TNF-α, which plays a role in many diseases. Today, the antibody, called adalimumab (Humira), treats rheumatoid arthritis, psoriasis, and inflammatory bowel disease, and it had sales of more than $18 billion last year, making it the top selling pharmaceutical. Dozens of other phage display–derived drugs are available or under development.

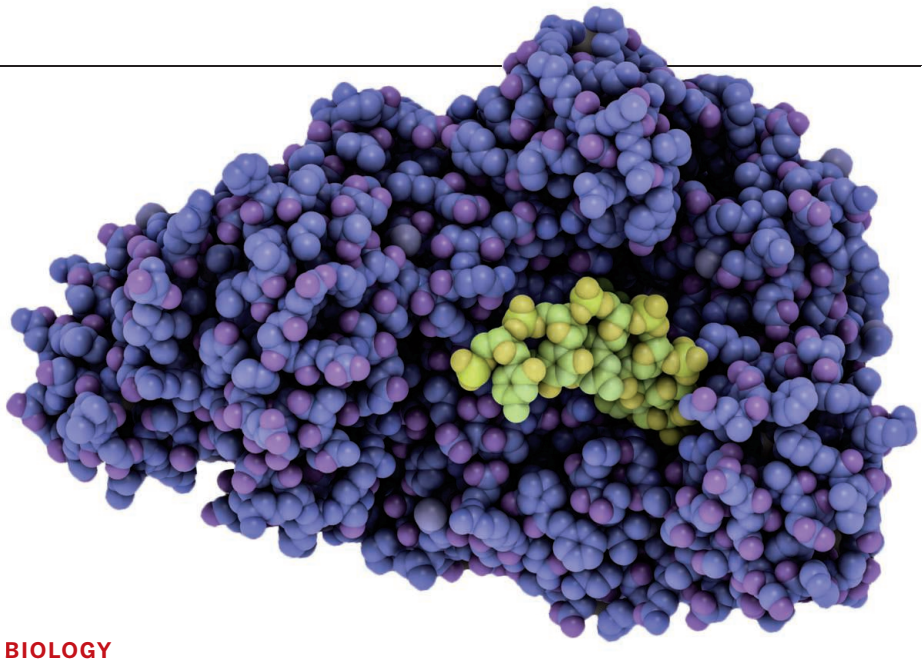
Thanks to the breakthroughs honored last week, evolution has become as powerful a force in technology as it is in nature. ■

With reporting by Lila Guterman and Daniel Clery.



Evolved enzymes help convert plant sugars to jet fuel.

PHOTO: AVIATION-IMAGES/NEWSCOM



A molecular structure shows an enzyme (purple) constructing a strand of DNA (green).

IMAGE: DNA SCRIPT

BIOLOGY

DNA printers poised to jump from paragraphs to pages

Enzymes that write DNA could turbocharge synthetic biology and data storage

By **Robert F. Service**

Scientists can read DNA sequences faster than ever before. But writing DNA hasn't kept pace. Those wanting made-to-order DNA for purposes such as synthetic biology make do with short strands, synthesized in a slow and expensive chemical process.

That appears ready to change. Last week, researchers from a biotechnology startup announced at a synthetic biology meeting in San Francisco, California, that by using close relatives of the DNA-writing enzymes in living things, they can build DNA strands as long as 150 "letters," or nucleotide bases. That's up from a record of 50 nucleotides just a few months ago, and nearly on par with the standard chemical approach.

"This is a significant milestone," says George Church, a geneticist at Harvard Medical School in Boston, who is working to develop similar techniques. The result, from Paris-based DNA Script, puts enzymatic DNA synthesis "on a curve that looks to grow exponentially." If it does, Church and others say, researchers wanting to create new genomes or use DNA as a storage medium that can archive vast troves of information will soon have longer DNA fragments made faster and more cheaply.

Although traditional chemical DNA synthesis has been miniaturized and automated, the underlying technique, known as phosphoramidite chemistry, is virtually unchanged since it was developed in the early 1980s. It involves adding nucleotide bases one at a time, each capped with a protecting group that stops it from reacting to extend the strand until scientists remove the cap and add the next base.

The approach isn't perfect. With each letter added, there's a 0.5% chance of an error. The longer the chain, the greater the chance that it will contain an error, effectively limiting DNA strands to about 300 bases long. As a result, researchers who hope to write genes containing thousands of letters must laboriously stitch fragments together. Enzymatic synthesis promises to do better by coopting polymerases, the enzymes living things use to fasten nucleotides together into long, virtually error-free sequences.

Although commercial enzymatic DNA synthesis efforts started only this decade, today about a half-dozen companies are working on the approach, according to Church. Others are pushing data storage and biomedical applications to be ready when DNA printing rates improve. According to Michael Kamdar, CEO of Molecular Assemblies, an enzymatic DNA synthesis company in San Diego, California, the market for traditional chemical DNA synthesis today is about $1 billion annually. The market for data storage is more than $14 billion, although DNA likely suits only a small fraction of it. "In the next 2 to 3 years you will see applications on the market with enzymatic DNA synthesis, if not by us then someone else," says Sylvain Gariel, co-founder and chief operating officer of DNA Script.

Shifting from chemical synthesis to polymerases has brought challenges. In living cells, most polymerases start with a template strand and create one that's complementary, pairing As with Ts and Gs with Cs. A specialty polymerase in immune cells, called terminal deoxynucleotidyl transferase (TdT), works without a template, making it scientists' most common choice for enzymatic synthesis. Yet because TdT adds new DNA letters randomly, researchers have had to find ways to force it to add just the desired letters, one at a time. Gariel's team does so by equipping each DNA base with a proprietary protecting group that, as in chemical synthesis, prevents TdT from adding more than one letter to the growing strand. After the correct letter is added and its protecting group is removed, the cycle repeats. According to Gariel, adding each letter takes only 5 minutes and occurs with an accuracy of 99.5%.

"That's great," says William Efcavitch, co-founder and chief scientific officer of Molecular Assemblies. That speed and accuracy, along with the 150-nucleotide-long strands, bring enzymatic synthesis nearly on par with the 5- to 10-minute cycle of conventional chemical DNA synthesis.

Many experts believe the enzymatic approach has plenty of room for improvement. "The potential of enzymatic synthesis far surpasses chemical synthesis," Kamdar says. Ultimately, DNA Script CEO Thomas Ybert says he expects his company will write 1000-base strands in a day. The company also hopes to begin to sell automated enzymatic DNA synthesizers by early 2020.

Such improvements could cut the cost of writing DNA by one to two orders of magnitude, Efcavitch predicts. That would make it easier and cheaper for synthetic biologists to design and test new genes for developing everything from new catalysts to medicines. It could also revolutionize DNA data storage, which, in theory, could capture all the world's information in a volume smaller than a suitcase.

For the past several years, such applications seemed fanciful, because fast and cheap enzymatic DNA synthesis lay over the horizon. Kamdar says: "The reality now is that the horizon is here." ■

A NEW LEAF

A decade ago, France launched an ambitious effort to cut pesticide use by half. It failed. Now, the country is trying again

By **Erik Stokstad**, *near Ancenis, France*

Mikael Fremont was up to his shoulders in rapeseed, looking pleased. The bright yellow flowers had faded, giving way to long slender pods that would yield valuable canola oil, as well as protein meal for livestock. While his spaniel darted about the farm here in May, the burly 41-year-old grower wanted to point out something nearly hidden: a mat of tattered, dead leaves covering the soil.

Months earlier, Fremont had planted this vetch and clover along with the rapeseed. The two legumes had grown rapidly, preventing weeds from crowding out the emerging rapeseed and guarding it from hungry beetles and weevils. As a result, Fremont had cut by half the herbicide and insecticide he sprayed. The technique of mixing plant species in a single field had worked "perfectly," he said.

This innovative approach is just one of many practices, now spreading across France, that could help farmers achieve an elusive national goal. In 2008, the French government announced a dramatic shift in agricultural policy, calling for pesticide use to be slashed

PHOTO: PIERRE HUGUET/SCIENCE SOURCE

Grapes are often sprayed with copper sulfate, a mildly toxic fungicide invented in France in 1884.

in half. And it wanted to hit that target in just a decade. No other country with as large and diverse an agricultural system had tried anything so ambitious. The goal "was very revolutionary," says Henriette Christensen of the Pesticide Action Network in Brussels, especially because France is the second largest consumer of pesticides in Europe.

Since then, the French government has spent nearly half a billion euros on implementing the plan, called Ecophyto. It created a network of thousands of farms that test methods of reducing chemical use, improved national surveillance of pests and plant diseases, and funded research on technologies and techniques that reduce pesticide use. It has imposed taxes on farm chemicals in a bid to decrease sales, and even banned numerous pesticides, infuriating many farmers.

The effort has helped quench demand on some farms. Overall, however, Ecophyto has failed miserably. Instead of declining, national pesticide use has increased by 12%, largely mirroring a rise in farm production. "We lost 10 years since 2008," says François Veillerette of Générations Futures, an environmental advocacy organization in Paris. "We can't afford to waste 10 more."

The French government agrees. Officials are now finalizing a revised plan dubbed Ecophyto 2+. It will boost research, add demonstration farms, increase taxes on pesticides, and prohibit more compounds. President Emmanuel Macron has even urged a ban of glyphosate, the world's best-selling weed killer and an important tool for many farmers.

Details of the revised plan, including funding levels, are still being decided. But some observers are already skeptical. Farmers fear burdensome rules and increased costs that will put them at a competitive disadvantage. Environmental organizations worry France will again fall short. "There are good ideas," says Carmen Etcheverry, formerly of France Nature Environnement in Paris. "But we don't know how they will be implemented."

There is also optimism. Despite Ecophyto's failure, it showed farmers have powerful options, such as mixing crops, planting new varieties, and tapping data analysis systems that help identify the best times to spray. With the right incentives and support, those tools might make a bigger difference this time around. And the fact that France isn't backing away from its ambitious goal inspires many observers. "You feel," says Robert Finger, an agricultural economist at ETH Zurich in Switzerland, "that something vivid is going on."

AFTER WORLD WAR II, synthetic herbicides, fungicides, and insecticides greatly boosted French farmers' harvests and profits. But the chemicals contaminated groundwater, lakes, and streams, and they harmed farm workers and wildlife. Consumers became wary, and by the 1970s public opposition to pesticides was growing.

In French fields of rapeseed, seen here in flower, planting companion crops is reducing the need to spray weed killers and insecticides.

During the 2000s, pesticide sales fell as farmers applied them with increasing efficiency and sometimes switched to more effective compounds that required smaller doses. But the ambition to do much better crystalized in 2007, when then–French President Nicolas Sarkozy convened a conference to set a 5-year environmental agenda. Ecophyto was the result, negotiated between environmentalists, farm unions, pesticidemakers, and others. It included a major political concession; the 50% cut would be reached "if possible," which meant that much of the plan was voluntary.

Still, Ecophyto served as a catalyst. Research funds were targeted at evaluating smarter ways to use pesticides. Approaches were tested on some 3000 farms that joined a demonstration network. Officials recruited observers around the country to scout for pests and plant diseases and provide weekly reports; the surveillance helps farmers decide when spraying might be a waste of resources. On average, farms in the demonstration network decreased their pesticide use by 18%, and most did it without sacrificing profits.

But France's overall use of chemical pest control went up (see chart, p. 147). Many factors contributed, analysts say. Taxes on chemicals, for example, weren't high enough to influence buying decisions. It was difficult to persuade some farmers to adopt new practices or technologies that might add to their costs or decrease yields. Ecophyto's funding—about €70 million a year since 2016—was too low and "out of all proportion to the challenge," France's inspector general concluded late last year. And market forces, such as high prices for cereals, may have created an incentive to spray more chemicals to protect unusually lucrative harvests.

Yet veterans of Ecophyto aren't discouraged. On many farms, analysts say, it appears that existing technologies and practices alone could cut chemical use by at least 20%. "Without underestimating the extent of the challenge, there are reasons to be optimistic," says Alain Tridon, director of the Ministry of Agriculture's plant health services in Paris.

ONE SOURCE OF THAT OPTIMISM can be found on the Ducourt Estate, 450 hectares of rolling vineyards in Ladaux, France, that produces about 3 million bottles of wine each year. In a long garage, massive four-wheeled tractor-sprayers stand 3 meters tall. Each carries a 2200-liter tank for fungicide. Their articulated arms, studded with nozzles, can spray chemicals on four rows of grapes in one pass, killing mildew and other plant pathogens.

Winemakers are France's biggest users of fungicides (see sidebar, p. 146), although most are based on sulfur and copper, rather than more toxic synthetic molecules. Still, the sight of Ducourt's yellow beasts trundling through the vineyards can unnerve estate neighbors worried about farm chemicals, says Jeremy Ducourt, who helps manage the family owned business. The machines are actually "a big part of the solution," he says. That's because they helped the estate reduce its use of fungicides by about 30%, thanks to nozzles that put more fungicide on the plant and less on the ground. The most advanced sprayers even collect and reuse any lingering mist.

PHOTO: GERARD SIOEN/GAMMA-RAPHO VIA GETTY IMAGES

Vive la resistant vines!

Winemaking is an art, but it couldn't happen in France without a lot of chemistry. The battle against two major diseases—powdery mildew and downy mildew—means vineyards account for about 14% of all pesticide sales nationwide. "If I don't spray, I don't have grapes," says Jeremy Ducourt, who helps manage his family's 450-hectare Ducourt Estate in Ladaux.

Many estates are adapting their practices to reduce spraying (see main story, p. 144). But Ducourt and a few others are also counting on the grapes for help. In 2014, Ducourt persuaded his family to plant a few hectares of two varieties, bred in Switzerland, that carry mildew-resistance genes added from wild relatives in North America and Asia. With this genetic protection, the vines require 80% to 90% less spraying.

The new breeds aren't perfect; branches can be brittle and the mildew resistance is only partial. But that didn't stop Ducourt from planting an additional two varieties of resistant vines this spring.

Even better vines are on the way. After decades of work, French plant breeders are testing grapes with more durable resistance to mildew. "It is a major breakthrough," says Francois Delmotte, a plant pathologist with the French National Agricultural Research Institute in Bordeaux, France, who leads an effort to monitor resistance efficacy. Breeders are now fine-tuning the vines to the regional needs of growers. Introductions could be slow, because growers replace just 3% of the long-lived grapevines each year.

Disease-resistant grapes helped vintner Jeremy Ducourt cut fungicide use.

Vintners also face marketing headaches. Ducourt, for example, must sell wine made from their resistant grapes as table wine, which fetches a lower price than bottles labeled as coming from Bordeaux. The grapes have not yet been officially recognized by a body that regulates France's wine growers. Many growers won't plant the new grapes until that official blessing is granted. But Ducourt isn't too worried. It's just part of the burden, he says, of being a pioneer. — *Erik Stokstad*

Similar high-efficiency sprayers are available for other crops, and just replacing older models with newer machines could make a dent in France's chemical use. But upgrades don't come cheap. The Ducourt Estate's sprayers, which double as harvesters, cost about €320,000. Add the fact that only 3% of the nation's 200,000 sprayers are replaced each year, and it could take decades to fully upgrade the fleet.

The Ducourt family has also cut fungicide by using decision support software. The program draws on timely weather, surveillance, and other information, such as the size of leaves, to advise when to spray. The tool can reduce fungicide use by about 20% in vineyards, and cereal growers have seen similar results. But such tools haven't yet spread to all farms. Potato farmers, who also spray copious fungicides, now use the tools on about half of their fields, but aim to increase that share to 90% within 5 years.

When it comes to insects, it's much more difficult for software to predict outbreaks in fields. So, farmers must diligently scout their fields so that they can apply insecticides before pests multiply out of control. Ecophyto 2+ aims to boost a noninsecticide approach called biocontrol. In this long-standing approach, farmers confuse pests with pheromones, for example, or seek to reduce populations by introducing the pest's natural predators. Advocates highlight the strategy's success in France's ample fields of maize. There, a tiny introduced parasitic wasp called *Trichogramma brassicae* has become a key weapon against the corn rootworm, a major pest. The wasps lay their eggs inside the eggs of the rootworms, shrinking populations just as effectively as insecticides when conditions are optimal.

The wasps are not a panacea. Although the costs are roughly comparable to insecticides, more labor is required to hang the cardboard cartons holding the wasps on maize plants. And insecticides remain more popular in southern France, where maize farmers face multiple pests the wasps don't attack. (In other nations, maize farmers control pests with less insecticide by planting genetically modified plants, but engineered crops are not allowed in France.) Despite such limitations, the wasps are now used on 23% of maize hectares where rootworms pose a threat.

THE MIXED CROP TECHNIQUE used by Fremont in his fields of rapeseed demonstrates another use of biology, in this case to control weeds. It's the kind of ancient technique that used to be commonplace. In August, one or more fast-growing legumes are planted between the rows of rapeseed. There's enough space that the legumes don't steal too much water or light, but they keep down weeds and, as a bonus, release nitrogen, a fertilizer. They also seem to minimize insect attacks, although this benefit hasn't been conclusively demonstrated. By the time frost kills the legumes, the rapeseed has grown thick enough that few weeds can challenge it.

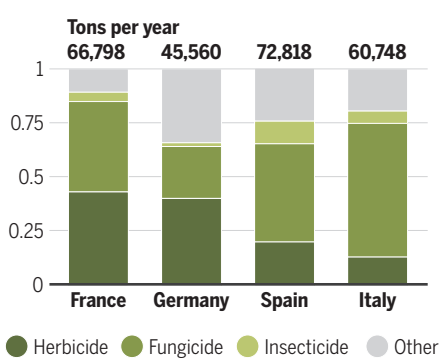
Such mixed cropping "is becoming very popular," says Marie-Hélène Jeuffroy, an agronomist with the French National Agricultural Research Institute (INRA) in Versailles. Nationwide, 3% to 5% of France's rapeseed hectares are now coplanted with legumes. That share could grow to 30% by 2030 under a pledge made in July by the French federation of oilseed producers.

One French seed company—Jouffray Drillaud, based in Cisse—sees enough potential in crop mixtures that 2 years ago it stopped selling herbicides, which generated 20% of its revenue. "When you have more diversity, you have more resilience," says Vincent Béguier, R&D director of the firm, which now focuses its weed control on mixed cropping and other nonchemical approaches. "Simplicity is the worst thing for agriculture."

So far, rapeseed growers appear to be reaping the biggest benefits in weed control from mixed cropping. But scientists are searching for other possibilities. Jeuffroy and other participants in ReMIX, a new

A mix of poisons

Pesticide sales in Europe's four leading markets differ because of factors including crop type, climate, soil, and regulations (average sales for 2011–15).



CREDITS: (PHOTO) DUCOURT; (GRAPHIC) N. DESAI/*SCIENCE*; (DATA) EUROSTAT

€5 million research collaboration among 13 European countries, are studying how to optimize mixtures, measure benefits, and remove obstacles to mixed cropping.

France's Ministry of Agriculture is moving to encourage greener approaches by requiring pesticide retailers to inform farmers about 36 alternatives to spraying. Instead of only touting insecticides to kill pests, for example, a dealer might recommend a crop mixture, or traps baited with sexual pheromones to confuse male insects, interfering with reproduction. The goal is to reduce the number of pesticide doses they sell by 20% by 2021. Dealers that miss the goal could face penalties.

To reach Ecophyto's goal of a 50% cut, however, many farmers will need to make more use of another practice—crop rotation. Alternating what's planted in a field, ideally over 5 or 6 years, is among the most effective ways to fight weeds, soil-borne pests and diseases. Switching between peas, wheat, and sugar beets, for example, can prevent pathogens from building up in the soil year after year, while swapping in a pasture grass hinders annual weeds.

Although simple in concept, it can be hard to increase the diversity of crops in rotation. That's because the whole system is locked: Farmers in many regions have specialized in certain crops—such as wheat or potatoes—and rely on finely tuned methods to produce high yields. There is often no nearby market for additional crops, because storage and processing facilities also tend to specialize in dominant crops—as do researchers, advisers, and policymakers. "Everything has been organized around major crops with high use of inputs," says Antoine Messéan, an agronomist with INRA. "It's difficult to get out of this self-reinforcing mechanism."

Crop diversification is not a top priority in the new version of the Ecophyto plan, but the Ministry of Agriculture has asked INRA for advice on how to encourage it. In a related effort, France hopes to double the amount of organic farming, which does not allow synthetic pesticides, to 15% of hectares by 2022. In May, the Ministry of Agriculture announced it will spend €1.1 billion to support organic expansion.

THE GOVERNMENT ALSO FACES growing pressure from environmentalists to ban more farm chemicals. The approach is controversial, and farmers complain that greener alternatives aren't always available. After an insecticide called dimethoate was banned in France in 2016, for instance, cherry growers had no effective way to fight an invasive fruit fly, says Eric Thirouin, deputy secretary general at the National Federation of Farmers' Unions in Paris. Meanwhile, the insecticide remains legal in Spain and Italy, he notes, putting French cherry growers at a disadvantage.

In other cases, banning one chemical can cause the use of others to spike, undermining reduction efforts. French wheat growers, for example, relied on neonicotinoids, which are coated on seeds, to protect the plants against aphids and leaf hoppers. Now that they are banned, some growers might increase applications of other insecticides. And there are other kinds of trade-offs. Some specialists fear banning the weed killer glyphosate could increase erosion or greenhouse gas emissions, if farmers start to till the soil to remove weeds. More research on such trade-offs is "urgently needed," Finger says.

In the meantime, it will be crucial to enlist France's farmer-owned cooperatives in making Ecophyto 2+ a success, observers say. French farming is dominated by a handful of these enormous agri-businesses. They buy and trade harvests, and most sell their members seeds, fertilizer, and pesticides. "You cannot reduce pesticides if you don't convince cooperatives that they should change their business model," Messéan says. A few have made strides in this direction, such as Terrena, the €5 billion cooperative headquartered near Ancenis that encouraged Fremont to adopt crop mixtures.

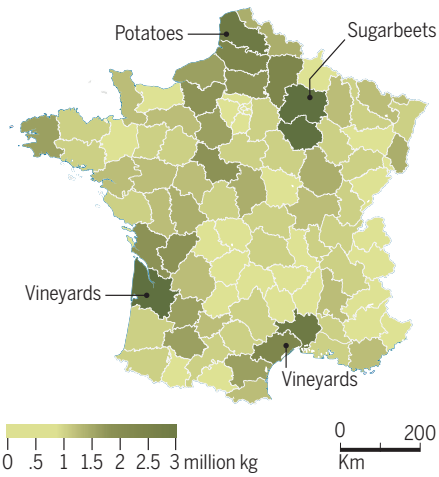
Although the majority of French farmers largely ignored or resisted Ecophyto, they are now showing signs of support. In July, more than 40 farmer organizations issued a "Contract for Solutions" that included pledges to reach specific reduction goals. The pledges represent a turning point, says Thirouin, as farm groups are no longer focused just on fighting pesticide bans. "The idea was to step aside from this defensive position and be proactive," he says. Tridon also sees it as a positive step. "We are really seeing a shift in mindset."

It's not only farmers who will have to adjust if France is to meet its ambitious goals. Reducing the cost of food production to the environment and public health will likely increase the cost to consumers and taxpayers. "Everything is possible," says Eugénia Pommaret, director of the Union of Plant Protection Industries, a pesticide trade group in Paris. "It's just a question of costs."

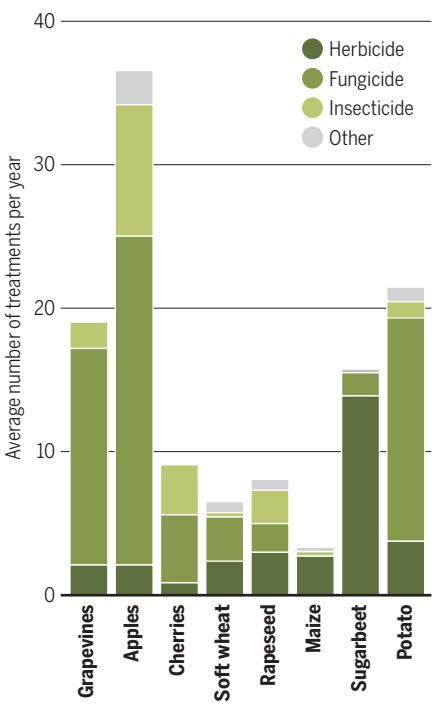
The key to change will be collaboration among all the players in the food system, adds Florence Leprince, an agronomist at Arvalis, a technical institute for arable crops in Montardon, France. "Solutions exist, but they are far from covering all the needs," she says. "It's more about increasing the commitment of everyone to change the way of producing." ■

France's pesticide challenge

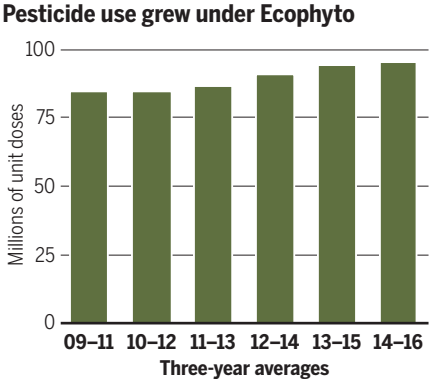
The use of pesticides (bottom) is most concentrated in a few growing regions (top) and certain crops (middle).



Pesticide intensity by crop



Pesticide use grew under Ecophyto



CREDITS: (GRAPHIC) N. DESAI/*SCIENCE*; (DATA, TOP TO BOTTOM) NATIONAL BANK OF DISTRIBUTOR SALES (BNV-D). TREATMENT: MINISTRY FOR AN ECOLOGICAL AND INCLUSIVE TRANSITION; BNV-D. TREATMENT: MINISTRY OF AGRICULTURE AND FOOD; MINISTRY OF AGRICULTURE AND FOOD

A lawn is dominated by decorative parterre in the formal (French style) garden of the Royal Domain of Drottningholm, Sweden.

PERSPECTIVES

SUSTAINABILITY

An alternative urban green carpet

How can we move to sustainable lawns in a time of climate change?

By **Maria Ignatieva**[1,2] *and* **Marcus Hedblom**[3,4]

Lawns are a global phenomenon. They green the urban environment and provide amenable public and private open spaces. In Sweden, 52% of the urban green areas are lawns (*1*). In the United States, lawns cover 1.9% of the country's terrestrial area (*2*) and lawn grass is the largest irrigated nonfood crop (*3*). Assuming lawn would cover 23% of cities globally [on the basis of data from the United States and Sweden (*1*)], it would occupy 0.15 million to 0.80 million km^2 (depending on urban definitions)—that is, an area bigger than England and Spain combined or about 1.4% of the global grassland area. Yet, lawns exact environmental and economic costs, and given the environmental and economic impacts of climate change, it is time to consider new alternative "lawnscapes" in urban planning as beneficial and sustainable alternatives.

Although lawns are widespread, their properties have received less attention from the scientific community compared to urban trees or any other types of green areas. Designers, urban planners, and politicians tend to highlight the positive ecosystem services provided by lawns. For example, lawns produce oxygen, sequestrate carbon, remove air pollution (although this has not been supported by good quantitative studies), reduce water runoff, increase water infiltration, mitigate soil erosion, and increase groundwater recharging (*4*). But perhaps the most important positive ecosystem service is the aesthetic and recreational benefits they provide. Aesthetics are a primary factor in modern urban planning and landscaping practice. For example, in developing countries located in arid zones, designers argue that lawns and irrigated turfs considerably enhance the quality of urban life (*5*).

Recent heat waves and an increasing prevalence of droughts have raised economic and environmental concerns about the effects of urban lawns on climate change. These circumstances have encouraged researchers and the public to reconsider the green-carpet concept and assess the controversial aspects of lawns. It has been argued that lawns moderate urban temperatures, but this when compared to the absence of any vegetation. In arid regions of the United States, lawn irrigation accounts for 75% of the total annual household water consumption (*6*). In Perth, Australia, the annual volume use of groundwater for irrigating public greenspaces is 73 gigaliters (Gl) and an additional 72 Gls of water in unlicensed backyard bores for private lawn irrigation (*7*). Another concern is the contamination of groundwater or runoff water due to overuse of fertilizers, herbicides, and pesticides. In 2012, the U.S. home and garden sector used 27 million kg of pesticides (*8*). The positive effect of soil carbon sequestration on the climate footprint of intensively managed lawns was found to be negated by greenhouse gas emissions from management operations such as mowing, irrigation, and fertilization (*9*, *10*). Gasoline-powered lawn mowers emitted high amounts of carcinogenic exhaust pollutants (*11*). Moreover, substituting degraded open green areas with plastic lawns eliminates real nature from cities and arguably reduces overall sustainability given that they

[1]*School of Design, University of Western Australia, 35 Stirling Highway, Perth, WA 6001, Australia.* [2]*Department of Urban and Rural Development, Swedish University of Agricultural Sciences, Post Office Box 7012, Uppsala SE-75007, Sweden.* [3]*Department of Ecology, Swedish University of Agricultural Sciences, Post Office Box 7044, Uppsala SE-75007, Sweden.* [4]*Department of Forest Resource Management, Swedish University of Agricultural Sciences, Umeå SE-90183, Sweden. Email: maria.ignatieva@uwa.edu.au; marcus.hedblom@slu.se*

PHOTO: MARIA IGNATIEVA

reduce habitats, decrease soil organisms, pollute runoff water, and may well have yet unknown negative consequences for human health through plastic particles.

However, the most noticeable constraint to new alternative lawn thinking may be the contribution of lawns to urban aesthetic uniformity (*12*) and urban ecological homogenization, where lawn plant communities become similar across different biophysical settings (*13*). From the vast variety of the grass genera, only a limited number of species are selected for lawns. In cold-season countries, *Festuca rubra*, *Lolium perenne*, *Poa pratensis*, and *Agrostis* spp. are the most frequently used grasses, and in warm-season countries, *Cynodon dactylon, Paspalum* spp., *Stenotaphrum secundatum*, and *Zoysia japonica* are the dominant species. In many regions, lawn grasses are becoming invasive. The main intent of lawn care is to keep the dominance of only a few species and eliminate the presence of others (all considered as lawn weeds).

The reason behind lawn uniformity may lie in its origin. Grass plots in ornamental gardens most likely appeared in medieval times and were probably obtained from the closest pastures and meadows. They were small and quite biodiverse (containing a large number of meadow herbaceous plants). In the 17th century, the role of lawns increased in the decorative grounds of geometrical gardens. For example, in iconic French Versailles, short-cut green grass was perfectly blended with the ideology of the power of man over nature. The English landscape, or the "natural" style of the 18th century, introduced an idealized version of urban nature: grazed grasslands with scarcely planted shade trees and the "pleasure ground" next to the mansion with a short, smoothly cut lawn. With the introduction of mowers and lawn-seed nurseries, the English pastoral vision flourished further in the public parks during the second half of the 19th century. In the 20th century, the modernistic prefabricated landscape was based on the same English picturesque model as the "natural" landscape, which was often mistaken for ecological quality. Consequently, monoculture and intensively managed lawnscapes dislodged the majority of native zonal plant communities in urban environments. At the beginning of the 21st century, perfect green lawns became part of the uptake by non-Western countries of the ideal Western lifestyle and culture (*12*).

Since the end of the 19th century, there have been numerous attempts to create alternatives to homogeneous short-cut conventional lawns. For several decades, British researchers tried to promote the idea of an urban meadow (infrequently mown) as an alternative to high-maintenance lawns. There are different typologies of such urban meadows: (i) Sheffield's "naturalistic herbaceous plantings" (mixed native-exotic meadows of native grasses and different herbaceous plants with added planted exotic forms from the Himalayas and East Asia, Caucasus, or the United States); (ii) pictorial meadows generated from seed mixtures of native and exotic colorful annuals [the most visible example of this realization (i and ii) is the Queen Elizabeth Olympic Park in London]; and (iii) the use of mostly native meadow (prairie and tussock grasslands) plants in urban parks and private gardens, popularized in the United Kingdom, Northern Europe, United States, and New Zealand that was inspired from rich zonal vegetation (*12*). Unfortunately, none of these ideas flourished because the lawn has been an unquestioned norm and people perceive lawns as urban nature.

An alternative grass-free (tapestry) lawn boasts native plants in the Ultuna Campus of the Swedish University of Agricultural Sciences, Uppsala, Sweden.

What is the state of alternative lawnscapes today? The most dramatic implementation of new lawn thinking is in Berlin, Germany, where spontaneous vegetation is accepted as a fundamental landscape design tool. In Gleisdreieck Park and Südgelände Nature Park (both established on abandoned railways), some areas were left to "go wild" and thus have been colonized by spontaneous vegetation. This has successfully challenged societal norms of conventional green lawns and has been accepted by local people. Recent research in the United Kingdom and Sweden showed that people desire to change the monotonous lawn to a more diverse environment (*12*, *14*). Grass-free lawn is one of the latest movements in both nations and is based on the use of low-growing native (in Sweden) or native and exotic (in the United Kingdom) herbaceous plants without the use of grasses. The goal is to create a dense biodiverse and low-maintenance mat, which can be used for recreation (*15*).

What about future research on urban lawns? Of course, new plant species that can survive heavy tramping or long drought, for example, are desirable. Beyond plant research, studies on planting design must go beyond the theoretical. Specific geographical, cultural, and social conditions of each country must be factored into such alternative lawn planning research. In Australia, South Africa, and Arizona, alternative lawns might be based on xeriscape local plants. In Chinese cities, historically proved groundcover species could provide sustainability and a sense of place. Lawns could become a universal model for experimental sustainable design and urban environment monitoring.

New alternative lawns represent a new human-made "wild" nature, which is opposed to the "obedient" nature of conventional lawns. Creating a new, ecological norm requires demonstration displays, public education, and the introduction of "compromised" design solutions. For example, a strip of conventional lawn can frame prairie or meadow and other "wild" vegetation and show a presence of culture in nature. The use of colors such as gray, silver, yellow, and even brown in lawnlike covers can add a feeling of real nature. Thus, one crucial challenge is how to accelerate people's understanding of sustainable alternatives and acceptance of a new vegetation aesthetic in urban planning and design. ■

REFERENCES AND NOTES

1. M. Hedblom, F. Lindberg, E. Vogel, J. Wissman, K. Ahrne, *Urban Ecosyst.* **20**, 1109 (2017).
2. G. L. Thompson, J. Kao-Kniffin, *Crop Sci.* **57**, S238 (2017).
3. M. Simmons, M. Bertelsen, S. Windhager, H. Zafian, *Ecol. Eng.* **37**, 1095 (2011).
4. J. A. Monteiro, *Urban For. Urban Green.* **26**, 151 (2017).
5. E. M. Elgizawy, *Procedia Environ. Sci.* **34**, 131 (2016).
6. C. Milesi *et al.*, *Environ. Manage.* **36**, 426 (2005).
7. D. Lieb, D. Brennan, D. McFarlane, "The economic value of groundwater used to irrigate lawns and gardens in the Perth metropolitan area" (Commonwealth Scientific and Industrial Research Organisation, National Research Flagships, 2007).
8. Environmental Protection Agency (EPA), "Pesticides industry sales and usage, 2008–2012 market estimates" (EPA, 2017); www.epa.gov/sites/production/files/2017-01/documents/pesticides-industry-sales-usage-2016_0.pdf.
9. C. Gu, J. Crane II, G. Hornberger, A. Carrico, *J. Environ. Manage.* **151**, 233 (2015).
10. P. Tidåker, T. Wesström, T. Kätterer, *Urban For. Urban Green.* **21**, 80 (2017).
11. J. Banks, "National lawn and garden equipment emissions," 2015 International Emission Inventory Conference, San Diego, CA, 16 April 2015; https://www3.epa.gov/ttn/chief/conference/ei21/session10/banks_pres.pdf.
12. M. Ignatieva, *in Urban Biodiversity: From Research to Practice*, A. Ossola, J. Niemelä, Eds. (Routledge, 2018), pp. 216–235.
13. C. Polsky *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **111**, 4432 (2014).
14. G. E. Southon, A. Jorgensen, N. Dunnett, H. Hoyle, K. L. Evans, *Landsc. Urban Plan.* **158**, 105 (2017).
15. L. S. Smith, M. D. E. Fellowes, *Urban For. Urban Green.* **13**, 433 (2014).

ACKNOWLEDGMENTS

We thank Formas, the Swedish Research Council for Environment, Agricultural Sciences and Spatial Planning (225-2012-1369, Lawn as ecological and cultural phenomenon-Search for sustainable lawns in Sweden) for financial support.

10.1126/science.aau6974

PHOTO: MARIA IGNATIEVA

POLYMER CHEMISTRY

Next-generation self-healing materials

Adjusting molecular structure tackles a long-standing problem of synthetic material longevity

By **Brent S. Sumerlin**

The long-term persistence of many synthetic materials and the resulting impact on the environment has made clear the importance of developing new routes to creating sustainable materials (*1*). This is especially true for man-made polymers, for which slow (or lack of) degradation is widely problematic, from plastic wastes generated by commodity packaging to high-tech electronics. Despite threats to human health, wildlife, oceans, and landfills, the fraction of polymeric materials that are recycled remains low. Polymers designed to degrade after their intended use represent a promising, chemistry-driven approach to minimize the impact of persistent, petroleum-derived materials (*2*). An alternative strategy for preparing sustainable materials is to design polymers that have even longer life spans and, as a result, need to be replaced less frequently. On page 220 of this issue, Urban *et al.* (*3*) advance the latter strategy through an approach to "self-healing" polymeric materials that stands out for its simplicity and potential scalability.

Achieving a polymeric material with a longer life span requires careful consideration of the myriad of external conditions to which the polymeric material will be exposed. Exposure to adventitious stimuli, such as light, solvents, or oxidation can cause discoloring, crazing (fine cracks on the surface), and swelling. Forces applied to the material that are continuous (such as gravity) or sudden (such as impact) can lead to bending, cracking, or fracture that compromise the material's integrity and shorten its intended lifetime. Therefore, substantial effort is being devoted to developing materials that heal themselves after damage (*4*).

In synthetic materials, healing can be accomplished by incorporating additional reagents. The components needed to effect healing (for example, monomer or catalyst) can be dispersed throughout a polymeric material so that physical damage will trigger the reactions that lead to repair. Indeed, some early examples of self-healing polymers were designed in this manner (*5, 6*). Although successful, a potential limitation of this approach is that healing can lead to depletion of the healing agents themselves, and self-repair in a single site may only be possible a limited number of times. Alternatively, materials may be rendered inherently healing by incorporating moieties within the polymer chains that promote reversible cross-linking. In this case, damage would ideally lead to dissociation or exchange of the reversible bonds such that healing is accomplished by these bonds reforming or undergoing additional exchange (*7, 8*).

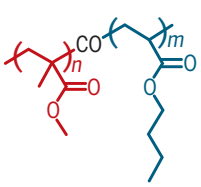
The reversible interactions needed for healing can be dynamic-covalent (for example, imine or boronic ester bonds) (*9*) or noncovalent [for example, hydrogen bonding (*10*), metal-ligand interactions (*11*)]. Although dynamic-covalent interactions often provide additional thermal stability (*12*), heating to relatively high temperatures (>100 °C) is often needed to bring about healing. On the other hand, most noncovalent bonds are readily exchangeable at much lower temperatures, which can facilitate healing under ambient conditions. However, the incorporation of moieties that, for example, can form hydrogen bonds or can chelate metals with the necessary affinity and concentration to produce robust materials with efficient healing, often requires additional synthetic steps or the design of polymers that are considerably different in expense and chemical makeup from common commercial materials.

Urban *et al.* describe a route to self-healing copolymers that relies on noncovalent van der Waals (vdW) interactions. Compared with hydrogen bonding, vdW interactions are present in essentially all materials. Although relatively weaker, they are particularly important in polymeric materials because their individual small attractive forces are summed over the length of chains that contain hundreds or thousands of units. Importantly, the copolymers considered in the study of Urban *et al.* were composed solely of methyl methacrylate (M) and butyl acrylate (B), two readily available and commercially relevant monomers. The authors demonstrated that scratches of 15 to 20 μm were healed without intervention in ~14 hours, whereas severed films (~200 to 300 μm thick) required up to ~80 hours for efficient self-repair.

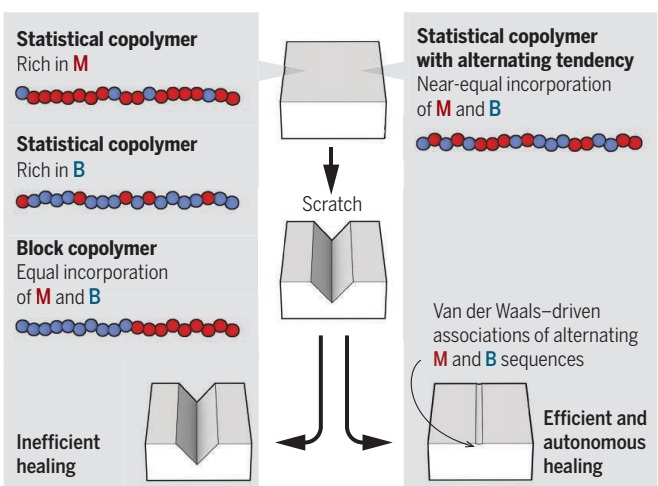
Interestingly, only copolymers with comonomer ratios (M/B ≈ 45/55 to 50/50) that were expected to lead to alternation—that is, methyl methacrylate and butyl acrylate units alternate in sequence in the polymer chain (such as –BMBMBM–), led to efficient healing (see the figure). This was true even when other ratios of comonomers resulted in polymers with a lower glass transition temperature (the temperature at which poly-

Next-gen self-healers

Copolymers of commodity monomers methyl methacrylate (M) and butyl acrylate (B), with runs of alternating sequences, can efficiently self-heal after mechanical damage, whereas copolymers of the same two monomers in different ratios or with alternative architectures do not lead to healing.



Poly(methyl methacrylate-*co*-butyl acrylate)
M B



George & Josephine Butler Polymer Research Laboratory, Center for Macromolecular Science and Engineering, Department of Chemistry, University of Florida, Post Office Box 117200, Gainesville, FL 32611, USA. Email: sumerlin@chem.ufl.edu

GRAPHIC: V. ALTOUNIAN/*SCIENCE*

mers transform from a brittle state to a viscous and rubbery state), a feature typically associated with increased healing efficiency. Other microstructures, including blocky arrangements of the monomers (–MMMMM-BBBBBB–), and copolymer compositions biased more strongly toward M or B, resulted in materials that healed inefficiently or not at all. Efficient healing was attributed to key-and-lock interactions of interdigitating alkyl pendant groups between chains, leading to a viscoelastic response that promoted self-recovery of neighboring chains upon separation. The experiments of Urban *et al.* suggest that this behavior resulted from the specific combination of both the extended helical conformation of chains within this narrow compositional window and interchain vdW forces. The role of helicity was supported by a modeling analysis that forced chains of all compositions into helical conformations. Such modeling led to increased cohesive energy densities between chains, irrespective of composition. Moreover, regardless of whether helical conformations were present, only when vdW interactions were considered did the modeled chains return to denser packing after physical separation.

It is surprising that such a narrow compositional range is required for this remarkable behavior, but Urban *et al.* attribute this to reduced interfacial fluidity in copolymers with higher contents of M and a reduced contribution of vdW interactions in B-rich regions. The healing behavior was also considered for copolymers in which B was replaced by other *n*-alkyl acrylates.

That healing of a polymer material can be achieved without resorting to complex synthetic effort or materials design serves as a reminder that careful consideration of molecular structure can often be sufficient to tackle long-standing problems of materials science. Moreover, the results of Urban *et al.* reinforce the importance of contemplating the role of the most fundamental features of macromolecules—in this case, chain conformation and microstructure—for upcycling and extending the lifetime of commodity polymers for next-generation self-healing materials. ■

REFERENCES

1. D. K. Schneiderman, M. A. Hillmyer, *Macromolecules* **50**, 3733-3750 (2017).
2. S. A. Miller, *ACS Macro Lett.* **2**, 550 (2013).
3. M. W. Urban *et al.*, *Science* **362**, 220 (2018).
4. Y. Yang, M. W. Urban, *Chem. Soc. Rev.* **42**, 7446 (2013).
5. K. S. Toohey *et al.*, *Nat. Mater.* **6**, 581 (2007).
6. S. R. White *et al.*, *Nature* **409**, 794 (2001).
7. C. N. Bowman, C. J. Kloxin, *Angew. Chem. Int. Ed.* **51**, 4272 (2012).
8. M. Rottger *et al.*, *Science* **356**, 62 (2017).
9. X. X. Chen *et al.*, *Science* **295**, 1698 (2002).
10. Y. Yanagisawa *et al.*, *Science* **359**, 72 (2018).
11. Q. F. Zheng, Z. Q. Ma, S. Q. Gong, *J. Mater. Chem. A* **4**, 3324 (2016).
12. W. K. Zoue *t al.*, *Adv. Mater.* **29**, 160100 (2017).

10.1126/science.aau6453

GRAPHIC: N. DESAI/*SCIENCE* FROM L. HERHAUS

CELL BIOLOGY

Dimerization quality control via ubiquitylation

A specialized ubiquitin ligase enzyme approves only functional dimers

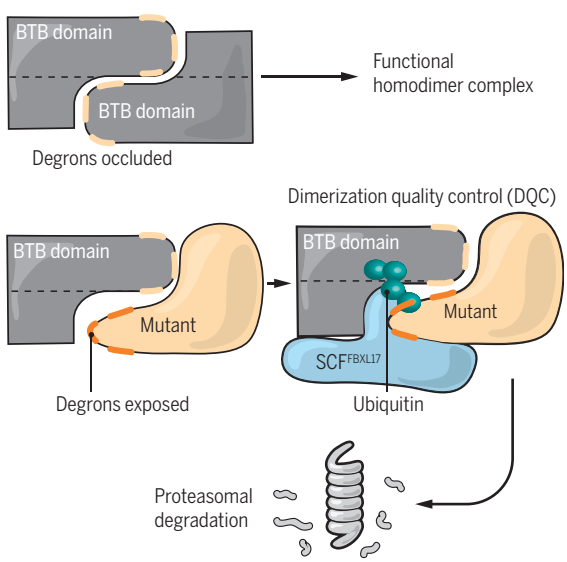
By **Lina Herhaus**[1] *and* **Ivan Dikic**[1,2]

Expanding the functional and structural repertoire of the proteome through protein dimerization is a fundamental concept in biology. Dimerization can occur in a homo- or heterotypic fashion with recurrent modular domains, such as BTB (BR-C, ttk, and bab) domains, leucine zippers, or coiled coils. Dimerization is a key factor in the regulation of enzymes, ion channels, receptors, and transcription factors. Indeed, proteomic studies indicate that a large fraction of mammalian proteins function only as dimers or multimers in cells (*1*). Thus, correct dimer formation is of fundamental importance for a healthy proteome and consequently for a healthy organism. However, whereas quality control systems for RNA, DNA, and the folding of nascent proteins are well-studied, whether and how correct composition of protein complexes is surveyed by cells has been unknown. On page 198 of this issue, Mena *et al.* (*2*) describe a quality control pathway that detects and eliminates nonfunctional dimer complexes. A failing dimerization quality control (DQC) pathway leads to the accumulation of aberrant dimers, which could contribute to the pathology of various diseases.

The authors investigated the large family of BTB proteins. Although the BTB domain can readily heterodimerize in vitro and upon overexpression in cells, endogenous BTB proteins mostly homodimerize to function as E3 ubiquitin (Ub) ligases, membrane channels, and transcription factors (*3, 4*). This observation prompted Mena *et al.* to ask, what prevents irregular heterodimerization under physiological conditions? To answer this question, the authors forced heterodimerization by overexpressing a number of BTB proteins in cells and then identified proteins that preferentially interact with heterodimeric BTB proteins. Intriguingly, one candidate that showed a high preference for aberrant heterodimers was the E3 Ub ligase SCF^{FBXL17} [a protein complex for which F-box and leucine-rich repeat protein 17 (FBXL17) confers specificity], an enzyme that mediates ubiquitylation, which targets proteins for degradation in the proteasome. Indeed, overexpression of SCF^{FBXL17} caused the proteasomal degradation of several endogenous BTB proteins. To un-

Removal of nonfunctional protein dimers

Functional BTB-containing proteins form homo- or heterodimers, occluding two out of three SCF^{FBXL17} degron residues. These degrons are exposed if BTB dimers are formed between mutant proteins and dimers cannot adopt a completely closed conformation, allowing SCF^{FBXL17} to bind and ubiquitylate the dimer and resulting in its proteasomal degradation.



[1]*Institute of Biochemistry II, Goethe University Frankfurt–Medical Faculty, University Hospital, Theodor-Stern-Kai 7, 60590 Frankfurt am Main, Germany.* [2]*Buchmann Institute for Molecular Life Sciences, Goethe University Frankfurt, Max-von-Laue-Straße 15, 60438 Frankfurt am Main, Germany. Email: dikic@biochem2.uni-frankfurt.de*

derstand how SCF^{FBXL17} targets BTB proteins, the authors searched for amino acids required for the interaction. They identified three "degron" amino acids located close to the edge of the dimer interface that were conserved in most of the 24 BTB proteins targeted for degradation by SCF^{FBXL17}. Strikingly, in functional homodimers, two of these degron residues seem to be buried by the dimer interface, thereby protecting them from recognition by SCF^{FBXL17} (see the figure). One can speculate that in aberrant heterodimers or inactive mutants or homodimers that fail to reach their mature conformation, these two residues are surface exposed and thereby complete the interaction site for SCF^{FBXL17}. Indeed, the behavior of 55 distinct artificial BTB hetero- and homodimers that were systematically tested in vitro for their sensitivity toward the E3 Ub ligase support this hypothesis. Notably, SCF^{FBXL17} does not recognize mono-, tetra-, or pentameric BTB proteins, nor BTB dimers the structure of which had been unfolded by heat. This further underlines the specific commitment of SCF^{FBXL17} to DQC. Additionally, many BTB domains do not only form homodimers but also engage in functional interactions with non-BTB proteins. How these mixed heterodimers of BTB–non-BTB proteins are protected against degradation is unclear. It will be very exciting to analyze the structural details of SCF^{FBXL17} bound to its substrates on the atomic level to understand how exactly SCF^{FBXL17} achieves this astonishing selectivity and to reveal whether other factors besides degron exposure might be involved.

"A failing dimerization quality control (DQC) pathway leads to the accumulation of aberrant dimers, which could contribute to the pathology of various diseases."

Knowing the structural basis of this interaction would also help to address another obvious question: What are the consequences of DQC failure? Even on the level of a single cell, this is a challenging question to answer because SCF^{FBXL17} has other functions besides surveying correct BTB dimer formation (*5, 6*). So, the cellular disturbances resulting from its full deletion or the inactivation of its Ub ligase activity can be only partially attributed to impaired DQC or a specific substrate of SCF^{FBXL17}. To dissect the role of SCF^{FBXL17} in DQC, tools such as biosensors need to be developed that specifically target its action on BTB proteins. Nevertheless, Mena *et al.* were able to show that differentiation, function, and survival of cells in the peripheral and central nervous system of *Xenopus laevis* (African clawed frog) embryos depend on correct BTB protein dimerization and fail in the absence of SCF^{FBXL17}. Besides playing a role in organismal development, defective BTB proteins have been linked to a variety of disorders including immune and neurodegenerative diseases as well as cancer (*7–9*). A deeper knowledge of SCF^{FBXL17}-mediated DQC might therefore have considerable implications for understanding disease pathology and for identifying therapeutic targets.

The discovery of SCF^{FBXL17}-mediated DQC is exciting as it provides the first example of a quality control mechanism that deals with the correct composition and function of protein complexes. Given the high abundance of dimers, trimers, and multimers with key functions in cells (*10*), it is very likely that more players acting in DQC will be discovered. If we look at just the "dimerome," there might be different DQC E3 Ub ligases in charge for coiled coil, leucine zipper, and other dimerization domains. The presence of SCF^{FBXL17} only in metazoans and not in basal eukaryotes, which contain a very limited number of BTB-containing proteins, indicates that dimerization modules and DQC might have coevolved. Thus, E3 Ub ligases that have emerged with or shortly after certain dimerization domains evolved could be candidates for DQC proteins. The discovery of DQC lifts research in quality control pathways to a new dimension and will certainly spark research into the quality control of multimeric complexes. ■

REFERENCES AND NOTES

1. E. L. Huttlin *et al.*, *Cell* **162**, 425 (2015).
2. E. L. Mena *et al.*, *Science* **362**, eaap8236 (2018).
3. P. J. Stogios *et al.*, *Genome Biol.* **6**, R82 (2005).
4. M. Zhuang *et al.*, *Mol. Cell* **36**, 39 (2009).
5. Y. Lai, J. Li, X. Li, C. Zou, *J. Cell Sci.* **130**, 3578 (2017).
6. M. Raducu *et al.*, *EMBO J.* **35**, 1400 (2016).
7. N. K. Hayward *et al.*, *Nature* **545**, 175 (2017).
8. S. Aharoni *et al.*, *BMC Med. Genet.* **17**, 82 (2016).
9. K. Igarashi *et al.*, *Nat. Rev. Immunol.* **17**, 437 (2017).
10. E. L. Huttlin *et al.*, *Nature* **545**, 505 (2017).

ACKNOWLEDGMENTS

The Dikic lab is supported by grants from the Deutsche Forschungsgemeinschaft (SFB 1177 on selective autophagy) and by the European Research Council under the European Union's Horizon 2020 research and innovation programme (grant agreement no. 742720). L.H. is supported by a European Molecular Biology Organization long-term postdoctoral fellowship (ALTF 1200-2014, LTFCOFUND2013, GA-2013-609409).

10.1126/science.aav1391

MICROBIOLOGY

A common trick for transferring bacterial DNA

A known virus-mediated process spreads bacterial genes more than expected

By **Alan R. Davidson**

A key reason that bacteria thrive is their ability to readily exchange pieces of their genomes both within and between species. This process of horizontal gene transfer allows them to rapidly adapt to diverse ecological niches and, unfortunately for humans, widely disperse genes encoding antibiotic resistance and virulence factors. Phages, the viruses that infect bacteria, are key players in horizontal gene transfer. Host bacterial DNA can be packaged into phage particles in a process called transduction (*1*). The resulting transducing particles can infect other cells and thereby deliver the bacterial DNA. Historically, transduction has been viewed as a rare event, caused by phages making "mistakes" and packaging the "wrong" DNA. However, on page 207 of this issue, Chen *et al.* (*2*) describe a mechanism that mediates transduction at frequencies 1000-fold higher than previously observed. This mode of transduction is a natural part of the phage life cycle and likely occurs often in diverse bacterial species, thus exerting a strong effect on horizontal gene transfer and bacterial evolution.

Most previous studies on transduction involved infecting cells with a phage of interest and measuring the percentage of phages present in the resulting lysate that can transfer an easily detected bacterial gene (for example, a gene that confers antibiotic resistance) to a different strain. The frequencies of transducing particles in such experiments were generally very low (10^{-5} to 10^{-6} of total phage particles in a lysate). A key difference in the Chen *et al.* study is that they measured transduction frequencies after induction of a prophage. Prophages form when intact phage genomes are integrated into the host genome. Most genes within a prophage are silenced by the activity of a phage repressor protein. However,

upon exposure to DNA damaging agents, prophages are induced. Subsequently, the phage DNA is excised from the host genome, and lytic viral replication is initiated with concomitant production and release of phage particles.

Excision relies on phage encoded proteins, excisionase (Xis) and integrase (Int), and has been presumed to be the first step in prophage induction such that prophage DNA excises first and then replication begins. However, Chen *et al.* found that for several *Staphylococcus aureus* phages, transcription of the *xis* gene is considerably delayed compared to that of the genes encoding the proteins required for DNA replication and phage-particle production. The result of this delay is replication of phage DNA in situ within the bacterial genome and subsequent encapsidation of this DNA, which is still attached to the adjacent bacterial DNA (see the figure).

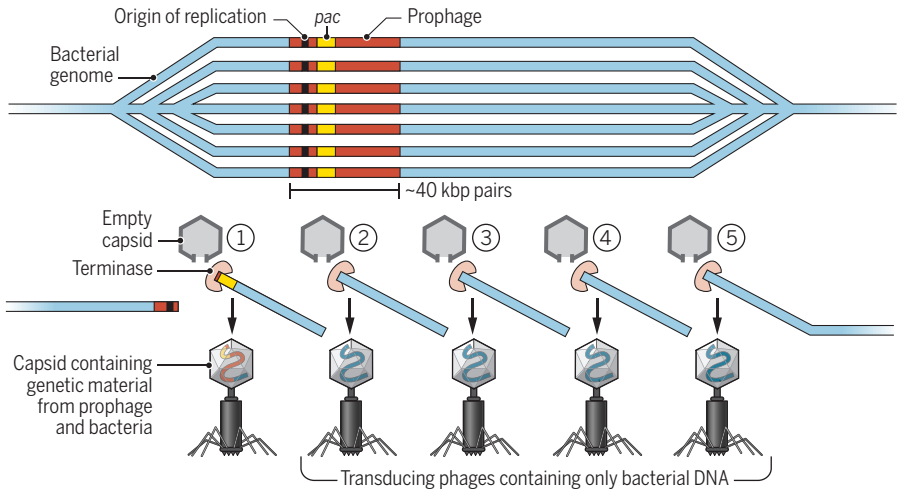
The packaging of phage DNA into empty capsids is mediated by the terminase enzyme. Terminase initiates packaging through cleavage near a specific *pac* site in the phage genome. Once started, DNA packaging continues until the capsid is full, at which time the DNA is cleaved nonspecifically. Remarkably, after this nonspecific cleavage, terminase remains bound to the DNA and begins another round of DNA packaging into another empty capsid. In this way, many phage genomes are packaged as the terminase moves processively along phage or bacterial DNA. This packaging mechanism is advantageous during the normal life cycle because phage DNA is replicated into long linear concatemers consisting of many copies of joined phage genomes; these concatemers are the substrate for packaging. By contrast, the result of this "headful" packaging mechanism, in the case of prophage induction, is the encapsidation of many phage genome lengths (~40 kilobase pairs) of bacterial DNA that is situated downstream of the *pac* site in the prophage genome. Chen *et al.* show that this "lateral transduction" (LT) mechanism results in the transfer of bacterial genes adjacent to prophages at a frequency 1000-fold greater than genes in other genomic locations. Remarkably, this increase in transduction frequency is still observed for a gene located six genome lengths away from the prophage, attesting to the high processivity of the packaging reaction.

LT is likely a widespread phenomenon. Approximately half of all bacteria contain prophages, and many contain multiple prophages (*3*). Although there are several different phage DNA packaging mechanisms, the headful mechanism is common (*4*). The only aspect of LT that has not been extensively investigated is excision timing. Even though most phage investigators, including myself, have assumed that prophages usually excise their DNA first and then replicate, there are no data to support this. In fact, even induction of an *Escherichia coli* λ prophage, which is the classical model system for studying prophage induction, results in considerable amplification of bacterial DNA near the prophage attachment site, indicating that replication initiates at the phage origin before excision (*5*). Similar behavior has been seen upon induction of *Salmonella* prophages (*6*). It is likely that none of the ingredients required for LT are rare. Thus, LT can be expected to occur in many bacterial species and particularly in pathogens, which have a higher than average prophage content (*3*). For *S. aureus*, a pathogen with a large number of prophages, genome analysis by Chen *et al.* indicates that LT has played an important role in promoting increased gene mobilization in regions downstream of phage attachment sites, which are often involved in virulence.

Perhaps the most important contribution of Chen *et al.* is in drastically altering our thinking about phage transduction. For example, all the phenomena associated with LT were previously recognized [for example, (*7*)], and a likely case of LT was recently described for a *Pseudomonas aeruginosa* prophage (*8*). But the broad importance of these findings was not appreciated. With the elegant description of LT by Chen *et al.*, we now realize that phages may frequently transfer large portions of bacterial DNA as a normal part of their life cycle. Accordingly, many observations may be reinterpreted. For example, large-scale sequencing studies of viral particles purified from sources of interest, such as human disease sites, have revealed relatively high numbers of antibiotic resistance genes and other bacterial genes, yet very few authentic phage genomes have been detected containing these genes (*9*). It now seems probable that the presence of these bacterial genes arises from LT. Additionally, treatment with ciprofloxacin, an antibiotic that causes prophage induction, led to an increase in antibiotic resistance genes within the mouse gut "phageome," and this resistance could be transferred to naïve microbiota (*10*). Another new concept that emerges is that high-frequency transduction may actually be an evolutionary adaptation for phages to increase the general fitness of their host. Thus, transduction may no longer be considered an accident but rather as another mechanism by which phages have evolved to manipulate their hosts for the sake of their own survival. ■

The mechanism of lateral transduction

Upon prophage induction, DNA replication causes amplification of the prophage DNA and surrounding bacterial DNA. Within the replicated DNA is a *pac* site, from which packaging into an empty capsid by terminase is initiated. These transducing phages infect other cells, thereby transferring the bacterial DNA.



GRAPHIC: KELLIE HOLOSKI/*SCIENCE* FROM A. DAVIDSON

Departments of Molecular Genetics and Biochemistry, University of Toronto, Toronto, Canada. Email: alan.davidson@utoronto.ca

REFERENCES AND NOTES

1. M. Touchon *et al.*, *Curr. Opin. Microbiol.* **38**, 66 (2017).
2. J. Chen *et al.*, *Science* **362**, 207 (2018).
3. M. Touchon *et al.*, *ISME J.* **10**, 2744 (2016).
4. S. R. Casjens *et al.*, *J. Bacteriol.* **187**, 1091 (2005).
5. T. Fukasawa *et al.*, *Mol. Gen. Genet.* **167**, 83 (1978).
6. J. G. Frye *et al.*, *J. Bacteriol.* **187**, 1485 (2005).
7. B. Kufer *et al.*, *Mol. Gen. Genet.* **187**, 510 (1982).
8. C. Pourcel *et al.*, *PLOS ONE* **12**, e0169684 (2017).
9. F. Enault *et al.*, *ISME J.* **11**, 237 (2017).
10. S. R. Modi *et al.*, *Nature* **499**, 219 (2013).

ACKNOWLEDGMENTS

A.R.D. is supported by grants from the Canadian Institutes of Health Research (FDN-15427) and the Natural Sciences and Engineering Research Council (RGPIN-2017-06858).

10.1126/science.aav1723

IMMUNOTHERAPY

T_{reg} cells—the next frontier of cell therapy

Will regulatory T cells be a frontline therapy for autoimmunity and other diseases?

By **Jeffrey A. Bluestone**[1,2] *and* **Qizhi Tang**[3]

In the past decade, effector T cells, engineered to express highly specific chimeric antigen receptors (CARs) or antigen-specific T cell receptors (TCRs) that recognize tumor antigens, have been shown to be highly effective adoptive cell therapies (ACTs) that have revolutionized certain cancer treatments. In 2017, the first two CAR–T cell therapies were approved by the U.S. Food and Drug Administration for the treatment of various $CD19^+$ B cell lymphomas, and extensive clinical trials are now under way to expand the therapies to multiple solid tumor settings (*1*, *2*). Here, we speculate on the next generation of immune cell therapies for non-cancer diseases. Specifically, we highlight the progress toward developing a new class of ACT using regulatory T cells (T_{regs}) to treat autoimmune diseases—including type 1 diabetes (T1D), rheumatoid arthritis, inflammatory bowel disease (IBD), graft-versus-host disease (GvHD) that can occur after bone marrow transplantation, and organ transplant rejection—and potentially to treat nonimmune diseases such as Alzheimer's disease, Parkinson's disease, heart disease, and type 2 diabetes (*3*).

Although T_{regs} constitute only 1 to 2% of peripheral blood lymphocytes, they are the master controllers of self-tolerance (whereby immune cells can recognize foreign substances and ignore self tissues, thus avoiding autoimmunity), tissue inflammation, and long-term immune homeostasis (*4*). In the most severe T_{reg} deficiency, IPEX (immunodysregulation, polyendocrinopathy, enteropathy, X-linked) syndrome, mutations in the key T_{reg} lineage–specific transcription factor FOXP3 (forkhead box P3) lead to T_{reg} defects resulting in lethal multi-organ inflammation and autoimmunity (*5*). Importantly, reduced numbers and/or function of T_{regs} have been implicated in the pathology of many common autoimmune diseases. Indeed, in some disease settings, defects in T_{reg} functions or reduced T_{reg} growth and survival factors such as interleukin-2 (IL-2), coupled with a local tissue environment of high concentrations of inflammatory cytokines such as IL-6, IL-12, and IL-1, results in T_{reg} instability, uncontrolled inflammation, and excessive tissue damage (*6*). Thus, a number of ongoing studies are repurposing approved drugs, such as T_{reg} growth factors (for example, IL-2) and T_{reg}-stabilizing factors (for example, rapamycin), to enhance T_{reg} function as an approach to controlling a variety of autoimmune diseases, GvHD, and organ transplant rejection (*7*). These results underscore the opportunity to directly alter pathological immune responses by boosting T_{reg} activity—which is also the goal of T_{reg}-based ACT.

The first preclinical proof-of-concept studies to treat autoimmune diseases were performed using polyclonal T_{regs}, then termed suppressor cells, isolated on the basis of the expression of multiple cell surface markers including CD4, CD25, and CD62L (*4*). The success in these efforts and in subsequent studies using purified $FOXP3^+$ T_{regs} has led to several clinical trials in a variety of disease settings including organ transplant rejection, GvHD, T1D, and autoimmune syndromes (*8*, *9*). The cells are purified from the peripheral blood of each patient, grown ex vivo in the presence of antibodies to CD3 and CD28 accompanied by high-dose IL-2 to expand the highly enriched T_{reg} population, and, after adequate characterization, adoptively transferred into patients. This polyclonal T_{reg} ACT has been shown to be safe in phase 1 studies, and some studies have suggested clinical activity (*8*, *9*). The first results from a phase 2 study in T1D are expected in 2019 (NCT02691247); other efforts to determine efficacy in GvHD are under way (for example, NCT01795573 and NCT01937468).

Moreover, it has become increasingly clear that nonimmune diseases such as cardiovascular diseases, obesity, type 2 diabetes, and degenerative diseases of muscle and brain are exacerbated by inflammation. Accumulating evidence in preclinical studies suggests that T_{regs} can quell inflammation and reduce morbidity in these diseases by contributing to tissue homeostasis and repair by producing the epidermal growth factor receptor (EGFR) ligand amphiregulin in damaged tissues including influenza-infected lungs, muscular dystrophy–affected muscle, nerve demyelination associated with multiple sclerosis (MS), and high-fat diet–induced obesity (*10*, *11*). A recent clinical trial in amyotrophic lateral sclerosis reported that T_{reg} ACT may reduce disease progression (*12*). In addition, T_{regs} have been shown to increase bone marrow engraftment, reduce immune responses to gene therapy, and facilitate wound healing (*12*).

So where will polyclonal T_{regs} be most effective? Preclinical data show that polyclonal T_{regs} are efficacious in controlling some diseases, such as lupus nephropathy and IBD, but far less so in others such as T1D and MS. For example, T_{regs} specific for a pancreatic islet antigen were 50 to 100 times more efficient in blocking T1D progression than polyclonal T_{regs} when transferred into mice with autoimmune diabetes (which is similar to T1D). Because T_{regs} have a TCR repertoire that is skewed toward self-antigens, we speculate that large tissues such as skin and gut with abundant and more diverse antigens may be recognized by a larger fraction of TCRs in the T_{reg} repertoire than smaller tissues such as pancreatic islets. Thus, larger tissues can activate sufficient T_{regs} in a polyclonal population to achieve therapeutic effect, whereas smaller tissues will require antigen-specific T_{regs} for ACT.

The use of polyclonal T_{regs} can potentially suppress protective immunity against tumors and infectious diseases by transferring a large number of T_{regs} of broad undefined specificity. Thus, developing antigen-specific T_{reg} therapy will likely provide a more effective and safer alternative. Importantly, in multiple preclinical models, T_{regs} that are selected ex vivo to be specific for limited tissue antigens or even single antigens can mediate localized dominant bystander suppression of effector T cells of diverse specificities. This concept of enriching target tissue-specific T_{regs} has already been translated into early-phase clinical trials. T_{regs} have been isolated from organ transplant recipients and stimulated with donor organ–derived antigen-presenting cells (APCs) to selectively expand donor-specific T_{regs}; these are more effective at suppressing organ rejection than polyclonal T_{regs} in humanized mouse transplant models (*13*). These alloantigen-reactive T_{regs} are being tested to prevent rejection in several clinical trials with the goal of reducing or even fully withdrawing immunosuppressive drugs (for example, NCT02474199).

Applying selective expansion approaches to tissue antigens other than transplant

[1]Diabetes Center, University of California, San Francisco, CA 94143, USA. [2]Sean N. Parker Autoimmune Research Laboratory, University of California, San Francisco, CA 94143, USA. [3]Department of Surgery, University of California, San Francisco, CA 94143, USA. Email: jeff.bluestone@ucsf.edu

antigens is far more challenging. This is because frequencies of alloantigen-reactive T_{regs} are as high as 10%, whereas frequencies of tissue antigen-specific T_{regs} are orders of magnitude lower. In addition, repeated stimulation of T_{regs} ex vivo can destabilize T_{regs} and lead to outgrowth of non-T_{regs}. Expression of engineered antigen-specific receptors (CARs and TCRs), a successful strategy in cancer immunotherapy, offers a solution to achieving antigen specificity of therapeutic T_{regs}. Alloantigen-specific CARs, reactive to human leukocyte antigen-A2 on donor tissue, can prevent allograft rejection in humanized mouse models (*14*). Similarly, T_{regs} engineered to express transgenic TCRs specific for autoantigens such as Factor IX in hemophilia, myelin oligodendrocyte glycoprotein in MS, and pancreatic islet–specific antigens in T1D have shown enhanced efficacy in preclinical models. Moreover, new approaches have been developed that allow T_{regs} to be "differentiated" to selectively control diseases mediated by subsets of T cells, including T helper 1 (T_H1)–like T_{regs} to treat T1D and organ transplant rejection, T_H2-like T_{regs} for asthma and allergy, and T_H17-like T_{regs} to treat MS (*15*). A combination of antigen receptor engineering and differentiation approaches to direct T_{regs} to specific sites of disease activity will lead to more robust and effective treatments.

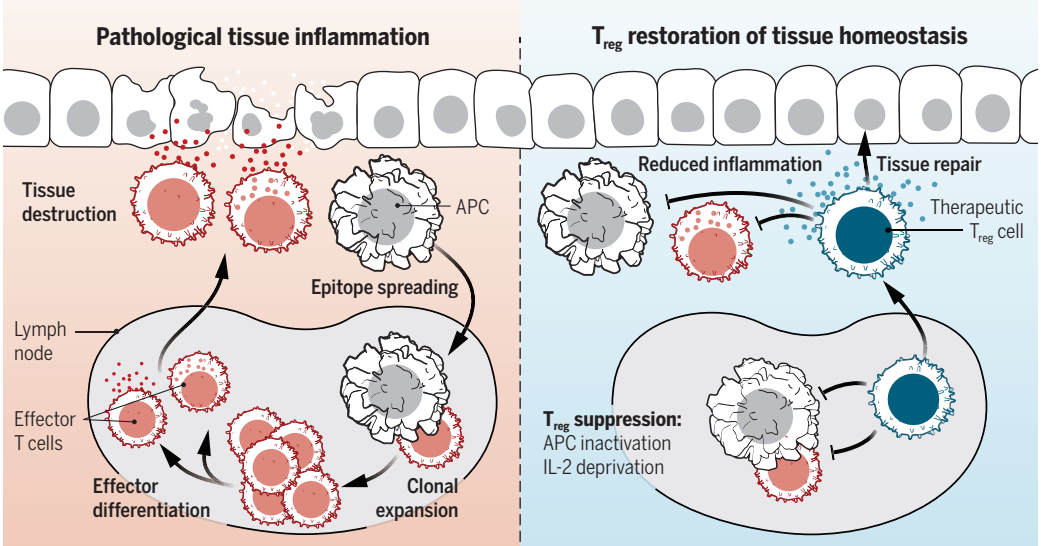
In this regard, two features of T_{regs} are worth emphasizing. First, unlike cancer-killing T cells that must bind to target cells directly, in many cases there is no requirement for therapeutic T_{regs} to directly contact effector T cells to inhibit function. T_{reg}-mediated immune suppression and tolerance typically occur through alteration of the local tissue microenvironment both in the draining lymph node and in the affected tissue (see the figure). Immune regulation by T_{regs} depends on antigen engagement normally provided by APCs. Once induced, these suppressive activities, carried out by soluble mediators and cell surface receptors, affect all cells and tissue in the vicinity in a paracrine fashion. Thus, it is possible to direct T_{regs} using engineered CARs and TCRs to any cells or even natural or synthetic multivalent ligands in the target tissue to achieve local immune suppression. Second, T_{regs} can mediate "infectious tolerance." T_{regs} not only dampen inflammation but also create a tissue microenvironment conducive to the emergence of additional immune-suppressive populations, including T_{regs} with additional specificities, myeloid-derived suppressor cells, and other subsets of suppressive T cells, such as type 1 regulatory T (Tr1) cells. Through this infectious spread of tolerance, the effects of therapeutic T_{regs} are amplified and prolonged, even if the adoptively transferred T_{reg} cells do not survive indefinitely.

With all the advantages of a living drug that can traffic to target tissue, automatically carry out therapeutic actions tailored to the tissue microenvironment, expand or contract in number depending on the therapeutic demands, and persist long-term for durable recovery, an inherent challenge of T_{reg} cell therapy is to ensure their lineage stability and prevent antigen-specific T_{regs} from turning into effector T cells that are capable of inflicting tissue injury (*6*). This risk may be mitigated by engineering T_{reg} cells to include suicide genes, to secrete autocrine IL-2, deleting receptors for proinflammatory cytokines, and stabilizing FOXP3 expression. It is also important to recognize that T_{reg} ACT alone may not be sufficient to induce tolerance because in many settings the preexisting effector T cells may be refractory to T_{reg} control or may outnumber the therapeutic T_{regs}. It is likely that durable remission will require a combination of drugs that reduce effectors and inflammation before administering antigen-specific T_{regs}, or perhaps T_{regs} may be genetically engineered to reduce inflammation, compromise effector T cells, and promote tolerance.

One of the challenges of ACT involves the cost and technical difficulty associated with personalized treatments. This challenge is being addressed for ACT in cancer treatments by developing universal T cells for off-the-shelf use. Using gene editing approaches, T cells from unrelated individuals are modified by deleting polymorphic human leukocyte antigens to protect the cells from being destroyed by the host immune system and by eliminating their endogenous TCRs to prevent nonspecific activities. The next few years will be critical in determining the efficacy and practicality of T_{reg} ACT. However, if the explosion of cell therapy approaches in cancer is any indication, we believe that T_{regs} will be the next frontier in the development of anti-inflammatory and tolerogenic therapies. ■

Sites of therapeutic T_{reg} action

In the lymph node, where primary immunity occurs, T_{regs} can shut down proliferation and differentiation of pathogenic T cells. In inflamed tissues, T_{regs} can alter effector T cell expansion and activation, cytokine production, innate immune cell activation, chemoattraction of inflammatory cells, and epitope spreading.



GRAPHIC: VERONICA FALCONIERI HAYS/*SCIENCE* FROM J. A. BLUESTONE

REFERENCES AND NOTES

1. C. H. June *et al.*, *N. Engl. J. Med.* **379**, 64 (2018).
2. J. Tang *et al.*, *Nat. Rev. Drug Discov.* **17**, 465 (2018).
3. A. L. Perdigoto *et al.*, *Front. Immunol.* **6**, 654 (2016).
4. S. Sakaguchi *et al.*, *Cell* **133**, 775 (2008).
5. R. Bacchetta *et al.*, *Ann. N.Y. Acad. Sci.* **1417**, 5 (2018).
6. M. DuPage *et al.*, *Nat. Rev. Immunol.* **16**, 149 (2016).
7. A. K. Abbas *et al.*, *Sci. Immunol.* **3**, eaat1482 (2018).
8. C. G. Brunstein *et al.*, *Blood* **127**, 1044 (2016).
9. J. A. Bluestone *et al.*, *Sci. Transl. Med.* **7**, 315ra189 (2015).
10. J. Li *et al.*, *Front. Immunol.* **9**, 585 (2018).
11. M. Panduro *et al.*, *Annu. Rev. Immunol.* **34**, 609 (2016).
12. J. R. Thonhoff *et al.*, *Neurol. Neuroimmunol. Neuroinflamm.* **5**, 465 (2018).
13. A. L. Putnam *et al.*, *Am. J. Transplant.* **13**, 3010 (2013).
14. N. A. J. Dawson *et al.*, *Front. Immunol.* **8**, 1460 (2017).
15. A. J. Levine *et al.*, *Nature* **546**, 421 (2017).

ACKNOWLEDGMENTS

J.A.B. and Q.T. acknowledge support from the JDRF, Caladrius Biosciences, Becton-Dickinson, Juno Therapeutics, and Pfizer. J.A.B. and Q.T. have patents and pending patents on polyclonal and antigen-specific T_{regs} and consult for Third Rock Ventures and Juno in this area. J.A.B. serves on the board of Pfizer.

10.1126/science.aau2688

MICROBIOLOGY

Anti-CRISPRs on the march

The diversity of anti-CRISPR proteins encoded by viruses is rapidly expanding

By **Eugene V. Koonin** *and* **Kira S. Makarova**

Viruses and other mobile genetic elements (MGEs) are ubiquitous to all cellular life-forms, and thus, evolution of life is a perennial host-parasite arms race (*1*). Nearly all cellular organisms have evolved diverse, multilayer defense systems against parasites (*2*). Arguably, adaptive immunity is the most elaborate, efficient, and specific form of such defense. In almost all archaea and about one-third of bacteria, adaptive immunity is mediated by CRISPR-Cas systems that incorporate fragments of viral genomes into CRISPR arrays and use processed transcripts of these inserted fragments (spacers) to recognize and destroy the cognate viruses (*3*). The efficiency of CRISPR-Cas is such that when a virus is recognized by the immune system, it has no chance to survive. And yet, viruses thrive in all microbial communities. Virus genomes mutate fast, so the most straightforward way to avoid adaptive immunity is mutational escape. However, the hosts keep adapting, so in addition to simply attempting to outpace immunity, viruses evolve active antidefense—in particular, multiple anti-CRISPR proteins (Acrs) (see the figure) (*4*, *5*). On pages 240 and 236 of this issue, Marino *et al.* (*6*) and Watters *et al.* (*7*), respectively, expand the collection of Acrs, which may have application in the regulation of gene editing.

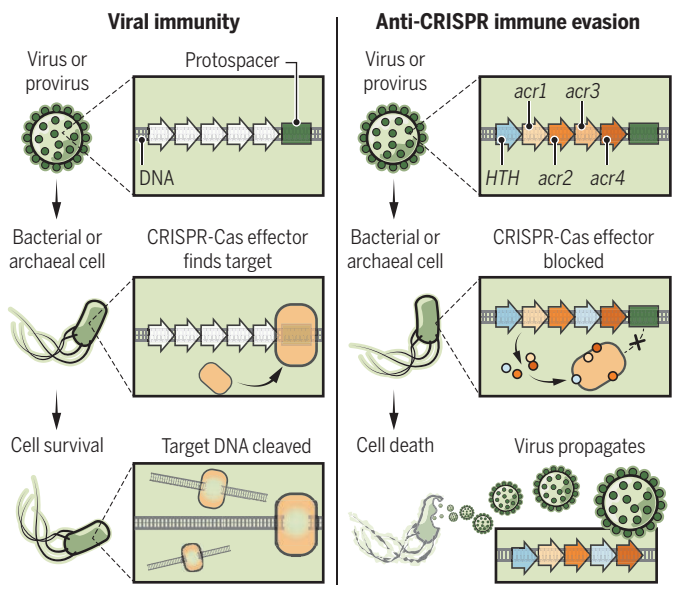
Although CRISPR-Cas are programmable immune systems that can provide resistance to any pathogen, they show remarkable diversity of molecular organization, with two distinct classes, six types, and about 25 subtypes discovered to date (*8*). The CRISPR-Cas types and subtypes differ primarily in the protein composition and structure of the effector complexes that are involved in interference—that is, recognition and cleavage of target nucleic acids by the CRISPR-Cas systems. Hence, there is a diversity of Acrs that appear to be specific for CRISPR-Cas subtypes or even to lower levels of classification. So far, all discovered Acrs have been shown to target several subtypes of class I or class II CRISPR-Cas systems. Marino *et al.* and Watters *et al.* identify previously unknown Acrs against subtypes I-C and V-A. The latter finding is of special interest considering both the distinct structure of the subtype V-A effector (known as Cas12a or Cpf1) and its increasing use as a genome-editing tool (*3*).

The approaches used to identify Acrs are both straightforward and elegant, seamlessly combining bioinformatics and biochemistry (see the supplementary materials). The clue that helped the original discovery of the Acrs was the ability of certain *Pseudomonas* bacteriophages (viruses that infect bacteria) to reproduce in bacterial strains that were supposed to be immune because of the presence of CRISPR spacers against those phages. This observation prompted genetic and biochemical experiments that led to the identification of the first Acrs (*9*). A related observation is "self-targeting," in which a provirus integrated in a bacterial genome is targeted by one or more CRISPR spacers from the same genome, but the provirus persists in the bacterial population (*7*, *10*). Such tolerance requires inactivation of the CRISPR-Cas system(s), and Acrs are the most likely solution. Indeed, experimental tests for Acr activity of proteins encoded by predicted proviruses (prophages) in the tolerant bacteria has led to the identification of several additional Acrs (*7*, *10*). A remarkable feature that became apparent once the first few Acrs were identified is that viruses often possess multiple *acr* genes that form clusters in virus genomes (*4*, *5*). Furthermore, although the Acrs themselves characteristically show little sequence conservation, the clusters of *acr* genes are typically flanked by more highly conserved genes that often encode proteins containing a helix-turn-helix (HTH) DNA-binding domain and are implicated in the regulation of *acr* gene expression. The presence of such conserved Acr-associated (*aca*) genes enables the use of the "guilt by association" approach to predict new Acrs (*5*, *11*). Together, these features provide for a consistent and relatively simple strategy for the identification of new Acrs that has been successfully used by Marino *et al.* and Watters *et al.*

The discovery of multiple Acrs against type V CRISPR-Cas in predicted MGEs from the bacterium *Moraxella bovoculi* brings the tally of Acrs to more than 30 families. These Acrs are encoded by diverse bacterial and archaeal viruses and target CRISPR-Cas systems of both classes, three types, and seven subtypes (see supplementary materials). The Acrs only share some generic features, such as gene clustering, small size (typically between 50 and 150 amino acids), and extreme sequence and structural diversity. In a sense, the major common feature

Anti-CRISPR proteins

In the presence of a cognate protospacer, CRISPR-Cas systems protect archaeal and bacterial cells from viruses. Virus-encoded Acrs prevent virus inactivation by CRISPR-Cas systems.



GRAPHIC: C. BICKEL/*SCIENCE*

National Center for Biotechnology Information, National Library of Medicine, NIH, Bethesda, MD, USA.
Email: koonin@ncbi.nlm.nih.gov

of the Acrs is that they have so little in common in sequence and structure. This diversity and the fast evolution of the Acr protein families are not surprising given that the Acrs are locked in the never-ending arms race with CRISPR-Cas. The mechanisms of action of those Acrs that have been studied in detail are also common in that they all directly bind the cognate effector Cas proteins and block their activity (*4*, *5*). However, different Acrs bind the same CRISPR effector, such as Cas9, at distinct, nonoverlapping sites or, in the case of the class I CRISPR-Cas system, bind to different subunits of the effector Cas complex.

Undoubtedly, exploration of Acr diversity is only just beginning. One can confidently predict that the Acrs encoded by bacterial and archaeal viruses are numerous, and multiple Acrs target each of the CRISPR-Cas subtypes. Thus, Acrs that target three types of CRISPR-Cas remain to be discovered, along with additional Acrs that target the widespread class I and II CRISPR-Cas systems. Numerous bacterial and archaeal viruses contain clusters of genes that encode small proteins, sometimes flanked by genes for HTH-containing proteins (see supplementary materials). The proteins encoded in such gene clusters are candidate Acrs, although activity against other host defense systems cannot be ruled out. Given that nearly all archaea have CRISPR-Cas systems, Acrs can be expected to be particularly abundant in archaeal viruses, and indeed, many archaeal virus genomes contain candidate gene clusters (*12*). Strikingly, such clusters account for up to 40% of some archaeal virus genomes, but their size and content are highly variable, even among closely related viruses. A notable example of this variation is the anti–type I-D Acr (AcrID1) family, one of the largest, relatively highly conserved Acr families, with more than 50 members encoded by archaeal viruses of three families (*13*). Among the *Sulfolobus islandicus* rod-shaped viruses (SIRVs), the number of AcrID1 genes varies from 0 to 12 (among about 50 virus genes), suggesting major, still not understood, differences in virus-host interactions.

The discovery of the Acrs came as a sensation, but in retrospect, this could have been readily anticipated. Indeed, the diversity of the Acrs reflects a general organizational principle of the virosphere: Apart from the structural and replication gene modules, virus genes encode proteins involved in virus-host interactions, of which the primary form is inhibition of host defenses. Perhaps the closest parallel with the Acrs are suppressors of gene silencing that are encoded by most plant viruses and inhibit the plant RNA interference (RNAi) machinery that is functionally analogous, albeit not homologous, to CRISPR-Cas (*14*). Furthermore, similar to the Acrs, plant virus RNAi suppressors are small, fast-evolving proteins.

Even if the general principles of Acr function and evolution can be considered understood, burning questions abound. Where do the Acrs come from? So far, not a single homolog of any of the Acrs with a different function has been discovered through sequence or structure comparison in either microbial or MGE genomes (apart from the presence of an HTH domain in AcrIIA1), despite extensive effort. Strikingly, all solved Acr structures seem to represent unique protein folds (*15*). Are there *acr* genes in MGEs other than viruses, such as plasmids and transposons? Such MGEs could have benefited from Acrs, but so far, none have been discovered. Do archaea and bacteria encode anti-Acrs? Again, from the arms-race perspective, one could predict the existence of such "defense against counter-defense." If so, how many turns does the arms-race spiral make? So far, all Acrs have been shown to inhibit Cas effector activity; do other types of Acrs exist—in particular, those inhibiting adaptation (spacer acquisition)? In the more practical vein, can Acrs be harnessed as useful regulatory tools for CRISPR-based genome editors? There is no doubt that these and other open questions will fuel fascinating Acr research for years to come. ■

"...can Acrs be harnessed as useful regulatory tools for CRISPR-based genome editors?"

REFERENCES

1. A. Stern, R. Sorek, *Bioessays* **33**, 43 (2011).
2. E. V. Koonin, K. S. Makarova, Y. I. Wolf, *Annu. Rev. Microbiol.* **71**, 233 (2017).
3. P. Mohanraju *et al.*, *Science* **353**, aad5147 (2016).
4. A. L. Borges, A. R. Davidson, J. Bondy-Denomy, *Annu. Rev. Virol.* **4**, 37 (2017).
5. A. Pawluk *et al.*, *Nat. Rev. Microbiol.* **16**, 12 (2018).
6. N. D. Marino *et al.*, *Science* **362**, 240 (2018).
7. K. E. Watters *et al.*, *Science* **362**, 236 (2018).
8. E. V. Koonin, K. S. Makarova, F. Zhang, *Curr. Opin. Microbiol.* **37**, 67 (2017).
9. J. Bondy-Denomy, A. Pawluk, K. L. Maxwell, A. R. Davidson, *Nature* **493**, 429 (2013).
10. B. J. Rauch *et al.*, *Cell* **168**, 150 (2017).
11. A. Pawluk *et al.*, *Nat. Microbiol.* **1**, 16085 (2016).
12. D. Prangishvili *et al.*, *Nat. Rev. Microbiol.* **15**, 724 (2017).
13. F. He *et al.*, *Nat. Microbiol.* **3**, 461 (2018).
14. A. S. Zvereva, M. M. Pooggin, *Viruses* **4**, 2578 (2012).
15. Y. Zhu, F. Zhang, Z. Huang, *BMC Biol.* **16**, 32 (2018).

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/362/6411/156/suppl/DC1

10.1126/science.aav2440

ORGANIC CHEMISTRY

Easy access to elusive radical reactions

Two mild approaches generate radical intermediates from masked aldehydes

By **J. Miles Blackburn** *and* **Jennifer L. Roizen**

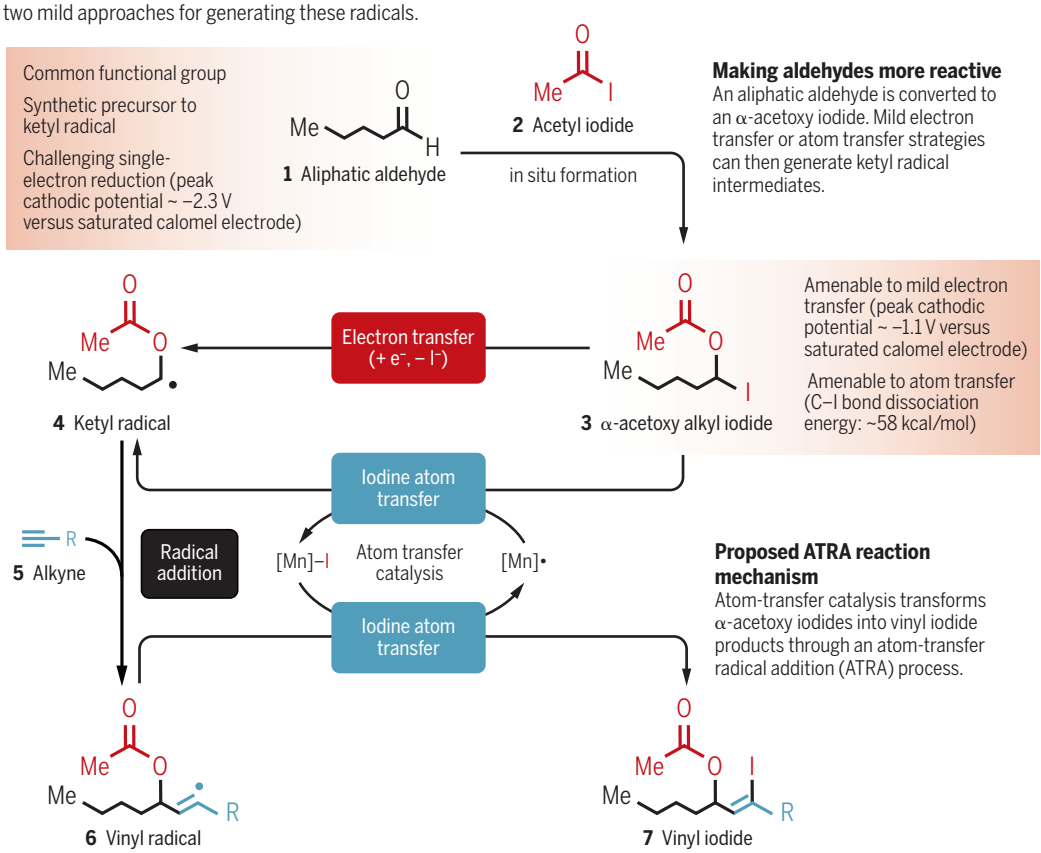
Radical intermediates are molecules that form transiently during a reaction and bear an unpaired electron. Molecules containing an unpaired electron are often viewed as challenging to access under mild conditions and accordingly have been underexploited as synthetic intermediates. Chemists are now developing catalytic methods to facilitate access to versatile radical intermediates (*1*). Despite recent advances, some types of potentially useful radical intermediates remain difficult to engage. On page 225 of this issue, Wang *et al.* (*2*) report a new approach for accessing elusive ketyl radicals. They leverage an in situ masking strategy of alkyl aldehydes and unlock opportunities for electron-transfer or atom-transfer catalysis.

Conceptually, the most straightforward approach to generating a radical intermediate from a nonradical reactant is to either add an electron to the substrate (substrate reduction) or remove an electron from the substrate (substrate oxidation). The elementary mechanistic step of electron transfer between an electron donor and an electron acceptor changes the formal charge of the reacting molecules. The thermodynamic feasibility of electron transfer is quantified by the oxidation (ability to lose an electron) and reduction (ability to gain an electron) potentials of the reactants.

Although electron-transfer processes engage a variety of functional groups (*3*–*5*), certain classes of molecules do not have redox potentials that allow for facile electron transfer. A complementary strategy to access a radical intermediate may be to instead move an atom (or group) that has

One intermediate, two strategies

The generation of ketyl radicals from aliphatic aldehydes generally requires harsh conditions. Wang *et al*. now report two mild approaches for generating these radicals.



an odd number of electrons. Unlike electron transfer that moves only one charged electron, atom transfer moves a neutral atom, resulting in no change to the formal charge of either of the reacting molecules. Instead, the connectivity of the atoms in the reacting molecules is altered, with bonds being formed or broken. The thermodynamic feasibility of an atom-transfer process is quantified by the analysis of the bond strengths of the bonds involved in the atom-transfer step.

Electron-transfer and atom-transfer processes can offer complementary approaches to accessing analogous radical intermediates, as is evident for aldehydes and ketones. Electron transfer to aldehydes and ketones is particularly challenging, with reduction potentials greater than –2.0 V versus the saturated calomel electrode (SCE) often lying outside the range of even the most reducing standard electron transfer or photoredox catalysts (*6*). To surmount this limitation, some ketyl-type radicals have been generated by exploiting an acid to activate the basic carbonyl moiety. In certain cases, acid activation of carbonyls can render an electron-transfer catalyst energetically capable of delivering an electron to the substrate and reducing the carbonyl substrate to furnish a ketyl radical. Unfortunately, this approach has been restricted to aromatic aldehydes and ketones, whereas aliphatic variants (molecule **1** in the figure) have proven to be less reactive.

Wang *et al*. note that although aldehydes and ketones may be difficult to engage in mild single-electron chemistry directly (see the figure), masking the carbonyl group as an α-acetoxy alkyl iodide (molecule ***3***) positions the substrate to potentially engage in either electron-transfer or atom-transfer processes. The α-acetoxy alkyl iodide has a much lower reduction potential (–1.1 V versus SCE) and can accept an electron from some of the standard photoredox catalysts. Alternatively, the weak carbon-iodine bond (calculated bond dissociation energy of 58 kcal/mol) of the α-acetoxy alkyl iodide can engage in atom-transfer processes. Indeed, the authors used manganese pentacarbonyl [$Mn(CO)_5$] as a catalyst, which is proposed to abstract the iodine atom and generate the desired aliphatic ketyl intermediate **4**. Although this masking strategy works for alkyl aldehydes, aliphatic ketones are not yet generally amenable to analogous approaches.

The ketyl radical **4** serves as a critical intermediate in an atom-transfer radical addition (ATRA) (*7*) reaction (see the figure). Formally, the carbon-iodine σ-bond of the α-acetoxy alkyl iodide substrate adds across a π-bond of an alkyne to afford two new σ-bonds: a carbon-carbon bond and a carbon-iodine bond. At a mechanistic level, Wang *et al*. propose that iodine atom transfer generates the ketyl radical **4**. This intermediate adds to the π-system of exogenous alkyne ***5*** in a radical addition step. The resultant vinyl radical **6** is poised to capture iodine from the [Mn]-I complex in the final atom-transfer process. Overall, the reaction furnishes the vinyl iodide product **7** in a net redox-neutral process. The presence of a carbon-iodine bond makes the products versatile compounds that can engage in subsequent two-electron reactions, which the authors highlight by elaboration in subsequent cross-coupling reactions.

Wang *et al*. provide valuable and mechanistically distinct electron- and atom-transfer platforms for the single-electron manipulation of aliphatic aldehydes and ketyl radical intermediates. Their work demonstrates that, despite the tendency of synthetic chemists to compare and contrast single-electron and two-electron processes, complementary tactics to access a single radical intermediate may offer previously unrealized synthetic potential. The approach of Wang *et al*. highlights a simple point, that the climate is now ripe for disruptive innovations in synthesis that allow for more facile access to previously elusive radical intermediates. ■

Department of Chemistry, Duke University, Durham, NC 27708, USA. Email: j.roizen@duke.edu

REFERENCES

1. M. Yan, J. C. Lo, J. T. Edwards, P. S. Baran, *J. Am. Chem. Soc.* **138**, 12692 (2016).
2. L. Wang, J. M. Lear, S. M. Rafferty, S. C. Fosu, D. A. Nagib, *Science* **362**, 225 (2018).
3. M. H. Shaw, J. Twilton, D. W. C. MacMillan, *J. Org. Chem.* **81**, 6898 (2016).
4. N. Zhang, S. R. Samanta, B. M. Rosen, V. Percec, *Chem. Rev.* **114**, 5848 (2014).
5. M. Yan, Y. Kawamata, P. S. Baran, *Chem. Rev.* **117**, 13230 (2017).
6. K. N. Lee, M.-Y. Ngai, *Chem. Commun.* **53**, 13093 (2017).
7. T. M. Williams, C. R. J. Stephenson, in *Visible Light Photocatalysis in Organic Chemistry* (Wiley-Blackwell, 2018), pp. 73–92.

10.1126/science.aav0574

POLICY FORUM

BIOETHICS

Return of results and data to study participants

A recent report urges progress but builds barriers to research participants' access

By **Susan M. Wolf**[1] *and* **Barbara J. Evans**[2]

Researchers conducting imaging, environmental health, and genetics studies have offered participants their research findings for years, publishing data on this experience and producing consensus guidelines (*1*–*5*). Research participants have articulated the value in the return of results and data (*6*). The National Academies recently issued a report at the request of the National Institutes of Health (NIH), the Centers for Medicare and Medicaid Services (CMS), and the Food and Drug Administration (*7*) that focuses on the return of results in studies with human biospecimens. It proclaims support for the interests of research participants, endorses the importance of offering participants their individual-specific results, and advocates assessing the value of results from the standpoint of participants. Unfortunately, in our view, the report's announced commitment to participant-centered progress is undercut by its actual recommendations. We believe the report creates major roadblocks to the return of data and results and would constrict participants' existing rights of access.

Some research results will meet clinical standards for quality, but many will not, because research seeks to advance understanding. Participants may value results related to themselves for a wide range of reasons (see the figure). As the report notes, empirical studies show that most participants want the option of return, and results are valued even when uncertain.

Researchers, in close consultation with ethicists and legal scholars, have developed three pathways for returning results when those results raise potential clinical concerns. The first is to perform research analyses in laboratories that comply with the Clinical Laboratory Improvement Amendments of 1988 (CLIA)—a federal statute that aims to ensure the safety and analytic quality of laboratory tests conducted for health care purposes—so that research results can be freely used in clinical care. A second pathway, for results from non-CLIA research laboratories, is for researchers to confirm results that raise clinical concerns in a CLIA laboratory before return. A third option for results from non-CLIA research laboratories is a clinical hand-off: return research results while advising the participant that clinical confirmation and follow-up are needed before clinical use. In this option, researchers maintain the line between research and clinical care by making a referral for clinical workup rather than venturing a diagnosis based on potentially uncertain research results (*8*). All three paths protect participants by requiring CLIA confirmation before clinical use of results and preserve participants' access.

Reasons for access to personal research results and data

- Preserve health (through diagnosis, treatment, and prevention)
- Seek clinical confirmation and follow-up (for the participant and possibly family members)
- Monitor developing scientific understanding of the results
- Form social networks with people who have similar test results
- Contribute data for other research uses, including citizen science
- Assess privacy risks posed by the collection and circulation of personal data

The Academies' report rejects this widely supported, legally sound approach. Although the report correctly states that research protocols planning to produce results for clinical use should use a CLIA laboratory, the report casts doubt on the feasibility of the accepted practice of seeking CLIA confirmation only for a subset of results to be returned. When protocols do not plan return of results for clinical use, the report recognizes that requiring CLIA compliance is unrealistic, as research laboratories may find the cost and regulatory burdens prohibitive, CLIA analytic requirements may not fit some research testing, and CLIA has not kept pace with rapid developments in advanced testing technologies. As an alternative, the report allows return from research laboratories approved under a "quality management system" (QMS) that does not yet exist. In the report, the NIH is urged to lead the massive effort required to develop the QMS. Unfortunately, the report counts on the success of that effort for return of non-CLIA results, although success is far from assured and such a system will take years to develop. As a last resort, the report allows Institutional Review Boards (IRBs) to approve return of results, but only after extensive and restrictive vetting.

The report thus constricts participant access to results, while providing no greater protection in clinical care, as accepted approaches already call for CLIA confirmation before using research results for diagnosis or treatment. The report expresses concern about possible sample mix-up in returning results but notes that research laboratories may have sound tracking systems (even better in some cases than CLIA laboratories) and details how researchers can communicate any uncertainty. The report's retreat from robust return of results is puzzling, given the report's acknowledgment that the risks of return have been overstated and the benefits underestimated.

Efforts to turn back the clock on return of results appear rooted in confusion about the law. The Academies' report states that CLIA "bars laboratories that are not CLIA certified from reporting individual research results" (*7*). However, neither the CLIA statute nor regulations impose this ban. They only prevent non-CLIA laboratories from reporting results "for the diagnosis, prevention or treatment of any disease or impairment of, or the assessment of…health"—that is, for clinical uses (*9*, *10*). When the purpose for return of results is to recommend that the participant seek clinical confirmation and evaluation, rather than for direct use in clinical care, CLIA does not apply (*8*). Nor does it apply when the goal of returning results is to respect the many non-

[1]*Law School, Medical School, Consortium on Law and Values in Health, Environment & the Life Sciences, University of Minnesota, Minneapolis, MN, USA.* [2]*Law Center, Department of Electrical and Computer Engineering, Center on Biotechnology and Law, University of Houston, Houston, TX, USA. Email: swolf@umn.edu*

GRAPHIC: N. CARY/*SCIENCE*

clinical reasons why participants want their results and data.

The report's error stems from giving credence to a highly restrictive CMS position that a laboratory reporting an individual's research results for any reason is doing so for clinical use and thus needs CLIA certification: "[T]he committee was advised that making any comments, analysis, or conclusions regarding the appropriateness of that interpretation would be beyond what is intended in the statement of task" (*7*). However, both the CLIA statute and regulations are clear—only laboratories that are reporting results for clinical use require CLIA certification. The mere act of reporting individual-specific results is not enough. CMS has no power to revise the statute, creating new rules and presumptions that would demand CLIA compliance by additional laboratories—only Congress can do that (*11*). Although the report ultimately recommends that CMS should allow somewhat broader return (under CLIA, the proposed QMS, or with IRB approval), the report impedes current return by wrongly assuming a CLIA laboratory is required by existing law.

"The report too often prefers paternalistic silence over partnership."

The report restricts access to research results and raw data, even when participants themselves request it. Under the Health Insurance Portability and Accountability Act (HIPAA) Privacy Rule, research participants have a legally protected right of access to their data and results in the "designated record set" (DRS) at HIPAA-covered clinical and research laboratories (*12*). HIPAA ensures a person's right of access to their information so that they can assess the level of privacy risk that circulation of that information entails.

In the Genetic Information Nondiscrimination Act (GINA), Congress made clear that these HIPAA access rights include genetic information. GINA recognizes that people can suffer invasion of privacy, discrimination, stigmatization, and other harms based on erroneous, low-quality genetic information, including information that is wrongly attributed to them. Accordingly, Congress placed genetic information under the protections of the HIPAA Privacy Rule, including HIPAA's access right (*13*).

The report advocates restricting these federal access rights. It calls on the HIPAA regulator (the Office of Civil Rights) to revise its definition of the DRS to exclude research data and results unless they meet the CLIA and QMS quality standards recommended by the report. Again, the report fails to trust participants with research data and results while misunderstanding the law. The report ignores a longstanding Congressional determination, reflected in privacy statutes for nearly 50 years [for example, (*14*)], that broad individual access is essential to privacy protection.

Given the problems in requiring compliance with CLIA or the proposed QMS, the report creates one last option—researchers may offer results if approved by an IRB. Here the report reverts to an old strategy with recognized drawbacks, by loading return-of-results decisions on IRBs to be made on a "study-by-study" basis. Although IRBs already play a role in return of results, the report goes further and calls on IRBs to decide with researchers "whether and how" to return results in each study; to determine whether the quality of results supports return through CLIA-certification, a QMS, or IRB approval; and to "develop policies and procedures that support the assessment of plans for…return" (*7*). This places substantial new burdens on IRBs, despite extensive literature on the limits of IRB decision-making and inconsistency in decisions across IRBs (*15*).

The report then maximizes the burden on IRBs by mischaracterizing existing consensus guidelines (*1*–*4*) and suggesting that IRBs start over. Those guidelines already embrace values endorsed by the report—assessing the value of results from the participant's perspective and "defin[ing] 'utility' to include information that a research participant is likely to find important," rather than focusing merely on what clinicians find actionable (*1*, *2*). These published guidelines distinguish results that should be returned, may be returned, and should not be returned and have offered a starting point for researchers and participants. Yet instead of building on that literature, the report repeatedly misstates the content of these guidelines, eroding progress already made.

The report ultimately recommends that IRBs reject return of results that current guidelines allow. The report urges returning only those results that have "high quality," calling on IRBs to evaluate the analytic validity, clinical validity, and value of the results, with in-depth review of laboratory quality. The report advocates that IRBs permit return only when "the probability of value to the participant is sufficiently high and the risks of harm are sufficiently low" and "the quality of laboratory analysis… provide[s] confidence in the result, as determined by a review process independent of the laboratory" (*7*). This contrasts with current guidelines that allow return when sufficiently important to participants, even when return may involve risks (*1*–*4*). Moreover, current guidelines would allow more generous return of results, based on the researcher's efforts to ensure the result's analytic validity and an assessment of the result's value—including its value in triggering clinical confirmation and follow-up—without insisting on the extensive review required by the report.

The Academies' report endorses the idea of participant access to results and data, but then builds daunting barriers. The report rejects established legal rights of access, two decades of consensus guidelines, and abundant data showing that participants benefit from access while incurring little risk. The report too often prefers paternalistic silence over partnership. Although the report acknowledges that research is transitioning to models involving participant engagement and leadership, the report creates roadblocks to this partnership.

True progress on return of results requires accepting participants' established rights of access and respecting the value that participants place on broad access to their data and results. The next step is not to build barriers but to promote transparency. Access is central to advancing participant-centered research practices and building successful collaboration with research participants. ■

REFERENCES AND NOTES

1. S. M. Wolf *et al.*, *J. Law Med. Ethics* **36**, 219 (2008).
2. R. R. Fabsitz *et al.*, *Circ. Cardiovasc. Genet.* **3**, 574 (2010).
3. S. M. Wolf *et al.*, *Genet. Med.* **14**, 361 (2012).
4. G. P. Jarvik *et al.*, *Am. J. Hum. Genet.* **94**, 818 (2014).
5. K. E. Boronow *et al.*, *Environ. Health Perspect.* **125**, A27 (2017).
6. A. Thorogood *et al.*, *Hum. Genomics* **12**, 7 (2018).
7. The National Academies of Sciences, Engineering, and Medicine, "Returning individual research results to participants: Guidance for a new research paradigm" (Consensus Study Report, 2018); http://nationalacademies.org/hmd/Reports/2018/returning-individual-research-results-to-participants.aspx.
8. W. Burke, B. J. Evans, G. P. Jarvik, *Am. J. Med. Genet. C Semin. Med. Genet.* **166C**, 105 (2014).
9. 42 U.S. Code (U.S.C.) § 263a (*a*).
10. 42 Code of Federal Regulations (C.F.R.) §§ 493.2, 493.3 (b) (*2*).
11. R. J. Pierce Jr., *Administrative Law Treatise* (and *2018 Cumulative Supplement*) (Wolters Kluwer, ed. 5, 2010), vol. 1, chap. 6.
12. 45 C.F.R. §§ 164.501, 164.524.
13. GINA §§ 102, 105.
14. Privacy Act of 1974, 5 U.S.C. § 552a.
15. C. Grady, *JAMA* **304**, 1122 (2010).

ACKNOWLEDGMENTS

This work was supported by the National Human Genome Research Institute and National Cancer Institute of the National Institutes of Health (NIH) under award number R01HG008605. The content is solely the responsibility of the authors and does not necessarily represent the official views of the funders or the views of other members of the funded project. S.M.W. is a member of the National Academy of Medicine, has presented at a workshop held by the committee producing the Academies' report on return of results, and served as a reviewer of the report, as the report acknowledges.

10.1126/science.aav0005

A health worker administers a polio vaccine to children in Karachi, Pakistan, in May 2018.

PHOTO: RIZWAN TABASSUM/AFP/GETTY IMAGES

BOOKS *et al.*

PUBLIC HEALTH

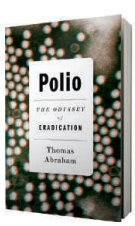
The persistence of polio

Inspired by the successful eradication of smallpox, efforts to eliminate poliovirus haven't achieved the same success

By **Pamela J. Hines**

In the mid-20th century, several successful vaccines against polio were developed, eventually leading to an international initiative in 1988 to eradicate the poliovirus from the planet. After the success of smallpox eradication in 1980, the optimism of the initiative's advocates and supporters seemed well placed. But some viruses are more unruly than others.

Is poliovirus eradication an achievable and worthwhile goal or a misapplication of public health efforts? Thomas Abraham poses this question in *Polio: The Odyssey of Eradication*, as he explores the complex and conflicting paths through the quest to see the end of polio. It's a tale of personalities as well as biology, of evolution and ecosystems.



Polio
The Odyssey of Eradication
Thomas Abraham
Oxford University Press, 2018. 280 pp.

Discovery of polio's key vaccines was first stalled by scientific missteps and later driven by scientists in competition. Many people dedicated their lives to the vaccination programs in hope of eradicating this disease, and others lost theirs to violent resistance against those same campaigns. Evolution and natural selection drove resurgence of new poliovirus strains in regions thought to be cleared. Poverty and remote geography enabled pockets of viral persistence. And the mismatch between where the motivation came from and where the action needed to happen weakened the end game.

Abraham comes to this topic as an outsider of sorts. Not an epidemiologist or medical historian, rather he is a professor who teaches health and science journalism. He extracts knowledge gained from analysis of documents of the World Health Organization and interviews with scientists. And he shares with the reader his own observations from remote parts of the world where poverty meets viral persistence.

The book contains good technical, but generally accessible, background on how the poliovirus works for its own interests at human expense. The biochemistry behind the vaccine discoveries and the molecular evolution that generated new strains are well explained. Abraham gets into more complex territory, though, when he delves into the motivations driving the eradication campaign and the campaign's social impacts.

The poliovirus spreads in environments where poor sanitation contaminates water supplies. Children who encounter the virus early in life tend to be less affected, whereas children who are protected from early virus exposure by good water and sanitation systems may be left susceptible to more severe disease when they encounter the virus later on. Thus, better water and sanitation systems may have paradoxically increased the risk of severe disease, offering a potential explanation for why the U.S. polio epidemic seemed to hit children in well-off suburban neighborhoods particularly hard. But regardless of what drove the disease's spread, the result was a hue and cry to stop the polio outbreaks from countries that had the scientific expertise and money to go after the problem.

Because of the efforts that ensued, large portions of the world are now free from the constant threat of polio. Not since 1979 has a case of polio originated in the United States, which is stunning progress given that some 58,000 cases of polio originated in the United States in 1952.

The 1988 initiative to completely eradicate polio worldwide, however, has failed to replicate this success. Abraham discusses the challenges to an eradication campaign in which motivation and application are not always pulling in tandem.

With polio cleared from wealthier nations, most of the action now occurs in developing areas, where public health systems must deal with a variety of pressing challenges, of which polio is only one. The virus persists in impoverished regions, where inadequate water and sanitation systems facilitate its spread. Weak health care systems struggle to sustain even minimal programs against scourges of childhood, ranging from measles to malnutrition.

The vaccines do work. But controlling the poliovirus has proven to be more difficult than expected. Abraham challenges us to apply lessons learned from the polio eradication campaign to inform future global public health endeavors. His book gives plenty for scholars to debate, particularly with regard to whether the passionate perseverance toward eradication defined by yesterday's technology has outlived its usefulness. Then again, with fewer than 200 cases of poliovirus reported in 2017 (*1*), perhaps it would be better to persist. ■

The reviewer is on staff at Science. *Email: phines@aaas.org*

REFERENCE

1. World Health Organization; http://polioeradication.org/polio-today/polio-now.

10.1126/science.aav3853

SCIENCE LIVES

A Nobel laureate in exile

The life and legacy of Muhammad Abdus Salam come into focus in a new film

By **Andrew Robinson**

The film *Salam: The First ****** Nobel Laureate* begins and ends with a gravestone in a quiet mass cemetery in Pakistan. We see the deceased's name: "Professor Muhammad Abdus Salam," followed by his birth and death dates, 1926 and 1996, and a puzzling inscription that declares that Salam "in 1979 became the first Nobel laureate for his work in physics." As the camera lingers, it becomes clear that a word has been painted out between "first" and "Nobel." But only after watching the film do we comprehend it: "Muslim."

Salam
The First **** Nobel Laureate**
Anand Kamalakar, director
Kailoola Productions, 2018. 76 minutes.

Salam shared the 1979 Nobel Prize in physics with U.S. physicists Steven Weinberg and Sheldon Glashow, who appear in the documentary along with other collaborators. The 1979 prize was awarded for work on the interaction of subatomic particles, out of which arose the "standard model" of particle physics. Salam independently developed a unified theory to explain electromagnetic interactions and the weak nuclear force.

The reviewer is the author of Einstein: A Hundred Years of Relativity *(Princeton Univ. Press, 2015) and* Satyajit Ray: The Inner Eye *(I. B. Tauris, 2004). Email: andrew@andrew-robinson.org*

Yet Salam was never widely accepted in his home country. He was a member of the minority Ahmadiyya Islamic community, which regards a 19th-century leader, Mirza Ghulam Ahmad, as a prophetic figure. This assertion clashes with the orthodox view that the Prophet Muhammad was Islam's final prophet. In 1953, anti-Ahmadi feeling led to riots in Pakistan. "It became quite clear to me that either I must leave my country or leave physics," Salam declares in an archival recording that appears in the film. "And with great anguish I chose to leave my country."

He returned to Britain, where in the 1940s he had been educated at the same Cambridge college as Paul Dirac, funded by a Small Peasants' welfare fund set up by the prime minister of the state of Punjab. But despite strained relations, he continued to be a key adviser to the Pakistani government on science, including its nuclear program.

In 1974—the year of India's first successful nuclear bomb test—the government constitutionally declared Ahmadis to be non-Muslims, provoking a shocked Salam to resign his advisory position. In 1984, under the military dictatorship of Zia-ul-Haq, Ahmadis were officially forbidden from practicing as Muslims: an ordinance enforced by a prison sentence. Since then, their violent persecution has continued to increase.

"We, the present generation, seem to have inherited a house which has no windows, and its walls are very high," Salam grimly observed. "It's very difficult to know whether we have inherited a house or a prison." At this point, we begin to understand the desecration of Salam's gravestone and the fact that not a single monument or university in Pakistan carries his name.

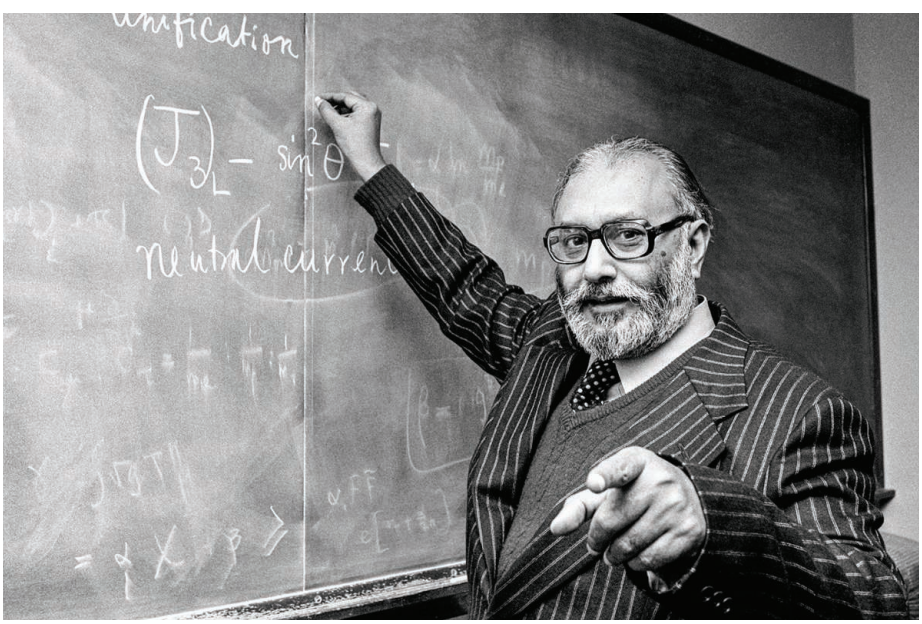
The film is frank about Salam's human failings. Science, and latterly Islam, dominated his being. The imam of his London mosque reveals how Salam would scribble fresh ideas in his scientific notebook during sermons. Both of his devoted wives—Amtul Hafeez Begum, a Pakistani, and Louise Johnson, a noted English biochemist—state that he had little time for ordinary family life. His long-time personal secretary at the International Centre for Theoretical Physics in Trieste, Italy, which Salam founded in 1964 to encourage young physicists from developing countries, says, "He was very charismatic, he was very humane, he was also very difficult."

Salam was also ambivalent about Pakistan's desire for nuclear weapons, lending his support until 1974, after which he became an advocate of science for peaceful purposes. Viewers may be reminded of Robert Oppenheimer, whom Salam knew personally, or even of Albert Einstein, who advocated the atomic bomb to President Franklin Roosevelt in 1939 and then changed his mind after the defeat of Nazi Germany in 1945.

The New York–based director of *Salam*, Anand Kamalakar, was born in India, whereas the film's two producers, Omar Vandal and Zakir Thaver, also based in the United States, were born in Pakistan. The producers met while studying at The College of Wooster in Wooster, Ohio. Their celebratory and tragic film, 14 years in the making, and plainly a labor of love, has been financed by 379 donors (duly listed in the credits), with no funds from the Ahmadiyya community, so as to preserve its independence.

What emerges most strikingly from *Salam* is the vital internationalism of science and the disturbing contrast between Salam's distinct intelligence and the religious fanaticism with which he had to contend. Let us hope that the film will inspire others in similar circumstances to follow in his footsteps. ■

10.1126/science.aav4296



Salam lectures at the Imperial College of London in 1979, the year he shared the Nobel Prize in physics.

PHOTO: PA IMAGES/CONTRIBUTOR/GETTY IMAGES

LETTERS

Edited by **Jennifer Sills**

Editor's note: Harassment policy

We have heard from readers with concerns about the publication of the Letter "Harassment charges: Injustice done?" (A. Moya *et al.*, 17 August, p. 655). At *Science*, we take harassment issues very seriously. We are working to develop policies that adhere to our editorial principle of airing a wide range of perspectives and that also prevent causing further harm to the targets of harassment.

Publication of a Letter does not represent an endorsement by the editorial staff at *Science*. Past practice has been based on the understanding that reader-submitted Letters are intended to reflect conversations taking place in the scientific community. The published Letter in question did that by raising concerns about the transparency of the investigatory process. This point touched on the challenges institutions face when determining how the processes and outcomes of harassment investigations should be shared, decisions that must weigh the benefits of transparency against important privacy concerns. However, the Letter also discussed the personal conduct and scientific accomplishments of the individual found guilty of harassment.

In the future, we will not publish Letters in which authors argue that an individual accused or found guilty of harassment is likely innocent because others have interacted with that person without incident; this argument is logically flawed. In addition, although some information about a person's scientific achievements is at times necessary to establish context, we will not publish Letters in which authors argue that professional achievements have any bearing at all on the likelihood that the individual engaged in harassment. Such arguments not only lack relevance to harassment behavior but also may result in further harm to the targets of harassment and exacerbate the already daunting process that targets face in coming forward publicly.

We are striving to increase our understanding of all facets of the issue of harassment and to review and modify our editorial processes accordingly.

Jeremy Berg
Editor-in-Chief

10.1126/science.aav6183

PPR virus threatens wildlife conservation

Peste des petits ruminants (PPR), a viral disease that affects domestic small ruminants with high morbidities and mortalities across more than 70 countries, engenders a global cost of US$2.1 billion and compromises livelihoods of some 900 million poor and low-income people (*1*, *2*). PPR has generally been regarded globally as only a livestock problem (*1*), but it is also a wildlife conservation challenge.

PPR has spread from its historic range of Africa, West Asia, and the Middle East to the vast and remote steppes and mountains of Eastern Asia, threatening even more wildlife. The mass mortality event affecting more than two-thirds of the critically endangered Mongolian saiga (*Saiga tatarica mongolica*) population in 2017 is one example of PPR's reach (*3*). PPR has also been diagnosed in mountain ungulate mortalities in the Middle East (*4*) and detected along the Himalayas (*5*) and the Tian Shan and Altai ranges (*6*).

PPR clearly threatens saiga populations, but its impact on other steppe and mountain ungulates of Asia is unknown. The inaccessibility of their habitats, together with the lack of wildlife health surveillance programs across large areas, contributes to underestimating and underreporting of PPR mortality events. The circulation of PPR in Asia may have grave consequences in wild populations that already struggle with overhunting, poaching, livestock competition, and stochastic climatic events. For example, in 2018, multiple PPR-related mortalities of the Siberian ibex (*Capra sibirica*) were recorded in Mongolia (*7*), and PPR-related mortalities of wild ungulates were found in

PHOTO: WILDLIFE CONSERVATION SOCIETY (WCS) MONGOLIA

The peste des petits ruminants (PPR) viral disease has decimated the Mongolian saiga population.

Flooding during hurricanes increases the chance of chemicals contaminating the environment.

Iran (*8*). PPR is also of special concern for the conservation of the susceptible markhor (*Capra falconeri*), argali (*Ovis ammon*), and goitered gazelle (*Gazella subgutturosa*), considered vulnerable by the IUCN (*9*), and for the survival of the snow leopard (*Panthera uncia*), which relies on prey abundance. A robust assessment will require better baseline information on population size and trends. More efforts to conduct PPR surveillance and integrate wildlife protection into control strategies are urgently required.

Xavier Fernandez Aguilar,[1]* Amanda E. Fine,[2] Mathieu Pruvot,[2] Felix Njeumi,[3] Christian Walzer,[2,4] Richard Kock,[1] Enkhtuvshin Shiilegdamba[2]

[1]Department of Pathobiology and Population Sciences, Royal Veterinary College, London NW1 0TU, UK. [2]Wildlife Conservation Society, Bronx, NY 10460, USA. [3]Food and Agriculture Organization of the United Nations, FAO, 00153 Rome, Italy. [4]Research Institute of Wildlife Ecology, University of Veterinary Medicine, A-1160 Vienna, Austria.
*Corresponding author.
Email: xfdezaguilar@gmail.com

REFERENCES

1. "Global strategy for the control and eradication of PPR" (Tech. Rep. FAO-OIE, 2015).
2. B. Jones *et al.*, *PLOS ONE* **11**, e0149982 (2016).
3. R. Kock, "Investigation of peste des petits ruminants (PPR) among wild animals and its potential impact on the current PPR situation in livestock" (Crisis Management Centre for Animal Health, Tech. Rep. FAO-OIE, 2017).
4. M. Marashi *et al.*, *Emerg. Infect. Dis.* **23**, 704 (2017).
5. J. Bao *et al.*, *China Res. Vet. Sci.* **90**, 238 (2011).
6. J. Li *et al.*, *Transbound. Emerg. Dis.* **64**, e43 (2017).
7. Science for Nature and People Partnership (SNAPP) team: Steppe Health (https://snappartnership.net/teams/steppe-health/).
8. ProMED-mail, "Peste des petits ruminants - Iran: (AL) wildlife" (Archive number 20180928.6058637, 28 September 2018); www.promedmail.org.
9. The IUCN Red List of Threatened Species, *Gazella subgutturosa* (www.iucnredlist.org/details/8976/0).

10.1126/science.aav4096

Preventing chemical release in hurricanes

With the promise of a dramatic hurricane season, the Perspective "Regulate to reduce chemical mixture risk" (A. Kortenkamp and M. Faust, 20 July, p. 224) provides additional context about the environmental risk from chemical mixtures. As a result of climate change, coastal cities around the world are predicted to experience severe storms and strong hurricanes with greater regularity, taxing infrastructure and the resources of response agencies (*1*, *2*). Storm-related flooding and coastal storm surges increase the risk of effluent release and transport (*3*), exposing populations to chemical and waste products and the potential for long-lasting health impacts (*4*, *5*). The combination of sewage overflow, petroleum runoff from roads, and discharge from chemical plants, such as those that occurred under Hurricane Harvey (*6*) and Hurricane Florence (*7*), can damage the health of first responders, the local population, and the physical environment (*7*–*10*).

As Kortenkamp and Faust explained, local, state, and national pre-disaster risk models have not yet comprehensively incorporated into their algorithms the location and potential for sudden effluent or chemical release during a natural disaster. We must work to update these models in order to protect both human and ecosystem health (*11*). Ensuring the stability of chemical and sewage infrastructure to prevent unregulated compound mixing in the environment is critical to securing health under an increasingly volatile climate regime. A coastal urban ecosystem already suffering from storm damage must be protected from uncontrolled pollution exposure, and this reality needs to be integrated into long-term planning for regional and national agencies.

Kimberley Miner,[1,2]* Nicole Wayant,[2] Heather Ward[2]

[1]Climate Change Institute, University of Maine, Orono, ME 04469, USA. [2]Geospatial Research Lab, U.S. Army Engineer Research and Development Center, Alexandria, VA 22315, USA.
*Corresponding author.
Email: kimberley.miner@maine.edu

REFERENCES

1. C. Rosenzweig, W. Solecki, *Glob. Environ. Chang.* **28**, 395 (2014).
2. U.S. National Climate Assessment, *Climate Change Impacts in the United States* (2014).
3. A. Kenward, N. Zenes, J. Bronzan, J. Brady, K. Shah, "Overflow: Climate change, heavy rain, and sewage" (2016); http://assets.climatecentral.org/pdfs/Overflow_sewagereport_update.pdf.
4. M. Nadal, M. Marquès, M. Mari, J. L. Domingo, *Environ. Res.* **143**, 177 (2015).
5. V. Kapoor, I. Gupta, A. B. M. T. Pasha, D. Phan, *Environ. Sci. Technol. Lett.* **5**, 322 (2018).
6. H. Tabuchi, S. Kaplan, "A sea of health and environmental hazards in Houston's floodwaters," *New York Times* (2017); www.nytimes.com/2017/08/31/us/houston-contaminated-floodwaters.html.
7. K. Pierre-Louis, K. Popovich, H. Tabuchi, "Florence's floodwaters breach coal ash pond and imperil other toxic sites," *New York Times* (2018): www.nytimes.com/interactive/2018/09/13/climate/hurricane-florence-environmental-hazards.html.
8. J. Cabral-Oliveira, J. Pratas, S. Mendes, M. A. Pardal, *Hum. Ecol. Risk Assess.* **21**, 135 (2015).
9. M. Al Aukidy, P. Verlicchi, *Sci. Total Environ.* **607**–**608**, 483 (2017).
10. R. Herbert *et al.*, *Environ. Health Perspect.* **114**, 1853 (2006).
11. J. Xie *et al.*, *Environ. Model. Softw.* **95**, 143 (2017).

10.1126/science.aav3822

TECHNICAL COMMENT ABSTRACTS

Response to Comment on "U-Th dating of carbonate crusts reveals Neandertal origin of Iberian cave art"

D. L. Hoffmann, C. D. Standish, M. García-Diez, P. B. Pettitt, J. A. Milton, J. Zilhão, J.J. Alcolea-González, P. Cantalejo-Duarte, H. Collado, R. de Balbín, M. Lorblanchet, J. Ramos-Muñoz, G.-Ch. Weniger, A. W. G. Pike

Slimak *et al.* challenge the reliability of our oldest (>65,000 years) U-Th dates on carbonates associated with cave paintings in Spain. They cite a supposed lack of parietal art for the 25,000 years following this date, along with potential methodological issues relating to open-system behavior and corrections to detrital or source water ^{230}Th. We show that their criticisms are unfounded.
Full text: dx.doi.org/10.1126/science.aau1736

TECHNICAL RESPONSE

PALEOANTHROPOLOGY

Response to Comment on "U-Th dating of carbonate crusts reveals Neandertal origin of Iberian cave art"

D. L. Hoffmann[1], C. D. Standish[2], M. García-Diez[3], P. B. Pettitt[4], J. A. Milton[5], J. Zilhão[6,7,8], J. J. Alcolea-González[9], P. Cantalejo-Duarte[10], H. Collado[11], R. de Balbín[9], M. Lorblanchet[12], J. Ramos-Muñoz[13], G.-Ch. Weniger[14,15], A. W. G. Pike[2]*

Slimak *et al*. challenge the reliability of our oldest (>65,000 years) U-Th dates on carbonates associated with cave paintings in Spain. They cite a supposed lack of parietal art for the 25,000 years following this date, along with potential methodological issues relating to open-system behavior and corrections to detrital or source water ^{230}Th. We show that their criticisms are unfounded.

Slimak *et al*.'s (*1*) supposed ~25,000-year (25-ka) hiatus in the production of parietal art comes from a misunderstanding of the logic of working with minimum ages. Our results (*2*) cannot be taken to imply the existence of such a hiatus. The minimum age of 45.9 ka for Ardales ARD16 and the minimum-maximum pair of 32.1 and 63.7 ka for ARD08, 09, and 06 bound painting episodes that could fall within Slimak *et al*.'s "hiatus," as could the El Castillo red disk dated to before 40.8 ka ago (*3*). Indeed, if dates older than 65 ka are excluded, the hundreds of minimum ages we have obtained are all consistent with dates in the 40- to 65-ka interval for the stratigraphically associated paintings. The origin of the red pigment at Ardales is also questioned, but its anthropogenic nature is backed by more than a century of research (*4*–*6*), and with careful inspection it is even possible to recognize technical characters linked to the execution processes used.

Slimak *et al*.'s methodological objections relate to (i) open-system behavior, (ii) nonradiogenic ^{230}Th in source water, and (iii) detrital contamination corrections. These topics have formed the focus of discussion in previous publications (*7*, *8*) and are thoroughly assessed in (*2*).

As Slimak *et al*. acknowledge, we use a sequential sampling methodology to test for open-system behavior. When dates for subsamples are in correct stratigraphic order (i.e., from younger to older systematically from the "outside" of a crust inward toward the pigment), we can be confident that the carbonate has remained a closed system. In an open system, preservation of the chronological order of subsamples is highly unlikely. We have published multiple sequences of three or more subsamples with ages in the expected stratigraphic order, including examples from all three caves under consideration; open-system behavior is not an issue here.

Concerning nonradiogenic ^{230}Th entering the carbonates from the source water, we have dated samples from all three sites to the very recent

[1]Department of Human Evolution, Max Planck Institute for Evolutionary Anthropology, 04103 Leipzig, Germany. [2]Department of Archaeology, University of Southampton, Avenue Campus, Southampton SO17 1BF, UK. [3]Faculty of Humanities and Social Sciences, University of Isabel I, 09003 Burgos, Spain. [4]Department of Archaeology, Durham University, Durham DH1 3LE, UK. [5]Ocean and Earth Science, University of Southampton Waterfront Campus, National Oceanography Centre Southampton, Southampton SO14 3ZH, UK. [6]Departament d'Història i Arqueologia (SERP), University of Barcelona, 08001 Barcelona, Spain. [7]Institució Catalana de Recerca i Estudis Avançats (ICREA), 08010 Barcelona, Spain. [8]Centro de Arqueologia da Universidade de Lisboa (UNIARQ), Faculdade de Letras, Campo Grande, 1600-214 Lisbon, Portugal. [9]Prehistory Section, University of Alcalá de Henares, 28801 Alcalá de Henares, Madrid, Spain. [10]Centro de la Prehistoria/Cueva de Ardales, 29550 Ardales (Málaga), Spain. [11]Quaternary-Prehistory Research Group, I-PAT Research Group, D.G. Bibliotecas, Museos y Patrimonio Cultural, Junta de Extremadura, Spain. [12]CNRS-Retraité, Roc des Monges, 46200 St. Sozy, France. [13]Departamento de Historia, Geografía y Filosofía, Universidad de Cádiz, Cádiz, Spain. [14]Neanderthal Museum, 40822 Mettmann, Germany. [15]Institute of Prehistory, University of Cologne, Cologne, Germany.
*Corresponding author. Email: a.w.pike@soton.ac.uk

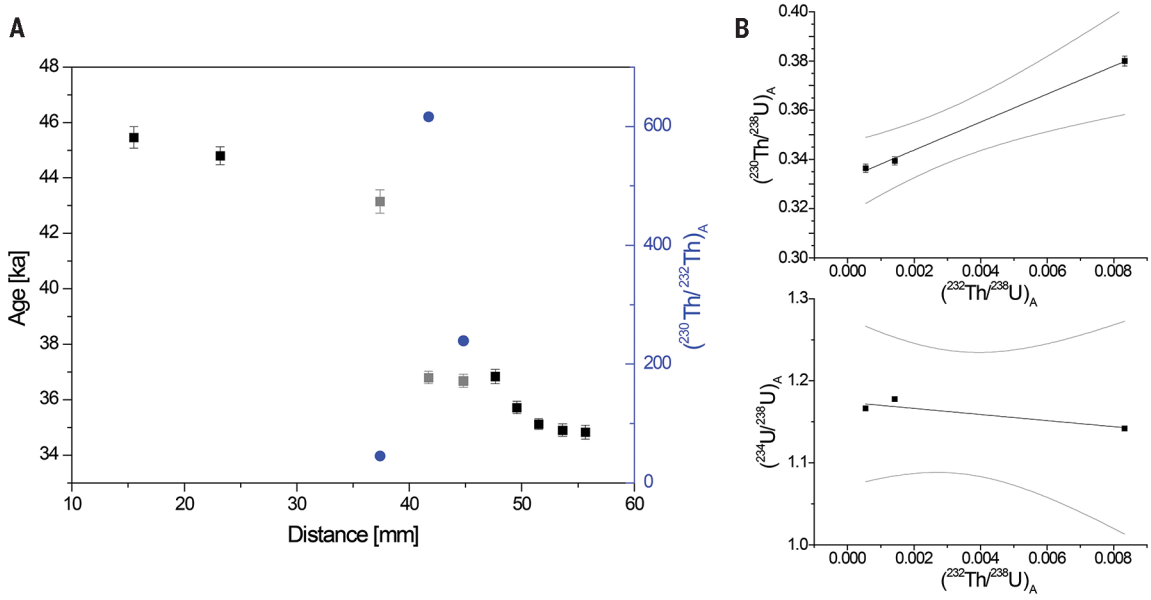


Fig. 1. A hypothetical example of erroneous isochron dating. (**A**) U-Th dating results across a 5-cm section of a flowstone (where distance is measured from the base). The results clearly reveal a 6-ka-long growth arrest at 40 mm. Dating results before and after the arrest are supported by the full dataset. (**B**) The three gray data points in (A) are used to obtain an Osmond-type pseudo-isochron, as done for PAS34a, -b, and -c by Slimak *et al*. The isochron gives an age of 36 ± 3 ka, which is clearly too young. Error bars denote 2 SD.

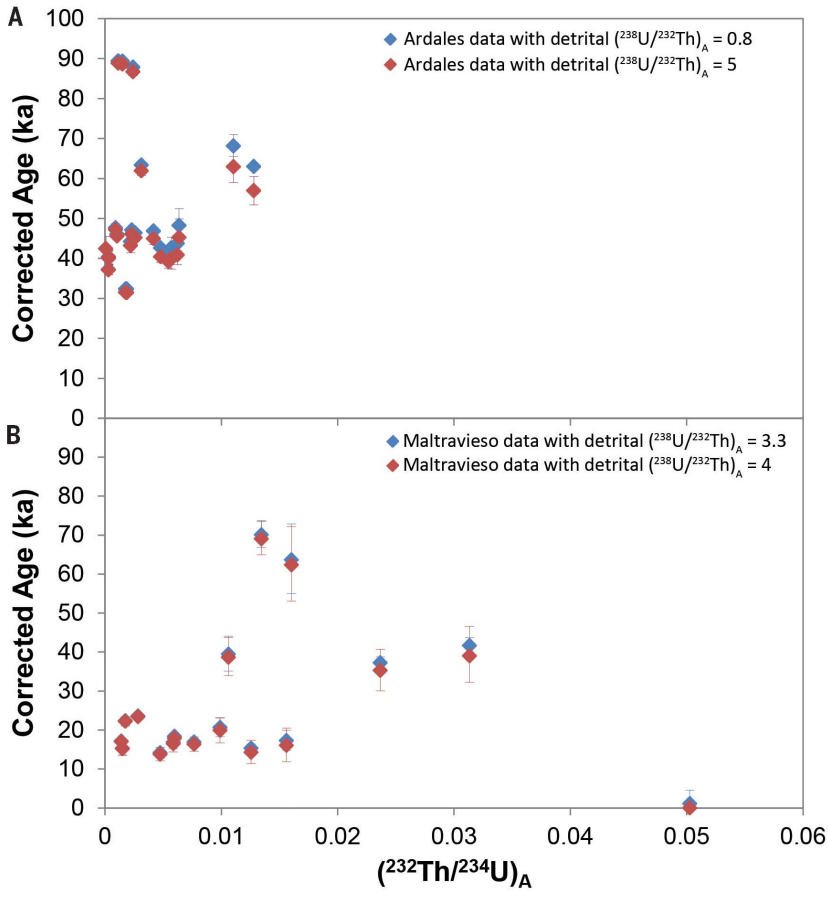


Fig. 2. Corrected ages and $(^{232}Th/^{234}U)_A$ for carbonate samples associated with art (i.e., as maximum or minimum ages). (**A**) Ardales cave; (**B**) panel GS3b in Maltravieso cave (*2*). Error bars denote 2 SD.

past, i.e., ~1 ka (e.g., PAS35a and -c) (*2*). This is entirely inconsistent with the hypothesis of high ^{230}Th drip water; dates as young as ~1 ka cannot be obtained by U-Th if the drip water has a high ^{230}Th content.

When considering detrital contamination corrections, it is true that La Pasiega PAS34c has large uncertainties due to the detrital Th correction; this was discussed at length in the supplementary materials of (*2*). There, we demonstrated that the chosen correction factor is appropriate by looking at the $^{234}U/^{238}U$ activity ratio, $(^{234}U/^{238}U)_A$, which is also affected by the detrital correction. A higher detrital $(^{238}U/^{232}Th)_A$ yields an initial $(^{234}U/^{238}U)_A$ for PAS34c inconsistent with all other samples from this cave. Slimak *et al.* propose using such an elevated detrital value, yet make no attempt to explain the effect this would have on the $(^{234}U/^{238}U)_A$. Furthermore, even if PAS34c is disregarded, PAS34a and PAS34b provide a minimum age of 53.0 ka, which still implies pre–Upper Paleolithic painting activity.

Slimak *et al.* argue for a younger minimum age for PAS34 based on an isochron derived from our results. However, deriving an isochron from three data points is not scientifically sound; a minimum of five would be needed. Furthermore, the assumption that these types of crust form within a short time is unsupported by previous results (*2*, *7*). A hypothetical example for flowstone dated sequentially by U-Th shows how misleading Slimak *et al.*'s pseudo-isochron is (Fig. 1A). The suite of dating results shows a 6-ka-long hiatus in growth at 40 mm, and the sample just below the hiatus is more contaminated than the two above. If we follow the same approach as Slimak *et al.* and use only three data points, one just below the hiatus and two above, to derive an "isochron" (i.e., assuming that all are of similar age and that the difference in detritus is the reason for the age difference), then we obtain an age of 36 ± 3 ka and a detrital $(^{238}U/^{232}Th)_A$ of 5.7 ± 0.5 (Fig. 1B). This age is clearly wrong for the sample below the hiatus, and the very high detrital correction is largely a result of the faulty assumption that the samples are coeval. The pseudo-isochron is biased by the pair of younger samples, which coincidentally are less contaminated, exactly as is the case for PAS34. Unless Slimak *et al.* can demonstrate that PAS34a, -b, and -c are contemporary, their approach is inappropriate.

All carbonate samples will be contaminated by detrital Th to some degree, and the threshold of reliability based on measured $(^{232}Th/^{238}U)_A$ or $(^{232}Th/^{234}U)_A$ that Slimak *et al.* suggest is entirely arbitrary. Of more importance is the sensitivity to the applied correction of the resulting corrected age. Figure 2 shows corrected ages and $(^{232}Th/^{234}U)_A$ for all the published Ardales and Maltravieso data (*2*) using our detrital $(^{238}U/^{232}Th)_A$ values and elevated ones. It is apparent that there is no clear positive correlation between age and $(^{232}Th/^{234}U)_A$ for either site, and the dates are relatively insensitive to the detrital correction. Critically, this means the shift in the two sets of corrected ages is not critical to our conclusion that some of the art is Neanderthal.

For Ardales (Fig. 2A), even with an unrealistic $(^{238}U/^{232}Th)_A$ value of 5, ARD13b still gives a minimum age of 59.0 ka. A highly unrealistic detrital $(^{238}U/^{232}Th)_A$ value of ≥11 is required before the corrected age of this sample is on the order of the ~47 ka that Slimak *et al.* prefer. When applying detrital Th corrections to relatively clean samples such as those from Ardales, using the bulk-earth value of $(^{238}U/^{232}Th)_A$ with a conservative error is adequate, and our applied detrital corrections are robust.

The samples from Maltravieso are characterized by higher detrital Th; thus, extra effort was made to characterize the detrital component directly. Sediment from the cave was collected and analyzed as a proxy for the samples' detrital fraction. A speleothem column was also sampled and a series of six growth layers dated to provide a control for this sediment-derived correction (*2*). In Fig. 2B, two detrital $(^{238}U/^{232}Th)_A$ values are used: 3.3 ± 0.2 (i.e., the sediment-derived correction) and an elevated value of 4 ± 2. Detrital $(^{238}U/^{232}Th)_A$ values of ≥4 are not possible, as beyond this limit the equivalent measured $(^{230}Th/^{232}Th)_A$ of one of the samples is exceeded. Shifting the detrital $(^{238}U/^{232}Th)_A$ to 4 ± 2 has very little effect on the corrected ages, giving, for example, a minimum age of 64.9 ka (instead of 66.7 ka) for MAL13a. This sample does contain a notable detrital component, but not enough to critically affect the corrected age. Finally, Slimak *et al.* cast further doubt by incorrectly claiming that the proposed Middle Paleolithic age of the Maltravieso hand stencil is based on a single sample. On the contrary, it is supported by a second sample, MAL17d ($63.6^{+9.6}_{-8.4}$ ka).

On the basis of present evidence, the most likely scenario is that in Europe, parietal art emerged prior to 65 ka ago and continued, perhaps episodically, throughout the remainder of the Paleolithic. Slimak *et al.*'s speculation that two technocomplexes dated to ~50 ka ago—the Bohunician and the Neronian—are possibly associated with modern humans sheds light on their willingness to accept a minimum age of 47 ka but not older. Their speculation is groundless. The earliest remains of modern humans in Europe, the Oase fossils from Romania, date to ~40 ka ago, and Neanderthal remains directly dated as recently as 40 to 50 ka are known across all of the then-inhabited Europe, east to west and north to south (*9*). There is no escaping the conclusion

that these temporal patterns imply Neanderthal authorship of Europe's earliest cave art.

REFERENCES AND NOTES

1. L. Slimak, J. Fietzke, J.-M. Geneste, R. Ontañón, *Science* **361**, eaau1371 (2018).
2. D. L. Hoffmann *et al*., *Science* **359**, 912–915 (2018).
3. A. W. G. Pike *et al*., *Science* **336**, 1409–1413 (2012).
4. H. Breuil, *L'Anthropologie* **XXXI**, 239–253 (1921).
5. J. L. Sanchidrián, in *Paleolítico da Península Ibérica*, A. Moure, R. de Balbín, Ed. (ADECAP, 2000), pp. 541–554.
6. P. Cantalejo *et al*., *La Cueva de Ardales: Arte Prehistórico y Ocupación en el Paleolítico Superior* (Diputación de Málaga, 2006).
7. D. L. Hoffmann, A. W. G. Pike, M. García-Diez, P. B. Pettitt, J. Zilhão, *Quat. Geochronol.* **36**, 104–119 (2016).
8. A. W. G. Pike, D. L. Hoffmann, P. B. Pettitt, M. García-Diez, J. Zilhão, *Quat. Int.* **432**, 41–49 (2017).
9. J. Zilhão, in *Dynamics of Learning in Neanderthals and Modern Humans*, T. Akazawa, Y. Nishiaki, K. Aoki, Eds. (Springer, 2013), vol. 1, pp. 21–57.

ACKNOWLEDGMENTS

Supported by Natural Environment Research Council (UK) grant NE/K015184/1, National Geographic Society grant EC0603-12, the Max Planck Society, a Royal Society Wolfson Research Merit Award (A.W.G.P.), and Research Group IT622-13 of the Basque government (M.G.-D.).

29 May 2018; accepted 5 September 2018
10.1126/science.aau1736

PRIZE ESSAY

NEUROBIOLOGY

Circuits for care

A small population of hypothalamic neurons orchestrates parenting behaviors

By **Johannes Kohl**

Raising a child to independence requires an estimated 13 million calories (*1*), near-constant attention, and the ability to survive on little sleep. Because parents perform this monumental task without any immediate benefit, it has been suspected that parental behavior relies on evolutionarily sculpted neural circuits. What do we know about the neural basis of parenting?

Classical lesion experiments in rodents have implicated many brain areas in this fascinating behavior (*2*–*4*). One region consistently identified as essential is the medial preoptic area (MPOA), nestled deep within an evolutionarily conserved part of the brain, the hypothalamus (*5, 6*). The identity of MPOA parenting neurons was elusive until a few years ago, when my mentor, Catherine Dulac, and her group found that MPOA neurons expressing the neuropeptide Galanin ($MPOA^{Gal}$ neurons) are crucial for parental behavior in both sexes (*7*).

This was a breakthrough, but a key question remained: How can a small population of neurons—10,000 out of a total of 100 million in the mouse brain—control such a complex behavior? Mouse parenting consists of stereotyped motor routines such as grooming pups, retrieving them to the nest and, in females, nursing. But parental animals also have an increased motivation to seek out infant stimuli, have distinct hormonal states, and engage less in nonparental behaviors such as mating (*3*).

We hypothesized that $MPOA^{Gal}$ neurons orchestrate these diverse behavioral components by assuming a "hub" position in a brain-wide, dedicated parenting circuit. My goal in the past 4 years has been to test this hypothesis.

Department of Molecular and Cellular Biology, Harvard University, Cambridge, MA, USA. Email: jkohl@fas.harvard.edu

A POTENTIAL PARENTING CIRCUIT EMERGES

I started my project by tracing the connections that $MPOA^{Gal}$ neurons form with the rest of the brain. An impressive palette of viruses has been developed for this purpose in recent years. Some have the ability to jump backwards through neural circuits, thereby visualizing a neuron's direct inputs. Others label the fine axonal arborizations and synaptic terminals of infected neurons. These tracing experiments revealed a staggering complexity: $MPOA^{Gal}$ neurons receive inputs from about 20 brain areas and send out a similar number of projections (*8*). However, on closer inspection, a simple organizational principle emerged: $MPOA^{Gal}$ neurons are organized in distinct pools, or subpopulations, each projecting to a different brain area. Intriguingly, each pool has access to incoming information from all 20 brain areas (see the figure).

DISCRETE NEURONAL CLUSTERS ARE ACTIVE DURING SPECIFIC PARENTING BEHAVIORS

We next asked which of these neuronal pools are crucial for parental behavior. In fact, most of the areas targeted by $MPOA^{Gal}$ neurons have been found to play a role in parenting (*2*–*4*). We therefore sought to determine which pools were most highly activated during pup interactions. Three candidate pools were identified for further investigation: those projecting to the periaqueductal gray (PAG), the ventral tegmental area (VTA) and the medial amygdala (MeA) (*8*).

Are these different pools active during specific aspects of parenting? Using fiber photometry, an imaging approach that can record population activity from genetically specified neurons in behaving animals (*9*), we found the entire $MPOA^{Gal}$ population to be activated during all components of parenting. Surprisingly, however, individual pools were tuned to discrete parenting episodes (*8*), suggesting that they might indeed represent functionally distinct modules. We tested this hypothesis by optogenetically manipulating each of the three candidate pools.

MANIPULATING SPECIFIC CLUSTERS INDUCES DISCRETE PARENTING BEHAVIORS

First, we turned to the PAG. Sexually inexperienced male mice typically attack pups, only becoming parental in the weeks following mating (*10*). Strikingly, activation of PAG-projecting $MPOA^{Gal}$ neurons suppressed

GRAND PRIZE WINNER

Johannes Kohl

Johannes Kohl received his undergraduate degree from the University of Magdeburg in Germany and his Ph.D. from the University of Cambridge. Dr. Kohl is currently a postdoctoral fellow at Harvard University and the Sainsbury Wellcome Centre for Neural Circuits and Behaviour in London where he has been exploring the neural circuits underlying parenting. In early 2019, Dr. Kohl will start his own group at the Francis Crick Institute in London.

FINALIST

Tomasz J. Nowakowski

Tomasz Nowakowski received his B.Sc (Hons.) and his Ph.D. from the University of Edinburgh. He completed his postdoctoral training at the University of California, San Francisco. Dr. Nowakowski is now an assistant professor at the University of California, San Francisco, where his research team seeks to identify the molecular mechanisms underlying cell fate specification and microcircuit formation in the developing cortex.

www.sciencemag.org/content/362/6411/169.1

FINALIST

Talia N. Lerner

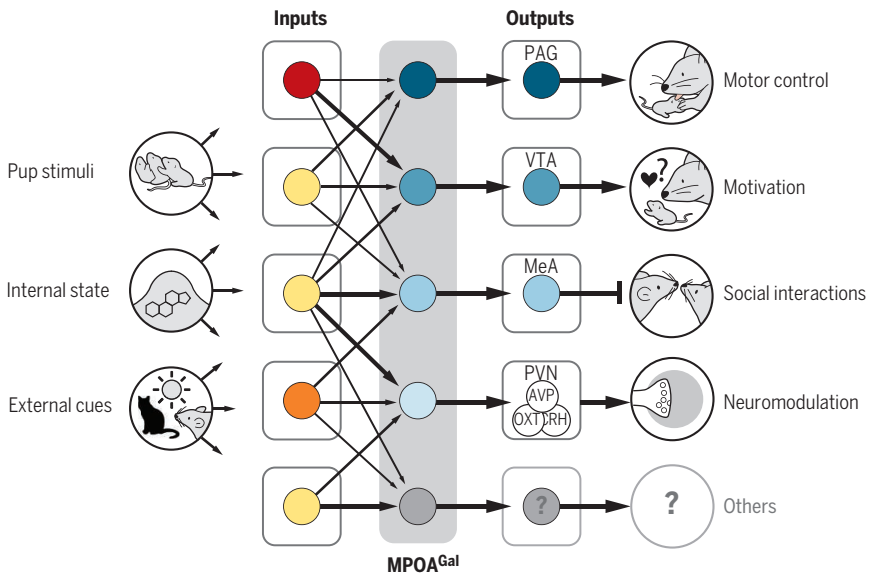
Talia Lerner received her B.S. from Yale University and her Ph.D. from The University of California, San Francisco. She then conducted postdoctoral research at Stanford University. Dr. Lerner is currently an assistant professor at Northwestern University, where she is continuing her research into how dopamine circuits regulate reward learning and habit formation and how individual differences in dopamine circuit architecture contribute to the risk for mental disorders.

www.sciencemag.org/content/362/6411/169.2

PHOTOS: (LEFT TO RIGHT) MCB GRAPHICS, HARVARD UNIVERSITY; SUSAN MERRELL/UCSF; NORTHWESTERN UNIVERSITY

Parental circuitry

A proposed model for how galanin-expressing neurons in the medial preoptic area (MPOA^{Gal}) orchestrate components of parental behavior. PAG, periaqueductal gray; VTA, ventral tegmental area; MeA, medial amygdala; PVN, paraventricular nucleus of the hypothalamus; AVP, vasopressin; OXT, oxytocin; CRH, corticotropin-releasing hormone.



pup-directed aggression in such males and increased pup grooming in both sexes, suggesting that this pool controls an important motor component of parenting. By contrast, activating the VTA-projecting neuronal pool did not directly affect pup interactions. The VTA has a well-established role in motivation and reward processing (*11*, *12*).

Because an increased motivation to interact with infants is a hallmark of parental animals (*13*), we inserted a climbable barrier between the test animal and pups. In this simple assay, activation of the VTA-projecting pool drastically increased the frequency with which animals crossed over to the pup compartment, suggesting that this circuit branch indeed controls the motivation to interact with infants.

Finally, activating those MPOA^{Gal} neurons projecting to the MeA affected neither pup interactions nor the motivation to interact with pups. However, we unexpectedly found that this manipulation suppressed interactions with adult mice in both males and females. This pool might therefore indirectly promote parenting by suppressing nonparental social behaviors.

CONCLUSIONS AND NEXT STEPS

Our work suggests a circuit motif in which projection-defined MPOA^{Gal} neuron pools each control specific aspects of parenting. This provides a novel model for how a small population of genetically defined neurons can orchestrate a complex behavior. But it also raises several intriguing questions. Do these pools interact with each other? If yes, how is their activity coordinated? If not, might moment-to-moment variations in sensory input determine which circuit element is active? And which pools control other motor aspects of parenting, such as pup retrieval or nest building?

We and others have made notable progress in uncovering how parenting and other social behaviors are wired into the brain (*8*, *14*, *15*). Such knowledge is crucial from both basic research and clinical perspectives. In humans, parental care is affected by stress and mental illnesses such as postpartum depression and anxiety, which together affect almost 20% of mothers in the United States (*16*). Addressing how physiological states and environmental factors interact with these circuits might therefore open new avenues for treatment of common mental illnesses. ■

REFERENCES

1. S.B. Hrdy, *Horm. Behav.* **77**, 272 (2016).
2. C. Dulac, L.A. O'Connell, Z. Wu, *Science* **345**, 6198 (2014).
3. J. Kohl, A.E. Autry, C. Dulac, *Bioessays* **39**, 1 (2017).
4. J. Kohl, C. Dulac, *Curr. Opin. Neurobiol.* **49** (2018).
5. A.E. Fisher, *Science* **124**, 228 (1956).
6. M. Numan, *J. Comp. Physiol. Psychol.* **87**, 746 (1974).
7. Z. Wu *et al.*, *Nature* **509**, 7500 (2014).
8. J. Kohl *et al.*, *Nature* **556**, 7701 (2018).
9. L.A. Gunaydin *et al.*, *Cell* **157**, 7 (2014).
10. F.S. vom Saal, *Physiol. Behav.* **34**, 1 (1985).
11. J.D. Salamone, M. Correa, *Neuron* **76**, 470 (2012).
12. J.A. McHenry *et al.*, *Nat. Neurosci.* **20**, 449 (2017).
13. H. Hauser, R. Gandelman, *Horm. Behav.* **19**, 454 (1985).
14. Y.Y. Fang *et al.*, *Neuron* **98**, 192 (2018).
15. P. Chen, W. Hong, *Neuron* **98**, 16 (2018).
16. K. L. Wisner *et al.*, *JAMA Psychiatry* **70**, 490 (2013).

10.1126/science.aav1249

GRAPHIC: ADAPTED FROM (*8*) BY N. CARY/*SCIENCE*

PRIZE ESSAY

FINALIST

Tomasz J. Nowakowski

Tomasz J. Nowakowski received his B.Sc (Hons.) and his Ph.D. from the University of Edinburgh. He completed his postdoctoral training at the University of California, San Francisco. Dr. Nowakowski is now an assistant professor at the University of California, San Francisco, where his research team seeks to identify the molecular mechanisms underlying cell fate specification and microcircuit formation in the developing cortex.

www.sciencemag.org/content/362/6411/169.1

NEUROBIOLOGY

Building blocks of the human brain

Area-specific excitatory neuron subtypes emerge before sensory experience

By **Tomasz Jan Nowakowski**

Classical computer software programs perform tasks in series: One algorithm has to be performed, and the output of that algorithm is used as an input into another. However, the most powerful computers in the world take advantage of the fact that complex computational tasks can be broken down into independent components. Distributing the computation over many parallel processes taking place simultaneously rapidly increases the overall speed. This paradigm is very similar to how the human brain performs its complex functions.

Our brains constantly receive information from the outside world; hundreds of thousands of sensory stimuli have to be processed in parallel and associated with each other to support a conscious perception of our environment. Unlike the engineered architecture of the computer circuit board, however, the biological hardware performing the complex operations in the brain is far from understood.

RADIAL GLIA AND CORTICAL DEVELOPMENT

The human cerebral cortex includes billions of neurons organized into six sheetlike layers. More than a century ago, Santiago Ramón y Cajal appreciated the astonishing diversity of cell types in the brain, but even today, it is unclear how the different neuronal cell types are assembled and distributed across the distinct anatomical areas of the cortex to support its diverse functions.

We know that excitatory neurons of the cortex are born from a uniform population of radial glia. During development, radial glia line the fluid-filled ventricles and extend long fibers that connect with the outer pial surface. Newborn neurons migrate along radial glia fibers to the outer layers of the tissue, where they form the cerebral cortex. This radial glia scaffold inspired the longstanding radial unit hypothesis (*1*), which posits that neurons born from the same group of radial glia migrate along the physical scaffold of the radial glia fibers to occupy a nearby position in the cerebral cortex.

The vestige of this developmental process is reflected in a tissue architecture feature called the "minicolumn" (*2*). The concept of an elementary functional unit, with a stereotyped microcircuit and a common input and output architecture, has resonated throughout studies of neuroanatomy, physiology, and developmental biology (*3*–*5*) because it provides a tractable framework for understanding the computational strategy of the human brain. Human cortical development involves a greatly expanded and diversified pool of radial glia compared with that of mice (*6*). I hypothesized that radial glia cell diversity in humans could influence the radial glia scaffold and the composition of neuronal cell types across cortical areas and layers.

MAPPING DEVELOPMENTAL TRAJECTORIES IN THE CORTEX

As a first step toward unraveling the diversity of cell types and developmental processes involved in cortical development, I pioneered the use of single-cell RNA sequencing for classifying diverse cell types mixed together in tissue samples (*7*–*9*). Using these transcriptomic signatures as biomarkers, I could selectively label distinct subtypes of radial glia and study their abundance, morphology, and distribution in the developing human brain (*10*).

Unexpectedly, I found that the radial glial scaffold undergoes a transformation from a physically continuous to a physically discontinuous structure, composed of radial fibers derived from two distinct types of radial glia, midway through human neurogenesis (*7*). This observation suggests that in primates, developmental histories of cells in the same cortical minicolumn may follow more complex trajectories than previously thought, particularly in the upper cortical layers, which are believed to support the higher-order cognitive functions. I synthesized these

Department of Anatomy and Department of Psychiatry, University of California, San Francisco, San Francisco, CA 94158, USA. Email: tomasz.j.nowakowski@gmail.com

eppendorf & Science PRIZE FOR NEURO BIOLOGY

PHOTO: SUSAN MERRELL/UCSF

findings into a revised model of human brain development (*7*) that provides a new framework for the study of lineage relationships during cortical minicolumn formation.

Cortical minicolumns are arranged serially throughout the cerebral cortical sheet and are composed of a relatively small number of major neuronal cell classes (*11*, *12*). This modular organization has inspired a uniformity hypothesis (*13*, *14*) that the cortex is composed of elementary units with similar cellular composition and a stereotyped connectivity pattern.

However, differences in cortical cytoarchitecture and gene expression across functional areas have long been appreciated (*15*–*18*). It is currently unclear whether these molecular differences are established by bona fide, area-specific neurodevelopmental programs or by neuronal activity. I reasoned that studying neuronal cells during their generation could reveal the molecular dynamics that regulate the emergence of diverse cell types in the brain and could address questions about cortical minicolumn composition, microcircuit development, and ultimately, brain function.

EXCITATORY NEURONS AND CONNECTIVITY

As a corollary to the radial unit hypothesis, I postulated that cortical areas would emerge from a primordial "protomap" encoded by radial glia. In this model, a mosaic of prespecified cortical area information would be projected onto the nascent cerebral cortex via lineage relationships between radial glia and neurons.

As part of my laboratory's effort to characterize the diversity of cell types in the developing brain by using single-cell RNA sequencing, we profiled cells from frontal and occipital lobes of the cortex at stages of neurogenesis, before sensory input shapes neuronal activity in the cortex. Whereas most cell types shared the same transcriptional signatures in prefrontal and visual areas, excitatory neurons segregated into molecularly distinct groups according to the area where they were generated (*19*). Many of the genes that distinguish excitatory neurons found in different cortical areas included factors previously shown to determine patterns of neuronal connectivity.

Although future experiments are needed to compare cortical microcircuit architecture in specific cortical areas, our findings suggest that in the developing human cortex, topographically distinct minicolumns consist of different classes of excitatory neurons. Importantly, similar area-specific neuronal signatures have been reported in the adult mouse brain (*20*), indicating that the topographic variation that we described reflects differences in neuronal identities, rather than transient developmental variation (*21*).

A COMPLEX SYSTEM EMERGES

Together, our studies show that the human radial glia scaffold transforms during development, which may increase the dispersion of upper-layer cortical neurons. In addition, even before sensory experience, cortical neurons are functionally distinct across cortical areas in addition to layers. These findings support a model of serial homology, in which topographic hierarchy of neurodevelopmental programs orchestrates differences in excitatory neurons across individual cortical areas and may contribute to interareal differences in microcircuit connectivity patterns to support higher-order information processing (*22*–*24*). This model could have important implications for the design of improved deep learning algorithms that seek to faithfully recapitulate computational processes in the brain (*25*, *26*) and has the potential to reveal developmental processes underpinning human cognition. ■

REFERENCES

1. P. Rakic, *Postgrad. Med. J.* **54**, 25 (1978).
2. N. R. De Lorente, *Physiology of the Nervous System*, 288-330 (1949).
3. D. H. Hubel, T. N. Wiesel, *J. Physiol.* **160**, 106 (1962).
4. P. Rakic, *Science* **241**, 170 (1988).
5. J. DeFelipe, S. H. Hendry, T. Hashikawa, M. Molinari, E. G. Jones, *Neuroscience* **37**, 655 (1990).
6. D. V. Hansen, J. H. Lui, P. R. Parker, A. R. Kriegstein, *Nature* **464**, 554 (2010).
7. T. J. Nowakowski, A. A. Pollen, C. Sandoval-Espinosa, A. R. Kriegstein, *Neuron* **91**, 1219 (2016).
8. A. A. Pollen *et al.*, *Cell* **163**, 55 (2015).
9. A. A. Pollen *et al.*, *Nat. Biotechnol.* **32**, 1053 (2014).
10. Z. Molnár *et al.*, *Eur. J. Neurosci.* **23**, 921 (2006).
11. A. Zeisel *et al.*, *Science* **347**, 1138 (2015).
12. B. Tasic *et al.*, *Nat. Neurosci.* **19**, 335 (2016).
13. O. D. Creutzfeldt, *Naturwissenschaften* **64**, 507 (1977).
14. V. B. Mountcastle, *J. R. Soc. Med.* **71**, 14 (1978).
15. G. N. Elston, *Cereb. Cortex* **13**, 1124 (2003).
16. S. A. Bayer, J. Altman, *Neuroscience* **45**, 391 (1991).
17. D. D. O'Leary, *Trends Neurosci.* **12**, 400 (1989).
18. M. B. Johnson *et al.*, *Neuron* **62**, 494 (2009).
19. T. J. Nowakowski *et al.*, *Science* **358**, 1318 (2017).
20. B. Tasic *et al.*, *bioRxiv* 229542 [Preprint]. 6 December 2017. https://doi.org/10.1101/229542.
21. H. Li *et al.*, *Cell* **171**, 1206 (2017).
22. N. T. Markov *et al.*, *Science* **342**, 1238406 (2013).
23. B. B. Averbeck, A. Battaglia-Mayer, C. Guglielmo, R. Caminiti, *J. Neurophysiol.* **102**, 1911 (2009).
24. K. D. Harris, G. M. Shepherd, *Nat. Neurosci.* **18**, 170 (2015).
25. G. J. Rinkus, *Front. Neuroanat.* **4**, 17 (2010).
26. Y. LeCun, Y. Bengio, G. Hinton, *Nature* **521**, 436 (2015).

10.1126/science.aav1252

PRIZE ESSAY

FINALIST

Talia N. Lerner

Talia Lerner received her B.S. from Yale University and her Ph.D. from The University of California, San Francisco. She then conducted postdoctoral research at Stanford University. Dr. Lerner is currently an assistant professor at Northwestern University, where she is continuing her research into how dopamine circuits regulate reward learning and habit formation and how individual differences in dopamine circuit architecture contribute to the risk for mental disorders.

www.sciencemag.org/content/362/6411/169.2

NEUROBIOLOGY

The effortless custody of automatism

Differentially controlled dopamine signaling in the striatum may be critical to habit formation

By **Talia N. Lerner**

We rely on habits to get us through life. Grab your keys, back your car out of the driveway, navigate to work—each of these daily actions is made more fluid, and less error prone, with repetition and habit. As the father of modern psychology, William James (*1*), put it:

"The more of the details of our daily life we can hand over to the effortless custody of automatism, the more our higher powers of mind will be set free for their own proper work. There is no more miserable human being than one in whom nothing is habitual but indecision, and for whom the lighting of every cigar, the drinking of every cup, the time of rising and going to bed every day, and the beginning of every bit of work, are subjects of express volitional deliberation."

But how do we transfer actions from express volitional control to automatism? Generally, a great deal of repetitive practice is required to perfect a new skill. Why? What about practice encourages the brain to consolidate a new action into an old habit?

DOPAMINE AND HABITUAL BEHAVIOR: WHAT'S THE CONNECTION?

Several major brain areas are implicated in the habit-formation process. Among the most prominent are the dorsal striatum and the midbrain dopamine system. As habits form, patterns of neural activity in the dorsal striatum shift. In the dorsomedial striatum, neural activity during the performance of a new task peaks in early acquisition and then fades as habits form, whereas in the dorsolateral striatum, activity emerges and solidifies over time, in sync with the habit's emergence (*2*). These habit-linked changes are likely caused by synaptic plasticity.

A large body of evidence points to the dopamine-dependent plasticity of cortical inputs onto striatal neurons as indispensable to habit formation (*3*–*5*), and my own graduate thesis focused on the molecular mechanisms by which dopamine acts as a master controller of the timing and direction of corticostriatal synaptic plasticity (*6*). The dopamine that controls striatal synaptic plasticity is supplied by the midbrain dopamine system. As a postdoc in the laboratory of Karl Deisseroth at Stanford University, I sought to examine the structure of the dopamine system as it relates to the control of habit formation. I hypothesized that dopamine circuits are structured to support the transfer of information between largely parallel corticostriatal systems, enabling the observed coordinated shifts in activity between striatal subregions. By understanding how dopamine signals to the dorsomedial and dorsolateral regions of the striatum might be differentially controlled, I could gain insight into the mechanisms by which habit formation circuitry is engaged to participate in action selection and how feedback from these habit circuits might then suppress volitional control.

DETERMINING DIFFERENTIAL CONTROL OF DOPAMINE SIGNALING

To begin, I used whole-brain circuit mapping and imaging techniques that were being developed at Stanford. Colleagues in Liqun Luo's lab were working on a rabies-mediated circuit-tracing strategy that allowed the mapping of whole-brain inputs to a cell type defined by its output. This technique, termed TRIO (*7*), was perfectly suited to determine whether dopamine signals to the dorsomedial and dorsolateral striatum could be generated by distinct combinations of inputs.

I combined TRIO mapping with a tissue-clearing method developed in the Deisseroth lab called CLARITY, which then enabled intact imaging of mapped dopamine circuits with light-sheet microscopy (*8*, *9*). From these studies, I concluded that dopamine neurons did indeed receive differential inputs depending on their output targets. In particular, there was a bias toward reciprocal connectivity of striatal subregions with the dopamine neurons that project to those subregions (*10*).

Department of Physiology, Northwestern University, Chicago, IL 60611, USA. Email: talia.lerner@northwestern.edu

PHOTO: NORTHWESTERN UNIVERSITY

A SURPRISING OBSERVATION HINTS AT AN UNEXPECTED MECHANISM

To confirm my anatomical observations, I performed functional studies on the connection probabilities and synaptic strengths of striatal inputs to dopamine neurons using optogenetics and slice electrophysiology. This alternative method of circuit mapping broadly confirmed my TRIO findings, but to my surprise, it also led to the startling new observation that dorsolateral striatal inputs to dopamine neurons are substantially stronger than inputs from the dorsomedial striatum (*10*). This observation suggests a potential route for the suppression of dopamine transients to the dorsomedial striatum after heightened activity in the dorsolateral striatum as habits emerge.

AVERSIVE EVENTS PROVOKE DIFFERENT DOPAMINE RESPONSES

Finally, I addressed the question of whether dopamine carries information differently to the dorsomedial than to the dorsolateral striatum. Using fiber photometry, another new technique I helped develop in the Deisseroth lab, in which the activity of a genetically defined population of neurons is recorded as a bulk fluorescence signal through a fiber optic brain implant, I tracked the activity of dopamine neurons during rewarding and aversive experiences. I compared the responses of dopamine neurons projecting to the dorsomedial striatum with those of dopamine neurons projecting to the dorsolateral striatum and found that aversive events, in particular, provoked profoundly different, indeed opposite, responses in the two populations (*10*).

In parallel experiments using and elaborating on the fiber photometry technology to include sampling signals from multiple brain regions simultaneously, my colleagues and I further showed that dopaminergic projections to additional output regions such as the prefrontal cortex can also carry distinct information (*11*). Together, these studies present a strong case for using circuit features to help define dopaminergic (and other) cell types in the brain (*12*).

UNRAVELING THE MYSTERY OF HABIT-FORMATION

Looking forward, I hypothesize that input-output–defined dopamine neurons are at the crux of the brain's habit engagement circuitry. As animals learn and explore, the shift from volitional to habitual control of their actions will depend on cost-benefit calculations made by the circuits and synapses I have studied and will continue to study in my own lab. As we unravel the mysteries of habit formation across many levels of neurobiological investigation—from molecules to behavior—we will learn how to more efficiently slip new skills into the effortless custody of automatism while also developing strategies to wrench back conscious control of our more counterproductive habits from their sometimes defiant, iron grip. ■

REFERENCES

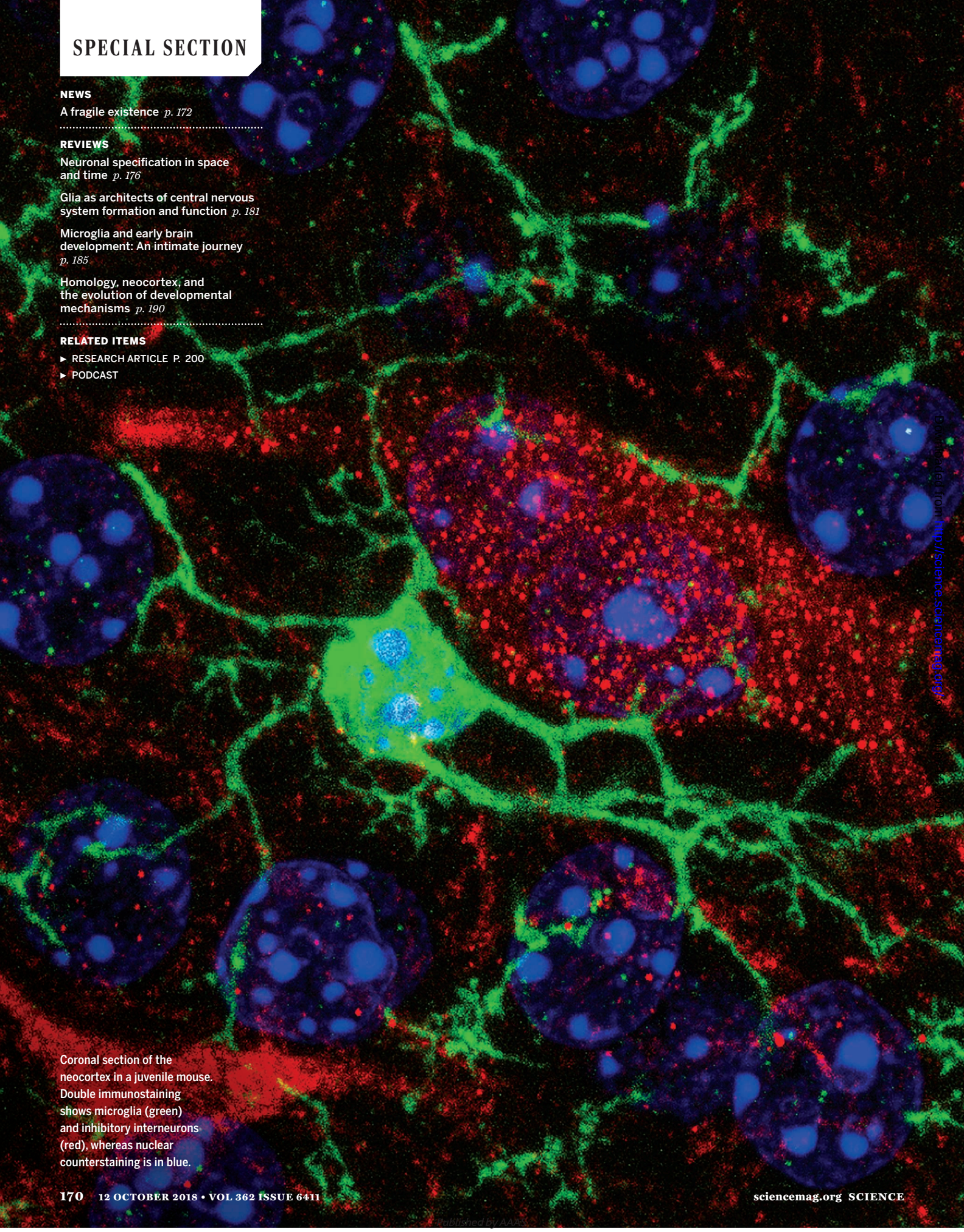
1. W. James, *The Principles of Psychology* (Henry Holt and Company, New York, 1890), vol. 1, chap. 4.
2. C. A. Thorn, H. Atallah, M. Howe, A. M. Graybiel, *Neuron* **66**, 781 (2010).
3. H. H. Yin *et al.*, *Nat. Neurosci.* **12**, 333 (2009).
4. A. C. Kreitzer, R. C. Malenka, *Neuron* **60**, 543 (2008).
5. J. N. Reynolds, B. I. Hyland, J. R. Wickens, *Nature* **413**, 67 (2001).
6. T. N. Lerner , A. C. Kreitzer, *Neuron* **73**, 347 (2012).
7. L. A Schwarz *et al.*, *Nature* **524**, 88 (2015).
8. K. Chung *et al.*, *Nature* **497**, 332 (2013).
9. R. Tomer, L. Ye, B. Hsueh, K. Deisseroth, *Nat Protoc.* **9**, 1682 (2014).
10. T. N. Lerner *et al.*, *Cell* **162**, 635 (2015).
11. C. K. Kim *et al.*, *Nat. Methods* **13**, 325 (2016).
12. T. N. Lerner, L. Ye, K. Deisseroth, *Cell* **164**, 1136 (2016).

10.1126/science.aav1250

SPECIAL SECTION

RELATED ITEMS

- ▶ RESEARCH ARTICLE P. 200
- ▶ PODCAST



Coronal section of the neocortex in a juvenile mouse. Double immunostaining shows microglia (green) and inhibitory interneurons (red), whereas nuclear counterstaining is in blue.

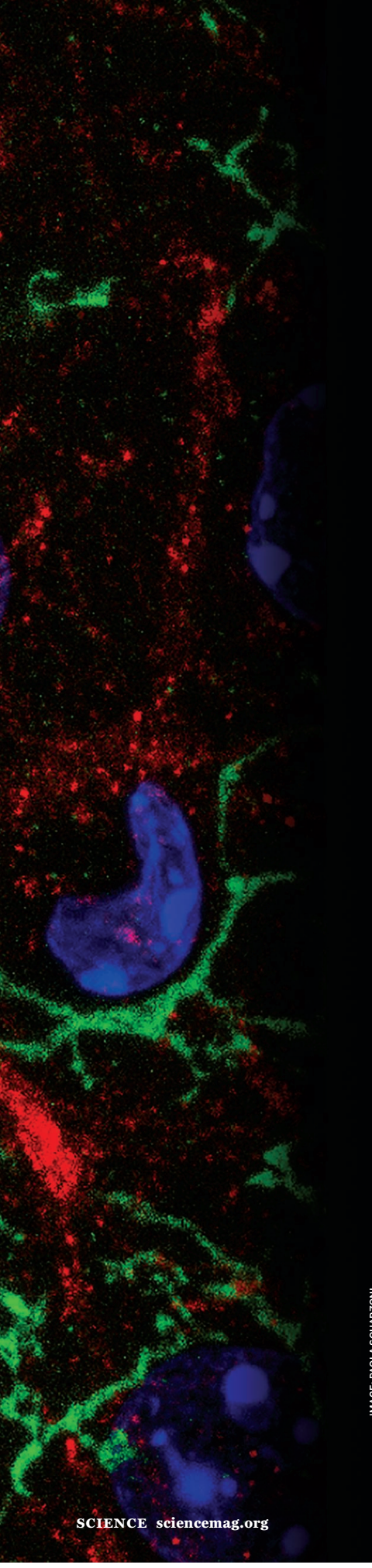


IMAGE: PAOLA SQUARZONI

MIND-BOGGLING BRAIN DEVELOPMENT

By **Pamela J. Hines**

The human brain contains billions of well-connected neurons. Neural neighborhoods perform different tasks: Some coordinate movement, whereas others hum along planning dinner. The mature brain is a complex assembly of networks, structures, and tracts. Like cities and their neighborhoods, however, the brain does not arise fully formed. Rather, operational patterns and developmental constraints guide the proliferating neurons that build the typical adult human brain. Just as cities are governed by both hard and soft infrastructure—e.g., highways channel traffic and laws define what sort of building can occur and where—the placement and function of neurons in the brain respond to multiple cues during development. In this special issue, we look behind the scenes of this elaborate process that places each neuron where it belongs or—sadly, in the cases of neurodevelopmental disorders—does not.

Generation of these neurons, along with the equally numerous accessory cells, requires enthusiastic progenitor cells. Early in development, straightforward proliferative programs morph to produce diverse cell types. With shifting cascades of transcription factors, each newborn neuron is shaped by its unique time and place.

Glia, originally viewed as bystanders in neurodevelopment, are now known to be quite the opposite. Radial glia are builders of both neural neighborhoods and highways along which neurons move. The immune system, once thought to be excluded from the brain, is also crucial to brain formation: Microglial cells regulate circuit formation and convey physiological information. Evolution constrains development according to successful precedents. Think about that with your marvelously formed brain.

A FRAGILE EXISTENCE

Why is a remote Colombian town a hot spot of an inherited intellectual disability?

By **Hannah Furfaro, Spectrum**, *in Ricaurte, Colombia; Photography by* **Juan Cristóbal Cobo**

Jair Triviño has fragile X syndrome and is nearing 60, but works exhausting hours collecting fruit in local fields.

It's late afternoon in this tiny town tucked into the Colombian Andes, when Mercedes Triviño, 82, lights the wood stove to start to prepare dinner. Smoke fills the two-bedroom home she shares with six of her adult children.

Francia, 38, one of the youngest, is the family's primary breadwinner. She brings home 28,000 Colombian pesos (roughly $10) a day harvesting papayas in the fields just outside town. "Really, what I earn is just enough for eating and nothing else," she says. Four of her siblings have fragile X syndrome, a genetic condition that causes intellectual disability, physical abnormalities, and often autism. Jair, 57, works alongside Francia when he can. Hector, 45, is also somewhat able to care for himself. Victor, 55, and Joanna, 35—who has both fragile X and Down syndrome—are less independent.

In fragile X syndrome, pieces of the X chromosome appear to be falling off.

As Mercedes serves coffee on this July afternoon, sweetening it with a hefty dose of sugar and offering her best cups to her guests, she talks about the condition that dominates the lives of her family and many others here. Her niece, Patricia, 48, who lives a few blocks away, cares for two adult sons and a nephew with fragile X. More distant kin in town, the Quinteros, also have grown children with the condition. Other neighbors are adults with fragile X who have no caretaker and look after one another.

In Colombia, this town has long been known as the home of *los bobos*, "the foolish ones"—thanks in part to a 1980s novel and a later TV series that depicted families like the Triviños. More recently, scientists have caught on that it is home to the world's largest known cluster of people with fragile X. One researcher, medical geneticist Wilmar Saldarriaga-Gil of the University of Valle (Univalle) in Cali, Colombia, has made Ricaurte the focal point of his scientific inquiry. Saldarriaga-Gil, who vacationed nearby as a child, says he has visited about a hundred times since the mid-1990s to trace how fragile X affected the town and its inhabitants—and to try to understand details of the syndrome's biology. "This is a history of scientific research, a history of my community, a history of my life," he says.

The payoff from research in this town could have global impacts. Caused by mutations in a gene called *FMR1* on the X chromosome, fragile X syndrome is the leading cause of inherited intellectual disability worldwide; it affects as many as one in 2000 men and one in 4000 women. And as a single-gene cause of autism—a recalcitrantly complex condition—fragile X has been the focus of efforts to develop drugs for autism. The proteins disrupted in people with the syndrome are also key players in brain development.

In March, Saldarriaga-Gil and his colleagues reported that at least 5% of residents here carry either the full-blown fragile X mutation or less severe "premutations" that can trigger the condition in future generations. Premutation carriers usually escape cognitive problems, but some develop physical symptoms, including tremors and fertility problems. The research here might explain such variability, which could reflect how the protein *FMR1* encodes, FMRP, interacts with other proteins and pathways.

The scale of Saldarriaga-Gil's investigations is small—Ricaurte only has 58 full mutation and premutation carriers, by his count—but the research benefits because the town's residents share the same environment and a similar genetic background, offering a natural control for some variables. "What you have [here] is something that certainly warrants a lot more intensive investigation," says Jim Grigsby,

IMAGE: MONICA SCHROEDER/SCIENCE SOURCE

a clinical health psychologist at the University of Colorado in Denver.

SALDARRIAGA-GIL'S OBSESSION with this town began in 1980. As a boy, he spent summers at a family home in Huasano, 10 kilometers away. When he attended church here, he couldn't help noticing the lanky men and women with large, flat ears who spoke very little or not at all. "Everyone who knows Ricaurte had curiosity," Saldarriaga-Gil says. "Why is it happening here?"

Growing up, he heard many stories. According to one, nearby magnesium mines had poisoned Ricaurte's groundwater, damaging the minds of people who drank it. Protestant missionaries to the region warned residents that God had sent "the foolishness" to punish them for worshiping "*El Divino*," an image of Jesus in Ricaurte's white A-frame church that draws Catholic pilgrims. "The other hypothesis was sorcery of some sort," Saldarriaga-Gil says. In that version, women in the town prepared a love potion that sometimes went wrong, producing intellectual disability instead of undying devotion. Saldarriaga-Gil's father warned him never to drink anything offered by a woman from Ricaurte.

Saldarriaga-Gil eventually set out to discover the truth as a medical student in the late 1990s. His adviser suggested the people here might have Down syndrome. But when Saldarriaga-Gil paged through a 1000-page medical textbook, he saw photographs of people who looked eerily similar to a boy he knew in Ricaurte—Patricia Triviño's nephew Ronald. The people in the textbook had fragile X syndrome.

To confirm that the resemblance was more than coincidence, in 1997 Saldarriaga-Gil took blood samples from 28 people in town who he suspected were affected, Ronald included. He analyzed each person's karyotype—the number and appearance of their chromosomes—by inspecting their blood cells under a microscope.

In most people, *FMR1* contains anywhere from six to 54 repeats of a specific set of three DNA "letters," or bases: CGG. In people with fragile X syndrome, however, the gene has

In the Quinteros family in Ricaurte, Colombia, Soledad (center) and Yeison (far right) have fragile X syndrome. Rosario (second from right), her daughter Sara (second from left), and Sara's son Juan Pablo (far left) are unaffected, but Sara had her fallopian tubes cut after learning she is a carrier of the full mutation causing the syndrome.

more than 200 repeats. The extra DNA disrupts the X chromosome; under the microscope, tiny islands appear to break away from the chromosome, making it look fragile. Of the 28 people whose karyotypes Saldarriaga-Gil analyzed, 19 showed those telltale islands.

Premutation carriers, however, have between 55 and 200 CGG repeats—too few to be obvious under a microscope. In 2012, Saldarriaga-Gil decided to try to identify those carriers by building a pedigree chart to trace the condition's inheritance through Ricaurte's families. Premutation carriers often have affected children or grandchildren because in fragile X—as in other "triplet repeat" conditions such as Huntington disease—the number of repeats typically increases with successive generations. Working backward from affected individuals, Saldarriaga-Gil tried to guess at who had passed the mutation on. That approach took him only so far, however, because he had no definitive test for premutations.

The next year, his karyotype research caught the attention of experts in fragile X, including Randi Hagerman, medical director of the Medical Investigation of Neurodevelopmental Disorders Institute at the University of California, Davis. She and her colleagues offered to help spot the premutation carriers by using a polymerase chain reaction (PCR) test—which Saldarriaga-Gil wasn't equipped to do in his own lab. PCR would make it possible to amplify and sequence the residents' DNA.

Hagerman recalls being struck by Ricaurte's promise for studying fragile X: "When I first visited this town, I was surrounded by individuals with fragile X syndrome, and I said, 'Oh, my God, this is like ground zero for fragile X.'"

THE TWO-LANE ROAD TO RICAURTE from Cali traverses sugarcane fields between the cloud-shrouded Andes that enclose the town. Saldarriaga-Gil estimates he has driven the route dozens of times in the past 5 years. Before 2010, Colombia's drug trade made the trip dangerous. The region is safer now, he says, but the mountains still teem with farmers secretly growing coca, the raw material of cocaine.

Saldarriaga-Gil checks in on residents with fragile X every 2 months or so, offering routine checkups and monitoring them for complications. Over multiple visits between 2015 and 2016, he and his students also collected blood samples from 926 people, about 80% of the population. Genetic analysis of the samples led to his recent finding that about 5% of Ricaurte's residents have either the full mutation or a premutation. He supplemented the genetic work by recording oral histories and digging up centuries-old land, marriage, and birth records with help from a local historian. Ultimately, Saldarriaga-Gil reconstructed much of the town's history of the syndrome.

An unwieldy pedigree chart now dominates one of his office walls, spanning nine generations and 420 names. Two big families—the Triviños and Gordillos—form its trunk. Saldarriaga-Gil slashes lines through the deceased and scrawls notes in looping handwriting where he is still guessing at kinship.

One name is circled, with sunlike rays extending out in every direction: Manuel Triviño, who may be Mercedes's great-grandfather. Saldarriaga-Gil says he suspects Manuel was one of the town's original settlers in the early 1880s and carried the premutation to Ricaurte. Everyone here with fragile X could be his direct descendant (although how the mutation spread to the Gordillos is still unclear). To confirm that "founder effect," Saldarriaga-Gil's team is conducting a haplotype analysis: The scientists are looking for other genetic variants shared by people with the condition, which would imply that they all share a common forebear.

Saldarriaga-Gil and colleagues from Univalle are also sequencing the exomes—the protein-coding portions of the genome—of the people from whom they took blood samples in 2015 and 2016. They hope to learn how genetic variability outside the *FMR1* gene influences how fragile X mutations manifest themselves—and why people with the same *FMR1* mutations can have such different outcomes.

Among women, "mosaicism"—in which a person's cells aren't all genetically identical—explains part of it. Because women have two X chromosomes, each cell turns off one of them at random. If most of a woman's cells turn off the mutated copy, she might show few outward signs of the mutation; if the normal copy is shut down more often, she might be more severely af-

PHOTO: JUAN CRISTÓBAL COBO

fected. Mosaicism emerges differently in men, who have a single X chromosome: Some of their cells may have the full *FMR1* mutation—200-plus CGG repeats—whereas others end up with the shorter premutation or with a complete deletion of *FMR1*.

The array of symptoms resulting from a mutation might also depend on how FMRP interacts with other proteins. FMRP is missing in people with the full mutation, which silences *FMR1*. Because FMRP controls the activity of nearly 1000 other proteins, many of which are crucial to the interactions between neurons, its loss can have far-reaching effects—particularly during brain development. But in people with the premutation, the impact of the reduced protein might be more or less severe depending on other genetic variations.

Saldarriaga-Gil and his colleagues predict that genetic analyses will reveal that people whose fragile X symptoms are similar have overlapping patterns of gene expression and protein interaction. "This type of population is ideal for this study because these people have a similar genetic background," says Univalle geneticist Julián Andrés Ramírez Cheyne, who leads the exome study.

The ultimate goal for fragile X researchers is to develop treatments. Because of its connection to intellectual disability and autism, fragile X has been the focus of an extensive—and so far, unsuccessful—drug development program. Several candidates that showed promise in early clinical trials fizzled out in larger trials. Researchers are seeking new proteins or pathways to target—and some of those may emerge from the work done here. "Most geneticists would say there are genetic modifiers in some of these families," says Eric Klann, director of the Center for Neural Science at New York University in New York City—clues, he says, to possible treatments.

Understanding the molecular underpinnings of fragile X might also explain why general anesthesia and some seizure medications are more toxic to premutation carriers than to typical people. Hagerman says she was struck by the number of premutation carriers here whose symptoms are unusually severe. Patricia Triviño's sister Rosaura, 60, for example, is deaf and mute; her sister Julieta, 58, has seizures and uses a wheelchair. Hagerman says pesticides, sprayed heavily in the nearby fields, might be to blame. "Looking at the environmental contaminants could tell us a lot about vulnerability" in people with the mutations, she says.

NO ONE HERE IS WAITING for radical new treatments. Even if the residents can help researchers develop drugs, they know they are likely to be among the last to receive them.

Mercedes Triviño's son Jair is one of the oldest workers in the papaya fields, but he has no complaints. Even though his work is arduous, living among others with similar symptoms has given him a degree of freedom he'd be hard-pressed to find anywhere else. At the end of his shift on a hot July day, Jair fills the back of a pickup truck with fruit. He swings the heavy crates one by one, his wiry arms flexing until he has stacked nearly 100. He stops to wipe sweat from his brow and secures the truck's door with an iron bolt. The driver starts the engine as Jair hops in the back. Jair says he looks forward to being back the next day "if it's God's wish."

Back in town, Patricia, who is Jair's cousin, says all she wants is a good drug to control the seizures of her sister Julieta. For years, Julieta has taken phenobarbital to control her convulsions, but the drug is risky for premutation carriers like her, who may be particularly vulnerable to its neurotoxic effects. Scans that Saldarriaga-Gil took show that portions of Julieta's brain have shrunk, and in the past few weeks, she has started to complain of headaches. Another of Patricia's sisters, Esperanza, 60, who also was a premutation carrier and relied on phenobarbital for years, died in 2015 after several massive seizures. Julieta could take safer medications, such as valproate, but getting a prescription hasn't been easy.

Wilmar Saldarriaga-Gil (right) comforts Julieta Triviño (left), who has fragile X syndrome.

PHOTO: JUAN CRISTÓBAL COBO

In July, a new doctor who serves Ricaurte and three other towns arrives to make her rounds. Rubbing Julieta's temples below her cropped black hair, the doctor explains that she doesn't have the correct paperwork to prescribe valporate. She suggests Patricia take Julieta to a doctor in Bolívar, about 7 kilometers away. Patricia doesn't own a car, and walking there would take her about an hour and a half.

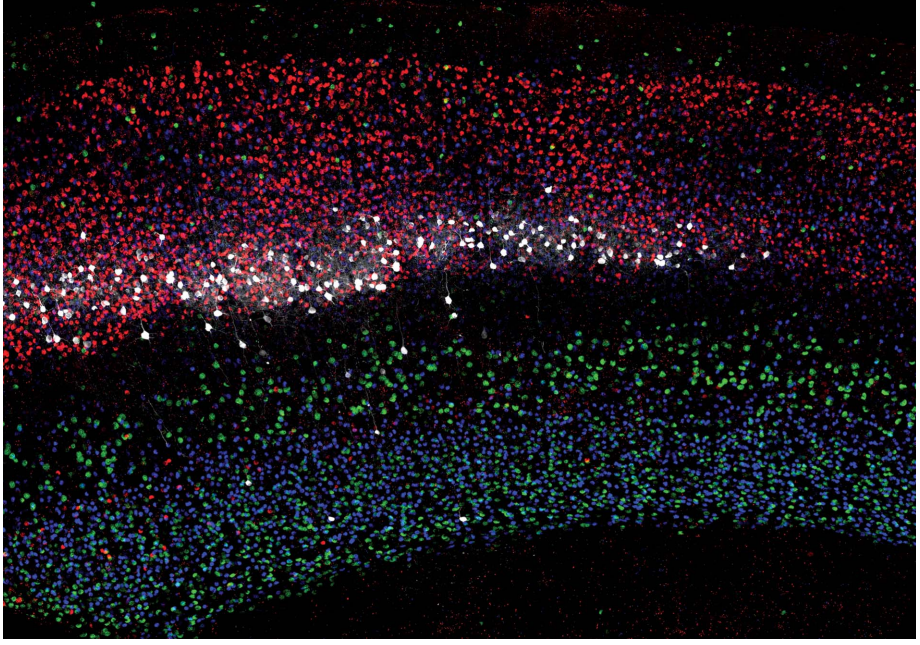
For months, Patricia has pleaded with a local health minister, Viviana Alvarez, to secure basic sanitary supplies and protein supplements for Julieta—but to no avail. Although the family is entitled to free care through the government, Alvarez says her hands are tied: "Health insurance takes its time; the problem is on the national level." The Hospital Santa Ana in Bolívar can't help much, either. With just eight doctors, it fielded 15,000 appointments and 5000 emergency room visits in 2017. The hospital director has hired a physical therapist to visit about 15 people here with fragile X every weekday but he says Ricaurte will probably never have its own physician.

"I always want them to do more," Saldarriaga-Gil says, although he understands the financial constraints. He tries to fill in the gaps during his visits and has enlisted a Colombian nonprofit to donate clothes and mattresses to many of the families.

Given the harsh realities of life here with fragile X, some residents have made difficult decisions about the future of their families. Rosario Quintero's daughter, Sara, has the full mutation but shows no signs of the syndrome. Before Sara learned that she carried the mutation, she had a son, who also seems unaffected. But afterward, she had her fallopian tubes cut so that she cannot have any more children. Another carrier, who chose to remain anonymous, also decided not to have children.

Over the past decade here, only three children with fragile X have been born, and many with the condition are older than 50. Trapped in this valley by economic hardship and unyielding geography, the population with fragile X could slowly die out, Saldarriaga-Gil says. He is racing to understand the syndrome's secrets before that happens. ■

This story was produced in collaboration with Spectrum (spectrumnews.org), where Hannah Furfaro is a staff reporter.



Cross section of primary somatosensory cortex in a 3-week-old wild-type mouse. The cortical layers are indicated by marker expression: layers II to IV, CUX1 (red); layer IV, RORβ (white); layer V/VI, CTIP2 (green); and layer VI, TBR1 (blue). The image was assembled from individual confocal sections acquired with a Zeiss LSM800 confocal microscope.

IMAGE: R. BEATTIE AND S. HIPPENMEYER/IST AUSTRIA

REVIEW

Neuronal specification in space and time

Isabel Holguera[1] and Claude Desplan[1,2]*

To understand how neurons assemble to form functional circuits, it is necessary to obtain a detailed knowledge of their diversity and to define the developmental specification programs that give rise to this diversity. Invertebrates and vertebrates appear to share common developmental principles of neuronal specification in which cascades of transcription factors temporally pattern progenitors, while spatial cues modify the outcomes of this temporal patterning. Here, we highlight these conserved mechanisms and describe how they are used in distinct neural structures. We present the questions that remain for a better understanding of neuronal specification. Single-cell RNA profiling approaches will potentially shed light on these questions, allowing not only the characterization of neuronal diversity in adult brains, but also the investigation of the developmental trajectories leading to the generation and maintenance of this diversity.

Nervous system development requires many cell types generated in the proper order, number, and location. This cellular diversity is generated from a small pool of progenitors initially defined by their spatial location and by sequential expression of temporal factors. Regulatory principles and some of the molecular players are conserved from insects to mammals. Here, we highlight these similarities and discuss promising research avenues that use single-cell transcriptomics during development to understand neuronal specification.

Defining cell types one cell at a time

Although neurons have been classified since the beginning of modern neuroscience on the basis of morphology, function, electrophysiological properties, or molecular markers (*1*), these criteria underestimate the diversity of neurons. Even in structures as well understood as the mammalian retina, morphology identifies 13 subtypes of bipolar cells but misses two additional types identifiable by transcriptional criteria (*2*). Although single-cell RNA sequencing (scRNA-seq) allows unbiased identification of mature neurons, further understanding of nervous system function and disease requires knowledge not only of factors that define neuronal identity, but also of molecules driving neuronal specification, migration, and connectivity during development.

Modes of progenitor division and generation of neuronal units

Neural progenitors use two axes of information, spatial and temporal, to generate the cellular diversity of the central nervous system (CNS) (*3*). The nervous system originates from simple neuroepithelial sheets where cells first proliferate by symmetric divisions and later become fate-restricted progenitors. We focus on two examples of these progenitors, the apical radial glia (aRG) in the mammalian cortex and the neuroblasts (NBs) in *Drosophila*.

aRG have long cellular processes that span from apical to basal sides of the developing cortex. aRG can give rise to neurons through either direct neurogenesis, in which the aRG cell divides asymmetrically to self-renew and generate a neuron, or indirect neurogenesis. In the latter, aRG give rise to intermediate progenitor populations, such as intermediate basal progenitors (bIPs), which usually divide once to produce neurons, and basal or outer radial glia (oRG), which have increased proliferative potential and can self-renew, amplifying the neuronal output of these lineages (*4*) (Fig. 1A). These two types of intermediate progenitors appeared during mammalian evolution and have different proliferative capacities depending on the species. bIPs establish a new germinal niche, the subventricular zone (SVZ), which contributes neurons to all cortical layers. In placental mammals, the SVZ appears to promote cortical expansion and the generation of new neuronal types, such as callosal projection neurons, which connect different areas of the cortex and correlate with increased associative capacities (*5*). In mice, oRG are present but are far less abundant than in primates, where they compartmentalize the SVZ by generating the outer SVZ. Thus, an increase in oRG may have allowed an expanded cortex (*4*). Single-cell transcriptomic studies of human aRG and oRG show that oRG preferentially express genes involved in growth factor signaling, cell migration, and self-renewal, suggesting an increased stemness of oRG as a mechanism underlying neocortex expansion in primates (*6*). Clonal analysis shows the extensive proliferative capacity of human oRG, which produce clones of neuronal and glial cells at mid-neurogenesis that are one to two orders of magnitude larger than mouse aRG clones (*6*, *7*). Gene expression differences underlying distinct cortical progenitor behavior from rodents to humans have been investigated and recently reviewed (*8*).

In *Drosophila*, different types of NB division occur: Type 0 NBs asymmetrically divide multiple times, and like aRG in direct neurogenesis, each time they self-renew they generate a single neuron. Type I NBs are the most abundant and also undergo multiple asymmetric divisions to self-renew, but they produce a ganglion mother cell (GMC) that divides once more to generate either two neurons or glial cells. Type II NBs have expanded lineages: They divide to self-renew and produce multiple intermediate neural progenitors (INPs) that themselves divide asymmetrically to produce four to six GMCs (*3*). Hence, type II NBs produce lineages with larger numbers of neurons

[1]Department of Biology, New York University, New York, NY 10003, USA. [2]Center for Genomics and Systems Biology, New York University Abu Dhabi, Abu Dhabi, United Arab Emirates.
*Corresponding author. Email: cd38@nyu.edu

in ways that resemble the indirect neurogenesis through oRG in primates (Fig. 1B).

Temporal patterning of neural progenitors

Most of our understanding of cortical neurogenesis comes from studies of excitatory projection neurons, which constitute ~70 to 80% of the neuronal population in the mammalian cortex. GABAergic inhibitory interneurons (representing ~20 to 30%) migrate into the cortex from their birthplaces, mainly in ganglionic eminences (*9*). The mature mammalian cortex comprises six neuronal layers that are generated in an inside-out manner: Early-born neurons occupy deep layers (DLs) (layers VI and V), closer to the neural progenitors, whereas later-born neurons progressively occupy upper layers (ULs) (layers IV, III, and II) after migrating along the processes of aRG through earlier-born neurons (Fig. 2A). Projection cortical neurons are thus organized in columns that are parallel to aRG processes and represent the functional units of each cortical area (*9*). A similar columnar structure is shared by the medulla neuropil in the fly optic lobe, in which each of the 800 columns generated by approximately the same number of NBs represents a functional processing unit (each column is composed of about 80 neuron types), allowing the retinotopic integration of visual information relayed from the 800 unit eyes (*10*). *Drosophila* medulla neurons are displaced away from the parent NBs as newly born neurons are generated, creating an early-to-late birth order–dependent axis in a manner opposite to that of the mammalian cortex (Fig. 2B).

The generation of neural diversity in *Drosophila* depends on temporal and spatial patterning of neural progenitors. The first described example of temporal patterning was in embryonic NBs of the *Drosophila* ventral nerve cord. As the NBs age, they sequentially express a series of temporal transcription factors (tTFs) [Hunchback (Hb) → Krüppel → Pdm → Castor → Grainy head], generating specific neuronal progeny in each of the different temporal identity windows defined by these factors (*3*). Expression of these tTFs in different lineages leads to different neuronal outcomes. For example, Hb expression specifies the U1-U2 motor neuron identity in the NB7-1 lineage but specifies Rp1-Rp4 motor neuron identity in the NB3-1 lineage (*3*), indicating that spatial positioning additionally patterns these NBs. Other fly lineages have since been shown to express a series of tTFs whose identity varies from structure to structure (*3*). The medulla NBs in the fly optic lobe sequentially express a series of six tTFs that are different from ventral nerve cord tTFs (*11*, *12*) (Fig. 2B). Expression of consecutive tTFs overlaps, generating additional temporal identity windows and expanding the diversity of neuronal progeny that can be produced. In addition, subtemporal genes act downstream of tTFs and further subdivide temporal windows (*3*). Each NB asymmetric division generates a GMC, which after its final division produces two different neurons, one of which activates the Notch pathway ($Notch^{ON}$) whereas the other does not ($Notch^{OFF}$), thus doubling the diversity of neural progeny (*11*). A different series of tTFs has also been found for the INP lineages that are orthogonal to the temporal series in type II central brain NBs (*13*) (Fig. 1B), further expanding the array of neuronal types generated from a single type II NB.

Two prevailing models posit how mammalian cortical layers are generated. In the first (Fig. 2A), a single multipotent progenitor has the competence to generate neurons for all the cortical layers. This progenitor undergoes asymmetric cell divisions and progressively becomes fate restricted, first producing DLs and then switching to produce more superficial layers. Supporting this model, transplantation experiments and retroviral lineage tracing showed that early-stage progenitors either transplanted into an old cortex or labeled early in corticogenesis can generate neurons for all the cortical layers whereas old-stage progenitors either transplanted into a young cortex or labeled late in corticogenesis can generate only neurons residing in superficial layers (*14*). This model is consistent with the notion that progenitors express a series of intrinsic temporal factors

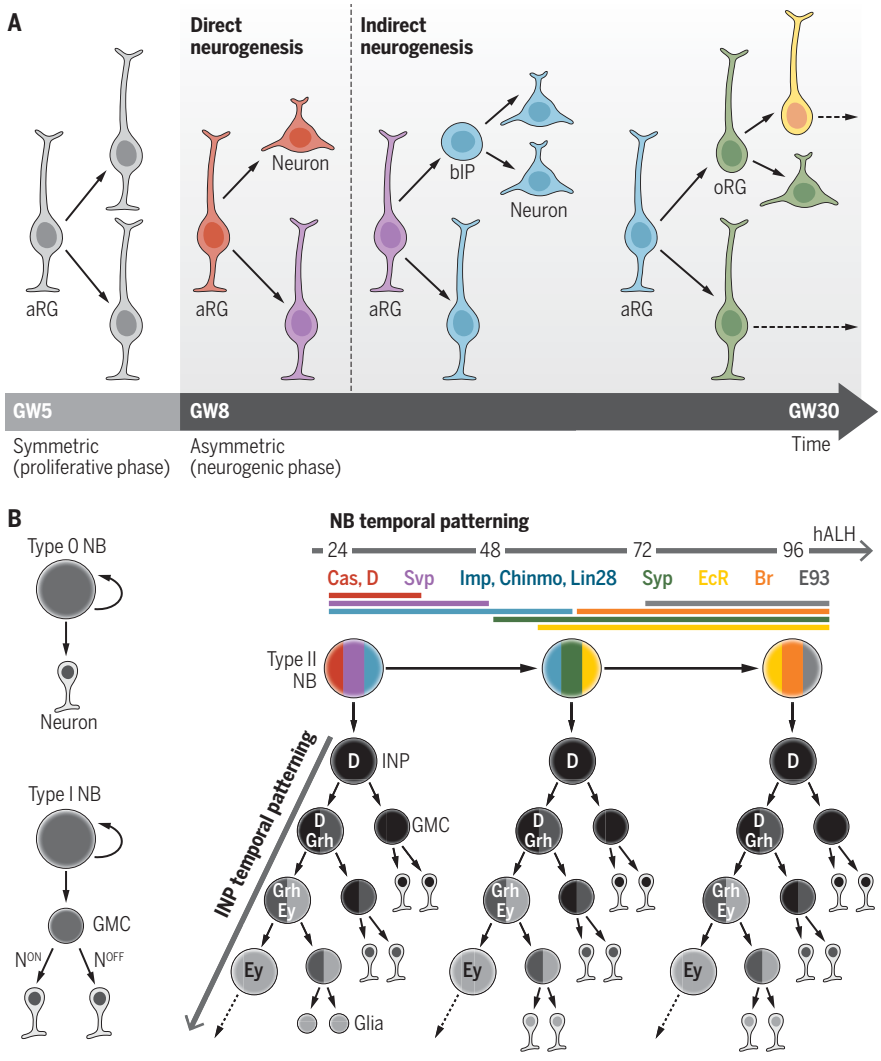


Fig. 1. Modes of division of neural progenitors. (**A**) Neurogenesis in the primate cortex. aRG undergo symmetric cell divisions to expand their pool (proliferative phase) and then transit to a neurogenic phase, where they generate neurons directly or indirectly through intermediate progenitors such as bIPs and oRG. Changes in TF expression in progenitors over time are shown as changes in cell color. GW, gestational week. (**B**) Different modes of NB division in *Drosophila*. Type 0 NBs self-renew and generate a single neuron at each division. Type I NBs self-renew and produce GMCs that divide once to produce one $Notch^{ON}$ (N^{ON}) and one $Notch^{OFF}$ (N^{OFF}) neuron. Type II NBs have an increased neuronal output because they generate INPs, which themselves asymmetrically divide multiple times to produce GMCs. Both NBs and INPs sequentially express series of tTFs as they age (examples of these and additional temporal factors in NBs are shown). Cas, Castor; D, Dichaete; Svp, Seven-up; Imp, IGF-II mRNA binding protein; Chinmo, Chronologically inappropriate morphogenesis; Syp, Syncrip; EcR, ecdysone receptor; Br, Broad; E93, Eip93; Grh, Grainy head; Ey, Eyeless; hALH, hours after larval hatching. [Adapted from (*3*)]

ILLUSTRATION: KELLIE HOLOSKI/*SCIENCE* BASED ON (*3*)

that induce with time the different neurons of successive layers. The second model of initially fate-restricted progenitors posits the existence of independent progenitors that each generate one of the different neuronal populations (*15*, *16*).

The behavior of the overall progenitor population in the studies described above does not explain how a single progenitor can generate all neuronal subtypes. MADM (mosaic analysis with double markers) lineage-tracing experiments labeling single progenitors in mice support the multipotent progenitor model for the generation of cortical projection neurons: aRG generate defined clones, with each aRG cell averaging eight to nine neurons during the neurogenic phase that span all layers of the cortex when aRG are induced early [embryonic day 10 (E10) to E13], and they shift to the production of only superficial-layer neurons when aRG are induced at a later time point (E15) (*7*) (Fig. 2A). Although these results support the multipotent progenitor model and imply the existence of temporal patterning in the cortex, this type of progenitor could coexist with other fate-restricted progenitors that remain to be identified.

Intrinsic temporal factors identified in vertebrate neurogenesis are often homologous to those found in *Drosophila*. Ikaros (the ortholog of the early tTF Hb in the fly ventral nerve cord) is expressed early in cortical and retinal progenitors and specifies early-born-neuron fates in both tissues, whereas Casz1 (the vertebrate ortholog of the late fly tTF Castor) is expressed in older retinal progenitors, specifying late-born fates (*17*, *18*) (Fig. 2, A and C). Induction of the FoxG1 transcription factor (TF) (the ortholog of *Drosophila* Slp1) induces DL neurogenesis through derepression of the TF Fezf2 (*19*). Brn1 and Brn2 are involved in the transition from early to mid-neurogenesis in the mouse cortex, as UL neurons fail to be generated in *Brn1-Brn2* double mutants (*20*), whereas their misexpression produces later-born neurons. Seven-up functions as a switching factor from early- to late-neuron generation in *Drosophila* (*3*), and knockdown of its CoupTFI/II ortholog in the developing mouse forebrain prolongs the generation of early-born neurons at the expense of late-born neurons (*21*). Therefore, tTFs generate neuronal diversity in both insects and vertebrates. However, in most cases, whether these factors act in vertebrate progenitors or in neurons and whether they are organized in temporal series as has been shown for fly NBs remain unknown.

The timing of fate switches in neural progenitors is essential for the generation of the proper number and identity of neuronal subtypes. In several cases, feed-forward mechanisms whereby an early tTF activates the expression of the next tTF in the series have been identified, whereas

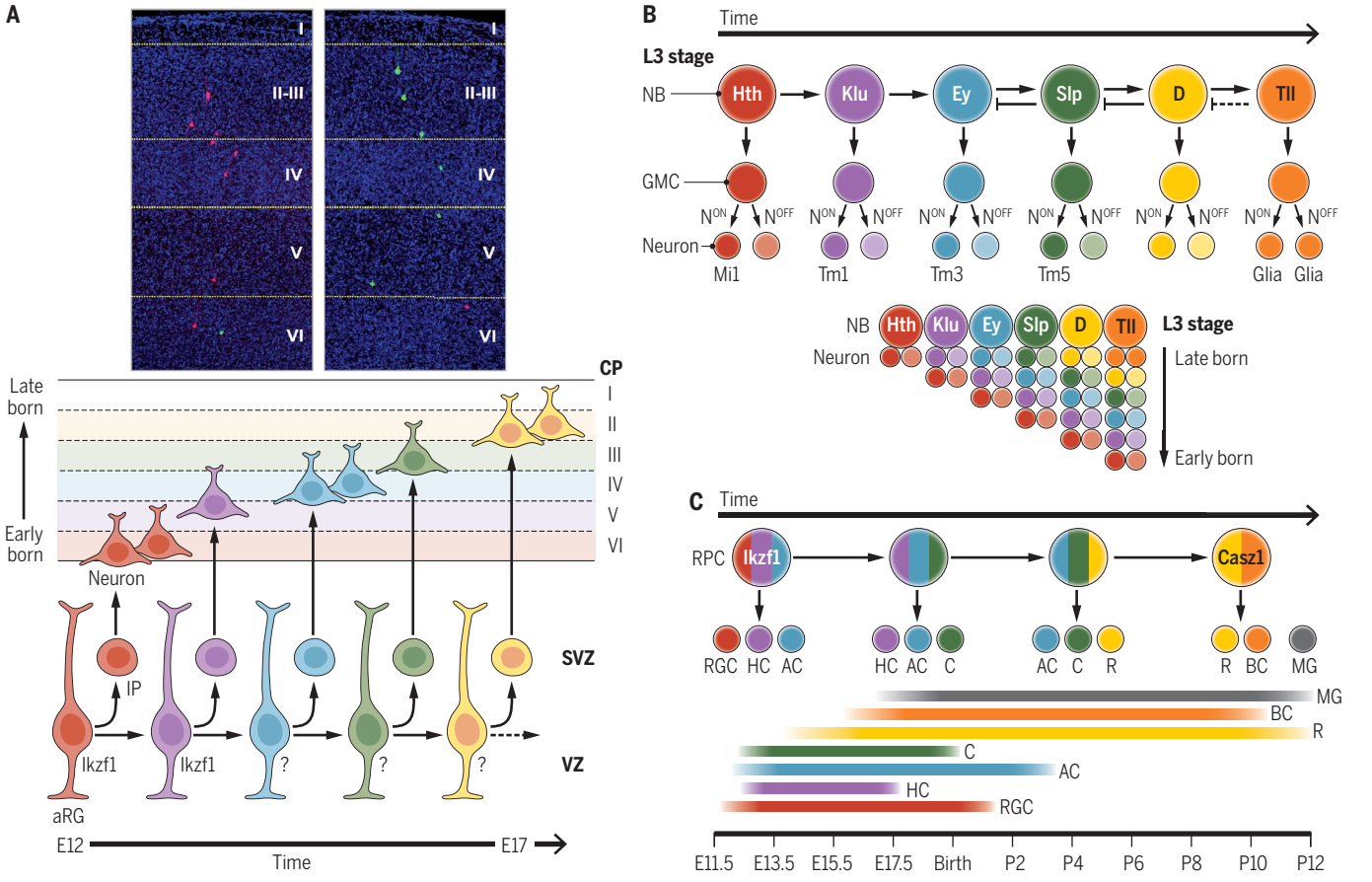


Fig. 2. Temporal patterning of progenitors. (**A**) Multipotent progenitor model of neurogenesis in the mammalian cortex. (Top) Asymmetric MADM clones showing the progeny of a single aRG cell that span each layer (II to VI) of the mouse cortex. [Image by R. Beattie and S. Hippenmeyer] (Bottom) A common progenitor generates neurons for the different cortical layers (VI through II) sequentially in an inside-out fashion. Ikaros (Ikzf1) is an example of a TF specifying deep-layer neuronal identity. Changes in TF expression in progenitors over time are shown with color changes. IP, intermediate progenitor; VZ, ventricular zone; CP, cortical plate. (**B**) Temporal patterning in *Drosophila* optic lobe NBs. (Top) Sequential expression of tTFs in *Drosophila* type I optic lobe NBs, specifying distinct neuronal (e.g., Mi1, Tm1, Tm3, and Tm5) and glial identities in each temporal window. Cross-regulatory interactions between tTFs are shown. Hth, Homothorax; Klu, Klumpfuss; Slp, Sloppy paired 1 and 2; Tll, Tailless. (Bottom) Newly born neurons displace older siblings away from the parent NB, generating a birth order–dependent layered neuronal arrangement in the medulla cortex in third-instar larvae (L3). (**C**) Vertebrate retina. RPCs sequentially generate the seven retinal cell types in overlapping waves. The TF Ikzf1 specifies early-born fates, whereas Casz1 specifies late-born fates. RGC, retinal ganglion cell; HC, horizontal cell; AC, amacrine cell; C, cone; R, rod; BC, bipolar cell; MG, Müller glia; P, postnatal day.

ILLUSTRATION: KELLIE HOLOSKI/*SCIENCE* BASED ON I. HOLGUERA AND C. DESPLAN

negative feedbacks allow later tTFs to repress earlier tTFs (*11*) (Fig. 2B). Extrinsic factors can also be involved in transitions: The hormone ecdysone regulates the transition from early to late tTFs in *Drosophila* central brain type II NBs (*22*) (Fig. 1B). The mechanisms by which vertebrate tTFs regulate one another and how these tTFs control gene expression to define cell fates are largely unknown (*18*).

Temporal patterning generates neuronal diversity and also allows the production of different neuronal types in an invariant order, which likely facilitates the assembly of neural circuits. In *Drosophila*, deterministic lineages generate hardwired circuits in an ordered and stereotypic manner (*23*–*25*). In the mammalian brain, lineage relationships between cortical neurons instruct both laminar organization and functional relationships: Excitatory neurons preferentially form chemical synapses between lineage-related neurons rather than nonsibling neurons (*26*). The formation of electrical synapses also occurs preferentially between lineage-related inhibitory interneurons (*27*). These examples highlight the instructive role of developmental time in the sequential generation and assembly of cortical circuitry, which is further refined in vertebrates by neuronal activity and stochastic processes. For instance, retinal progenitor cells (RPCs) generate the seven distinct retinal cell classes in overlapping temporal sequences following the multipotent progenitor model (*28*) (Fig. 2C). However, clones derived from single RPCs are variable in size and composition, suggesting that stochastic factors also control lineage progression. Time-lapse imaging of single labeled RPCs suggests that RPCs are equipotent but have certain probabilities of dividing or differentiating such that each cell type is biased but not deterministically specified in a restricted temporal window (*29*). In the end, integration of stochastic and deterministic mechanisms at the population level allows the mouse retina to reach a defined size and cellular composition (*28*).

Spatial patterning of neural progenitors

Intrinsic temporal patterning of neural progenitors, although a conserved strategy, cannot explain the diversity of neurons generated. Neural progenitors are additionally patterned by other mechanisms on the basis of their spatial location. Temporal patterning and spatial patterning intersect to generate neural diversity in the *Drosophila* optic lobe. The neuroepithelium that generates optic lobe NBs is regionalized into different compartments by the restricted expression of spatial homeodomain TFs such as Vsx, Optix, and Rx, as well as by the signaling molecules Decapentaplegic (Dpp), Wingless (Wg), and Hedgehog (Hh), generating eight NB domains along the dorsoventral axis (Fig. 3A) (*30*). The NBs generated in

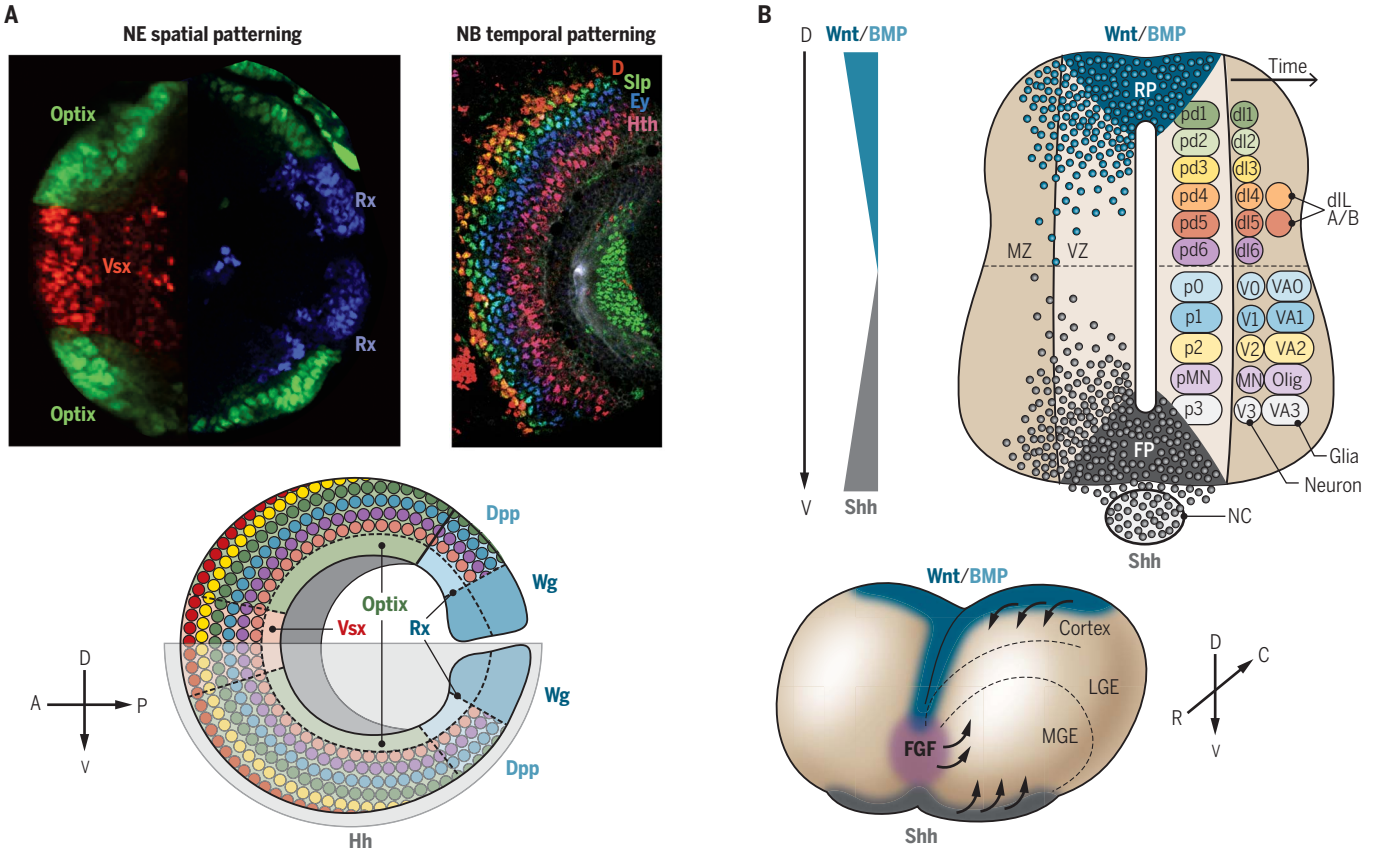


Fig. 3. Spatial patterning of progenitors. (**A**) Integration between spatial and temporal patterning in the *Drosophila* optic lobe. (Top left) Regionalization of the *Drosophila* optic lobe neuroepithelium (NE) by the expression of the TFs Optix (green), Vsx (red), and Rx (blue). (Top right) Sequential expression of the tTFs Hth (pink)–Ey (blue)–Slp (green)–D (red) in optic lobe NBs of different ages. (Bottom) Schematic of spatial and temporal factors acting in optic lobe NBs. The neuroepithelium is additionally patterned by the signaling molecules Dpp and Wg and by Hh in the ventral part, creating eight compartments along the dorsoventral axis. D, dorsal; V, ventral; A, anterior; P, posterior. (**B**) Morphogens spatially pattern the mammalian spinal cord and telencephalon. (Top) Dorsoventral patterning of the neural tube. Wnt, BMP, and Shh regulate the expression of TFs in progenitors, which produce different types of neurons. Early-born neurons are different from those produced at later stages. The progenitor domains and the neurons generated in each one of them are indicated. NC, notochord; FP, floor plate; RP, roof plate; MZ, mantle zone. (Bottom) Schematic of the mouse embryonic telencephalon that is similarly patterned by the morphogens Wnt and BMPs from the dorsal hem and ventrally patterned by the secretion of Shh. FGFs are secreted from a rostral signaling center (violet). GABAergic interneurons are generated from the lateral and medial ganglionic eminences (LGE and MGE) in the ventral telencephalon and migrate tangentially to reach the cortex. R, rostral; C, caudal.

ILLUSTRATION: KELLIE HOLOSKI/*SCIENCE* BASED ON I. HOLGUERA AND C. DESPLAN

each domain (except the Wg domain) transit through the same series of tTFs as they age, but the combinatorial input of spatial and temporal factors allows the generation of an increased diversity of neuronal types whereby spatial information modifies the type of neurons produced by tTFs. Unicolumnar neurons (in a 1:1 ratio with medulla columns), generated from all spatial compartments, appear to ignore spatial information, whereas multicolumnar neurons (which are fewer in number because they arborize in more than one column) are made from fewer NBs in restricted spatial compartments (*30*).

In the vertebrate neural tube, progenitors are patterned by spatial cues in the form of Hox gene expression along the rostrocaudal axis in response to morphogenetic gradients, including fibroblast growth factors (FGFs) and retinoic acid, whereas Sonic Hedgehog (Shh), bone morphogenetic proteins (BMPs), and Wnt pattern progenitors in the dorsoventral axis (Fig. 3B). Opposing signaling gradients confer positional information to neural progenitors and direct the expression of distinct TFs, which respond differentially to signaling inputs. Moreover, cross-repression between TFs expressed in neighboring progenitor domains generates sharp gene expression boundaries (*31*). Each progenitor compartment then generates different neuronal types. Temporal patterning also appears to be involved in spinal cord neurogenesis, as early-born motor and interneurons differ from those produced later, although clonal relationships remain to be determined (*32*, *33*) (Fig. 3B).

Additional factors controlling neuronal specification

Both spatial factors and tTFs act in progenitors to specify neuronal progenies. In postmitotic neurons, additional TFs termed terminal selectors act individually or in combination to activate the expression of terminal differentiation genes that define neuronal attributes such as neurotransmitter expression, neurite morphology, and electrophysiological properties (*34*). Spatial and temporal factors acting combinatorially regulate the expression of terminal selectors (*35*, *36*). In addition, extrinsic mechanisms are involved in the generation of cortical diversity; these mechanisms include transforming growth factor–β signaling, as well as thalamic input. Feedback cues from previously generated DL neurons influence the production of UL cortical neurons or promote the switch from neurogenesis to gliogenesis (*14*). These extrinsic mechanisms provide an additional level of regulation and plasticity to corticogenesis that would not be possible through reliance on only intrinsic temporal and spatial factors.

Shifting to single-cell approaches during development

Although much of what is known about spatial and temporal patterning of progenitors derives from molecular genetic techniques, much will be learned using scRNA-seq technologies. scRNA-seq has amplified our perception of neural diversity, redefining the concept of cell type. It has allowed validation of previously proposed cell types and uncovered new ones (*37*). However, the relationship of transcriptomic cell types to bona fide cell types has often not been established. Accurate cell type classification will require correlations between molecular and morphological, physiological, and connectomic characteristics. The transcriptional signature of a neuron should be predictive of its function. Indeed, a transcriptomic study of neuronal diversity in the fly optic lobe identified TFs explaining much of the neuronal diversity in this brain region (*38*), where the same neurotransmitter identity can be assigned by distinct TF combinations in different neuronal types, suggesting phenotypic convergence and supporting principles described in *Caenorhabditis elegans* (*34*).

Unresolved, however, is how the adult transcriptional profile arises from the establishment of neuronal identity during development, when most gene expression changes occur, often coincident with circuit assembly (*39*, *40*). Because scRNA-seq yields snapshots of expression data whereas neuronal specification through development is a continuum of different states, computational methods such as trajectory inference algorithms are necessary to temporally order these states and to allow the identification of branching points during developmental trajectories (*41*). CRISPR mutation–based high-throughput lineage-tracing approaches (*41*), although still technically challenging, are beginning to allow the reconstruction of lineage trees. Methods for reconstructing the positional information of a given cell in the tissue being subjected to transcriptomic analysis (spatial transcriptomics) are also being implemented (*41*). Parallel improvements in computational methods applied to scRNA-seq studies are necessary to interpret the increasing complexity of data generated by these approaches. Additionally, functional validation is essential to test hypotheses emerging from scRNA-seq studies. scRNA-seq has allowed researchers to begin a comprehensive characterization of cell types in the CNS, to discover new cell types, to identify new markers that enable their manipulation, to trace developmental cellular decisions, and to uncover disease-related genes [reviewed in (*37*)]. Integration of transcriptomic, epigenomic, and proteomic analyses at the single-cell level is within reach and will reveal the cellular diversity and development of systems as complex as the human brain.

Outlook

In both invertebrates and vertebrates, conserved mechanisms of temporal and spatial patterning allow neural progenitors to generate a broad diversity of cell types in the CNS. As the brain evolved with greater complexity, new progenitor types were generated. Extrinsic influences and developmental plasticity acquired increased roles in cellular patterning, allowing the integration of increased numbers and diversity of cells.

Single-cell analyses have allowed the identification of cell type–specific markers and drivers, which will be invaluable for testing hypotheses regarding the development of neuronal diversity. Insight into temporal and spatial factors that specify neuronal diversity will be crucial for producing the desired neuronal types for cell replacement therapy. Finally, comparison of neural diversity and development across species will contribute to our understanding of the evolution of the brain.

REFERENCES AND NOTES

1. S. Ramón y Cajal, *Histologie du Système Nerveux de L'homme et des Vertébrés* (Maloine, 1909).
2. K. Shekhar *et al.*, *Cell* **166**, 1308–1323.e30 (2016).
3. C. Q. Doe, *Annu. Rev. Cell Dev. Biol.* **33**, 219–240 (2017).
4. E. Taverna, M. Götz, W. B. Huttner, *Annu. Rev. Cell Dev. Biol.* **30**, 465–502 (2014).
5. R. M. Fame, J. L. MacDonald, J. D. Macklis, *Trends Neurosci.* **34**, 41–50 (2011).
6. A. A. Pollen *et al.*, *Cell* **163**, 55–67 (2015).
7. P. Gao *et al.*, *Cell* **159**, 775–788 (2014).
8. M. Florio, V. Borrell, W. B. Huttner, *Curr. Opin. Neurobiol.* **42**, 33–44 (2017).
9. T. Kumamoto, C. Hanashima, *Neurosci. Res.* **86**, 37–49 (2014).
10. N. Nériec, C. Desplan, *Curr. Top. Dev. Biol.* **116**, 247–271 (2016).
11. X. Li *et al.*, *Nature* **498**, 456–462 (2013).
12. T. Suzuki, M. Kaido, R. Takayama, M. Sato, *Dev. Biol.* **380**, 12–24 (2013).
13. O. A. Bayraktar, C. Q. Doe, *Nature* **498**, 449–455 (2013).
14. L. Adnani, S. Han, S. Li, P. Mattar, C. Schuurmans, *Int. Rev. Cell Mol. Biol.* **336**, 223–320 (2018).
15. S. J. Franco *et al.*, *Science* **337**, 746–749 (2012).
16. C. Guo *et al.*, *Neuron* **80**, 1167–1174 (2013).
17. J. M. Alsiö, B. Tarchini, M. Cayouette, F. J. Livesey, *Proc. Natl. Acad. Sci. U.S.A.* **110**, E716–E725 (2013).
18. P. Mattar, J. Ericson, S. Blackshaw, M. Cayouette, *Neuron* **85**, 497–504 (2015).
19. K. Toma, T. Kumamoto, C. Hanashima, *J. Neurosci.* **34**, 13259–13276 (2014).
20. M. H. Domínguez, A. E. Ayoub, P. Rakic, *Cereb. Cortex* **23**, 2632–2643 (2013).
21. H. Naka, S. Nakamura, T. Shimazaki, H. Okano, *Nat. Neurosci.* **11**, 1014–1023 (2008).
22. M. H. Syed, B. Mark, C. Q. Doe, *eLife* **6**, e26287 (2017).
23. D. J. Brierley, K. Rathore, K. VijayRaghavan, D. W. Williams, *J. Comp. Neurol.* **520**, 1629–1649 (2012).
24. F. Pinto-Teixeira *et al.*, *Cell* **173**, 485–498.e11 (2018).
25. H. Li, S. A. Shuster, J. Li, L. Luo, *Neural Dev.* **13**, 5 (2018).
26. Y. C. Yu *et al.*, *Nature* **486**, 113–117 (2012).
27. X. J. Zhang *et al.*, *Nat. Commun.* **8**, 16091 (2017).
28. C. Cepko, *Nat. Rev. Neurosci.* **15**, 615–627 (2014).
29. J. He *et al.*, *Neuron* **75**, 786–798 (2012).
30. T. Erclik *et al.*, *Nature* **541**, 365–370 (2017).
31. J. Briscoe, S. Small, *Development* **142**, 3996–4009 (2015).
32. M. Sugimori *et al.*, *Development* **134**, 1617–1629 (2007).
33. L. R. Hernández-Miranda, T. Müller, C. Birchmeier, *Dev. Biol.* **432**, 34–42 (2017).
34. O. Hobert, *Curr. Top. Dev. Biol.* **116**, 455–475 (2016).
35. M. Baumgardt, D. Karlsson, J. Terriente, F. J. Díaz-Benjumea, S. Thor, *Cell* **139**, 969–982 (2009).
36. H. Gabilondo *et al.*, *PLOS Biol.* **14**, e1002450 (2016).
37. B. Tasic, *Curr. Opin. Neurobiol.* **50**, 242–249 (2018).
38. N. Konstantinides *et al.*, *Cell* **174**, 622–635.e13 (2018).
39. H. Li *et al.*, *Cell* **171**, 1206–1220.e22 (2017).
40. B. T. Kalish *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **115**, E1051–E1060 (2018).
41. J. A. Griffiths, A. Scialdone, J. C. Marioni, *Mol. Syst. Biol.* **14**, e8046 (2018).

ACKNOWLEDGMENTS

We thank the members of the Desplan lab for discussion and editing of the manuscript; we are also very grateful to S. Hippenmeyer and R. Beattie for providing images of cortical MADM clones. **Funding:** This work was supported by NIH grant EY13012 to C.D. I.H. was supported by a Human Frontier Science Program postdoctoral fellowship (LT000757/2017-L). **Competing interests:** The authors declare no conflict of interest.

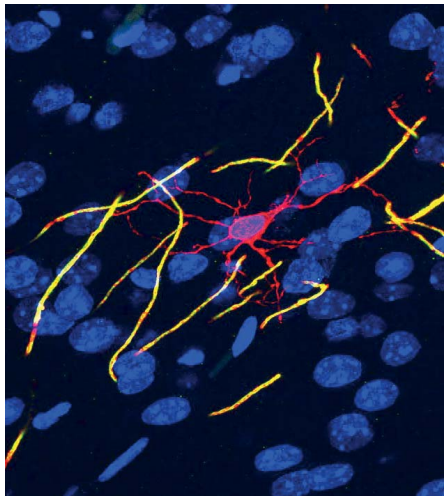
10.1126/science.aas9435

REVIEW

Glia as architects of central nervous system formation and function

Nicola J. Allen[1]* and David A. Lyons[2]*

Glia constitute roughly half of the cells of the central nervous system (CNS) but were long-considered to be static bystanders to its formation and function. Here we provide an overview of how the diverse and dynamic functions of glial cells orchestrate essentially all aspects of nervous system formation and function. Radial glia, astrocytes, oligodendrocyte progenitor cells, oligodendrocytes, and microglia each influence nervous system development, from neuronal birth, migration, axon specification, and growth through circuit assembly and synaptogenesis. As neural circuits mature, distinct glia fulfill key roles in synaptic communication, plasticity, homeostasis, and network-level activity through dynamic monitoring and alteration of CNS structure and function. Continued elucidation of glial cell biology, and the dynamic interactions of neurons and glia, will enrich our understanding of nervous system formation, health, and function.



Myelinating oligodendrocyte in the mouse cortex

Cursed perhaps by the incorrect assertion of Virchow in 1846 that the brain contains a connective structure called "nervenkitt," the mundane notion of glia as glue (derived from Greek) emerged and stuck [(*1*) and references therein]. This is despite the fact that soon after Virchow's "discovery" of glia, the famous neuroanatomists of the late 19th and early 20th centuries had not only identified the major glia of the central nervous system (CNS) but also speculated with keen foresight on their potential diverse functions (*1*). Here we provide an overview of insights that are now addressing the question posed by Ramon y Cajal in 1909, "What is the function of glia?" [(*1*) and references therein]. We focus on the function of glia of the vertebrate CNS: radial glia, astrocytes, oligodendrocyte progenitor cells (OPCs, also called NG2 cells), oligodendrocytes, and microglia (Fig. 1). For example, we now know that radial glia are CNS progenitors that generate the majority of our neurons and glia, either directly or via intermediate progenitors (Fig. 1). Astrocytes, star-shaped cells with thousands of processes that interact with likely all cell types of the CNS, exhibit a range of functions that help drive nervous system development and sculpt its activity (Fig. 2). OPCs are the most proliferative cells in the CNS, generate myelinating oligodendrocytes throughout life, and likely serve additional yet-to-be-identified roles in circuit formation and function (Fig. 3). Mature oligodendrocytes generate myelin sheaths that speed up nerve impulse conduction and provide metabolic support to axons (Fig. 3), and it now appears that dynamic regulation of myelination may regulate the precise timing of information propagation and communication across functional circuits. Finally, microglia are the guests of the CNS, best known as the brain's resident macrophages, but with increasingly evident roles during multiple stages of nervous system development and activity (Fig. 4).

One of the most quoted "facts" about glia is that they outnumber neurons by 10 to 1 in the human brain, but it is now clear that we have a roughly equal proportion of neurons and glia [(*2*) and references therein]. The relative proportions vary by region (e.g., gray matter versus white matter), developmental stage, and species, but as a very general rule, human brains are composed of 20% astrocytes (*2*), 3 to 10% OPCs (*3*), 25% oligodendrocytes (*2*), and 5 to 15% microglia (*4*). We have known of neuronal diversity for more than 100 years, and it is now emerging that glia also exhibit important functional diversity [(*5*–*7*) and references therein], likely to be driven by both intrinsic programs based on developmental origins and extrinsic interactions with their environment. In this review, we take a chronological perspective that aims to chart how glia first drive neurogenesis and subsequently regulate neuronal migration, axon growth, synapse formation, and, ultimately, circuit function. We focus on the role of glia and their reciprocal interactions with neurons in regulating healthy nervous system formation and function and refer interested readers to other excellent sources on the development of glia (*3*, *8*–*10*) and the myriad roles of glia in disease (*4*, *11*–*13*).

Glia in CNS formation

Glia are the principle regulators of cell number in the CNS

During embryonic development, neuroepithelial progenitors take on characteristics of astrocytes, including the expression of factors such as glutamate transporters, glycogen granules, and intermediate filaments. These neuroeiptheial cell–derived "radial glia" go on to generate neurons during embryonic stages of mammalian development [(*14*) and references therein] and then switch to generating

[1]Molecular Neurobiology Laboratory, Salk Institute for Biological Studies, 10010 North Torrey Pines Road, La Jolla, CA 92037, USA. [2]Centre for Discovery Brain Sciences, University of Edinburgh, 49 Little France Crescent, Edinburgh EH16 4SB, UK.
*Corresponding author. Email: nallen@salk.edu (N.J.A.); david.lyons@ed.ac.uk (D.A.L.)

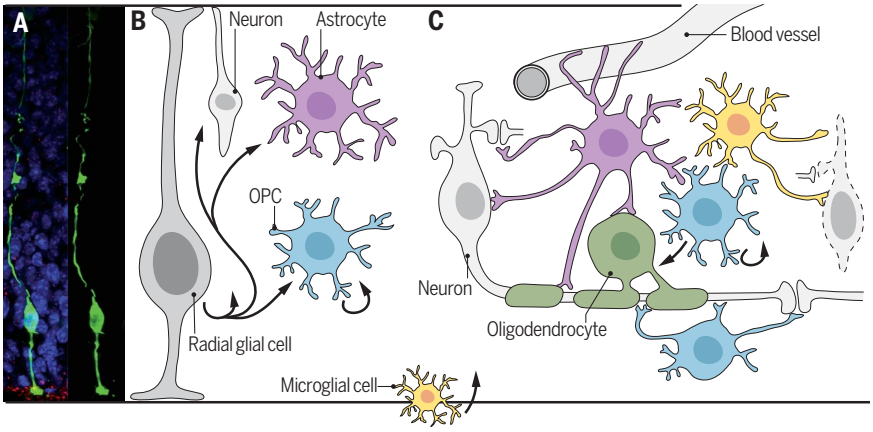


IMAGE: MATTHEW SWIRE/UNIVERSITY OF EDINBURGH; ILLUSTRATION: N. DESAI/*SCIENCE*

Fig. 1. Origin and overview of CNS glial cells. (**A**) Confocal image of a radial glial cell in an E14 mouse cortex, visualized after in utero electroporation of green fluorescent protein (GFP, green) (right) and costaining for apical centrosomes (pericentrin, red) and cell nuclei [4′,6-diamidino-2-phenylindole (DAPI), blue] (left). [Photo credits: Sven Falk and Magdalena Goetz, Helmholtz Centre, Munich] (**B**) Radial glial cells are the principal neuroepithelial progenitor cells of the central nervous system and generate the majority of CNS neurons and glia, either directly (e.g., neurons) or indirectly through intermediate progenitors (e.g., OPCs). Microglia (yellow) enter the CNS during embryonic development. (**C**) Neurons and glia interact in a myriad of ways (explained in text and subsequent figures). Dashed line indicates dying neuron.

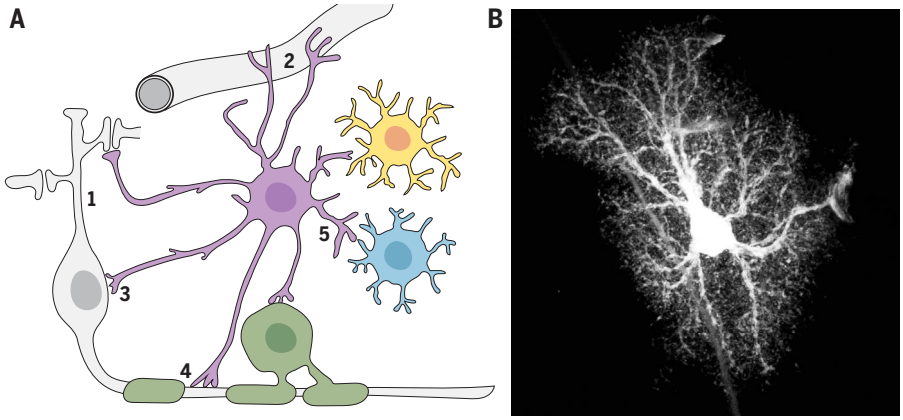


Fig. 2. Astrocytes. (**A**) Astrocytes (purple) have a characteristic star-like morphology and send out multiple branches that terminate in thousands of fine processes that interact with synapses, blood vessels, and other cells. Astrocytes regulate synapse formation, elimination, and function (1); have endfeet that ensheath CNS vasculature (2); contribute to metabolic support and homeostatic function; and have reciprocal interactions with neurons regulating circuit function (3). Astrocytes also make contact with nodes of Ranvier (4), the function of which is unclear, and interact bidirectionally with OPCs, oligodendrocytes, and microglia (5), the relevance of which is emerging. (**B**) Astrocyte confocal image. [Photo credit: The Cell Image Library, image CIL 48001; available at www.cellimagelibrary.org/images/48001]

mature glia, which they do either directly by generating proliferating astrocytes or through generation of intermediate progenitors, such as OPCs [(*3*) and references therein]. Toward the end of gliogenesis, many radial glia differentiate directly into mature astrocytes, whereas in anamniotic vertebrates (e.g., fish and amphibians), mature radial glia retain a radial morphology but may subserve the functional roles of astrocytes (*15*). Subtypes of radial glia with distinct neurogenic and gliogenic potentials have been described, and radial glia–derived progenitors persist through adulthood and give rise to adult neural stem cells [(*16*) and references therein]. In addition to cells derived from the neuroepithelium, microglia, which are derived from erythromyeloid progenitors, enter the vertebrate CNS early in development at the onset of neurogenesis and subsequently proliferate and migrate to colonize the entire brain and spinal cord [(*10*) and references therein]. Microglia regulate neuronal survival, phagocytose excess neurons undergoing apoptosis during early development and in neurogenic regions of the adult brain [(*17*) and references therein], and have multiple roles in refining CNS formation and function [(*18*) and references therein]. Although glia regulate neuron number (*14*), neurons can, in turn, regulate glial cell number. For example, the programmed cell death of specific neurons influences initial microglial cell entry to the CNS [(*18*) and references therein], and neuronal activity can regulate OPC proliferation and oligodendrocyte generation [(*19*) and references therein]. As will become clear, many aspects of nervous system formation and function are underpinned by dynamic bidirectional interactions of neurons and glia.

Glia influence neuronal migration, axon specification, and growth

Following their birth, neurons in the CNS migrate, either radially through the neuroepithelium—for example, establishing specific cortical layers—or tangentially, as exemplified by interneuron migration from the ganglionic eminences to the cortex. The role of glia in regulating tangential migration has not been intensively studied, although microglia have been implicated [(*10*) and references therein]. By contrast, it is well known that radial glia support radial migration of neurons, and the underlying mechanisms have been extensively investigated [(*14*) and references therein]. For example, disruption of radial glial cell polarity dysregulates neuronal migration to such an extent that a duplicated cortex can form underneath a normal one (*14*). Radial glia also exhibit polarized mRNA transport and localized protein translation along their processes (*20*), suggesting that they may locally regulate additional aspects of neuronal development. Indeed, axon specification occurs on the opposite side of the neuron to that of radial glial contact, at least in vitro (*21*). In addition, microglia (*10*), astrocytes, and OPCs (*22*) have been shown to subsequently help guide axons to their destinations, and analogous roles for glia in axon specification (*23*) and guidance have also been identified in *Caenorhabditis elegans* (*24*).

Glia can coordinate circuit-wide neuronal differentiation

Our understanding of the mechanisms that coordinate neuronal differentiation across circuits is rudimentary, but it is now clear that glial cells can coordinate the timing of axon-driven target neuron differentiation. In the *Drosophila* visual system, photoreceptor neurons first secrete a signal to their associated wrapping glia (analogous to myelinating glia), which then secrete cues that induce target neuron differentiation. This indirect signaling via glia introduces a delay that controls the timely differentiation of target neurons, because the wrapping glia arrive at specific target locations with a delay relative to axons (*25*). The mechanisms that might coordinate differentiation of neurons distributed across local and long-range circuits in vertebrates are unclear. However, given that invertebrate models have served as mechanistic pathfinders for many aspects of vertebrate neural development [(*26*, *27*) and references therein], the prospective role of glia in regulating circuit-level neuron differentiation warrants further investigation.

Glia regulate synapse formation and pruning

After neurogenesis, neuronal migration, and axon guidance are complete, prospective partner neurons begin to make synaptic connections. Synaptogenesis involves a series of formation, strengthening, and remodeling steps, with glia playing distinct roles at each stage. The timing of initial synaptogenesis coincides with the generation of astrocytes, putting them in the right place at the right time [(*28*) and references therein]. First, astrocytes contact immature neurons, making them competent to form synapses (*28*). Then, neurons become receptive to astrocyte-secreted signals that instruct formation of immature synapses, including thrombospondins, which induce silent synapses, and glypicans, which induce functionally active synapses (*28*). Synapses are then matured by additional glial signals, including Sparcl1 from astrocytes, which induces synaptic strengthening and stabilization (*18*). As a balance to the prosynaptogenic signals, astrocytes also produce signals (e.g., Sparc) that negatively regulate synapse formation and function [(*18*) and references therein]. In addition to astrocytes, microglia have been implicated in synapse formation [(*18*) and references therein]. The various roles of distinct glia in regulating synapse formation and the extent to which they do so throughout the nervous system and over time in vivo remains to be fully determined.

Toward the end of the synaptogenic period, weak and inappropriate synapses are eliminated by astrocytes and microglia, ultimately leaving neurons with their adult connectivity. Astrocytes phagocytose synapses via Megf10- and Mertk-driven phagocytic pathways, whereas microglia prune synapses tagged with components of the classical complement pathway [(*18*) and references therein] or through the Trem2 pathway (*29*). Interleukin-33 signaling from astrocytes to microglia has also been shown to regulate synapse elimination (*30*), highlighting the complexity of cell-cell interactions in the CNS.

A role for glia in regulating synapse development appears broadly conserved across evolution. For example, human astrocytes transplanted into mice can regulate multiple aspects of the structure and function of host synapses [(*5*) and references therein]. Invertebrate glia also regulate synapses, with glia in *C. elegans* determining the precise location of specific synapses and promoting synaptogenesis (*27*) and *Drosophila* glia regulating synapse maturation (*26*). The distinct roles of glia in synapse formation and remodeling is likely to be regulated in part by their regional identity (*8*) and complex interactions with their environment. For example, neurons can regulate astrocyte maturation and diversity

ILLUSTRATION: N. DESAI/*SCIENCE*

in vivo (*28*), and neural activity can regulate hundreds of astrocyte genes in vitro (*31*). Therefore, distinct neurons may be able to influence and use specialized glial functions to sculpt circuit-specific architecture. Determining how interactions between neurons and glia regulate different synaptic classes in different brain regions in vivo will be an important area of exploration.

Glia in regulating CNS function

Once neuronal circuits are assembled, glia regulate numerous aspects of nervous system function. They do so by affecting synaptic function, ion homeostasis, and metabolism and likely also through yet-to-be-identified roles.

Glia adjust synaptic communication and plasticity

Glia regulate many aspects of synaptic function. Astrocytes regulate neurotransmitter uptake, for example, via glutamate transporters [(*32*) and references therein], which is dynamically regulated; glutamate transporters on astrocyte processes are up-regulated by high activity and down-regulated during prolonged periods of low activity, providing a way to modulate the concentration of transmitter available at synapses (*32*). Furthermore, increased synaptic activity can also lead to a greater ensheathment of synapses by astrocyte processes, decreasing the chance of neurotransmitters diffusing and activating a neighboring synapse [(*33*) and references therein]. Astrocytes express receptors for multiple neurotransmitters and neuromodulators. Many of these are G protein–coupled receptors associated with intracellular signaling cascades that converge on intracellular calcium ion (Ca^{2+}) signaling [(*33*) and references therein]. Astrocytes can, in turn, release gliotransmitters, thus signaling back to neurons and other glial cells [(*34*) and references therein]. Whether gliotransmitter release primarily regulates the function of individual synapses or groups of synapses (given that an individual astrocyte can contact thousands of synapses) is likely to depend on the synapse type and gliotransmitter in question. Gliotransmitters include adenosine triphosphate, which, when converted to adenosine, acts on presynaptic adenosine receptors to alter neurotransmitter release probability and short-term plasticity. In addition, D-serine, released from astrocytes in a calcium-dependent manner, is a coagonist for the *N*-methyl-D-aspartate (NMDA) glutamate receptor, through which it can also regulate synaptic plasticity [(*32*–*34*) and references therein]. Astrocyte activation has even been shown to enhance both neuronal plasticity and memory in vivo (*35*).

In addition to astrocytes, OPCs and microglia have also been shown to regulate synaptic function. For example, OPCs have been implicated in controlling postsynaptic NMDA and AMPA receptor-mediated currents and plasticity through cleavage of the transmembrane proteoglycan NG2, and NG2 mutant mice exhibit deficits in sensorimotor function [(*19*) and references therein]. Dysregulation of microglial signaling also affects synaptic and circuit function. For example, cell type–specific knock out of purinergic receptors leads to disruption of ocular dominance plasticity in the visual cortex (*4*), and tumor necrosis factor–α, a cytokine secreted by microglia that regulates homeostatic synaptic scaling, influences ocular dominance plasticity (*36*). In addition, depletion of microglia from the juvenile mouse brain inhibits motor learning, with microglial brain-derived neurotrophic factor proposed to mediate the effect. However, the elimination of microglia at adult ages does not affect behavior (*4*, *18*), further indicating that stage- and region-specific neuron-glial interactions regulate neural circuit function.

Glial regulation of ion homeostasis affects circuit function

Glia can affect neuronal excitability by regulating ion homeostasis, for example, by clearing potassium (K^+) ions from the extracellular space, which facilitates neuronal membrane repolarization after an action potential and thus continued firing (*37*). The glial cells responsible for K^+ uptake can vary by brain region, and these include astrocytes, OPCs, and oligodendrocytes [(*37*) and references therein]. Along myelinated axons, most K^+ channels are localized under the myelin sheath, suggesting that K^+ clearance might occur at the interface of the axon and the myelin sheath, rather than at the node of Ranvier where astrocytes and OPCs make contact. Indeed, knock out of the inward-rectifying potassium channel Kir4.1 from mature myelinating oligodendrocytes leads to poor K^+ clearance and consequent conduction defects (*38*), revealing a function for mature oligodendrocytes in facilitating circuit function.

Glial regulation of extracellular K^+ can influence diverse aspects of neural circuit function. For example, in the rat lateral habenula (LHb), Kir4.1 is expressed at low levels in astrocytes but is up-regulated in models of depression. Regulation of astrocyte Kir4.1 can alter both the bursting activity of LHb neurons and associated depression-like symptoms (*39*). By contrast, in the spinal cord, the astrocytes surrounding large, fast motor neurons express a very high level of Kir4.1, which is necessary to maintain the high firing rate of these specific neurons (*40*). In the cortex, specialized satellite oligodendrocytes tightly associate with the cell bodies of cortical neurons and regulate extracellular K^+ concentration and

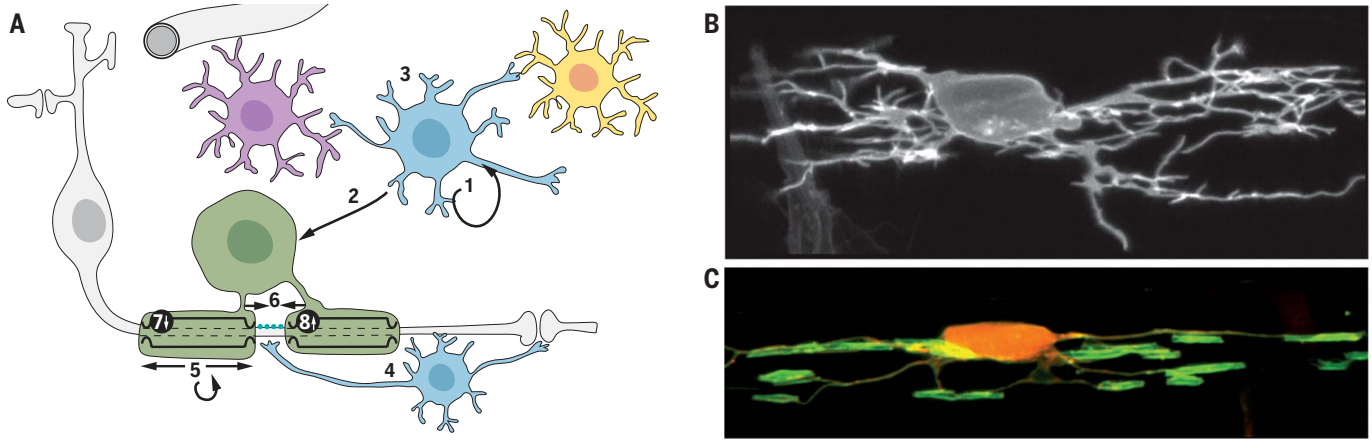


Fig. 3. OPCs and oligodendrocytes. (**A**) OPCs (blue) are the most proliferative cells of the CNS (1) and generate mature myelinating oligodendrocytes throughout life (2). OPCs interact with many other cells of the CNS (3), particularly in disease. OPCs extend processes that contact nodes of Ranvier, receive synapses from axons, and regulate synaptic function (4). Whether there are distinct subtypes of OPCs or simply different functional states remains unclear. Oligodendrocytes (dark green) produce lipid rich myelin sheaths that wrap around axons and regulate action potential conduction velocity. Dynamic regulation of myelination in response to neuronal signals (5) may be a fundamental mechanism by which neural circuit function is fine-tuned. Myelinating oligodendrocytes organize axonal domains (6), including nodes of Ranvier, where the voltage-gated Na^{2+} channels that mediate action potential propagation are localized; provide metabolic support to axons (7); and facilitate ion homeostasis that is essential to normal action potential conduction, for example, the uptake of K^+ ions by myelin (8). (**B**) Confocal image of an OPC expressing membrane-tethered fluorescent protein visualized in a living transgenic zebrafish larvae at 3 days postfertilization. (**C**) Confocal image of an oligodendrocyte expressing membrane-tethered GFP (green) and cytoplasmic red fluorescent protein (orange) visualized in a living transgenic zebrafish larvae at 4 days postfertilization. [Photo credit: Marion Baraban, Lyons lab]

ILLUSTRATION: N. DESAI/*SCIENCE*

neuronal activity through formation of a glial syncytium with neighboring astrocytes and oligodendrocytes [(*19*) and references therein]. It remains unknown to what extent region-specific glial phenotypes (e.g., high or low Kir4.1 expression) are driven by intrinsically distinct developmental programs or by dynamic interactions with the extrinsic environment.

Neuro-glial-vascular coupling provides metabolic support

Interactions between the vasculature and cells of the CNS are increasingly implicated in brain health. CNS blood vessels have specialized barrier properties, known as the blood-brain barrier, which limit movement of substances, including nutrients, between the blood and the brain (*41*). However, astrocytes are ideally positioned to actively supply neurons with nutrients because they extend processes that contact both blood vessels and neurons (*42*). Astrocytes are not just relays but also process precursor molecules for delivery to neurons. For example, astrocytes are enriched in lipid synthesis pathways and provide lipids to neurons, including cholesterol, which is essential for maintaining presynaptic function (*28*). Astrocytes are also enriched in glycolytic pathways and make glycogen, which can be converted into lactate and transferred to neighboring neurons by monocarboxylate transporters (*43*). Although there has been debate about the specific role of lactate as an energy source for neurons, mice lacking monocarboxylate transporters do show impairments in long-term memory formation (*43*).

Along myelinated axons, myelin sheathes create a barrier between the axon and the extracellular environment, limiting direct metabolic support routes to the axon to the short unmyelinated gaps between myelin sheaths at nodes of Ranvier. Instead, myelinating oligodendrocytes generate glycolytic products, including lactate, which can be transported to the axon via monocarboxylate transporters at the myelin-axon interface (*44*). It has been proposed that high-frequency axonal firing activates NMDA receptors on oligodendrocytes, triggering increased glucose uptake and the subsequent transport of lactate to the axon to support the energy requirements associated with maintaining high levels of activity (*44*). Visualizing this metabolic support system in real time and in response to neural activity and investigating how astrocytes and oligodendrocytes collaborate on metabolic support are important areas of ongoing investigation.

Glia regulate network-level function

Astrocytes dynamically alter brain states

There is now compelling evidence of glial involvement in dynamic regulation of circuit-level function. For example, networks of astrocytes in the mouse visual cortex can be activated by the neuromodulator norepinephrine in line with dynamic changes in the arousal state of the animal. This increased astrocyte activity can in turn enhance the ability of astrocytes to detect altered neuronal activity (*5*). The regulation of circuit function via astrocytes has also been demonstrated in *Drosophila*, where octopaminergic neurons signal through astrocytes to control downstream neurons and behavior (*28*). Furthermore, dynamic regulation of astrocyte population-level Ca^{2+} activity has been shown to alter ensemble-level oscillatory states in mice (*5*). More than a century ago, Ramon y Cajal proposed that astrocytes may regulate sleep state, a suggestion now supported by extensive evidence (*5*).

Myelination dynamically alters circuit function

Myelination of axons by oligodendrocytes has long been known to increase conduction velocity. In addition to whether an axon is myelinated or not, changes to the number, distribution, length, and thickness of myelin sheaths all, in principle, affect conduction velocity and thus the timing of information propagation, prompting speculation that dynamic regulation of myelination in response to neural activity may fine-tune the timing of communication between neurons in a circuit. Indeed, neural activity can regulate many aspects of oligodendrocyte lineage progression in developing, juvenile, and adult animals [(*19*) and references therein]. OPCs receive bona fide synapses from neuronal axons, which may mediate the effects of neural activity on their proliferation and/or differentiation (*19*). Mature myelinating oligodendrocytes also express a range of neurotransmitter receptors, ion channels, and transporters, whose diverse functions remain to be fully elucidated but which likely mediate the effects of neural activity on myelination itself. That only some neuronal subtypes regulate myelination via activity [(*19*) and references therein] suggests that activity-regulated myelination may be circuit specific.

Direct evidence that activity-regulated myelination, in turn, affects circuit function comes from several studies. Social isolation of mice, for a distinct postnatal period or a prolonged period in adulthood, disrupts myelination in the prefrontal cortex [(*19*) and references therein]: In addition, the behavioral traits associated with social isolation can be phenocopied by loss of oligodendrocyte ErbB receptor signaling during the corresponding postnatal period and reversed by chemically promoting myelination in adulthood [(*19*) and references therein]. These results indicate that the changes in neural activity associated with isolation lead to behavioral deficits via changes in myelination. In addition, learning new motor tasks promotes new oligodendrocyte generation in adult mice, and the prevention of such oligodendrogenesis impairs learning [(*19*) and references therein]. Nonetheless, it remains to be determined how dynamic regulation of myelination affects the timing of conduction, synaptic communication, and the function of neuronal populations within intact circuits in vivo (*19*). Evidence that neurons regulate myelinating glia, and that myelinating glia, in turn, regulate circuit function, further exemplifies how reciprocal interactions between neurons and glia play fundamental roles in circuit function.

Conclusions

Here we have highlighted integral roles of glial cells in many aspects of CNS formation and function, and we refer interested readers to our more extensive supplementary reading list for further information. Given the emerging evidence of diverse subtypes and dynamic states of glia, we expect that further new functions will be revealed, and their mechanistic underpinnings elucidated, in years to come. This will require interrogation using a variety of models and approaches, bridging of scales from molecule to system, and investigating mechanisms from development through the entire life course. The study of glia reminds us that nervous system development does not end at birth, nor indeed at adolescence, and that the dynamic behavior of glial cells and ongoing neuron-glial interactions actively sculpt and remodel the CNS throughout life. We expect that further insight into how the brain forms, functions, and ages in a healthy

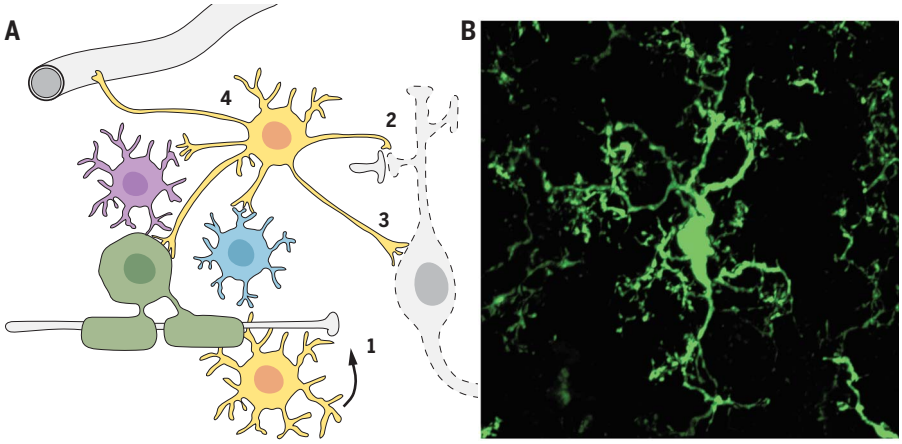


ILLUSTRATION: N. DESAI/*SCIENCE*

Fig. 4. Microglia. (**A**) Microglia (yellow) are the resident immune cells of the brain, entering during early development from the periphery (1). In addition to immune surveillance roles (not shown), microglia interact with multiple cell types of the CNS and regulate numerous developmental and functional processes, including synaptic pruning (2), clearing apoptotic neurons (3), and interacting with multiple CNS cell types, in health and disease (4). (**B**) Confocal image of microglia expressing GFP in mouse cortex (Cx3cr1-GFP). [Photo credit: Youtong Huang and Greg Lemke, Salk Institute]

manner will derive from integrating our understanding of how neurons interact with glia, how different glial cells interact with each other, and how the brain interacts with the rest of the body.

REFERENCES AND NOTES

1. G. G. Somjen, *Glia* **1**, 2–9 (1988).
2. C. S. von Bartheld *et al.*, *J. Comp. Neurol.* **524**, 3865–3895 (2016).
3. L. Dimou, M. Götz, *Physiol. Rev.* **94**, 709–737 (2014).
4. M. W. Salter, B. Stevens, *Nat. Med.* **23**, 1018–1027 (2017).
5. K. E. Poskanzer, A. V. Molofsky, *Annu. Rev. Physiol.* **80**, 143–157 (2018).
6. G. S. Tomassy *et al.*, *Trends Cell Biol.* **26**, 135–147 (2016).
7. S. A. Wolf *et al.*, *Annu. Rev. Physiol.* **79**, 619–643 (2017).
8. O. A. Bayraktar, L. C. Fuentealba, A. Alvarez-Buylla, D. H. Rowitch, *Cold Spring Harb. Perspect. Biol.* **7**, a020362 (2014).
9. D. E. Bergles, W. D. Richardson, *Cold Spring Harb. Perspect. Biol.* **8**, a020453 (2015).
10. F. Ginhoux, M. Prinz, *Cold Spring Harb. Perspect. Biol.* **7**, a020537 (2015).
11. E. Dossi *et al.*, *Brain Res. Bull.* **136**, 139–156 (2018).
12. R. Tognatta, R. H. Miller, *Neuropharmacology* **110** (pt. B), 539–547 (2016).
13. K.-A. Nave, H. Ehrenreich, *JAMA Psychiatry* **71**, 582–584 (2014).
14. E. Taverna *et al.*, *Annu. Rev. Cell Dev. Biol.* **30**, 465–502 (2014).
15. D. A. Lyons, W. S. Talbot, *Cold Spring Harb. Perspect. Biol.* **7**, a020586 (2014).
16. S. Falk, M. Götz, *Curr. Opin. Neurobiol.* **47**, 188–195 (2017).
17. K. Kierdorf, M. Prinz, *J. Clin. Invest.* **127**, 3201–3209 (2017).
18. K. Reemst *et al.*, *Front. Hum. Neurosci.* **10**, 566 (2016).
19. R. G. Almeida, D. A. Lyons, *J. Neurosci.* **37**, 10023–10034 (2017).
20. L.-J. Pilaz *et al.*, *Curr. Biol.* **26**, 3383–3392 (2016).
21. C. Xu *et al.*, *J. Neurosci.* **35**, 14517–14532 (2015).
22. S. Minocha *et al.*, *Nat. Commun.* **6**, 6887 (2015).
23. L. Meng, A. Zhang, Y. Jin, D. Yan, *eLife* **5**, e19510 (2016).
24. G. Rapti *et al.*, *Nat. Neurosci.* **20**, 1350–1360 (2017).
25. V. M. Fernandes *et al.*, *Science* **357**, 886–891 (2017).
26. M. R. Freeman, *Cold Spring Harb. Perspect. Biol.* **7**, a020552 (2015).
27. S. Shaham, *Cold Spring Harb. Perspect. Biol.* **7**, a020578 (2015).
28. N. J. Allen, C. Eroglu, *Neuron* **96**, 697–708 (2017).
29. F. Filipello *et al.*, *Immunity* **48**, 979–991.e8 (2018).
30. I. D. Vainchtein *et al.*, *Science* **359**, 1269–1273 (2018).
31. P. Hasel *et al.*, *Nat. Commun.* **8**, 15132 (2017).
32. N. J. Allen, *Annu. Rev. Cell Dev. Biol.* **30**, 439–463 (2014).
33. B. S. Khakh, K. D. McCarthy, *Cold Spring Harb. Perspect. Biol.* **7**, a020404 (2015).
34. A. Araque *et al.*, *Neuron* **81**, 728–739 (2014).
35. A. Adamsky *et al.*, *Cell* **174**, 59–71.e14 (2018).
36. M. Kaneko *et al.*, *Neuron* **58**, 673–680 (2008).
37. B. R. Larsen *et al.*, *Glia* **62**, 608–622 (2014).
38. V. A. Larson *et al.*, *eLife* **7**, e34829 (2018).
39. Y. Cui *et al.*, *Nature* **554**, 323–327 (2018).
40. K. W. Kelley *et al.*, *Neuron* **98**, 306–319.e7 (2018).
41. R. Daneman, A. Prat, *Cold Spring Harb. Perspect. Biol.* **7**, a020412 (2015).
42. R. Nortley, D. Attwell, *Curr. Opin. Neurobiol.* **47**, 80–85 (2017).
43. P. J. Magistretti, I. Allaman, *Nat. Rev. Neurosci.* **19**, 235–249 (2018).
44. A. S. Saab, K.-A. Nave, *Curr. Opin. Neurobiol.* **47**, 104–112 (2017).

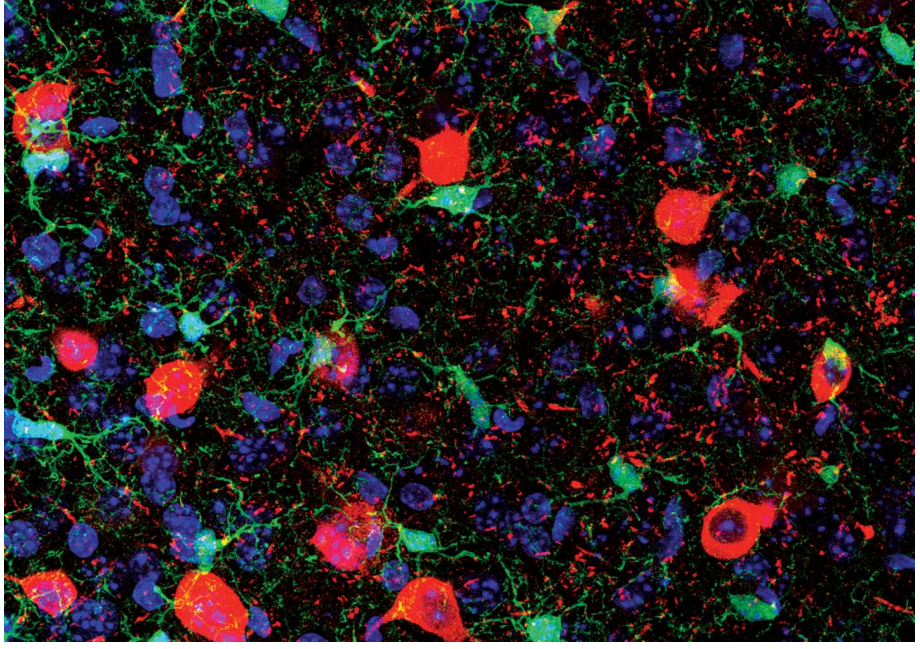
ACKNOWLEDGMENTS

We dedicate this piece to Ben A. Barres, a pioneer in glial cell biology, who challenged neuroscientists to think of glia not as glue but as active participants in nervous system development, function, and disease. We thank members of the Allen and Lyons laboratories for helpful feedback on the manuscript. We apologize to colleagues for omitting original references, owing to space constraints, and refer interested readers to our supplementary reading list. **Funding**: Work in the Allen lab is supported by NIH-NINDS NS105742 and NS089791 and the Hearst, Pew, Dana, CART, and LouLou Foundations; and in the Lyons lab by a Wellcome Trust Senior Research Fellowship (102836/Z/13/Z), a MRC Project Grant (MR/P006272/1), and a Strategic Research Agreement with Biogen. **Competing interests:** The authors declare no competing interests.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/362/6411/181/suppl/DC1
Supplementary Reading List

10.1126/science.aat0473



Coronal section of the embryonic mouse somatosensory neocortex. Green, microglia; red, inhibitory interneurons; blue, nuclear staining.

IMAGE: MORGANE S. THION

REVIEW

Microglia and early brain development: An intimate journey

Morgane S. Thion[1]*, Florent Ginhoux[2,3], Sonia Garel[1]*

Cross-talk between the nervous and immune systems has been well described in the context of adult physiology and disease. Recent advances in our understanding of immune cell ontogeny have revealed a notable interplay between neurons and microglia during the prenatal and postnatal emergence of functional circuits. This Review focuses on the brain, where the early symbiotic relationship between microglia and neuronal cells critically regulates wiring, contributes to sex-specific differences in neural circuits, and relays crucial information from the periphery, including signals derived from the microbiota. These observations underscore the importance of studying neurodevelopment as part of a broader framework that considers nervous system interactions with microglia in a whole-body context.

The standard classification of central nervous system (CNS) lineages separates neurons that process information from glial cells, which support and modulate neuronal activity. Glial cells are now beginning to emerge as major contributors to CNS development and homeostasis. In particular, brain-resident macrophages termed "microglia" appear to play major roles at the interface between the immune and nervous systems (*1*–*5*), each of which is crucial for host perception of the external environment. Both the nervous and immune systems share a number of key attributes, including the need for a critical training period during which baseline levels of input are defined, such that subsequent atypical and potentially dangerous signals can be detected efficiently. Despite these common features, the nervous and immune systems have classically been studied independently. However, there is now increasing evidence of a strong neuroimmune interplay in adults that determines how the brain modulates tissue immunity and normal processes of aging, as well as inflammation and neurological disorders including Alzheimer's disease, Parkinson's disease, multiple sclerosis, and various autoimmune pathologies (*1*–*4*). In addition, activation of the immune system during pregnancy or early life has been shown to exert long-term effects on the wiring of neural circuits and may contribute to the etiology of neurodevelopmental disorders. Microglia develop in close proximity to neurons and other glial cells from early embryonic stages, and multiple studies have now established a pivotal role for these cells in CNS development and homeostasis. This Review attempts to highlight important recent advances in this field, focusing on the developmental roles of microglia at the interface between brain circuits and the periphery while also identifying key goals for future research.

[1]Institut de Biologie de l'École Normale Supérieure (IBENS), École Normale Supérieure, CNRS, INSERM, PSL Université Paris, 75005 Paris, France. [2]Singapore Immunology Network (SIgN), Agency for Science, Technology and Research (A*STAR), 8A Biomedical Grove, IMMUNOS Building no. 3-4, BIOPOLIS, 138648, Singapore. [3]Shanghai Institute of Immunology, Shanghai JiaoTong University School of Medicine, 280 South Chongqing Road, Shanghai 200025, China.
***Corresponding author. Email: thion@biologie.ens.fr (M.S.T.); garel@biologie.ens.fr (S.G.)**

Ontogeny and developmental trajectories of CNS immune cells

Microglia represent 5 to 15% of adult brain cells, with densities varying between distinct brain regions. They constitute by far the largest population of immune cells in the brain, and under physiological conditions they are unique in being located within the brain parenchyma, where they lie in direct contact with neural progenitors, neurons, and other glial cells (namely astrocytes and oligodendrocytes). The genesis of microglia remained obscure for many years, until landmark studies in mammals identified a very early embryonic origin for this population within the yolk sac (YS) (*1*, *2*, *5*). This earliest primitive wave of hematopoiesis arises at embryonic day (E) 7.5 in mice, generating nucleated red blood cells and macrophages that colonize the entire embryo. Macrophages start to enter the CNS at E9 by using the blood vasculature before closure of the blood-brain barrier (BBB), which occurs around E13 in mice, restricting access to immune cells that arise later in development (*5*) (Fig. 1). At these early stages, neural progenitors are actively dividing and generate the first neurons, whereas they will give rise to oligodendrocytes and astrocytes only at late embryonic time points. Consequently, microglia constitute the main glial population during a large part of fetal life.

In most tissues located outside of the CNS, YS-derived macrophages are progressively replenished from circulating monocytes, which are generated later in development from fetal and definitive bone-marrow hematopoiesis (*1*, *2*, *5*). In contrast, microglia, which are located behind the BBB, self-renew for the entire life of the animal, at least under steady-state conditions. It is important to note that this situation is different in zebrafish, in which secondary waves of hematopoiesis contribute to the microglial pool. YS-derived macrophages also colonize the meninges and other surrounding CNS tissues to produce border-associated macrophages (BAMs), including meningeal, choroid plexus, and perivascular macrophages (PVMs) (*6*) (Fig. 1). However, future studies will be needed to determine whether these arise from common precursors, with the goal of characterizing their potential differentiation pathways and divergent functions. In other words, it remains unclear whether microglia versus BAM lineage commitment depends on specific precursor populations generated in the YS or whether these distinct fates are instructed by local cues encountered by macrophages as they travel through the developing CNS. In contrast, choroid plexus macrophages are known to be replaced over time by cells generated during adulthood, supporting the notion that the choroid plexus lacks a discernable BBB and thus remains accessible to circulating bone marrow–derived cells during later life (*6*).

In adult mice, other immune populations including monocytes and lymphocytes, which lie outside of the parenchyma, have been shown to regulate neural function via the production of long-range signals, including interleukin-4 (IL-4), tumor necrosis factor–α (TNF-α), and interferon-γ (IFNγ) (*7*–*9*). However, there are only a few studies to suggest that equivalent populations influence the developing brain. In parallel, recent evidence has shown that the neonatal mouse parenchyma can be infiltrated by B cells even in physiological conditions, with implications for oligodendrogenesis (*10*). Future studies will therefore be essential to decipher the potential impact of different immune cell populations on brain development, whether acting indirectly via effects on microglia or capable of directly influencing neurons.

In any event, the developing CNS clearly has a distinctive immune status associated with the presence of long-lived resident macrophages that originate in early embryogenesis (Fig. 1). These data indicate that the microglial population has a lifelong history, potential memory of previous interactions, and a peculiar symbiotic relationship with the developing brain.

Neural cues and systemic signals shape microglial development

Colonization of the brain parenchyma by YS-derived macrophages is a long-lasting process that spans embryogenesis and early postnatal stages (*1*, *11*, *12*). In the mouse, the first macrophages enter around E9, and the adult pattern of homogeneous tiling is reached during the second postnatal week (*1*), paralleling the acquisition of a progressively more ramified microglial morphology. Transcriptomic and functional studies have revealed that, like other tissue-resident macrophages, microglia undergo distinct phases of differentiation (*13*–*16*) that rely on signals derived from the maturing CNS. These signals include but are not restricted to colony-stimulating factor 1 (CSF1), IL-34, and transforming growth factor–β (TGF-β) (*1*) (Fig. 2). At the molecular level, transition between a brain macrophage signature and a microglial profile can be identified by the expression of core genes, including genes encoding the key transcription factor SALL1 and purinergic receptor P2Y12 (*1*, *15*, *16*). Deciphering the full range of signals that regulate this maturation process is extremely challenging, given that microglia rapidly lose transcriptional and epigenetic identity upon separation from their niche (*15*, *16*). In addition, the adult CNS environment is not sufficient to induce full microglial differentiation of bone marrow–derived monocytic cells (*17*, *18*), indicating that both their early origin and interactions with the CNS are critical for the acquisition of a bona fide microglial identity.

In addition, local neural signals likely regulate the brain colonization pattern and heterogeneity of microglial populations in both time and space. Microglia enter the CNS via sequential waves (*11*, *12*) and transiently show a stereotypical and heterogeneous pattern of localization (*11*, *19*, *20*). For instance, microglia undergo timely invasion into the deep layers of the neocortex, but they also associate with distinct axonal tracts, including the corpus callosum and dopaminergic midbrain axons, and populate specific niches known to contain neurogenic progenitors (*20*). This distinctive pattern is thought to be regulated by transient CNS cell expression of cues including IL-34, CSF1, CXCL12, and CX3CL1 (also known as fractalkine) (*1*) (Fig. 2). In addition, studies in zebrafish have indicated that developmental programmed cell death in the CNS regulates microglial colonization (*1*). Following early colonization, there is increasing evidence that adult microglia display some heterogeneity across brain regions, particularly in the corpus callosum, prefrontal cortex, cerebellum, or basal ganglia (*1*, *12*, *21*, *22*). Indeed, transcriptomic profiling, genetic labeling, and analysis of cell dynamics have revealed a degree of heterogeneity suggesting that regional differences may arise, in part, because of local signals produced by neural cells (*1*, *12*, *21*, *22*). Future single-cell transcriptomic studies will be required to unravel this heterogeneity and determine the underlying mechanisms. Nevertheless, the studies conducted to date indicate that CNS-derived signals have a major role to play in the specification, maturation, and colonization of microglial populations (Fig. 2).

In addition to local cues, the developmental trajectory of microglia is known to be influenced by sexual identity and systemic signals such as those derived from the microbiota or inflammation at pre- and postnatal stages (*14*, *23*). Several reports have highlighted a distinct transcriptomic signature of microglia in males versus female animals that begins postnatally and is maintained in adults (*14*, *23*, *24*). This sexual dimorphism appears to be conserved after grafting and may occur independently of circulating sex hormones, suggesting long-lasting programming by early-life hormones or genetic factors (*24*). Barrier organs, including the gut, skin, and lungs, are also populated by a diverse range of bacteria and other microorganisms termed the "microbiota," which play essential roles in the development and training of both innate and adaptive immune systems (*25*). Despite their relative isolation behind the BBB, microglia are able to respond to microbiota-derived signals both pre- and postnatally (*13*, *14*, *26*), as germ-free mice display incomplete development of microglia both during embryogenesis and at birth (*13*, *14*). Whereas in adults the microbiota has been shown to act on microglia in part via systemic production of short-chain fatty acids (*26*), the mechanisms involved remain unexplored in embryos. Conversely, inflammation is also known to change microglial developmental trajectory, from early embryonic stages (*13*, *23*) into adulthood, in particular via the production of cytokines that can act over long distances. Microglial responses to inflammation and microbiota-derived signals differ between females and males (*14*), although the underlying mechanisms remain to be defined. As the list of signals that can influence microglial development continues to expand, it becomes increasingly clear that these cells are capable

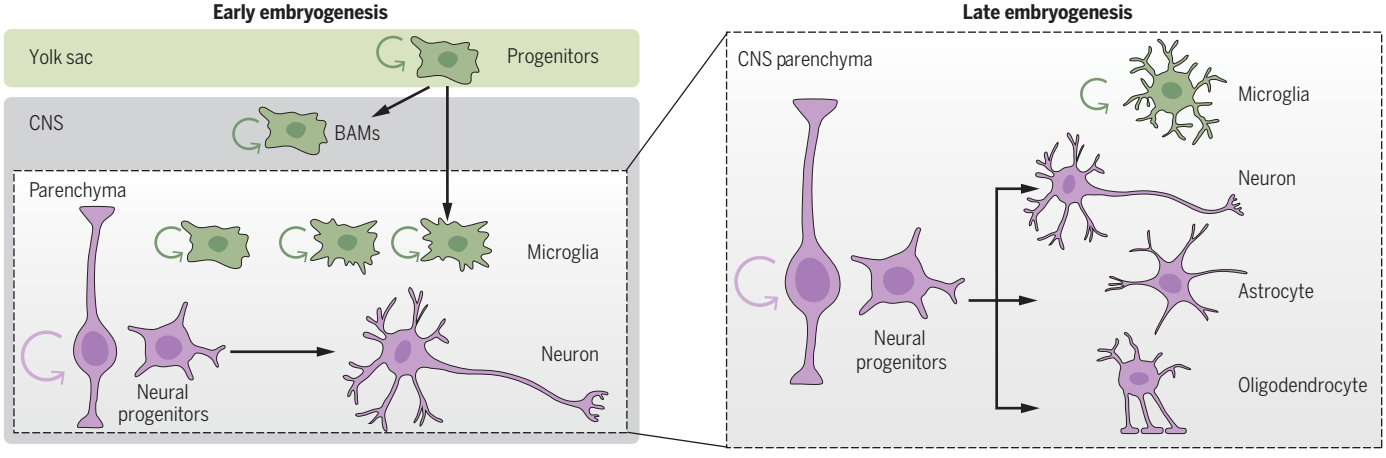


Fig. 1. Parallel development of neural and immune cells of the CNS. Schematic showing early CNS colonization by YS-derived macrophages that give rise to microglia and BAMs, whereas neural progenitors generate neurons (**left**). At later embryonic stages (**right**), neural progenitors additionally produce oligodendrocytes and astrocytes in the parenchyma while the microglia undergo local differentiation.

of integrating a diverse range of systemic and local cues that likely have important implications for neuronal function.

At a mechanistic level, it will be essential for future studies to determine how microglia can integrate such a wide variety of influences. Recent analyses of conditional HDAC1 and -2 mutants have shown that embryonic epigenetic modifications are essential for full microglial differentiation (*27*). Because the microglial population self-renews throughout life, somatic mutations occurring in the lineage or epigenetic modifications within this population, induced either by local or systemic signals, could have major physiological or even pathological consequences in adults (*15*, *16*, *28*). For instance, somatic mutations in primitive hematopoiesis in mice were shown to induce neurodegeneration in adults (*28*). It will therefore be important to assess whether and how genetic and epigenetic changes during the critical developmental period regulate the emergence of correctly tuned microglia and brain-resident immune cells.

The multifaceted roles of microglia in early CNS development

Although the overall morphogenesis of brain structures and their assembly do not require microglia, a number of recent studies have reported a central role for these cells in specific aspects of brain development, homeostasis, and disease. In particular, microglia were shown to eliminate apoptotic cell debris and contribute to activity-dependent synaptic reorganization, with a role in presynaptic nibbling and promotion of postsynaptic spine formation (*2*–*4*, *29*, *30*). This later process has already been the focus of landmark reviews (*1*–*4*) and is regulated by the complement cascade, TREM2 (*31*), or the CX3CL1-CX3CR1 receptor pathway, depending on the stage and structure concerned. Interestingly, these studies link microglia dysfunction with behavioral deficits in adults as well as neuropsychiatric disorders.

Consistent with their distinctive early ontogeny and uneven pattern of brain colonization, several additional functions of microglia during neuronal development are now beginning to emerge (Fig. 3). Microglia have been linked with the regulation of neuronal numbers, the early wiring of neural circuits, and the acquisition of sex-specific features. Indeed, microglia show a precise localization to several neurogenic niches in the rat brain, including the postnatal subventricular zone and neocortex, where they regulate nonapical progenitor cell numbers via selective engulfment (Fig. 3A) (*20*). This situation has also been observed in the mouse but appears less pronounced, suggesting that these events may display species specificity (*1*). In addition, microglia are also emerging as regulators of early circuit assembly via the CX3CL1-CX3CR1 pathway. In normal physiology, microglia transiently associate with specific axonal tracts, including dopaminergic axons and corpus callosum axons, where they then regulate axon progression and fasciculation (Fig. 3B) (*19*, *32*). Microglia can also regulate the migration of inhibitory cortical interneurons, which are essential regulators of the excitation-inhibition balance and have been implicated in autism spectrum disorders (ASDs) and schizophrenia (Fig. 3C) (*19*). Independently of this early role in wiring, a subset of white matter microglia have been reported to promote survival of layer V pyramidal neurons via the production of insulin-like growth factor 1 (IGF-1) (Fig. 3D) (*33*), thus supporting the idea that microglia can influence neuronal numbers at distinct steps and through multiple mechanisms. In addition, early postnatal microglia have also been reported to contribute to masculinization of the preoptic area, releasing prostaglandins and regulating connectivity, thereby inducing male-specific copulatory behavior in the offspring (Fig. 3E) (*34*). These findings reveal a noteworthy interplay between steroid hormones, microglia, and the early acquisition of sexually dimorphic brain circuitry and behavior. Whether these features also apply to other brain regions or time periods such as adolescence remains entirely unknown.

In parallel with their roles in early neuronal development, microglia cross-talk with other glial cells that modulate neuronal function. Microglia have been reported to lie in close apposition with blood vessels and to modulate endothelial tip cell fusion (*1*). Because microglia and PVMs are difficult to discriminate at these early stages, further experiments will be required to clearly assess the relative contributions of each population. In addition, microglia are key regulators of oligodendrocyte precursor differentiation and myelin formation. In particular, subsets of microglia expressing CD11c promote myelination via IGF-1 release, revealing a substantial coordination of both neuronal survival and myelination by a single soluble factor (Fig. 3D) (*21*, *22*). Lastly, astrocyte-derived IL-33 was recently shown to regulate microglial phagocytosis of neuronal spines (Fig. 3E) (*35*), clearly demonstrating a neuronal output to astrocyte-microglial interactions.

Together, these studies establish microglia as central mediators of cross-talk between neuronal and glial cells, where they regulate multiple aspects of early brain wiring in addition to their well-characterized roles in synaptic remodeling. Importantly, though specific molecular pathways control microglial activity and live imaging studies have revealed dynamic physical interactions with their CNS neighbors, we still have a limited understanding of how microglia act and exert specific functions. Combining in vivo imaging, single-cell spatial information, and identification of the key molecules involved via transcriptomic studies will bridge the gap between the molecular and cellular levels and provide a much-needed conceptual framework for microglial functions.

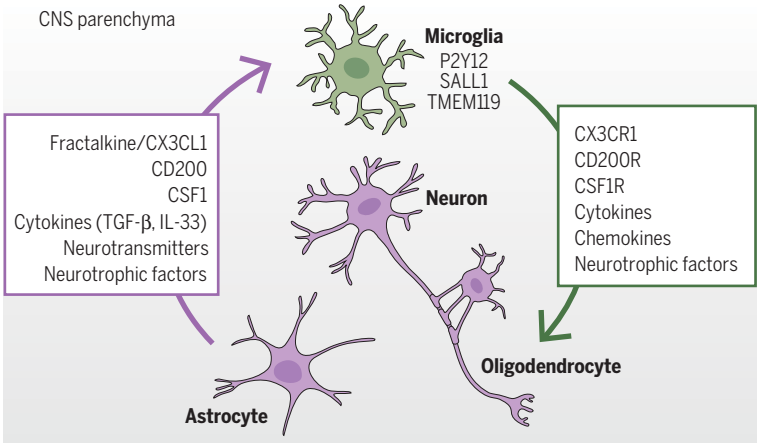


Fig. 2. Cross-talk in the development of microglia and CNS cells. Expression of secreted or membrane-bound factors (boxes) is essential for neuroimmune dialogue and is instrumental in the acquisition of a microglia-specific signature, as characterized by expression of CSF1R, Sall1, P2Y12, and Tmem119.

Impact of inflammation on microglial and CNS development

In addition to local interactions between microglia and their CNS neighbors, long-range signaling from other embryonic tissues and even the maternal environment can modulate directly or indirectly developing CNS cells. For example, it is now well recognized that systemic inflammation during pregnancy is associated with defective brain wiring. Epidemiological studies have revealed potential links between bacterial or viral infection during the first and second trimesters with an increased risk of ASDs or schizophrenia in the offspring (*36*). Importantly, animal models of this maternal immune activation (MIA) can be induced by injections of polyI:C (a synthetic analog of double-stranded RNA present in some viruses) or lipopolysaccharide (a component of the membrane of Gram-negative bacteria) in pregnant dams to respectively mimic viral or bacterial infections. These procedures are sufficient to induce a set of core behavioral deficits related to neurodevelopmental disorders in the adult offspring (*36*, *37*). It is important to note that there are multiple MIA models described in the literature that employ different stimuli and injection protocols (dose, timing, and frequency) to induce distinct behavioral alterations. In the well-studied MIA-polyI:C model, the Patterson lab performed seminal work demonstrating an essential role for long-range signaling by IL-6

cytokine in MIA pathology (*37*). More recently, it was shown that IL-6–induced differentiation of maternal T helper 17 (T_H17) cells and subsequent production of IL-17 in the mother can signal directly to the placenta and embryo (*38*, *39*). These two cytokines in particular appear to play key roles in mediating the deleterious effects of prenatal inflammation on embryonic wiring of the mouse brain. In humans, levels of IL-6 production in pregnant mothers have been correlated with functional connectivity and working memory in babies (*40*), suggesting a potential conservation of this mechanism between species.

How might circulating cytokines affect fetal brain wiring? Although it has been proposed that IL-17 could act directly on neurons (*39*), it is also well established that microglia are perturbed by MIA and could thereby contribute to neuronal dysfunction (*3*, *13*, *23*, *36*). A consistent finding is that MIA models and embryonic-stage depletion of microglia induce phenotypic similarities in the affected animals, in particular the distinctive dysregulation of cortical inhibitory interneurons (*19*), which is a hallmark of

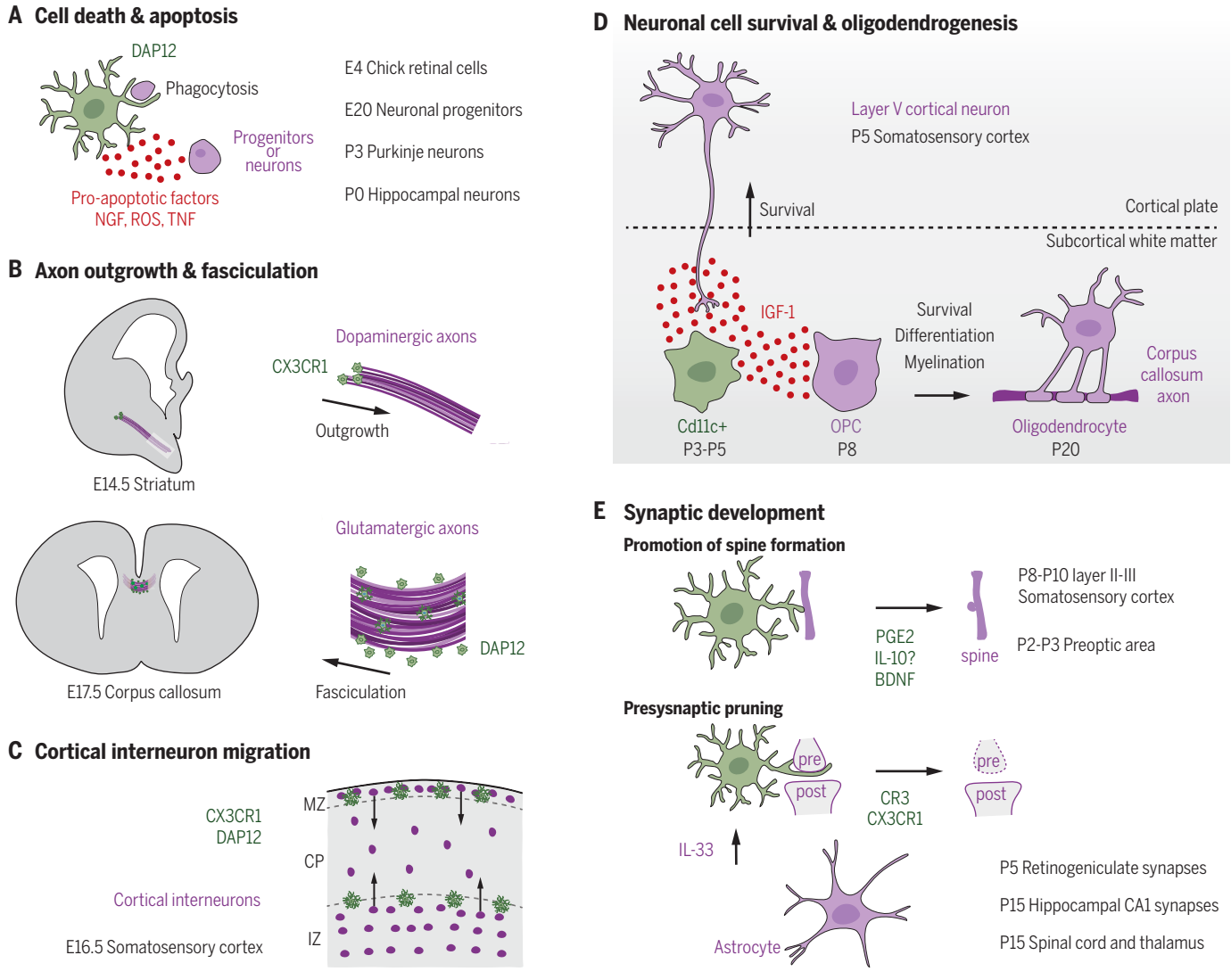


Fig. 3. Main cellular functions of embryonic and postnatal microglia. Schematic representation of microglial functions during pre- and postnatal development. (**A**) Microglia act as regulators of neuronal cell number by actively phagocytosing progenitor cells and promoting apoptosis of differentiated cells via mechanisms including nerve growth factor (NGF) secretion. ROS, reactive oxygen species. (**B**) During pre- and postnatal development, microglia also participate in the emergence of connectivity by promoting outgrowth or fasciculation of axonal tracts such as dopaminergic and corpus callosum axons. (**C**) Microglia can further influence neuronal migratory processes and regulate the development of inhibitory interneurons in the embryonic somatosensory cortex. MZ, marginal zone; CP, cortical plate; IZ, intermediate zone. (**D**) During the first postnatal week, microglial release of IGF-1 is required to regulate layer V pyramidal cell survival in the somatosensory cortex. Similarly, in the cerebellar white matter and corpus callosum of rodents, a neonatal subpopulation of $CD11c^{high}$ microglia expresses large amounts of IGF-1 and thereby regulates oligodendrocyte precursors and myelination. OPC, oligodendrocyte progenitor cell. (**E**) After birth, microglia are additionally involved in the promotion of excitatory synapses in the neocortex, either by promoting the formation of spines or regulating synaptic transmission. Microglia are well-described mediators of synapse elimination through engulfment of presynaptic inputs, which rely on activity-dependent and complement tagging in the visual cortex and retinogeniculate system, as well as on CX3CR1 expression in the hippocampus. This role of microglia in synaptic pruning is closely regulated by neuronal activity and by astrocyte secretion of IL-33, as recently shown in the spinal cord and thalamus. Throughout the figure, microglia are depicted in green, released factors in red, and targets in purple. For each panel, the developmental stage [embryonic (E) or postnatal (P)] and anatomical location are indicated. PGE2, prostaglandin E2; BDNF, brain-derived neurotrophic factor.

several neurodevelopmental disorders. In addition, MIA is known to induce long-term changes in the developmental trajectories of microglia (*23*). Because microglia are a long-lived population, early-life inflammatory events induce not only acute effects but potentially also long-term modifications in brain wiring, plasticity, and neurodegeneration (*27*). Mechanistically, it will be essential to assess the roles of epigenetic modifications, which regulate microglial differentiation and are susceptible to environmental signals (*13*, *14*).

It is important to note here that maternal T_H17 cells are known to be regulated by the composition of the microbiota, and hence MIA-induced morphological and behavioral phenotypes are highly variable between offspring generated under different housing conditions. Indeed, host T_H17 cells appear strictly dependent on the presence of species including segmented filamentous bacteria (SFB) within the gut microbiota, but SFB are not present in all animal facilities used to generate MIA mouse models. These results highlight the urgent need to clarify the effects of various different MIA protocols, including the nature and timing of the microbial stimulus employed, the mouse strain used, and the microbiota composition in different animals. They furthermore reveal the tight circular relationship between microbiota composition, systemic immune response, and brain miswiring.

Future directions

The study of neuron-microglia interactions has led to substantial progress in our understanding of synaptic homeostasis and has major implications for knowledge of the aging process and neurodegenerative diseases (*1*–*3*, *5*). Additional key functions of microglia are emerging during early neurodevelopment, albeit the underlining mechanisms and their direct contribution to neurodevelopmental disorders remain to be fully explored. Building novel tools and models to dissect these functions and assess their roles in neurodevelopmental disorders remains a major goal for the field. In particular, combining in vivo studies in animal models, brain organoids, induced pluripotent stem cell–derived microglia (*1*), and CRISPR-CAS9 gene editing technology will be instrumental for our comprehension of microglia-neuron-glia cross-talk in physiological and pathological conditions. Furthermore, deciphering the roles of microglia versus other immune populations lying inside and outside the parenchyma during CNS development will be important to grasp the logic of neuroimmune interactions and the deleterious effects of inflammation. Finally, whereas most prior studies have focused on the prenatal or early postnatal stages of life, little is known about neuroimmune interactions during adolescence, which constitutes a major period of brain circuit rearrangements and a critical window of susceptibility to disorders, including schizophrenia. In light of the sexually dimorphic properties of immune cells and their potential to modulate brain wiring, it will be of particular interest to study this key developmental event. Not only do early neuroimmune interactions forge circuits in the context of the whole body, sexual identity, and dialogue with the environment, they also program immune cells that persist in the brain throughout life. Defining the cellular and molecular basis of these interactions will be essential to advance our current understanding of brain homeostasis and pathology.

REFERENCES AND NOTES

1. Q. Li, B. A. Barres, *Nat. Rev. Immunol.* **18**, 225–242 (2018).
2. M. W. Salter, B. Stevens, *Nat. Med.* **23**, 1018–1027 (2017).
3. S. A. Wolf, H. W. Boddeke, H. Kettenmann, *Annu. Rev. Physiol.* **79**, 619–643 (2017).
4. D. P. Schafer, B. Stevens, *Cold Spring Harb. Perspect. Biol.* **7**, a020545 (2015).
5. G. Hoeffel, F. Ginhoux, *Cell. Immunol.* **330**, 5–15 (2018).
6. T. Goldmann *et al.*, *Nat. Immunol.* **17**, 797–805 (2016).
7. N. C. Derecki *et al.*, *J. Exp. Med.* **207**, 1067–1080 (2010).
8. J. M. Garré, H. M. Silva, J. J. Lafaille, G. Yang, *Nat. Med.* **23**, 714–722 (2017).
9. A. J. Filiano *et al.*, *Nature* **535**, 425–429 (2016).
10. S. Tanabe, T. Yamashita, *Nat. Neurosci.* **21**, 506–516 (2018).
11. N. Swinnen *et al.*, *Glia* **61**, 150–163 (2013).
12. S. De *et al.*, *Development* **145**, dev152306 (2018).
13. O. Matcovitch-Natan *et al.*, *Science* **353**, aad8670 (2016).
14. M. S. Thion *et al.*, *Cell* **172**, 500–516.e16 (2018).
15. I. Amit, D. R. Winter, S. Jung, *Nat. Immunol.* **17**, 18–25 (2016).
16. I. R. Holtman, D. Skola, C. K. Glass, *J. Clin. Invest.* **127**, 3220–3229 (2017).
17. J. C. Cronk *et al.*, *J. Exp. Med.* **215**, 1627–1647 (2018).
18. F. C. Bennett *et al.*, *Neuron* **98**, 1170–1183.e8 (2018).
19. P. Squarzoni *et al.*, *Cell Reports* **8**, 1271–1279 (2014).
20. C. L. Cunningham, V. Martínez-Cerdeño, S. C. Noctor, *J. Neurosci.* **33**, 4216–4233 (2013).
21. A. Wlodarczyk *et al.*, *EMBO J.* **36**, 3292–3308 (2017).
22. N. Hagemeyer *et al.*, *Acta Neuropathol.* **134**, 441–458 (2017).
23. R. Hanamsagar *et al.*, *Glia* **65**, 1504–1520 (2017).
24. A. Villa *et al.*, *Cell Reports* **23**, 3501–3511 (2018).
25. Y. Belkaid, O. J. Harrison, *Immunity* **46**, 562–576 (2017).
26. D. Erny *et al.*, *Nat. Neurosci.* **18**, 965–977 (2015).
27. M. Datta *et al.*, *Immunity* **48**, 514–529.e6 (2018).
28. E. Mass *et al.*, *Nature* **549**, 389–393 (2017).
29. L. Weinhard *et al.*, *Nat. Commun.* **9**, 1228 (2018).
30. A. Miyamoto *et al.*, *Nat. Commun.* **7**, 12540 (2016).
31. F. Filipello *et al.*, *Immunity* **48**, 979–991.e8 (2018).
32. L. Pont-Lezica *et al.*, *Eur. J. Neurosci.* **39**, 1551–1557 (2014).
33. M. Ueno *et al.*, *Nat. Neurosci.* **16**, 543–551 (2013).
34. K. M. Lenz, B. M. Nugent, R. Haliyur, M. M. McCarthy, *J. Neurosci.* **33**, 2761–2772 (2013).
35. I. D. Vainchtein *et al.*, *Science* **359**, 1269–1273 (2018).
36. M. L. Estes, A. K. McAllister, *Science* **353**, 772–777 (2016).
37. S. E. Smith, J. Li, K. Garbett, K. Mirnics, P. H. Patterson, *J. Neurosci.* **27**, 10695–10702 (2007).
38. S. Kim *et al.*, *Nature* **549**, 528–532 (2017).
39. G. B. Choi *et al.*, *Science* **351**, 933–939 (2016).
40. M. D. Rudolph *et al.*, *Nat. Neurosci.* **21**, 765–772 (2018).

ACKNOWLEDGMENTS

We thank members of the Garel laboratory for stimulating discussions. N. McCarthy of Insight Editing London assisted with manuscript preparation. We apologize to those whose articles are not cited in this Review because of space constraints. **Funding:** S.G.'s laboratory is supported by INSERM, CNRS, Investissements d'Avenir (ANR-10-LABX-54 MEMO LIFE, ANR-11-IDEX-0001-02 PSL* Research University) and the ERC Consolidator NImO 616080. S.G. is part of the École des Neurosciences de Paris Île-de-France network. F.G. is an EMBO YIP awardee and is supported by Singapore Immunology Network (SIgN) core funding, as well as Singapore National Research Foundation Senior Investigatorship (NRFI) NRF2016NRF-NRFI001-02. **Competing interests:** None declared.

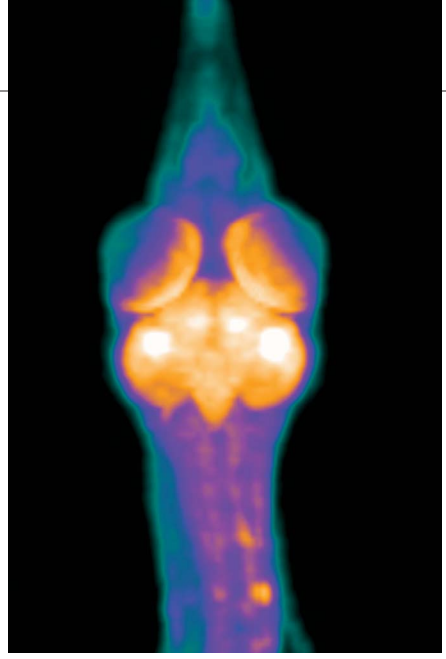
10.1126/science.aat0474

REVIEW

Homology, neocortex, and the evolution of developmental mechanisms

Steven D. Briscoe[1]* and Clifton W. Ragsdale[2,3]

The six-layered neocortex of the mammalian pallium has no clear homolog in birds or non-avian reptiles. Recent research indicates that although these extant amniotes possess a variety of divergent and nonhomologous pallial structures, they share a conserved set of neuronal cell types and circuitries. These findings suggest a principle of brain evolution: that natural selection preferentially preserves the integrity of information-processing pathways, whereas other levels of biological organization, such as the three-dimensional architectures of neuronal assemblies, are less constrained. We review the similarities of pallial neuronal cell types in amniotes, delineate candidate gene regulatory networks for their cellular identities, and propose a model of developmental evolution for the divergence of amniote pallial structures.



PET scan of crow's brain.

The six-layered neocortex is shared by all extant mammals but is absent from non-mammals, including birds and non-avian reptiles (*1*). The question of where it came from has been disputed for more than 150 years (*2*). Neuroanatomists in the early 20th century proposed that the neocortex represents a phylogenetic addition to a primitive reptilian brain and that it was responsible for those cognitive abilities thought to be unique to mammals (*2*). This scenario implies the rapid evolutionary genesis of the neocortex, seemingly from nothing, while leaving open the question of how nonmammals operate without a neocortex equivalent. However, de novo addition is not a mechanistic explanation for the origin of evolutionary novelty. To understand neocortex origins, we must search for homologous neocortical features shared by mammals and their relatives.

Biological structures are organized and evolve at multiple levels: cell types, neural circuits, molecules, and embryological territories (*3*–*5*). These levels can evolve independently of one another, and homology at one level does not require conservation at others. Although the neocortex is present only in mammals, some of its components may evolutionarily predate the common ancestor that mammals share with birds and reptiles. If true, the evolutionary origin of the neocortex could be explained, at least in part, by reorganization of old components at other hierarchical levels. Such a mechanism would resolve the apparent conflict between morphological novelty and the historical continuity required of evolving biological systems.

A current evolutionary developmental perspective focuses on cell types for considerations of homology (*6*, *7*). The claims of this view are that (i) cell types are elemental units of animal organization, (ii) cell identity is specified during development by evolutionarily stable collections of transcription factors, (iii) homologous cell types can exist in distantly related animals, and (iv) cell types can acquire species-specific states through evolutionary modifications that do not compromise their identities. Here, we review recent and classical findings indicating that core neuronal cell types of neocortical circuits are shared across mammals, reptiles, and birds. Lineage-specific developmental mechanisms may differentially arrange these conserved cell types into a wide variety of derived structural architectures, modify their morphologies and developmental origins, and add novel cell types.

Homologous cell types and circuitry in the amniote pallium

The neuroanatomical structures in the dorsal telencephalon, or pallium, of amniotes (mammals, birds, and non-avian reptiles) look very little alike (Fig. 1). Neuronal cell types of the mammalian neocortex are organized into layers. In contrast, the pallium of birds and other reptiles (the sauropsids) features a vast territory, the dorsal ventricular ridge (DVR), of neuronal cell body clusters called nuclei. In non-avian reptiles, a three-layered dorsal cortex lies above the

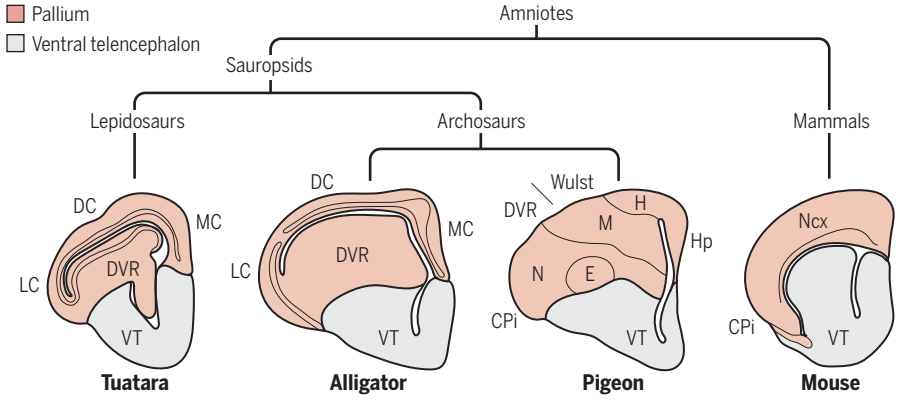


Fig. 1. The amniotes and their pallial anatomies. Amniotes include mammals and sauropsids (birds and non-avian reptiles). Tuataras, together with snakes and lizards (not shown), form the lepidosaurs. A second major group of sauropsids, the archosaurs, includes birds and crocodilians. Schematic tracings of telencephalon anatomies are shown as left-side coronal cross sections with medial to the right and dorsal at the top. All amniotes have a ventral telencephalon (VT, gray shading) and a dorsal telencephalon, or pallium (peach shading). The mammalian pallium includes the neocortex (Ncx), piriform cortex (CPi), hippocampus (not shown), and amygdala (not shown). The pallium in non-avian reptiles includes a dorsal ventricular ridge (DVR) and a cerebral cortex with medial (MC), dorsal (DC), and lateral (LC) divisions. The bird DVR contains the ventral part of the mesopallium (M), the nidopallium (N), the entopallium (E), and the arcopallium (not shown). Birds have a medial hippocampus (Hp), a lateral piriform cortex (CPi), and a dorsally located Wulst that includes the dorsal mesopallium and the hyperpallium (H). Drawings are not to scale.

[1]Max Planck Institute of Molecular Cell Biology and Genetics, Dresden, Germany. [2]Department of Neurobiology, University of Chicago, Chicago, IL, USA. [3]Department of Organismal Biology and Anatomy, University of Chicago, Chicago, IL, USA.
*Corresponding author. Email: briscoe@mpi-cbg.de

DVR. Birds likely descended from reptiles with a dorsal cortex; if so, this cortex became modified into a second nuclear complex known as the Wulst (*8*). These anatomical disparities have confounded efforts to establish homologies based on morphology alone—an approach that has generated incompatible interpretations (*2*, *4*, *5*). Comparisons made at the level of neuronal cell types have yielded a more coherent story.

The avian pallium does not contain cortical layers, but its excitatory neurons form input and output connections that are very similar to neocortical circuitry. In both the mammalian neocortex and the avian pallium, distinct populations of neurons either receive direct sensory input from the thalamus or convey motor output through projections to the brainstem. One hypothesis proposes that the connections are similar because the pallial input and output neurons were inherited from a common ancestor (Fig. 2, green input cells and red output cells) (*9*). Comparative molecular studies confirmed key aspects of this cell type homology hypothesis. The selective expression of conserved neocortical layer markers distinguishes pallial input and output neurons in a wide range of amniotes, rendering convergent evolution an unparsimonious explanation for the connectional similarities (*10*–*12*).

The lesson from the cell type homology hypothesis, and from its later molecular affirmation, is that natural selection preserves the core cell types and circuits of the amniote pallium, with less constraint on the morphology of pallial structures. Pallial input and output neurons in amniotes have maintained their defining connectional and molecular characteristics but assemble into taxon-specific architectures, forming stacked cortical layers in the neocortex, nuclei in the DVR, and, in the turtle cortex, spatially segregated cortical fields (*10*).

Intratelencephalic neurons and the evolution of higher cognitive abilities

Behavioral studies established the cognitive prowess of birds, some of which possess impressive skills in problem solving, memory, and tool use. The evolution of avian cognitive abilities is attributed to their massive, cell-dense pallium and to two large pallial territories, the mesopallium and the nidopallium (Fig. 1) (*13*). These territories serve as integrative centers and can be viewed as functionally analogous either to neocortical upper-layer neurons or to neocortical association areas such as the prefrontal cortex (*14*). The mesopallium and nidopallium receive axons from the pallial primary sensory input nuclei, interconnect with one another, and project in turn to the pallial motor output nuclei (*15*). They do not structurally resemble any district of the mammalian pallium, and their relationships even to non-avian reptile structures were until recently obscure (*1*).

Insight into the evolution of the mesopallium has been gained through molecular studies in multiple bird and non-avian reptile species. RNA sequencing in chickens and single-cell transcriptomics in turtles and lizards identified multiple transcription factors as diagnostic of conserved pallial cell populations (*16*, *17*). Six transcription factors characterize the avian mesopallium (*17*), which, as defined by Jarvis *et al.* (*11*), includes a dorsal division associated with the Wulst and a ventral division in the DVR. Gene expression experiments in alligators revealed a similar bipartite organization, with mesopallium-like cells in the dorsal cortex and a mesopallium in the DVR (*17*, *18*). Cell populations in turtle and lizard cortices also express several of the mesopallium transcription factors, indicating that mesopallium-like cell types are shared across sauropsids (*16*).

The mesopallium transcription factors are expressed in the mammalian neocortex, but not by any specific layer or area. Rather, they identify neocortical intratelencephalic (IT) neurons, found in both upper and deep neocortical layers in all neocortical areas (*17*, *19*). Moreover, two mesopallium transcription factors, *Satb2* and *Bcl11a*, are genetically required in mice to specify an IT connectional phenotype (*20*–*22*), which suggests that the mesopallium transcription factors may form a cell type gene regulatory network (Fig. 2, blue IT cells). Thus, the IT neurons, like those in the avian mesopallium, are a class of excitatory neurons linking the input with output populations and having axons that remain within the telencephalon (*15*, *19*).

Connectional and molecular data together support a model that pallial IT neurons, along with input and output neurons, were present in the amniote last common ancestor (LCA) where they formed a conserved circuit motif (Fig. 2). We propose that a limited ancestral stock of pallial IT neurons independently diversified to give rise to neurons in the reptile cortex, the DVR, the Wulst, and the IT neurons of the neocortex. The evolutionary expansion of pallial IT neurons may be a common neuroanatomical substrate for the evolution of higher cognitive abilities in mammals and birds (*17*). That is, brainy amniotes such as great apes and crows have disproportionately expanded pallial IT cell populations, and this extra integrative circuitry may have facilitated the evolution of complex cognition (*14*).

Specification and evolution of pallial neuronal cell types

Cell type identity is established during development by networks of transcriptional regulators that are relatively stable across phylogeny (*7*). Comparative studies suggest that pallial neuronal cell types may develop and evolve by the same general principles as other metazoan cell types and, in particular, they reveal a striking evolutionary correlation between transcription factor expression and cell type defined by connectivity (Fig. 2). For example, the paralogous genes *RORA* and *RORB*, both of which regulate the differentiation of neocortical input neurons (*23*, *24*), are expressed by pallial input cells of birds and reptiles (*10*, *11*, *16*, *18*). These two transcription factors, together with *SATB1* (*16*, *18*), may participate in a conserved gene regulatory network for the identity of pallial input cells, including their connections and physiological properties (Fig. 2, ancestral input cell). This possibility demands further study in bird and reptile embryos where in ovo gene manipulation experiments are straightforward (*17*, *25*–*27*).

Whereas the transcription factor profiles of input and IT cell types are similar across amniotes,

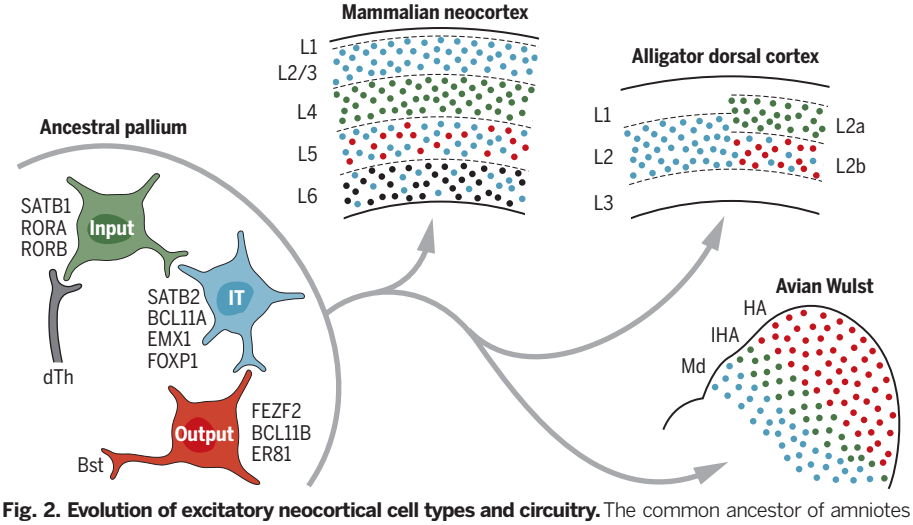


Fig. 2. Evolution of excitatory neocortical cell types and circuitry. The common ancestor of amniotes is hypothesized to have had input (green), output (red), and intratelencephalic (IT, blue) pallial neurons. Input neurons receive primary sensory information from the dorsal thalamus (dTh), whereas output neurons extend axons from the pallium to the brainstem (Bst). IT neurons serve as a relay between input and output neurons and additionally project to the striatum. These three principal pallial cell types were reorganized into the divergent architectures of the neocortex, dorsal cortex, Wulst, and DVR (not shown). Candidate transcriptional regulators of cell identity are listed beside the ancestral cell types (*10*, *16*–*18*). The evolutionary origin of neocortical corticothalamic neurons (black cells in L6) is not known (*18*). The pallial inhibitory interneurons, derived from the ventral telencephalon, are not indicated [but see (*16*)]. HA, hyperpallium apicale; IHA, interstitial nucleus of the hyperpallium apicale; L, layer; Md, dorsal mesopallium.

KELLIE HOLOSKI/*SCIENCE* BASED ON (*18*)

other features of pallial neurons are clearly divergent. Many excitatory neurons of the mammalian neocortex are pyramidal in morphology, with a layer-spanning apical dendrite, whereas most avian pallial neurons are stellate. The developmental mechanisms regulating taxon-specific neuronal morphologies are not known, nor are the consequences of these divergent morphologies for information processing (*10*, *28*). Substantial group differences exist also at the molecular level, as some sauropsid pallial neurons express transcription factors in combinations not observed in the neocortex (*16*). Moreover, between sauropsids and mammals, there remain pallial cell types for which molecular and connectional data do not suggest clear homologies (Figs. 2 and 3) (*17*). Reconfiguration of transcription factor expression patterns, protein interactions, and transcriptional targets may be a means by which neuronal cell types evolve and diversify (*6*, *7*, *16*).

A model for the developmental evolution of amniote pallial structures

There is no fossil record for the internal organization of vertebrate brains. To propose mechanisms of anatomical transformation, we can only construct models of the ancestral anatomy by comparing extant forms. Evidence reviewed here suggests that the amniote LCA had input, output, and IT cells (Fig. 3, green, red, and blue cells, respectively) (*17*). Moreover, recent work establishes that hippocampal dentate gyrus– and CA field–like neurons are also ancestral (Fig. 3, orange cells) (*16*, *18*). Our model of anatomical transformations incorporates these findings and makes two assumptions: (i) The pallial anatomy of the amniote LCA was simpler than in extant forms, and (ii) the spatial distribution of cell types resembled that of extant sauropsids more than it resembled the neocortical organization.

The amniote LCA

Pallial morphology in amphibians, the closest living outgroup to amniotes, is simple (*29*). Neurons are typically distributed next to the ventricle with few structural specializations. There is no DVR. We propose that the amniote LCA possessed an architecturally simple pallium, similar to that of extant amphibians, but one with extensive primary sensory inputs and local integrative circuitry mediated by IT neurons (Fig. 3, Amniote LCA).

All extant amniotes, including mammals, receive somatosensory and visual information in the dorsal pallium and visual and auditory information in the lateral pallium (*30*, *31*). The model amniote LCA therefore has two populations of input cells: a dorsal group that gave rise to input cells in the sauropsid dorsal cortex and Wulst, and a lateral group that gave rise to input populations in the DVR. We speculate that a population of IT cells was intercalated between the dorsal and lateral input groups, much as an IT cell population separates dorsal and lateral input cell groups in birds and alligators (*17*, *18*). Output cells have been described in the avian Wulst, the avian DVR, and the reptile dorsal cortex, so we suggest that output cells were intermingled with the input cells in the amniote LCA.

The sauropsid LCA

The first developmental transformation would be the acquisition of the DVR in the sauropsid lineage. At this transition, the targets of the dorsal and lateral ascending sensory pathways became individualized into morphologically distinct territories: a dorsal cortex-like structure and a lateral DVR, each with different modes of development. This dorsal and lateral dichotomy differs from the mammalian pallium, where dorsal and lateral sensory pathways both target a morphologically uniform structure, the neocortex.

The sauropsid LCA may have resembled the modern tuatara, which, unlike any other studied reptile, has a thin cortex-like structure extending from the dorsal cortex through the DVR (Fig. 3, Sauropsid LCA) (*32*). The DVR may have originated through a tangential expansion of ventricular zone progenitors analogous to neocortical progenitor expansion in mammals, perhaps accompanied by the origin of a vast, associative nidopallium (Fig. 3, brown arrow). This expansion may have resulted in more neurons and more neural processing power for the laterally targeted sensory channels.

The archosaur LCA

We suggest that the archosaur LCA closely resembled modern alligators by pallial morphology (Fig. 3, Archosaur LCA). The archosaur LCA differs from the sauropsid LCA by possessing many more neurons that pack the DVR from the ventricle to the brain surface, but the spatial relationships of the principal neuronal classes are broadly similar. Evolutionary mechanisms of cell number increase are well understood from comparative studies in mammals (*26*, *33*, *34*). Archosaurs might have similarly increased neuronal production by prolonging symmetric stem cell divisions or by further elaborating neuronal stem and progenitor cells into distinct ventricular and subventricular zones (*27*, *35*).

The bird LCA

The transition to the LCA of birds involved another burst of increased neurogenesis (*13*, *26*) and a major developmental transformation: the modification of the laminated dorsal cortex into the nuclear Wulst (Fig. 3, Bird LCA). The dorsal cortex and Wulst contain the same core set of cell types but are organized into very different structures (*17*, *18*). For example, alligator dorsal cortex input cells form a superficial cortical layer, but in birds the input population is sandwiched between IT cells and output cells. The morphogenetic mechanisms generating these divergent architectures are unknown.

The mammal LCA

Efforts to describe neocortex origins have invoked the transformation of a reptile-like pallium into the neocortex (*9*). However, mammals are not thought to have evolved from animals resembling modern sauropsids (*29*). The transformations leading to the mammalian neocortex need to be

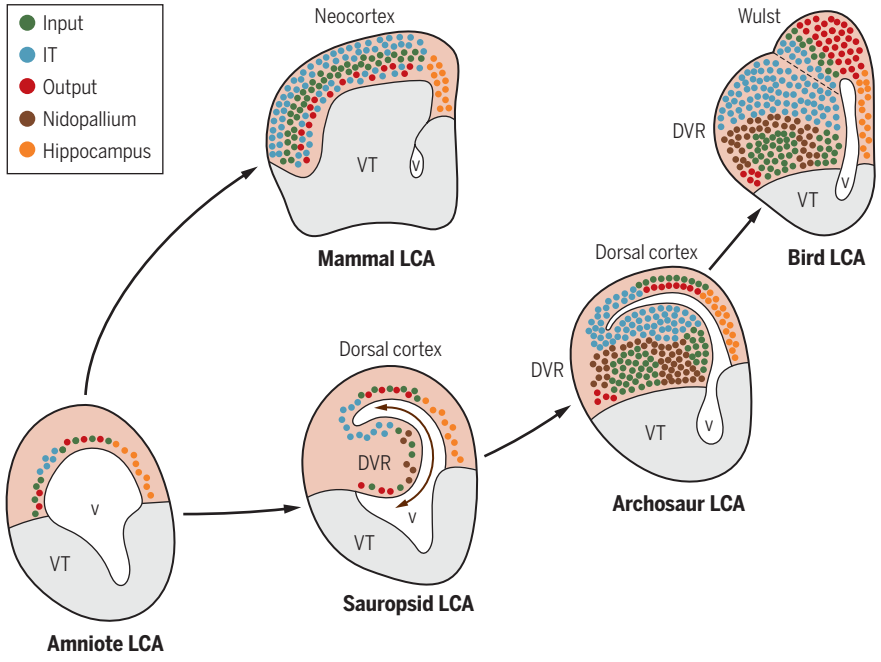


Fig. 3. Anatomical transformations during the evolutionary diversification of amniote pallial structures. The hypothetical amniote last common ancestor had an architecturally simple pallium, which contained input (green), output (red), IT (blue), and hippocampal (orange) cell types. The ancestral pallium underwent independent evolutionary transformations to give rise to the mammalian neocortex and the sauropsid DVR, dorsal cortex, and Wulst. The brown arrows denote an expansion of the lateral pallium with the origin of the DVR in early sauropsids. This expansion may have accompanied the origin of non-input nidopallium cells (brown). DVR input nuclei are embedded within the nidopallium. v, ventricle; VT, ventral telencephalon.

KELLIE HOLOSKI/*SCIENCE* BASED ON S. D. BRISCOE AND C. W. RAGSDALE

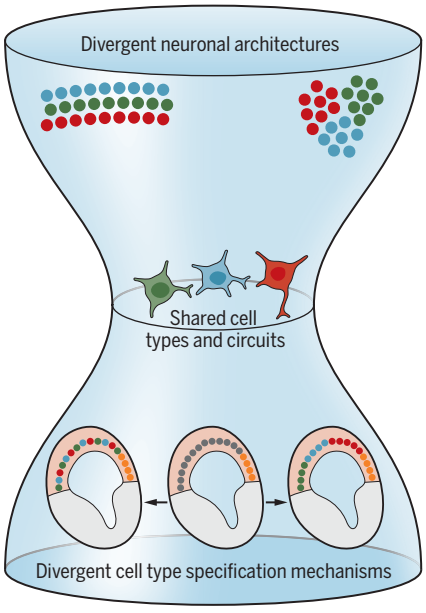


Fig. 4. Pallium evolution resembles an hourglass. A field of embryonic neuronal progenitors (gray cells, bottom center) can be developmentally patterned such that pallial cell types arise from all parts of the neuroepithelium, as in mammals (bottom left), or from restricted parts of the neuroepithelium as in sauropsids (bottom right). Pallial cell types are specified to form conserved connections (middle of hourglass), possibly through conserved transcriptional regulators (see Fig. 2), but migrate to form clade-specific cortical or nuclear architectures (top).

considered in the context of the amniote LCA and not the sauropsid lineage. The pallium in the amniote LCA may not have had either a DVR or a cortex, but would later give rise to both architectures (*36*).

Cell type distribution across the tangential extent of the neocortex is highly uniform, with major cell classes distributed in thin layers extending through the neocortical plate (Fig. 3, Mammal LCA) (*19*). In contrast, neuronal cell types in the DVR and dorsal cortex are found in spatially segregated territories (*10*, *16*–*18*). These anatomical differences suggest that the mechanisms of cell type specification may differ between mammals and sauropsids. Sauropsid pallial cell types may be specified by spatial information and arise from fate-restricted ventricular zone territories, as demonstrated for the mesopallium and the nidopallium (*17*, *25*). In mammals, a temporal inside-out progression generates the neocortical excitatory neurons (*37*). The evolutionary origin of the neocortex likely entailed reorganization of the pallial cell type developmental lineages and was not accomplished through elaboration of the reptile cortex.

Conclusions

Natural selection acts on behavior, the functional output of neural circuitry. In this light, the emerging picture from comparative neuroanatomical studies makes sense: The core circuits of the amniote pallium are conserved, but pallial architectures and the spatiotemporal regulation of cell type specification are not (Fig. 4). These latter levels of biological organization can shift over evolutionary time, so long as sensory information gets to the pallium, gets integrated, and elicits an appropriate output. Some transcription factor molecules, such as the hypothesized input cell determinants, exhibit a conserved correlation with defined neuronal cell types across species. We suggest that these factors are evolutionarily maintained together at the cell type level because they collaborate to control the circuits essential for behavior.

What is the avian homolog of the six-layered neocortex? Put simply, there isn't one because the neocortex, as a structure, is an evolutionary novelty of mammals. This statement does not imply any paradox. Every morphological character must have a beginning and lack a homolog in other lineages, because the alternative is that the character has always existed. For this reason, it is essential that we refer to levels of homology and recognize that nonhomologous structures can share homologous features (*3*, *4*). The mammalian neocortex and avian pallium share such features: their cell types and circuitries.

An additional level of homology, that of embryological territories, is often invoked to describe the relationships of amniote pallial structures. One hypothesis states that structures are homologous on the basis of their regional developmental origins in the embryo, and that the sauropsid DVR is homologous with the mammalian claustrum and amygdala rather than with the neocortex (*5*). Several recent studies have embraced this alternative view of pallial evolution (*16*, *25*, *26*). We emphasize that developmental origin is one of many features that can be weighed to infer homology of two similar characters. Homologous characters can be generated by nonhomologous mechanisms and from nonhomologous progenitors (*3*, *4*, *7*). Nonetheless, concordance between two or more levels, such as between connections and gene expression, can strengthen a claim of homology.

The idea that pallial cell types are rearranged across amniotes has met with some resistance (*38*). However, there are examples outside of amniotes in which conserved neuronal cell types were driven to adopt different architectures by behavioral and ecological pressures. Goldfish provide one example. In their hindbrains, gustatory secondary sensory neurons and motor neurons of the oral musculature form a multilayered vagal lobe, rather than separate nuclei as in amniotes (*39*, *40*). This highly derived architecture, which places sensory input and motor output neurons in a columnar registration, likely mediates the specialized ability of these scavengers to sort food from gravel in their mouths (*39*).

By analogy, clade-specific behavioral adaptations may be responsible for the striking diversification of amniote pallial anatomies. Even among reptiles, extensive pallial diversity is evident. In tuataras, the dorsal cortex transitions into a cortex-like DVR (*32*). The cerebral cortex in some snakes and lizards is splintered into short, overlapping segments (*18*). Alligator and turtle cortices possess some degree of cell type layering, but the relative arrangement appears to be reversed (*10*, *16*, *18*). Finally, there is the variability across DVR organizations (*1*, *16*, *18*). Relative to other brain structures, pallial architecture is unusually plastic during evolution. Reptiles provide a venue to investigate the selective pressures and developmental mechanisms that reorganize pallial cell types and circuitries in evolution.

REFERENCES AND NOTES

1. P. S. Ulinski, *Dorsal Ventricular Ridge: A Treatise on Forebrain Organization in Reptiles and Birds* (Wiley, 1983).
2. G. F. Striedter, *Principles of Brain Evolution* (Sinauer Associates, 2005).
3. B. K. Hall, *Homology: The Hierarchical Basis of Comparative Biology* (Academic Press, 1994).
4. J. Dugas-Ford, C. W. Ragsdale, *Annu. Rev. Neurosci.* **38**, 351–368 (2015).
5. L Puelles *et al.*, in *Evolution of Nervous Systems*, G. Striedter, Ed. (Academic Press, ed. 2, 2017), vol. 1, pp. 519–555.
6. D. Arendt *et al.*, *Nat. Rev. Genet.* **17**, 744–757 (2016).
7. G. P. Wagner, *Homology, Genes, and Evolutionary Innovation* (Princeton Univ. Press, 2014).
8. A. Reiner *et al.*, *J. Comp. Neurol.* **473**, 377–414 (2004).
9. H. J. Karten, *Ann. N.Y. Acad. Sci.* **167**, 164–179 (1969).
10. J. Dugas-Ford, J. J. Rowell, C. W. Ragsdale, *Proc. Natl. Acad. Sci. U.S.A.* **109**, 16974–16979 (2012).
11. E. D. Jarvis *et al.*, *J. Comp. Neurol.* **521**, 3614–3665 (2013).
12. J. J. Rowell, A. K. Mallik, J. Dugas-Ford, C. W. Ragsdale, *J. Comp. Neurol.* **518**, 3272–3289 (2010).
13. S. Olkowicz *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **113**, 7255–7260 (2016).
14. J. Mehlhorn, G. R. Hunt, R. D. Gray, G. Rehkämper, O. Güntürkün, *Brain Behav. Evol.* **75**, 63–70 (2010).
15. Y. Atoji, J. M. Wild, *J. Comp. Neurol.* **520**, 717–741 (2012).
16. M. A. Tosches *et al.*, *Science* **360**, 881–888 (2018).
17. S. D. Briscoe, C. B. Albertin, J. J. Rowell, C. W. Ragsdale, *Curr. Biol.* **28**, 686–696.e6 (2018).
18. S. D. Briscoe, C. W. Ragsdale, *J. Comp. Neurol.* **526**, 1613–1646 (2018).
19. K. D. Harris, G. M. Shepherd, *Nat. Neurosci.* **18**, 170–181 (2015).
20. M. B. Woodworth *et al.*, *Cell Rep.* **15**, 999–1012 (2016).
21. E. A. Alcamo *et al.*, *Neuron* **57**, 364–377 (2008).
22. O. Britanova *et al.*, *Neuron* **57**, 378–392 (2008).
23. T. Vitalis *et al.*, *Cereb. Cortex* 10.1093/cercor/bhx262 (2017).
24. D. Jabaudon, S. J. Shnider, D. J. Tischfield, M. J. Galazo, J. D. Macklis, *Cereb. Cortex* **22**, 996–1006 (2012).
25. F. García-Moreno *et al.*, *Cell Rep.* **22**, 96–109 (2018).
26. A. Cárdenas *et al.*, *Cell* **174**, 590–606.e21 (2018).
27. T. Nomura *et al.*, *Development* **143**, 66–74 (2016).
28. L. Maler, *Curr. Biol.* **28**, R213–R215 (2018).
29. R. G. Northcutt, *Annu. Rev. Neurosci.* **4**, 301–350 (1981).
30. L. Medina, A. Reiner, *Trends Neurosci.* **23**, 1–12 (2000).
31. L. Krubitzer, *Neuron* **56**, 201–208 (2007).
32. A. Reiner, R. G. Northcutt, *Brain Behav. Evol.* **55**, 26–36 (2000).
33. A. A. Pollen *et al.*, *Cell* **163**, 55–67 (2015).
34. M. Florio, W. B. Huttner, *Development* **141**, 2182–2194 (2014).
35. V. Martínez-Cerdeño *et al.*, *J. Comp. Neurol.* **524**, 433–447 (2016).
36. A. B. Butler, *Brain Res. Brain Res. Rev.* **19**, 66–101 (1994).
37. P. Gao *et al.*, *Cell* **159**, 775–788 (2014).
38. J. F. Montiel, F. Aboitiz, *Brain Behav. Evol.* **91**, 59–64 (2018).
39. T. E. Finger, *J. Comp. Physiol. A* **194**, 135–143 (2008).
40. E. Coppola, F. D'autréaux, M. Nomaksteinsky, J. F. Brunet, *J. Comp. Neurol.* **520**, 3633–3649 (2012).

ACKNOWLEDGMENTS

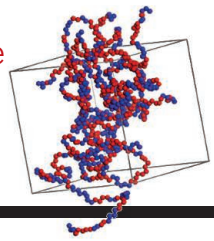
We thank M. Coates for comments. **Funding:** The Molecular and Cellular Biology Training Program (T32 GM007183) and the Developmental Biology Training Program (HD055164). **Competing interests:** None declared.

10.1126/science.aau3711

RESEARCH

Self-healing in simple vinyl polymers

Urban et al., *p. 220*



IN *SCIENCE* JOURNALS

Edited by **Stella Hurtley**

PLANT SCIENCE

Keeping the doors open for symbiosis

Nitrogen fixation by legumes results from a symbiotic partnership between plant and microbes. These together elaborate nodules on the plant roots that house the bacteria. Tsikou *et al.* identified a microRNA made in the aboveground shoots of *Lotus japonicus* that translocates to the plant's roots. In the roots, the microRNA posttranscriptionally regulates a key suppressor of symbiosis, thus keeping the uninfected root susceptible to productive infection by symbiotic bacteria. —PJH

Science, this issue p. 233

Legumes like this *Lotus japonicus* send microRNAs from shoot to root to promote root nodule synthesis.

MATERIALS SCIENCE

Modeling collapse with soaked cereal

Naturally brittle and porous media, like ice shelves and rockfill dams, are collapse-prone when they encounter chemically active fluids, or pressure. What happens when these materials are subject to both forces at once is not well understood. Einav and Guillard analyzed a simpler version of a rockfill dam: a vertical cylinder filled with puffed rice. After applying pressure and then injecting liquid into the bottom of the cylinder, they observed "ricequakes." These abrupt collapses of wetted puffed rice began shortly after liquid injection and were marked by continuous clicking sounds. A model based on these observations could be applied to analyze the effects of geological pressures over long time scales in crustal rocks and ice sheets. —PJB

Sci. Adv. 10.1126/sciadv.aat6961 (2018).

SUPERNOVAE

Explosive origin of a binary neutron star

Some types of core-collapse supernovae are known to produce a neutron star (NS). A binary NS merger was recently detected from its gravitational wave emission, but it is unclear how such a tight binary system can be formed. De *et al*. discovered a core-collapse supernova with unusual properties, including the removal of the outer layers of the star before the explosion. They interpret this as the second supernova in an interacting binary system that already contains one NS. Because the explosion probably produced a second NS (rather than a black hole) in a tight orbit, it could be an example of how binary NS systems form. —KTS

Science, this issue p. 201

MOLECULAR BIOLOGY

Cas12 inhibitors join the anti-CRISPR family

Bacteria and their phages continually coevolve in a molecular arms race. For example, phages use anti-CRISPR proteins to inhibit the bacterial type I and II CRISPR systems (see the Perspective by Koonin and Makarova). Watters *et al*. and Marino *et al*. used bioinformatic and experimental approaches to identify inhibitors of type V CRISPR-Cas12a. Cas12a has been successfully engineered for gene editing and nucleic acid detection. Some of the anti-Cas12a proteins identified in these studies had broad-spectrum inhibitory effects on Cas12a orthologs and could block Cas12a-mediated genome editing in human cells. —SYM

Science, this issue p. 236, p. 240; see also p. 156

MULTIPLE SCLEROSIS

Move over myelin

Although it is well established that autoreactive lymphocytes induce demyelination in multiple sclerosis, the exact

Published by AAAS

antigenic targets that initiate disease are undefined. Planas *et al*. studied CD4$^+$ T cells from the cerebrospinal fluid of patients with multiple sclerosis. One CD4$^+$ T cell clone was reactive to the human enzyme guanosine diphosphate (GDP)–L-fucose synthase; T cells from other patients were then identified, as well as myelin-reactive cells. Intriguingly, some of the GDP-L-fucose synthase–reactive cells could also be stimulated by a bacterial version of the enzyme. These results identify an autoantigen and suggest that one possible trigger of disease could be cross-reactivity to microbiota-derived peptides. —LP

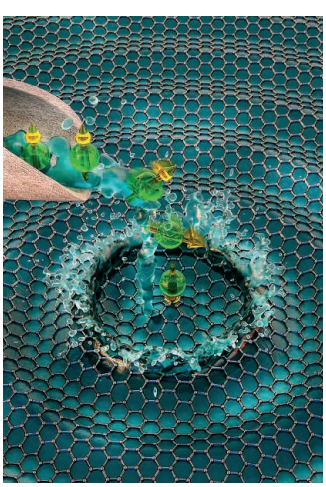
Sci. Transl. Med. **10**, eaat4301 (2018).

MAGNETISM

Magnons propagating in graphene

At sufficiently low temperatures, a two-dimensional electron system placed in an external magnetic field can exhibit the so-called quantum Hall effect. In this regime, a variety of magnetic phases may occur, depending on the electron density and other factors. Wei *et al*. studied the properties of these exotic magnetic phases in graphene. They generated magnons—the excitations of an ordered magnetic system—that were then absorbed by the sample, leaving a mark on its electrical conductance. The magnons were able to propagate across long distances through various magnetic phases in the bulk graphene. —JS

Science, this issue p. 229



Artist's view of magnon absorption by graphene

CREDITS: (LEFT) SECOND BAY STUDIOS; (RIGHT) EYE OF SCIENCE/SCIENCE SOURCE

VIROLOGY

Pathologizing *Staphylococcus*, fast

Bacteriophages are the main vehicle for gene swapping in bacteria, notoriously of pathogenicity islands and antibiotic resistance genes. Chen *et al*. noticed that the *Staphylococcus aureus* prophages do not excise from their host's genome until very late in their life cycles (see the Perspective by Davidson). Thus, the phage DNA is amplified while embedded in the bacterial chromosome. The resulting concatemers are processively packed into virus capsules while still integrated in the host chromosome. Each virion is only set loose when the capsule has reached physical capacity—a process called "headful" packaging. In situ amplification maximizes viral replication, and the headful mechanism means adjacent bacterial-host DNA also gets grabbed to fill the capsule. This process ensures that host genes are transmitted along with the phage. —CA

Science, this issue p. 207; see also p. 152

NANOMATERIALS

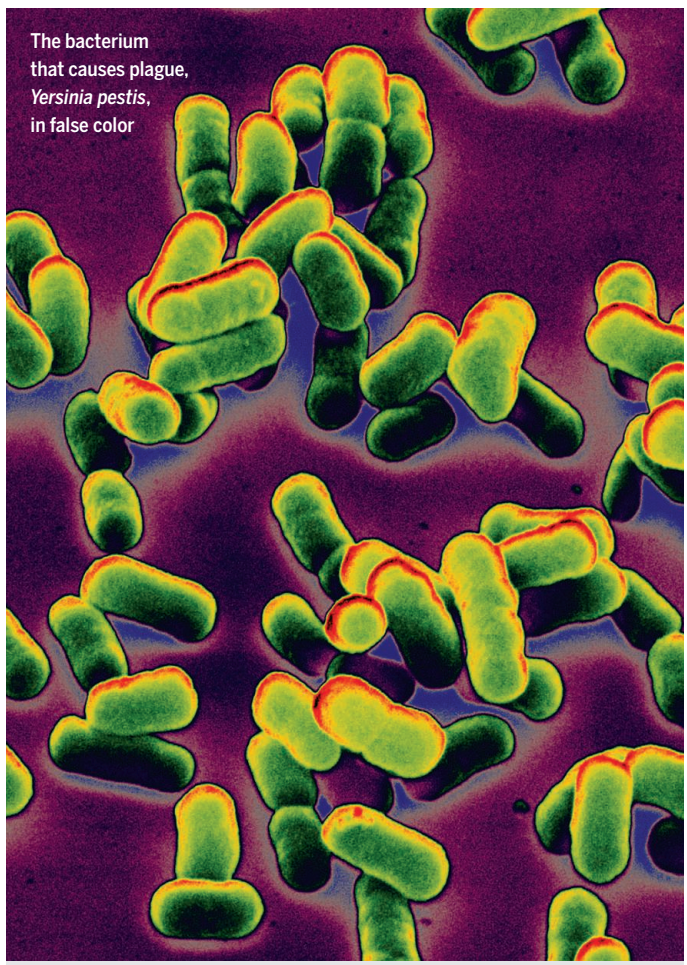
The twisted carbon nanotube story

Despite progress in growing single-walled carbon nanotubes of specific size and chirality, the factors that control their growth are still not fully known. Magnin *et al*. developed a thermodynamic model for the growth of single-walled carbon nanotubes. The model explains the origin of nanotube chirality in terms of the configurational entropy of the nanotube edge. The model should be useful in helping to guide nanotube growth parameters to enhance selectivity. —MSL

Science, this issue p. 212

IN OTHER JOURNALS

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



The bacterium that causes plague, *Yersinia pestis*, in false color

MICROBIOLOGY

Plague, one lymph node at a time

Swollen lymph nodes, or buboes, are the hallmark of plague, which is caused by the pathogen *Yersinia pestis*. Buboes result from a massive influx of immune cells into draining lymph nodes (DLNs). Recently, *Y. pestis* was shown to disseminate by carriage within immune cells migrating from one DLN to the next. Arifuzzaman *et al*. investigated how *Y. pestis* exploits the features of buboes to promote pathogenesis. Infiltration of infected monocytic cells into tightly packed buboes coincided with cytolysis triggered by *Yersinia* outer protein J, resulting in the release of intracellular bacteria and extensive infection of neighboring cells. The dying monocytic cells released sphingosine 1-phosphate (S1P), which attracted yet more cells, and up-regulated expression of the S1P receptor promoted the exit of newly infected monocytic cells from buboes. Preventing necrotic cell death protected mice from otherwise lethal infection. Thus, *Y. pestis* commandeers cell-death and immune-cell trafficking programs to convert the host's DLNs into specialized hubs for dissemination. —SMH

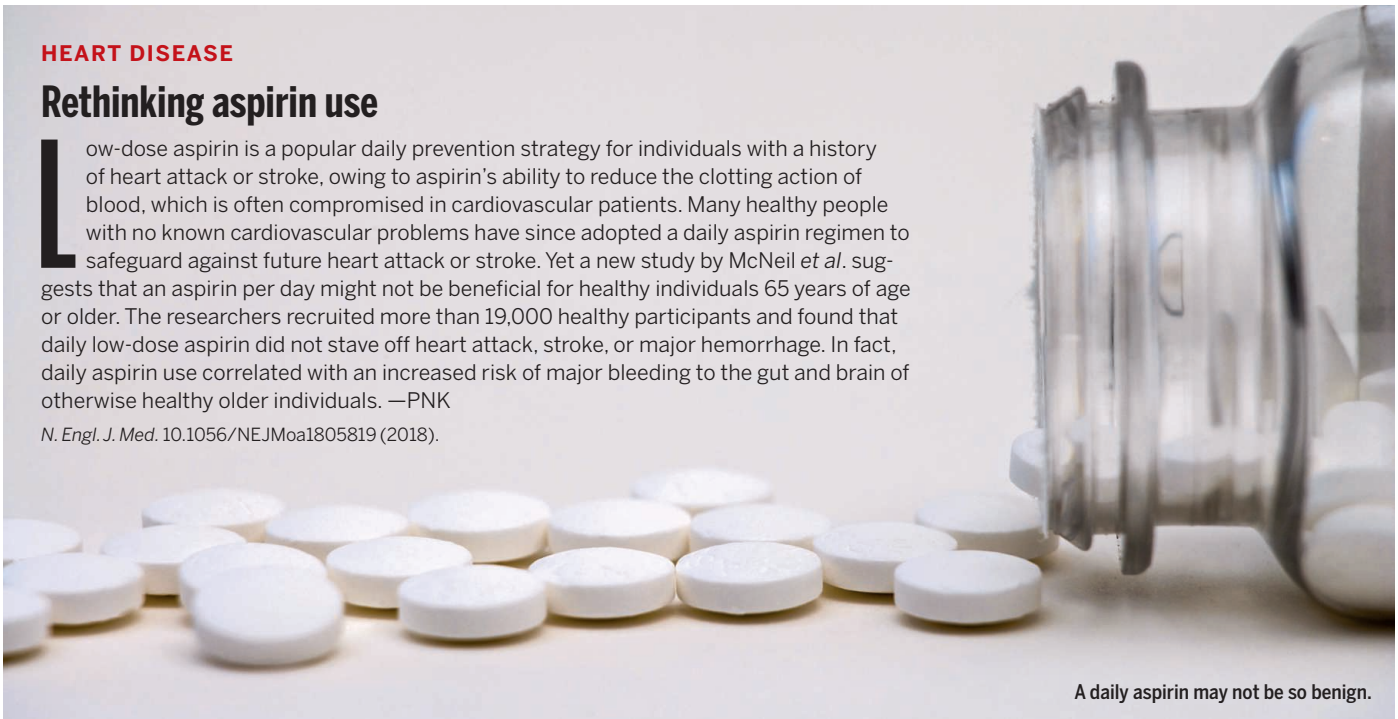
JCI Insight **3**, e122188 (2018).

HEART DISEASE

Rethinking aspirin use

Low-dose aspirin is a popular daily prevention strategy for individuals with a history of heart attack or stroke, owing to aspirin's ability to reduce the clotting action of blood, which is often compromised in cardiovascular patients. Many healthy people with no known cardiovascular problems have since adopted a daily aspirin regimen to safeguard against future heart attack or stroke. Yet a new study by McNeil *et al*. suggests that an aspirin per day might not be beneficial for healthy individuals 65 years of age or older. The researchers recruited more than 19,000 healthy participants and found that daily low-dose aspirin did not stave off heart attack, stroke, or major hemorrhage. In fact, daily aspirin use correlated with an increased risk of major bleeding to the gut and brain of otherwise healthy older individuals. —PNK

N. Engl. J. Med. 10.1056/NEJMoa1805819 (2018).



A daily aspirin may not be so benign.

SPINTRONICS

Making graphene useful for spintronics

Thanks to graphene's weak spin-orbit coupling, spin currents flow through it unimpeded. However, this also means that the spin currents are hard to manipulate—a drawback for using graphene in spintronics. Two groups now show that this needn't be the case. Leutenantsmeyer *et al*. and Xu *et al*. studied spin transport in heterostructures of bilayer graphene with hexagonal boron nitride. In the presence of an electric field, the spin lifetimes in the directions parallel to the heterostructural layers and perpendicular to them were markedly different. The long and anisotropic lifetimes may lead to useful spintronics applications. —JS

Phys. Rev. Lett. **121**, 127702, 127703 (2018).

CARBON CYCLE

Deciphering atmospheric CO_2 change

Many different processes and carbon sources can affect the concentration of atmospheric carbon dioxide (CO_2), so to understand past, present, and likely future variability of CO_2, those mechanisms and reservoirs must be identified. Bauska *et al*. present a record of the stable carbon isotopic composition of CO_2 over the interval between 50,000 and 35,000 years ago, which reveals that the primary source of atmospheric carbon during millennial-duration events probably was organic carbon residing in the deep ocean, whereas more abrupt events occurring over centennial time scales were associated both with hydrological change in the tropics and rapid increases in Northern Hemisphere temperature. —HJS

Geophys. Res. Lett. **45**, 7731 (2018).

PROTEIN DESIGN

Learning from diminutive ligand design

One strategy for understanding the origin of life is proposing simple replacements for the complex biomolecules that have developed through billions of years of evolution. Ferredoxins are small proteins that contain simple, cubic clusters of iron and sulfur atoms and act as mobile electron carriers in cells. Kim *et al*. designed a 12-residue peptide with alternating D and L amino acids that can replicate the placement of cysteine ligands found in many natural ferredoxins. After reconstitution with iron and sulfur, the peptides bound a single iron-sulfur cluster. The resulting minimal, artificial ferredoxin exhibited a redox potential compatible with some biological processes. —MAF

J. Am. Chem. Soc. **140**, 11210 (2018).

MICROBIOTA

The long reach of the gut

How does the gut microbiota shape the composition and function of distal host organs, despite being segregated in the gut? Uchimura *et al*. used stable isotope tracing to show that microbial metabolites penetrate host tissues and fluids to influence host immunological and metabolic signaling networks. However, metabolite impact is modulated by a high rate of urinary excretion of microbial products. Furthermore, secretory immunoglobulin A antibodies limit bacterial dwell times in the small intestine, which also ameliorates host exposure to microbial metabolites. The joint effect contributes to resolving gut function as both nutrient gateway and barrier. —STS

Immunity **49**, 545 (2018).

NEUROSCIENCE

Memory recirculation and integration

Hippocampal pattern separation minimizes interference between experiences and is critical to episodic memory (a person's collection of memories of an event). However, the hippocampus is also critical for the integration of information across episodes. These two roles are apparently in conflict. Using high-resolution brain scanning, Koster *et al*. investigated information flow through the layers of the entorhinal cortex that are the inputs and outputs of the hippocampus. Rather than the output of the system being the end product of hippocampal processing, it is recirculated as a new input. This result indicates big-loop recurrence, predicted by computational theories, which states that episodic memory and the integration of information across experiences need not conflict. —PRS

Neuron **99**, 1342 (2018).

ALSO IN *SCIENCE* JOURNALS

Edited by **Stella Hurtley**

CANCER BIOMARKERS

Mining immunotherapy clinical trials

Clinical trial data can provide a wealth of information about how drugs work. Yet such information often belongs to pharmaceutical companies and is rarely accessible to the scientific community at large. Cristescu *et al*. provide exploratory analysis of a cancer genomics dataset, collected from four separate clinical trials of Merck's PD-1 immunotherapy drug, pembrolizumab. This informative public resource examines more than 300 patient samples representing 22 different tumor types. Two widely used signatures that currently predict immunotherapy response are tumor mutational burden and a "hot" T cell–inflamed microenvironment. The study analyzed these two proposed biomarkers in combination to see what predictive clinical utility they may hold. —PNK

Science, this issue p. 197

PROTEIN HOMEOSTASIS

A way to prevent deadly interaction

Many metazoan proteins form oligomers, which is often mediated by modular domains such as BTB domains. Mena *et al*. now describe a quality control pathway they term dimerization quality control (DQC) (see the Perspective by Herhaus and Dikic). DQC monitors and prevents aberrant dimerization of BTB domain–containing proteins. The system relies on FBXL17, an adaptor protein that recruits an E3 ligase that specifically ubiquitylates nonfunctional BTB heterodimers, triggering their degradation. FBXL17 accesses a degradation signal at the BTB dimer interface in nonphysiological, nonfunctional complexes. The loss of DQC from *Xenopus laevis* embryos leads to lethal neurodevelopmental defects. —SYM

Science, this issue p. 198; see also p. 151

STRUCTURAL BIOLOGY

From DNA unwrapping to histone exchange

The yeast SWR1 complex, a member of the INO80 family of nucleosome remodelers, exchanges the H2A-H2B histone dimer for the Htz1 variant–containing dimer. Unlike all other remodelers, SWR1 does not translocate the nucleosome. Willhoft *et al*. applied structural and single-molecule analyses to show that the interaction between SWR1 and the nucleosome destabilizes the DNA wrapped around the histone core. This SWR1-catalyzed partial unwrapping of the DNA was regulated by adenosine triphosphate (ATP) binding but did not require ATP hydrolysis. —SYM

Science, this issue p. 199

NEURODEVELOPMENT

Thyroid hormone in color vision development

Cone photoreceptors in the eye enable color vision, responding to different wavelengths of light according to what opsin pigments they express. Eldred *et al*. studied organoids that recapitulate the development of the human retina and found that differentiation of cone cells into their tuned subtypes was regulated by thyroid hormone. Cones expressing short-wavelength (S) opsin developed first, and cones expressing long- and medium-wavelength (L/M) opsin developed later. The switch toward development of L/M cones depended on thyroid hormone signaling through the nuclear thyroid hormone receptor. —PJH

Science, this issue p. 200

ORGANIC CHEMISTRY

An acid inaccessible to aldol products

The aldol reaction is a venerable and widely applicable method for making carbon-carbon bonds. Ironically, it is most challenged by the simplest substrates. The trouble is that the product looks a lot like one of the reactants, and so it can latch onto the coupling partner instead. Schreyer *et al*. report that a bulky phosphorus-based acid catalyst alleviates this problem. The acidic site is buried in a pocket that is too small to activate the product for further reaction. The chiral geometry of the catalyst also induces high enantioselectivity. —JSY

Science, this issue p. 216

POLYMERS

Simple routes to self-healing

Biology provides many routes for self-healing or repair, but this trait is hard to endow into engineering materials. Although self-repair has been demonstrated for some polymers, it usually required specialized monomers. Urban *et al*. demonstrate that for a very narrow range of compositions, simple vinyl polymers based on methyl methacrylate and *n*-butyl acrylate show repeatable self-healing properties (see the Perspective by Sumerlin). A key characteristic of this system is that it relies on van der Waals interactions rather than the reformation of hydrogen or covalent bonds for repair. —MSL

Science, this issue p. 220; see also p. 150

ORGANIC CHEMISTRY

Iodine smooths the way to ketyl radicals

Chemists typically transform carbonyl compounds through polar two-electron reactions. It is also possible to pursue radical coupling strategies by adding just one electron to form a ketyl group. However, the strong reductant supplying that electron often limits the reaction's versatility. Wang *et al*. report a mild means of forming ketyls by first adding acetyl iodides across the C=O bond (see the Perspective by Blackburn and Roizen). A photoactivated manganese catalyst then temporarily pulls the iodine away, leaving a ketyl to couple with alkynes. The iodine then returns to one of the alkyne's carbons, stabilizing the product but remaining poised for further transformations. —JSY

Science, this issue p. 225; see also p. 157

IMMUNOTHERAPY

The next step for cell therapy?

Considerable advances have been made in engineering T cells for adoptive cell transfer to treat cancer by manipulating the patients' immune system. In a Perspective, Bluestone and Tang discuss the potential for use of regulatory T cells (T_{regs}) to treat patients with autoimmunity and other diseases that involve pathological inflammation. Preclinical research has provided encouraging results with T_{reg} therapies, and numerous clinical trials have been initiated to test adoptive T_{reg} transfer in patients with, for example, type I diabetes, organ transplant rejection, and amyotrophic lateral sclerosis. —GKA

Science, this issue p. 154

CYTOKINES

Stabilizing the IL-17 response

The inflammatory cytokine interleukin-17 (IL-17) can stimulate both antifungal host defense and autoimmunity by promoting the stability of target messenger RNAs (mRNAs). Amatya *et al*. found that IL-17 increased the abundance of the RNA binding protein Arid5a in mouse cells (see the Focus by Puel and Casanova). Arid5a promoted the cellular response to IL-17 by increasing the mRNA stability of

a selection of IL-17–stimulated transcripts. For other transcripts, Arid5a interacted with the translation initiation factor eIF4G to augment their translation. —ERW

Sci. Signal. **11**, eaat4617, eaau8876 (2018).

INFLAMMATION

Brushing up on lung inflammation

Inhaled environmental allergens elicit type 2 lung inflammation and lead to an increase in the risk of developing allergies and asthma. Bankova *et al*. found that one step along this pathway depends on lipid mediator leukotriene E_4 signaling. This occurred via a receptor on respiratory epithelial cells that increased the number of brush cells. These brush cells represent a rare population of chemosensory cells in the lung epithelium that express receptors shared by taste bud cells. The brush cells were identified as the major pulmonary source of interleukin-25, a proinflammatory protein increased in diseases associated with type 2 inflammation. —IW

Sci. Immunol. **3**, eaat9453 (2018).

RESEARCH ARTICLE SUMMARY

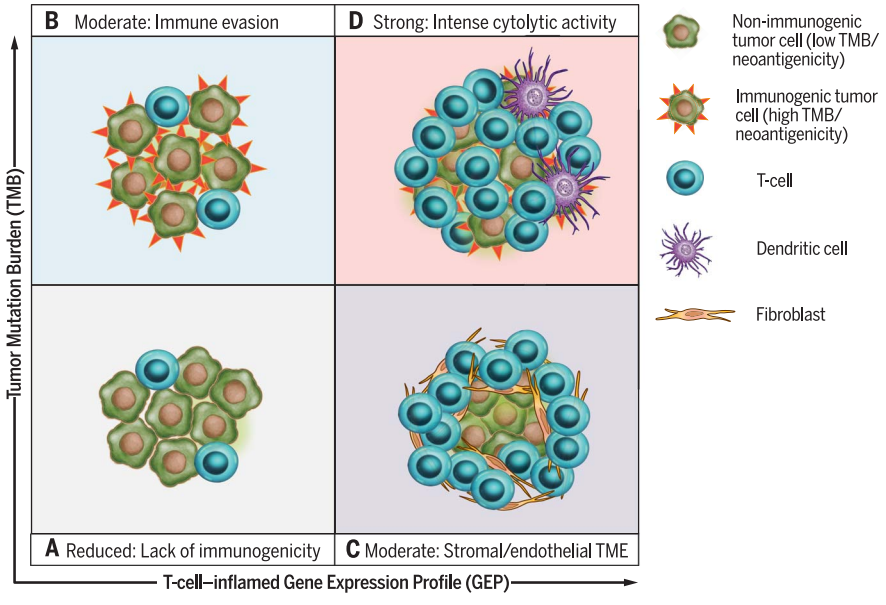
CANCER BIOMARKERS

Pan-tumor genomic biomarkers for PD-1 checkpoint blockade–based immunotherapy

Razvan Cristescu*, Robin Mogg, Mark Ayers, Andrew Albright, Erin Murphy, Jennifer Yearley, Xinwei Sher, Xiao Qiao Liu, Hongchao Lu, Michael Nebozhyn, Chunsheng Zhang, Jared Lunceford, Andrew Joe, Jonathan Cheng, Andrea L. Webber, Nageatte Ibrahim, Elizabeth R. Plimack, Patrick A. Ott, Tanguy Seiwert, Antoni Ribas, Terrill K. McClanahan, Joanne E. Tomassini, Andrey Loboda, David Kaufman

INTRODUCTION: Immunotherapy targeting the programmed cell death protein–1 (PD-1) axis elicits durable antitumor responses in multiple cancer types. However, clinical responses vary, and biomarkers predictive of response may help to identify patients who will derive the greatest therapeutic benefit. Clinically validated biomarkers predictive of response to the anti–PD-1 monoclonal antibody pembrolizumab include PD-1 ligand 1 (PD-L1) expression in specific cancers and high microsatellite instability (MSI-H) regardless of tumor type. Tumor mutational burden (TMB) and T cell–inflamed gene expression profile (GEP) are emerging predictive biomarkers for pembrolizumab. Both PD-L1 and GEP are inflammatory biomarkers indicative of a T cell–inflamed tumor microenvironment (TME), whereas TMB and MSI-H are indirect measures of tumor antigenicity generated by somatic tumor mutations. However, the relationship between these two categories of biomarkers is not well characterized.

RATIONALE: This study assessed the potential for TMB and a T cell–inflamed GEP to jointly predict clinical response to pembrolizumab in >300 patient samples with advanced solid tumors and melanoma across 22 tumor types from four KEYNOTE clinical trials. To assess the individual and joint clinical utility of TMB and GEP, patients were stratified in four biomarker-defined clinical response groups [GEP low and TMB low (GEP^{lo} TMB^{lo}), GEP low and TMB high (GEP^{lo} TMB^{hi}), GEP^{hi} TMB^{lo}, and GEP^{hi} TMB^{hi}] based on predefined cutoffs for TMB and GEP. These patient-defined biomarker groups were further used to guide transcriptome and exome analyses of tumors in a large molecular database [The Cancer Genome Atlas (TCGA)] (n = 6384 tumors) to identify targetable patterns of biology that may modulate response and resistance.



Biomarker-defined responses to pembrolizumab monotherapy identify targetable-resistance biology. (**A**) Tumors have low TMB and low neoantigenicity and lack a T cell–inflamed TME. (**B**) Tumors can evade the immune response despite high TMB and high neoantigenicity. (**C**) Although T cells are present, stromal and/or endothelial factors in the TME, low TMB, and low neoantigenicity impede their activity. (**D**) Tumors have high TMB, high neoantigenicity, and a T cell–inflamed TME, typified by activated T cells and other immune cells with cytolytic roles.

ON OUR WEBSITE

Read the full article at http://dx.doi.org/10.1126/science.aar3593

RESULTS: TMB and GEP exhibited only modest correlation and were independently predictive of response across the KEYNOTE clinical datasets. We found that objective response rates were strongest in patients with GEP^{hi} TMB^{hi} (37 to 57%), moderate in those with GEP^{hi} TMB^{lo} (12 to 35%) and GEP^{lo} TMB^{hi} (11 to 42%), and reduced or absent in those with GEP^{lo} TMB^{lo} (0 to 9%) (see the figure). Additionally, longer progression-free survival times were seen in patients with higher levels of both TMB and GEP. Findings were comparable when TMB and PD-L1 expression were jointly assessed. Within TCGA database, GEP and TMB again had a low correlation, demonstrating the potential to jointly stratify transcriptomic and genomic features across cancer types. Specific gene expression patterns reflective of TME biology showed significant associations with TMB, GEP, or both. In particular, gene set enrichment analysis identified proliferative and stromal, myeloid, and vascular biology corresponding to specific TMB-defined subgroups within GEP^{hi} tumors. In TMB^{hi} tumors, indication-dependent somatic DNA alterations in key cancer driver genes showed a strong negative association with GEP.

CONCLUSION: This analysis shows that TMB and inflammatory biomarkers (T cell–inflamed GEP and PD-L1 expression) can jointly stratify human cancers into groups with different clinical responses to pembrolizumab monotherapy and identify patterns of underlying, targetable biology related to these groups. TMB and inflammatory biomarkers independently predict response and may capture distinct features of neoantigenicity and T cell activation, respectively. This approach may provide a precision medicine framework for rationally constructing and evaluating anti–PD-1– and/or –PD-L1–based combination therapy regimens. ■

The list of author affiliations is available in the full article online.
*Corresponding author. Email: razvan_cristescu@merck.com
Cite this article as R. Cristescu *et al.*, *Science* **362**, eaar3593 (2018). DOI: 10.1126/science.aar3593

RESEARCH ARTICLE

CANCER BIOMARKERS

Pan-tumor genomic biomarkers for PD-1 checkpoint blockade–based immunotherapy

Razvan Cristescu[1]*, Robin Mogg[1]†, Mark Ayers[1], Andrew Albright[1], Erin Murphy[1], Jennifer Yearley[1], Xinwei Sher[1], Xiao Qiao Liu[1], Hongchao Lu[1], Michael Nebozhyn[1], Chunsheng Zhang[1], Jared K. Lunceford[1], Andrew Joe[1], Jonathan Cheng[1], Andrea L. Webber[1], Nageatte Ibrahim[1], Elizabeth R. Plimack[2], Patrick A. Ott[3], Tanguy Y. Seiwert[4], Antoni Ribas[5], Terrill K. McClanahan[1], Joanne E. Tomassini[1], Andrey Loboda[1], David Kaufman[1]†

Programmed cell death protein–1 (PD-1) and programmed cell death ligand–1 (PD-L1) checkpoint blockade immunotherapy elicits durable antitumor effects in multiple cancers, yet not all patients respond. We report the evaluation of >300 patient samples across 22 tumor types from four KEYNOTE clinical trials. Tumor mutational burden (TMB) and a T cell–inflamed gene expression profile (GEP) exhibited joint predictive utility in identifying responders and nonresponders to the PD-1 antibody pembrolizumab. TMB and GEP were independently predictive of response and demonstrated low correlation, suggesting that they capture distinct features of neoantigenicity and T cell activation. Analysis of The Cancer Genome Atlas database showed TMB and GEP to have a low correlation, and analysis by joint stratification revealed biomarker-defined patterns of targetable-resistance biology. These biomarkers may have utility in clinical trial design by guiding rational selection of anti–PD-1 monotherapy and combination immunotherapy regimens.

Emerging immune-relevant biomarkers for checkpoint blockade immunotherapy response can be placed broadly into two categories: those related to tumor neoepitope burden, such as microsatellite instability (MSI) or high tumor mutational burden (TMB), and those indicative of a T cell–inflamed tumor microenvironment (TME). The latter include programmed cell death ligand–1 (PD-L1) protein expression on tumor and immune cells, which in many cases is up-regulated in response to local T cell–derived interferon-γ (IFN-γ), and gene signatures of activated T cells (*1*–*3*). TMB is correlated with clinical response to cytotoxic T lymphocyte–associated antigen–4 blockade in advanced melanoma (*4*–*6*) and with anti–programmed cell death protein–1 (PD-1) and/or PD-L1 blockade in melanoma (*7*), non–small cell lung cancer (NSCLC) (*8*, *9*), colorectal and gastric cancers (*10*, *11*), and urothelial cancer (*12*). Similarly, tumors with MSI that have high levels of both single-nucleotide and frameshift mutations [high MSI (MSI-H)] are responsive to anti–PD-1 therapy in colorectal cancer and other malignancies (*10*, *11*). Expression of genes related to immune cytolytic activity have also been shown to be associated with clinical response to checkpoint blockade in certain tumors (*13*, *14*). Recently, a T cell–inflamed gene expression profile (GEP) was shown to predict response to anti–PD-1–directed therapy (*15*). However, the interplay between these two distinct categories of biomarkers has not been well characterized across cancer types with respect to their ability either to independently or jointly predict response to immunotherapy or to reveal underlying genomic and/or transcriptomic features of tumor antigenicity and TME.

We evaluated the relationship between somatic TMB and clinical response to anti–PD-1 immunotherapy with pembrolizumab. Twenty-two cancer types were included in the discovery and validation cohorts and were analyzed for the independent and joint predictive values of TMB and T cell–inflamed GEP. Additionally, by using large molecular databases [e.g., The Cancer Genome Atlas (TCGA) (*16*)], we explored transcriptomic and genetic features associated with the presence or absence of either of these two markers.

Study cohorts and tumor and mutation types

The predictive values of TMB and the T cell–inflamed GEP were first assessed separately by rigorous stepwise testing in four cohorts of patients across the pembrolizumab clinical development program (one discovery, one pan-tumor validation, and two single-indication summary cohorts). TMB was evaluated by whole-exome sequencing (WES) of germline and tumor DNA, and the T cell–inflamed GEP was analyzed by targeted gene expression profiling of tumor RNA (with the NanoString platform) from formalin-fixed, paraffin-embedded (FFPE) pretreatment samples. The initial discovery cohort for TMB comprised patients with PD-L1–positive head and neck squamous cell carcinoma (HNSCC) from a phase 1b clinical trial (KEYNOTE-012 B1 cohort; n = 34 patients), and the pan-tumor validation cohort consisted of patients with PD-L1–positive advanced solid tumors (n = 119 patients) from two multicohort phase 1b trials across 20 cancer types [KEYNOTE-028 (17 cohorts; n = 80 patients) and KEYNOTE-012 (A, C, and D cohorts; n = 39 patients)]. The HNSCC single-indication cohort (n = 107 patients) included patients in the phase 1b KEYNOTE-012 B1 cohort and additional patients with PD-L1–unselected HNSCC (n = 73 patients) from the KEYNOTE-012 B2 cohort. The melanoma single-indication cohort included patients with advanced melanoma from the phase 1b (KEYNOTE-001; n = 30 patients) and the phase 3 (KEYNOTE-006 pembrolizumab arm; n = 59 patients) trials. The clinical characteristics of each cohort are listed in table S1, and the characteristics of all patients included in this study are listed in table S2.

The distribution of tumor mutational signatures across the study cohorts largely reflected recognized cancer subtype–dependent determinants of mutagenesis (*17*) (table S3 and fig. S1). The dominant mutational signatures varied across tumor types in the pan-cancer cohort, with higher TMB associated with tissue-specific signatures, such as smoking in small cell lung cancer; apolipoprotein B mRNA editing enzyme, catalytic polypeptide–like (APOBEC) in genitourinary tumors; and mismatch repair (MMR) in gastrointestinal cancer. Within the pan-cancer validation cohort, the DNA polymerase epsilon catalytic subunit (PolE) signature and the Val^{411} mutation in *POLE* were observed in an endometrial carcinoma tumor that had the highest TMB (5464). Dominant signatures in the single-indication cohorts were more homogenous, with an APOBEC signature in the HNSCC cohort (61% of tumors) and an ultraviolet (UV) light exposure signature in melanoma (in 78% of the tumors, >30% of mutations were UV light induced).

Association of TMB and T cell–inflamed GEP with clinical response

Clinical response associations were assessed on the basis of best overall response (BOR) and progression-free survival (PFS) by RECIST 1.1. BOR and PFS associations with TMB and the T cell–inflamed GEP were assessed in all patients who had WES and transcriptomic data available.

We first assessed the predictive value of each individual genomic biomarker separately across the different cohorts. In the HNSCC B1 discovery cohort, higher TMB predicted a greater frequency of clinical response (BOR) (P = 0.0123). This was validated by using the pan-tumor cohort, in which

[1]Merck & Co., Kenilworth, NJ 07033, USA. [2]Fox Chase Cancer Center, Philadelphia, PA 19111, USA. [3]Dana-Farber Cancer Institute, Boston, MA 02215, USA. [4]University of Chicago, Chicago, IL 60637, USA. [5]University of California, Los Angeles, Los Angeles, CA 90095, USA.
*Corresponding author. Email: razvan_cristescu@merck.com
†Present address: Bill and Melinda Gates Medical Research Institute, Cambridge, MA 02139, USA.

TMB was again associated with BOR ($P < 0.001$) (Fig. 1A). Higher T cell–inflamed GEP scores were also positively associated with BOR in the pan-tumor cohort ($P < 0.01$) (Fig. 1B), showing that a T cell–activated tumor environment also affects response in addition to TMB. Similarly, both TMB and T cell–inflamed GEP scores were positively associated with BOR in the single-indication cohorts of HNSCC ($P < 0.05$ and $P < 0.001$, respectively) and melanoma ($P < 0.05$ for both) patients (Fig. 1, A and B). In this study, we did not evaluate the effect of human papillomavirus (HPV) antigens on the association of TMB with response in the HNSCC cohort; however, we have previously described the association of TMB with clinical outcome in a larger, overlapping group of HNSCC patients (KEYNOTE-012 B1 and B2 cohorts) stratified by HPV status (*18*). Although we found that TMB was more strongly associated with BOR in HPV-negative patients than in HPV-positive patients, those exploratory findings await validation in larger, independent studies.

The clinical utility of TMB in predicting BOR was generally high, and degrees of utility were similar across cancer types, with areas under the receiver operating characteristic curves (AUROCs) of 0.740, 0.617, and 0.602 in the pan-tumor, HNSCC, and melanoma cohorts, respectively. Similar results were observed for the T cell–inflamed GEP across the cohorts (AUROCs = 0.782, 0.768, and 0.638, respectively) (Fig. 1C). The potential performance of a targeted sequencing–based TMB assay was simulated by using the genes in the Foundation Medicine targeted sequencing platform (*19*). The corresponding AUROC across the cohorts was comparable to that observed by using WES (0.721), suggesting potential translatability to a targeted panel diagnostic. Taken together, these data imply that both TMB and the T cell–inflamed GEP have comparable performance characteristics and potential diagnostic utility.

We next evaluated the joint utility of the two genomic biomarkers in predicting response. The correlation between TMB and GEP was low in the pan-tumor and melanoma cohorts (Spearman correlation coefficient $r = 0.221$, $P < 0.05$, and $r = 0.252$, $P < 0.05$, respectively), and there was no correlation in the HNSCC cohort ($r = -0.020$, $P = 0.841$) (Fig. 2A). This lack of correlation, combined with the observed individual predictive values, suggested that TMB and the T cell–inflamed GEP are independent predictive measures of response to pembrolizumab. When tested in a multivariate model adjusted for each measure, both TMB and T cell–inflamed GEP retained significant predictive value in the pan-tumor ($P = 0.0028$ and 0.0051, respectively) and HNSCC ($P = 0.0013$ and 0.0004) cohorts, whereas only GEP remained significant in the melanoma cohort ($P = 0.1644$ and 0.026). Although a portion of the patients in this study were PD-L1 selected, these relationships were observed even in those cohorts of patients that were not PD-L1 selected.

We evaluated the association of the genomic biomarkers with PD-L1 immunohistochemistry (IHC) scores (fig. S2). TMB was significantly but moderately correlated with PD-L1 in the pan-tumor cohort [combined positive score (CPS), $r = 0.330$; $P = 0.0038$] and showed no association with PD-L1 in the HNSCC cohort (CPS, $r = 0.020$; $P = 0.8084$) or in the melanoma cohort [melanoma (MEL) score, $r = 0.049$; $P = 0.6473$]. In contrast, GEP was more significantly correlated with PD-L1 in the pan-tumor, HNSCC, and melanoma cohorts ($r = 0.49$, 0.51, and 0.53, respectively; all P values < 0.001), consistent with the known regulation of PD-L1 gene expression by T cell–derived IFN-γ (*1*–*3*). This correlation suggests that a PD-L1 IHC–based assay is relevant in assessing a T cell–inflamed TME. As seen with high TMB (TMB^{hi}) and high GEP scores (GEP^{hi}), responses in patients who had both TMB^{hi} and greater PD-L1 expression ($PD\text{-}L1^{+}$; CPS ≥ 1) were greater than those in patients who had low levels of both TMB and PD-L1 expression.

We next studied the potential joint utility of TMB and GEP for patient stratification and treatment outcome prediction. Clinical response was evaluated on the basis of cut points associated with the Youden Index (derived from the AUROCs for TMB in each cohort) and a discovery cutoff of −0.318 for the T cell–inflamed GEP score (selected via analysis of pan-cancer data) (*15*). Rates of response to pembrolizumab were greater in patients with TMB^{hi} (greater than or equal to Youden Index cut points) than in those with low TMB (TMB^{lo}) (less than Youden

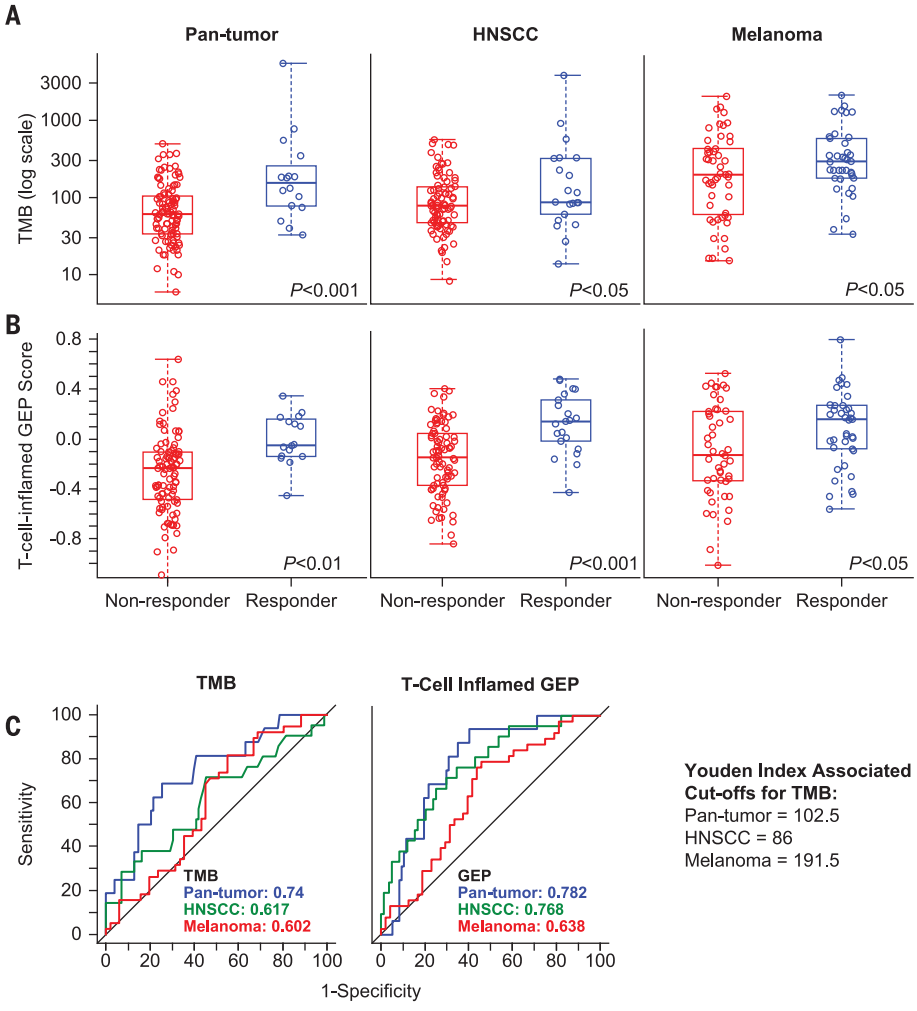


Fig. 1. Individual association of TMB or T cell–inflamed GEP with anti–PD-1 response across multiple patient cohorts. (**A** and **B**) The association of (A) TMB, defined as the sum of somatic nonsynonymous mutations, and (B) T cell–inflamed GEP with BOR was assessed in pan-tumor, HNSCC, and melanoma cohorts by central radiology review for all-patients-as-treated populations in all cohorts. A responder is defined as having a partial response (PR) or a complete response (CR); a nonresponder is defined as having no PR or CR. Nonresponders and responders for TMB, respectively, were $n = 103$ and $n = 16$ for pan-tumor, $n = 86$ and $n = 21$ for HNSCC, and $n = 51$ and $n = 38$ for melanoma cohorts. For GEP score analysis, nonresponders and responders were $n = 97$ and $n = 16$ for pan-tumor, $n = 84$ and $n = 21$ for HNSCC, and $n = 48$ and $n = 38$ for melanoma cohorts. For both (A) and (B), raw data are displayed in standard box plots with medians and interquartile ranges. (**C**) AUROCs for TMB and T cell–inflamed GEP in the three patient cohorts. Youden Index–associated cutoffs for TMB in each cohort are shown.

Index cut points) and were similarly greater for those with higher T cell–inflamed GEP scores (greater than or equal to the cutoff of −0.318) than for those with lower scores (less than the −0.318 cutoff) (Fig. 2B). The highest objective response rate was observed for patients within each cohort who had both TMB^{hi} and GEP^{hi}. Additionally, among patients with both TMB^{lo} and low T cell–inflamed GEP scores (GEP^{lo}), no responses were observed in the pan-tumor and HNSCC cohorts and only one response was observed in the melanoma cohort, suggesting greater sensitivity for the combination of biomarkers. Patients who had high scores for only one of the biomarkers (TMB^{lo} GEP^{hi} and TMB^{hi} GEP^{lo}) had moderate responses (Fig. 2B). These data suggest the potential for greater positive and negative predictive value when these biomarkers are used together in the setting of PD-1–directed monotherapy.

Patient stratification by TMB and GEP was also differentially associated with PFS. In all three cohorts, hazard ratios associated with PFS were <1.0 (implying PFS benefit) among patients with high versus low TMB and high versus low T cell–inflamed GEP scores. The most pronounced PFS-associated hazard ratios were observed for TMB^{hi} GEP^{hi} tumors in the pan-tumor (Fig. 3A), HNSCC (Fig. 3B), and melanoma cohorts (Fig. 3C). The greatest differential was observed in each cohort for patients with TMB^{hi} GEP^{hi} versus patients with TMB^{lo} GEP^{lo}. Patients who had greater levels of either TMB or GEP (TMB^{hi} or GEP^{hi}) versus low levels of these biomarkers (TMB^{lo} or GEP^{lo}) also had longer PFS.

We also explored the feasibility and potential clinical value of identifying a pan-cancer threshold for TMB across our cohorts that maximizes its joint predictive utility with GEP by using a method similar to that of Panda *et al.* (*20*). A TMB cutoff of ≥123 mutations per exome maximized the effect size of the difference in GEP distributions between tumors having TMB less than and greater than the cutoff. The response rates to pembrolizumab in the TMB-GEP–defined groups of each clinical cohort were comparable to those observed by using the cohort-specific cut points for TMB reported above (fig. S3). The hazard ratios observed for PFS were also generally similar with the use of the TMB cutoff of ≥123 mutations per exome (fig. S4). A pan-tumor threshold may be further optimized with the availability of additional data beyond those in our study. For example, a pan-tumor TMB threshold of ≥175 mutations per exome was recently reported for response to pembrolizumab (*21*).

Association of other DNA-based measures with response

The predictive value of other DNA-based measures of mutation status in relation to response was also evaluated in these cohorts, including predicted neoantigen signature, smoking status, APOBEC-driven mutations, UV light exposure, DNA transversions, homologous recombination deficiency, and MSI. Aside from MSI, none of these specific measures of genetic alteration provided additional meaningful improvement in predictive value over TMB assessment alone. The predicted neoantigen load was highly correlated with TMB in the pan-tumor, HNSCC, and melanoma cohorts (r = 0.87, 0.83, and 0.90, respectively), as expected (fig. S5). In the pan-tumor cohort, most measures of mutagenic processes were significantly associated with BOR (e.g., predicted neoantigen load and smoking; both P values = 0.001), with similar relevant trends toward significant association with PFS (table S4). By using a WES-based method to infer MSI (*22*), two patients with MSI-H tumors (gastric and biliary tract carcinomas) were identified, and both were responders; the MSI status of these patients was confirmed with standard MSI polymerase chain reaction (PCR) methods. In the melanoma cohort, the percentage of UV light–induced mutations correlated with TMB (r = 0.77; $P < 1 \times 10^{-10}$) (fig. S1) and was significantly associated with response (P = 0.02). These data suggest that nonsynonymous mutations arising from a wide variety of mutagenic processes are capable of enhancing the antigenicity of tumors, with comparable effects on the response to PD-1 checkpoint blockade.

Somatic mutation clonality and copy number variation (CNV) have previously been reported to positively and negatively associate, respectively, with response to PD-1 checkpoint blockade (*23*, *24*). In an analysis of clonal versus nonclonal tumors (clonality of 1 versus <1, respectively), the treatment response rates were numerically higher in clonal tumors in the pan-tumor cohort (18% versus 10%) but not different in the HNSCC (21% versus 23%) or melanoma (44% versus 41%) cohort. A low and nonsignificant overall correlation was observed between clonality and TMB (r = 0.05; $P > 0.05$) in the pooled dataset, suggesting a potential utility of including clonality assessment in the application of a TMB-based biomarker. Higher levels of CNV trended toward negative associations with response but approached statistical significance only in the HNSCC and melanoma cohorts (AUROCs = 0.48, 0.35, and 0.42; P = not significant, 0.1, and 0.1 for the pan-tumor, HSNCC, and melanoma cohorts, respectively). Correlations between TMB and CNV load were low in the pan-tumor (r = −0.03), HNSCC (r = 0.16), and melanoma (r = −0.12) cohorts ($P > 0.05$

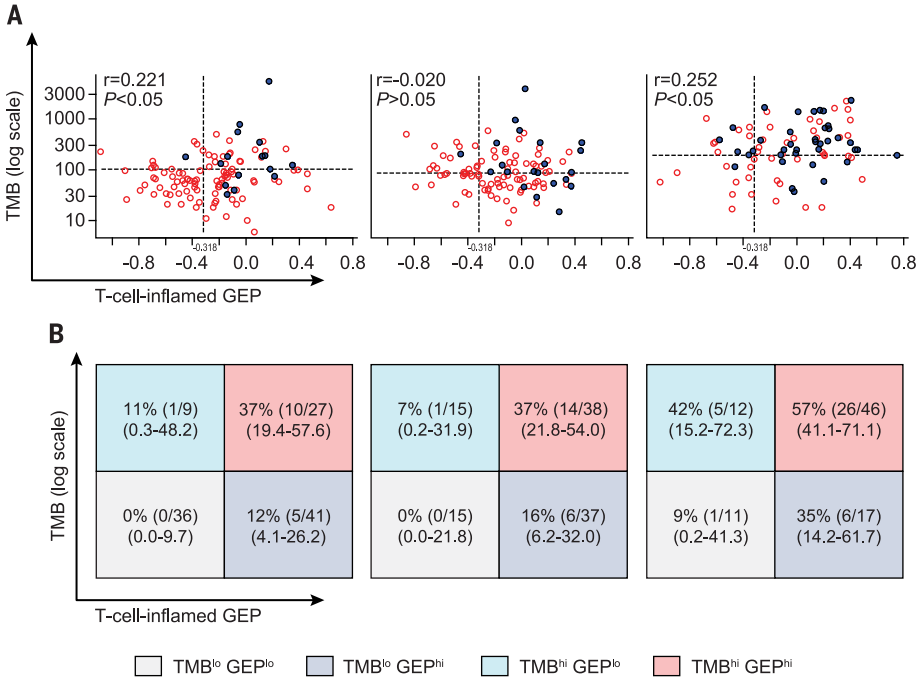


Fig. 2. Joint relationship of TMB or T cell–inflamed GEP with anti–PD-1 response across multiple patient cohorts. (**A**) Relationships of both TMB and T cell–inflamed GEP signatures with BOR. A responder is defined as having a PR or CR (filled circles); a nonresponder has no PR or CR (open circles). Dashed horizontal lines represent the Youden Index–associated cutoffs for TMB in each cohort as derived from AUROCs in Fig. 1C. Dashed vertical lines represent a discovery cutoff for the T cell–inflamed GEP selected via analysis of pan-cancer data. (**B**) Response (PR or CR) rates [expressed as a percentage calculated as the number of responders divided by the number in the cutoff-defined group, with 95% confidence intervals (CI)] per TMB cutoff status and T cell–inflamed GEP cutoff status as designated in (A). TMB^{hi} and TMB^{lo} response groups are defined by values greater than or equal to and less than Youden Index–associated cut points (102.5, 86, and 191.5 for pan-cancer, HNSCC, and melanoma cohorts, respectively); GEP^{hi} and GEP^{lo} groups are defined by cutoffs greater than or equal to and less than −0.318, respectively.

for all), suggesting a potential complementary role of CNV in biomarker-based prediction of responders versus nonresponders.

TMB and T cell–inflamed GEP relationships can be applied to a wide range of tumor types across genomic databases

To explore the generalizability of our findings and the utility of our stratification schema across tumor types, the relationship among TMB, T cell–inflamed GEP, and related genomic features was further explored in TCGA (n = 9963 patients with transcriptomic data, 6384 of which also had WES data) (*16*). Patients were stratified by TMB (WES score ≤ 100 mutations per exome) and T cell–inflamed GEP score (below the top tertile of data) by using cutoffs equivalent in terms of prevalence to those that were used to define the clinical response groups in the pan-tumor cohort (Fig. 4A). Consistent with our clinical data, TMB and the T cell–inflamed GEP were found to have low but significant correlations (r = 0.30; $P < 1 \times 10^{-4}$), as did TMB and PD-L1 gene expression (r = 0.16; $P < 1 \times 10^{-4}$) and TMB and PD-L2 gene expression (r = 0.22; $P < 1 \times 10^{-4}$). By contrast, both PD-L1 expression and PD-L2 expression, which are induced by IFN-γ from activated Th1 and cytotoxic T cells (*1*–*3*), were highly correlated with the T cell–inflamed GEP (r = 0.61 and 0.72; $P < 1 \times 10^{-10}$). MSI-H tumors made up a subset of tumors with TMB^{hi} in both T cell–inflamed and noninflamed tumors. Even in these tumors, which exhibit very high mutational burdens, the modest correlation between GEP and TMB was preserved. The frequency of the TMB^{hi} GEP^{hi} subgroup, which was identified as the most clinically responsive population in our datasets, varied across cancer types (Fig. 4B), with enrichment among patients with tumors that are generally more responsive to pembrolizumab, such as melanoma and NSCLC (*25*, *26*), and underrepresentation among patients with tumors such as prostate cancer and glioblastoma that are typically more resistant to immunotherapy (*27*, *28*).

Rooted in the well-studied field of T cell inflammation and cytolytic process (*13*, *29*–*31*), the T cell–inflamed GEP signature was derived by a stepwise process of discovery, validation, and refinement of candidate gene sets associated with patient response to pembrolizumab across multiple solid tumors with the use of a NanoString platform enriched in immune genes (*15*) and thus represents a universal signature. Notably, in TCGA dataset, we observed a strong correlation (r > 0.9) between the GEP and several other previously published transcriptional signatures reflective of a T cell–inflamed TME associated with cytolytic processes (Fig. 5A).

Stratification of additional genomic features by TMB and T cell–inflamed GEP

The patient groups defined by TMB and GEP status show notable differences in clinical response to pembrolizumab. In particular, the two groups with only one positive biomarker indicative of potential for pembrolizumab response (TMB^{hi} GEP^{lo} or TMB^{lo} GEP^{hi}) have markedly lower response rates than the TMB^{hi} GEP^{hi} group, suggesting that mechanisms of resistance to pembrolizumab may exist that are specific to each respective group. In order to identify potential mechanisms of resistance, we assessed molecular differences among tumors that belong to different TMB- and T cell–inflamed GEP–defined groups through analyses in TCGA molecular database.

First, we compared the correlation of genes in the transcriptome with GEP in TMB^{hi} and in TMB^{lo} tumors separately. Both distributions of correlations diverged from a normal distribution because of a pattern of significant skewing toward positive correlations with the T cell–inflamed GEP, consistent with robust coregulation of gene expression markers of cell types present in a cytolytic TME. However, there were no major differences in the correlations of individual genes with the T cell–inflamed GEP between TMB^{hi} (TMB > 100 mutations per exome) and TMB^{lo} (TMB ≤ 100 mutations per exome) tumors (r = 0.76; $P < 1 \times 10^{-20}$) (Fig. 5B), suggesting a lack of qualitative difference in T cell inflammation markers as a function of tumor neoantigenicity. Notably, much smaller deviations from a normal distribution were observed in the negative range of correlations with GEP in both TMB^{hi} and TMB^{lo} tumors, suggesting the absence of major pan-cancer transcriptional signatures strongly associated with T cell exclusion.

To understand the origin of the skewness toward positive correlations with the T cell–inflamed GEP, genes positively correlated with the T cell–inflamed GEP (r > 0.15) were classified into two sets by using cutoffs defined by deviations from a normal distribution of the correlation with the T cell–inflamed GEP at 83% and 98% quantiles, respectively (Fig. 5C). Set 1 comprised genes that had a Spearman correlation r > 0.6 with the T cell–inflamed GEP (the lower bound for the correlation of individual genes in the signature with the signature as a whole), whereas set 2 genes had correlations with GEP that ranged between 0.15 and 0.6. Additionally,

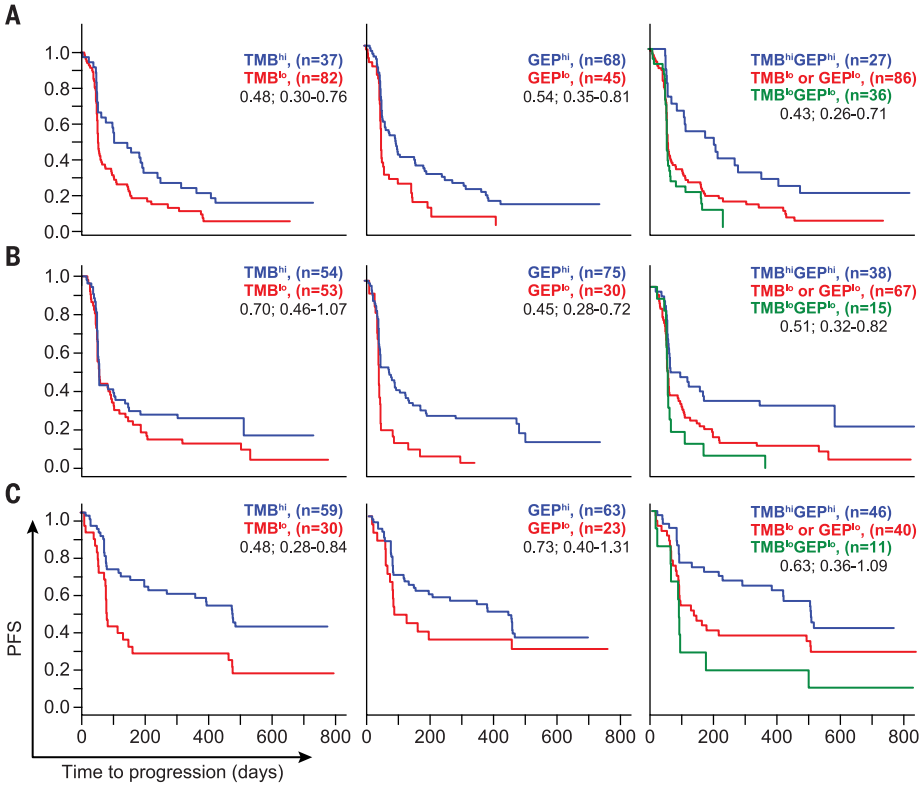


Fig. 3. Relationship between TMB and T cell–inflamed GEP signatures and PFS after anti–PD-1 treatment across multiple patient cohorts. Relationships of TMB and T cell–inflamed GEP with PFS in all patients as treated per TMB cutoff and GEP cutoff as described in the legend to Fig. 2. Median PFS times in days for (**A**) pan-tumor, (**B**) HNSCC, and (**C**) melanoma cohorts for TMB^{hi} versus TMB^{lo} were 115 versus 59 (hazard ratio, 0.48; 95% CI, 0.30 to 0.76), 64 versus 64 (0.70; 0.46 to 1.07), and 502 versus 85 (0.48; 0.28 to 0.84); those for GEP^{hi} versus GEP^{lo} were 96 versus 57 (0.54; 0.35 to 0.81), 103 versus 57 (0.45; 0.28 to 0.72), and 418 versus 90 (0.73; 0.40 to 1.31); those for TMB^{hi} GEP^{hi} versus TMB^{lo} GEP^{lo} or TMB^{lo} GEP^{lo} were 189 versus 59 (0.43; 0.26 to 0.71), 110 versus 62 (0.51; 0.32 to 0.82), and 504 versus 123 (0.63; 0.36 to 1.09). Kaplan-Meier plots are shown, and median survival was estimated on the basis of Kaplan-Meier estimates. Hazard ratios with 95% CI were derived from a Cox proportional model fit, with adjustment for baseline ECOG score and protocol where relevant.

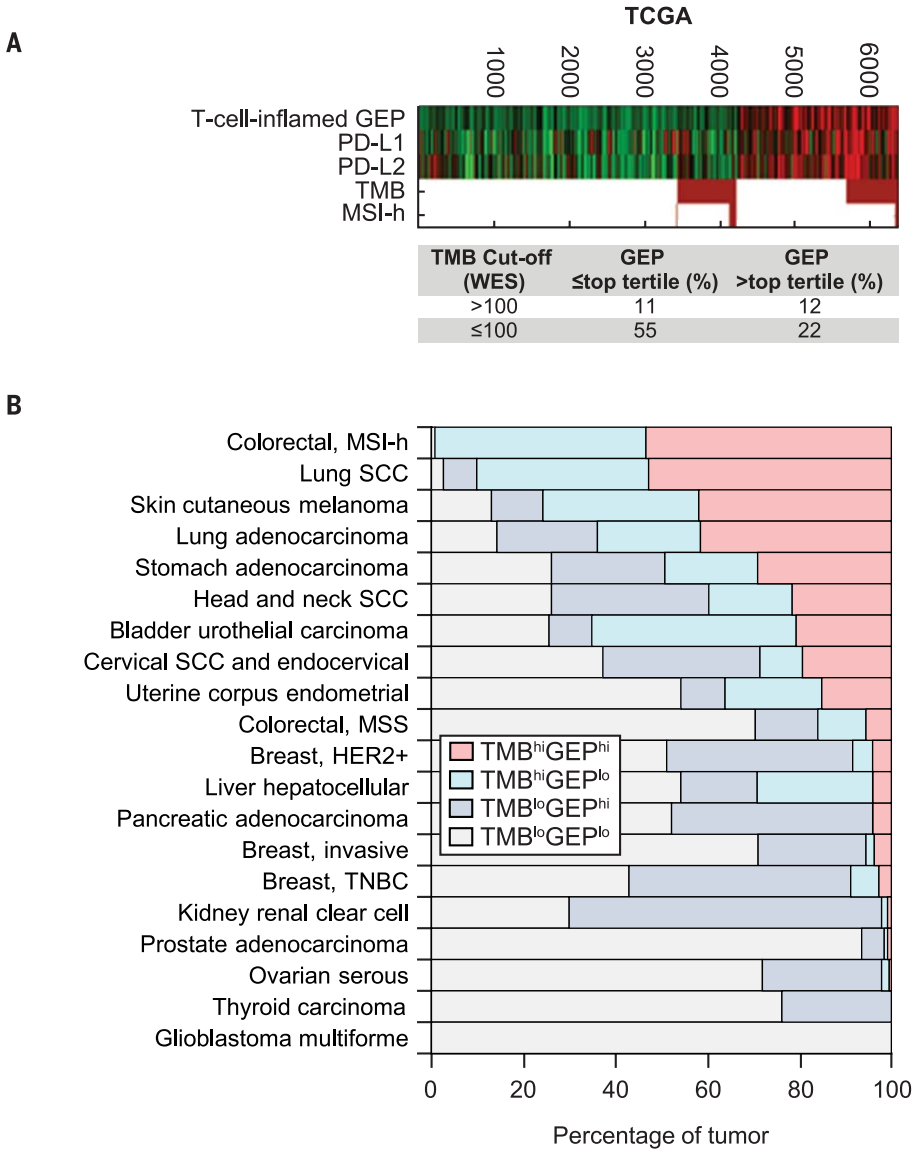


Fig. 4. Relationships of TMB, GEP, and other key biomarkers with gene expression across tumor types in TCGA. (**A**) Data are stratified by TMB and GEP cutoffs, which are equivalent in terms of prevalence to those that define the clinical response groups in the pan-tumor cohort of patients treated with pembrolizumab from the KEYNOTE studies. The WES cutoff of >100 mutations per exome for TMB was chosen to match the Youden Index–associated TMB cutoff defined for the pan-tumor cohort. The GEP cutoff was chosen as the top pan-cancer tertile value. Columns represent individual tumors, and rows represent genomic features. Red and green represent elevated and decreased expression, respectively (versus the median, in black), for continuous variables, and red and white represent true and false for Boolean (binary) variables. In the absence of MSI evaluation across cancer types, MSI-H status was determined by loss of *MLH1* gene expression by using cutoffs determined by the bimodality in the distribution of expression. (**B**) Percentages of tumors in each cancer type in biomarker-defined response groups as defined in (A) in TCGA database. SCC, squamous cell carcinoma; MSS, microsatellite stable; TNBC, triple-negative breast cancer.

genes negatively correlated with the T cell–inflamed GEP and divergent from a normal distribution ($r < -0.15$ at 14% quantile) were grouped in set 3.

As expected, a strong enrichment of genes related to T cell–inflamed cytolytic processes was observed in set 1 (table S5). By contrast, set 2 showed enrichment in genes specific to other cell types in the TME, including vascular endothelium and myeloid infiltrate, but did not show enrichment of genes for T cell–inflamed cytolytic processes or tumor cell–intrinsic pathways. Genes in set 1 and set 2 were further grouped as modules of gene coexpression by *K*-means clustering ($K = 10$ for set 2, and $K = 4$ for set 1). Modules in set 1 did not show a strong association with TMB, consistent with the weak associations between TMB and the T cell–inflamed GEP described above. However, several modules in set 2 (table S6) displayed distinct patterns of correlation or anticorrelation with TMB. Annotation of the genes in the modules that were most strongly correlated and anticorrelated with TMB (modules 4 and 5, respectively), revealed enrichment in biology related to cell proliferation (module 4) and vasculature (module 5). These data suggest that distinct patterns of underlying biology can be identified by using TMB and the T cell–inflamed GEP to categorize tumors (Fig. 5D). The association of the average expression of these gene modules (modules 4 and 5) with TMB and T cell–inflamed GEP is represented in Fig. 5D in the upper left and lower right panels, respectively, by using the cytolytic module 1 from set 1 in the upper right panel as a reference.

The group of genes in set 3 that were anticorrelated with the T cell–inflamed GEP ($r < -0.15$) was also investigated; however, the biological annotation of the resulting coexpression modules was less informative than that for genes positively correlated with the T cell–inflamed GEP. However, some modules in this group were anticorrelated with TMB as well as with T cell–inflamed GEP. In particular, a module enriched in stromal and Wnt signaling elements was identified in tumors with both TMB^{lo} and T cell–inflamed GEP^{lo} (Fig. 5D, lower left panel).

An additional analysis was performed by interrogating the entire transcriptome for genes associated with TMB in T cell–inflamed tumors, independently of the GEP-based clustering approach described above. Similar to the analysis of modules, this analysis showed that genes that positively correlated with TMB were enriched for proliferation whereas those that were anticorrelated with TMB were related to vascular and stromal biology (table S7). Consistent with these analyses, the distribution of previously identified signatures of stromal biology, proliferation, cytolytic activity, and Wnt signaling (*13*, *32*–*34*) also showed similar patterns of association with TMB and the T cell–inflamed GEP (fig. S6). However, in this analysis, we were not able to identify a gene expression signature of TMB^{hi} that was as predictive as TMB itself for response to pembrolizumab.

A complementary approach was used to identify genomic determinants of low cytolytic transcriptomic activity (absence of a T cell–inflamed GEP) in tumors with TMB^{hi} as potential drivers of immune evasion in a mutagen-rich context. As described above, the transcriptomic correlation of the T cell–inflamed GEP in TMB^{hi} tumors (Fig. 5B) showed a distribution that skewed toward positive correlation with GEP, suggesting the absence of a robust transcriptome signal in tumors with TMB^{hi} and GEP^{lo}. Therefore, DNA alterations in TCGA were explored to reveal potential negative associations of somatic mutations with GEP by using a previously reported approach (*13*) but focusing specifically on tumors with TMB^{hi}. Among known cancer drivers, serine-threonine kinase 11 (*STK11*) [also known as liver kinase B1 (*LKB1*)] mutation in lung adenocarcinoma, Kelch-like ECH-associated protein 1 (*KEAP1*) mutation in lung adenocarcinoma and lung

squamous cell carcinoma, and adenomatous polyposis coli (*APC*) mutation in colorectal cancer showed highly significant negative associations with the T cell–inflamed GEP (Fig. 6). Notably, none of these associations passed the nominal significance level ($P < 0.01$) in the pan-cancer analysis, suggesting a potential cancer type–specific role for these somatic alterations. Other genes demonstrating negative associations with the T cell–inflamed GEP were either of low frequency or were not known cancer drivers (Fig. 6B).

Discussion

Several studies have shown that either TMB^{hi} or cytolytic elements of the TME are associated with clinical response to checkpoint blockade immunotherapy in some tumor types (*4*–*9*, *11*–*13*, *15*). However, the relationship between these two central aspects of tumor immunobiology and their combined association with clinical response to checkpoint blockade immunotherapy has not been well-studied across multiple cancer types. Here, we show that TMB and a T cell–inflamed GEP are tissue-agnostic measures of distinct aspects of tumor immunobiology and independently predict response to anti–PD-1 therapy in multiple tumors. In particular, limited clinical responses to pembrolizumab occurred in patients with low levels of both TMB and T cell–inflamed GEP, whereas the greatest response rates were seen in patients with high levels of both biomarkers. Similarly, improved responses were seen in patients who had high levels of both PD-L1 IHC expression and TMB, reflective

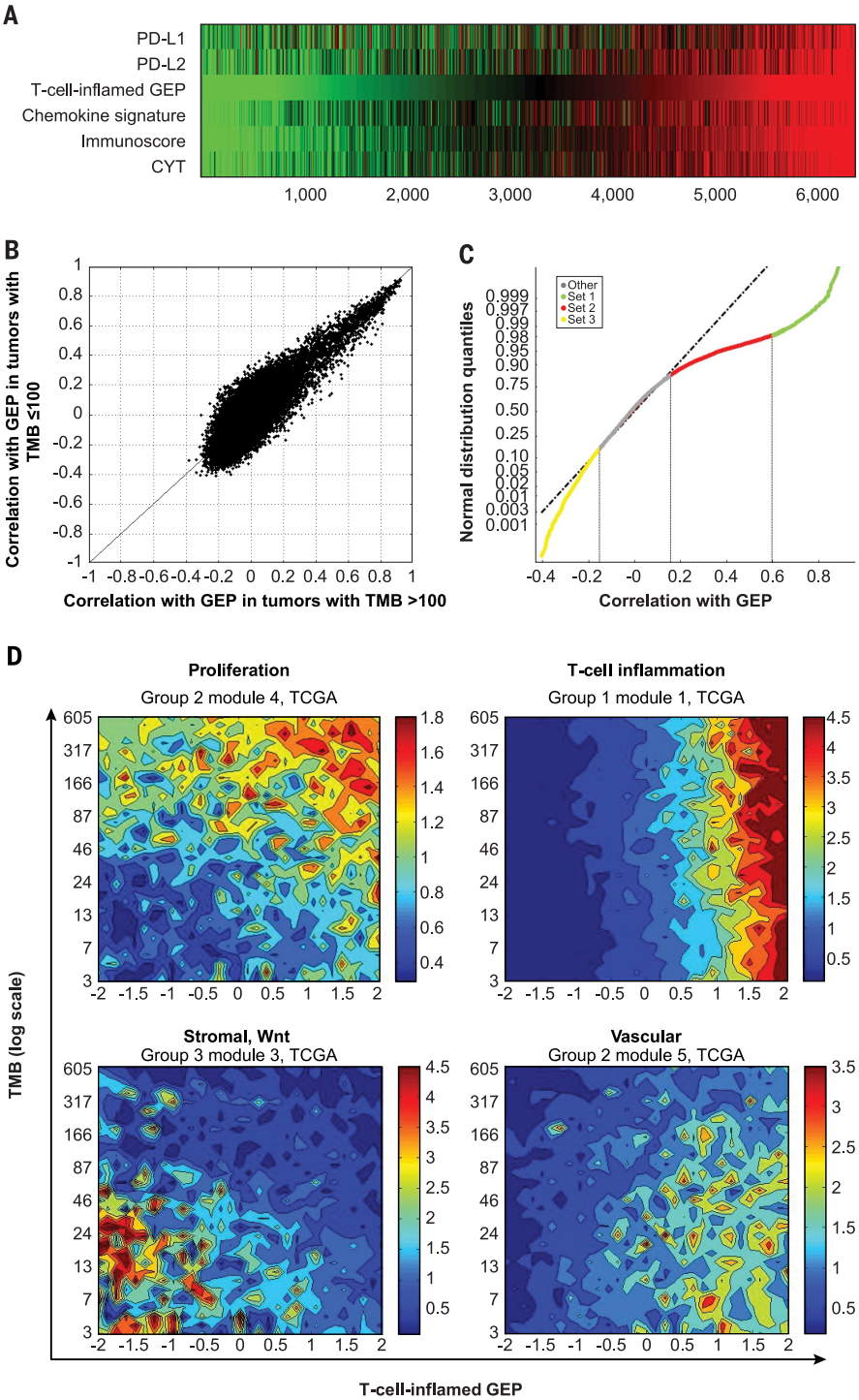
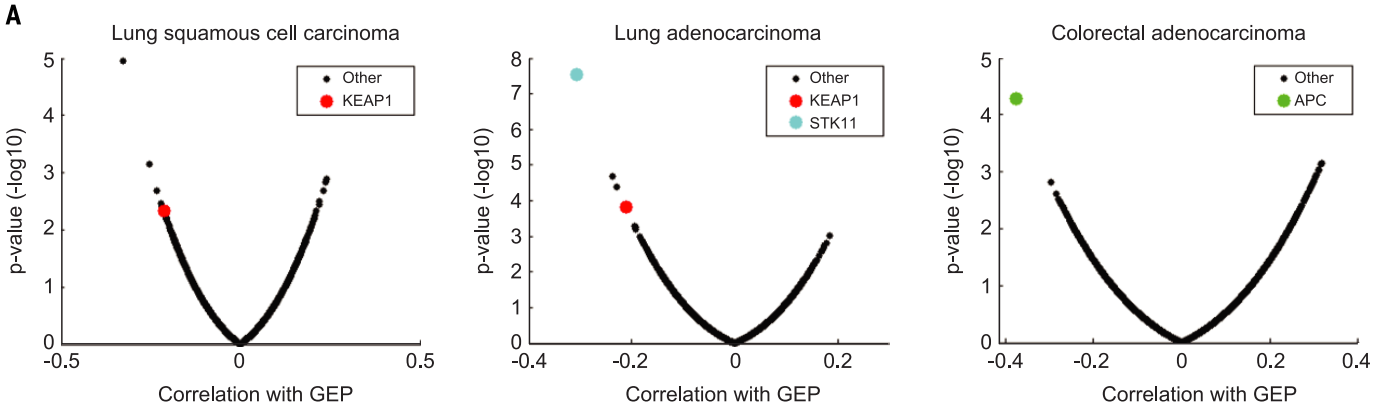


Fig. 5. Transcriptomic and genomic features defined by the GEP and TMB biomarker–based stratification in TCGA database. (**A**) Association of T cell–inflamed GEP (*15*) with other key markers and expression signatures representative of T cell inflammation and a cytolytic environment, including chemokine signature (*29*), Immunoscore (*30*), and cytolytic activity (CYT) (*13*). (**B**) Association between T cell–inflamed GEP and expression of each gene in TCGA for tumors with a TMB of >100 mutations per exome (*x* axis) and in tumors with a TMB of ≤100 mutations per exome (*y* axis). (**C**) Each gene in the transcriptome is assigned to one of four clusters determined by cutoffs obtained from the distribution of correlation with the T cell-inflamed GEP. The cutoffs used were the inflection point where the distribution deviates from normal on the positive side (0.15; 83rd quantile), the cut point that selects T cell–inflamed GEP genes (0.6; 98% quantile), and the inflection point where the distribution deviates from normal on the negative side (−0.15; 15th quantile). Vertical lines represent cutoffs for gene sets 1, 2, and 3 ($r > 0.6$, $r = 0.15$ to 0.6, and $r < -0.15$, respectively); gene sets are color coded on the regression line. (**D**) Gene set annotation in each cluster suggested enrichment for biological patterns with distinct relevance for the individual biomarker-based groups. Contour plots illustrate the association with TMB and GEP of selected patterns of TME and cellular biology represented by gene expression modules formed by genes coexpressed in TCGA database. Blue and red represent under- and overexpression, respectively.

A



B

Indication	***STK11***	***APC***	*MYBPC2*	*ALMS1*	*DYNC2H1*	*OR4N3P*	*PTPRB*	*L1TD1*	*BANK1*	*ZNF423*	***KEAP1***	*COL11A1*	*GLTPD2*	*CSMD1*	*TJP3*	*CARD11*
Lung adenocarcinoma	-5.28	0.54	-0.54	0.14	0.25	0.48	0.01	0.21	-0.98	-2.13	-2.1	-0.06	0	-2.01	0	-1.18
Lung SCC	0.04	-0.02	0	-2.51	0.07	0	-1.71	0.97	0.08	0.17	-1.87	0.23	-0.14	0.33	0.13	-0.03
Colorectal adenocarcinoma	1.09	-4.9	0.46	0.17	0.24	0	0.44	-2.18	-0.26	0.62	0.76	0.02	-2.02	0.67	-2.01	0.57
Skin cutaneous melanoma	-0.69	0.11	-3.07	1.23	0.06	-2.22	-2.18	0.15	-2.16	-0.4	0.81	0.17	0.48	-0.16	-0.22	0.41
Liver hepatocellular carcinoma	0	-1.42	0.3	-0.85	-2.35	0	-0.11	0.69	-0.62	-1.44	0.25	-0.87	0	-0.18	0	-2.01
Head and neck SCC	0.03	-0.03	1.07	0.18	0.37	-0.07	0.65	0	0.52	0.1	-0.25	-2.04	0	0.09	0.43	-0.03
Uterine corpus endometrial carcinoma	0.45	0.4	0.13	1.41	0.52	0	1.08	0.78	-0.07	0.36	1.5	0.64	0.11	0.33	1	0.45
Stomach adenocarcinoma	-0.1	0.14	-0.18	0.62	0.66	0	0.13	0.62	-0.22	0.42	0.23	0.19	0	0.79	1.61	-0.03
Cervical squamous cell carcinoma	-0.4	0.14	0.51	-0.09	0.59	-0.32	-0.4	0.43	-0.01	-0.48	0	-1.02	0.39	-0.17	-0.32	-0.62
Bladder urothelial carcinoma	0	0.54	0.33	0.39	0.09	0	-0.3	0	0.92	0.38	0.44	-0.2	0	-0.42	0.68	0.12
Pan-cancer	-0.48	-0.9	-0.91	1.21	0.56	0.03	-0.27	0.54	-1.05	0.07	0.14	0.41	-1.64	1.13	0.17	0.02

Fig. 6. Cancer driver genes associated with immune evasion in selected tumor types. (**A**) Volcano plots of AUROC and rank sum *P* values illustrating the association of somatic SNV mutations with GEP in lung squamous cell carcinoma, lung adenocarcinoma, and colorectal adenocarcinoma in TCGA database. Analysis was restricted to cancer types having >20% of tumors with TMB^{hi} (>100 mutations per exome). For each cancer type, the negative $\log_{10}$-transformed rank sum *P* value between GEP and mutations was calculated for each gene. (**B**) Rank sum *P* values of association between GEP and mutations in selected genes. The selection was made on the basis of a nominal *P* value of <0.01 for negative association with GEP in any cancer type and an alteration frequency of ≥10% in that cancer type. Negative and positive associations are represented in blue and red, respectively. Negative associations for known cancer driver genes are shown in boxes.

of the relationship of PD-L1 and GEP to a T cell–inflamed TME. These observations suggest that using inflammatory biomarkers such as the T cell–inflamed GEP or PD-L1 jointly with TMB may help to identify patients who are responsive to anti–PD-1 therapies. Additional IHC assays have been developed that measure protein markers of a cytolytic T cell environment, and evaluating their performance characteristics in conjunction with TMB in future studies may be useful (*14*, *35*). More broadly, our study demonstrates the orthogonal relationship between universal measures of tumor antigenicity and tumor infiltration that can occur by activated T cells (*14*, *36*–*38*). Although these are upstream and downstream components, respectively, of a robust antitumor T cell response, there is sufficient intervening biology such that biomarkers for each process can provide complementary information.

As an increasing number of PD-1– and PD-L1–based combination regimens show clinical benefit, it will become challenging to determine the relative utility of each regimen for an individual patient. A refined set of biomarker tools that can stratify underlying patterns of tumor immunobiology may enable rational and biology-driven personalization of these various treatment regimens, such as selection of patients with tumors typically less responsive to immunotherapy. Our data demonstrate that TMB and a T cell–inflamed GEP can be used to categorize tumors into discrete subgroups that exhibit distinct patterns of potentially targetable biology to enhance clinical response. These patterns include tumor type–agnostic signatures of proliferative, vascular, myeloid, and stromal biology, as well as tumor type–specific dysregulation of tumor cell–intrinsic signaling pathways. Although the utility of TMB, T cell–inflamed GEP, and PD-L1, as well as other emerging tumor-agnostic biomarkers, will need to be prospectively validated for use in predicting response to various immunotherapy regimens, including combination therapies, the findings reported here suggest a rationale for further exploring the utility of these biomarkers as guides for precision cancer immunotherapy.

Materials and methods

Clinical tumor samples

Associations of TMB and the T cell–inflamed GEP with BOR and PFS were evaluated by using tumor samples from subgroups of patients treated with pembrolizumab in clinical trials who had WES data available. These included a discovery cohort of patients with HNSCC (KEYNOTE-012 B1), a pan-tumor validation cohort (KEYNOTE-012/028), and single-indication cohorts of patients with HNSCC (KEYNOTE-012 B1+B2) and melanoma (KN001 and 006). The discovery cohort included 34 of 297 total enrolled patients with PD-L1–selected (≥1%, modified proportion score or interface pattern, QualTek IHC) (*39*) HNSCC (B1 cohort). The pan-tumor cohort comprised patients with PD-L1–positive (≥1%, modified proportion score or interface pattern, QualTek IHC) (*39*) advanced solid tumors pooled from two multicohort trials, including 39 of 297 total enrolled patients in KEYNOTE-012 (cohorts A, C, and D: triple-negative breast cancer, urothelial cancer, and gastric cancer, respectively) and 80 of 450 total enrolled patients in KEYNOTE-028 (17 of 20 cohorts with anal, biliary, carcinoid, cervical, colorectal, endometrial, esophageal, estrogen receptor–positive human epidermal growth factor receptor-2–negative breast, pancreatic, salivary gland, prostate, small cell lung, thyroid, and vulvar cancers and neuroendocrine tumors, mesothelioma, and leiomyosarcoma). Single-indication cohorts included 107 HNSCC patients from the KEYNOTE-012 PD-L1–positive (≥1%, modified proportion score or interface pattern, QualTek IHC) (*39*) B1 (n = 34) and

PD-L1–unselected B2 (n = 73) cohorts (*40*, *41*) and patients with advanced melanoma from the pembrolizumab arms of the KEYNOTE-001 (n = 30 of 668 total enrolled patients) and KEYNOTE-006 (n = 59 of 834 total enrolled patients) studies (*26*, *42*). Tissue specimens were obtained with the approval of the institutional review boards, and patients provided informed consent [clinical trial registration: KEYNOTE-012 (NCT01848834); KEYNOTE-028 (NCT02054806); KEYNOTE-001 (NCT01295827); KEYNOTE-006 (NCT01866319)].

Clinical end points

BOR was assessed in the discovery HNSCC, pan-tumor, and HNSCC cohorts by central radiology review and in the melanoma cohort by integrated radiology and oncologist assessment. For BOR, a responder was defined as a patient with a partial response (PR) or complete response (CR), and PFS was defined as the time from the start of treatment to documented evidence of progressive disease or death. BOR and PFS were both assessed in the all-patients-as-treated populations, defined as those who had received ≥1 dose of study drug, in each cohort.

Processing of tissue samples

DNA sequencing (WES) and RNA analysis (gene expression profiling) were performed by using FFPE sections of pretreatment tumor samples from the above-listed studies. WES was performed on both germline and tumor samples, and gene expression profiling was performed on tumor samples. With a fresh scalpel, the tissue was either macrodissected from the marked tumor area (tissue containing <20% tumor) or scraped from the entire section and transferred to a 1.5-ml tube containing 200 μl of 100% ethanol.

Gene expression (RNA) profiling: NanoString methodology

The previously described T cell–inflamed GEP was derived by using a stepwise derivation process of discovery, validation, and refinement of candidate gene sets across a wide variety of solid tumors (*15*). The GEP was composed of 18 inflammatory genes related to antigen presentation, chemokine expression, cytolytic activity, and adaptive immune resistance, including *CCL5*, *CD27*, *CD274* (*PD-L1*), *CD276* (*B7-H3*), *CD8A*, *CMKLR1*, *CXCL9*, *CXCR6*, *HLA-DQA1*, *HLA-DRB1*, *HLA-E*, *IDO1*, *LAG3*, *NKG7*, *PDCD1LG2* (*PDL2*), *PSMB10*, *STAT1*, and *TIGIT*. For GEP analysis, total RNA was isolated from 5-μm-thick FFPE sections of tumor tissue fixed on positively charged slides (Ambion RecoverAll total nucleic acid isolation kit for FFPE; catalog no. AM1975) at ALMAC, United Kingdom. Total RNA concentrations were measured using the NanoDrop ND1000 (Thermo Fisher Scientific) in 1.5 μl of test sample.

Gene expression analysis was conducted on the NanoString nCounter gene expression platform (NanoString Technologies, Seattle, WA) as described previously (*15*). Per sample, 50 ng of total RNA was mixed in a final volume of 5 to 7 μl with a 3′-biotinylated capture probe and 5′-reporter probe tagged with a fluorescent barcode, from the desired custom gene expression codeset (HUIMR680_V2_C2406+PLS_SPIKE80_C2765 for Batch 1 and HUIMR800_C3176 for Batch 2), containing probes designed to function as positive and negative hybridization controls. Probes and target transcripts were hybridized overnight at 65°C for 14 to 18 hours as per manufacturers' recommendations. Hybridized samples were run on the NanoString nCounter preparation station by using a high-sensitivity protocol where excess capture and reporter probes were removed and transcript-specific ternary complexes were immobilized on a streptavidin-coated cartridge. The cartridge samples were scanned at maximum resolution by using the nCounter digital analyzer. GEP scores were calculated as a weighted sum of normalized expression values for the 18 genes. Quality control of the gene expression data followed an approach similar to that of the NanoString clinical-grade assay, with the use of joint criteria that assessed the relationships between housekeeping genes and the negative control probes plus a weighted score evaluating the GEP gene counts versus background-subtracted counts. For housekeeping normalization, raw counts for the individual genes were $\log_{10}$ transformed and then normalized by subtracting the arithmetic mean of the $\log_{10}$ counts for a set of 11 housekeeping genes.

WES pipeline

Somatic single-nucleotide variant (SNV) calling

Whole-exome sequence reads were aligned to reference human genome GRCh37 by using bwa mem (*43*) followed by preprocessing steps including duplicate marking, indel realignment, and base recalibration with Picard (v1.114) and GATK (Genome Analysis Toolkit, v2) (*44*) to generate analysis-ready BAM files. MuTect was used to generate somatic SNV calls using default parameters by comparing BAM files from tumor and matched normal samples (*45*). MuTect-called SNVs present in the Single Nucleotide Polymorphism Database (dbSNP, v141) (*46*) but not in the Catalogue of Somatic Mutations in Cancer (COSMIC, v68) (*47*) were filtered out. The SNVs with mutant reads of <4 in tumor samples were also eliminated. TMB for a subject was defined as the sum of somatic nonsynonymous SNVs that passed all the filters described.

HLA class I typing

HLA-I major loci, A, B and C, were typed at four-digit resolution by using OptiType (v1.0) (*48*).

For output typed alleles not found in the NetMHC (v3.4) (*49*) input list, the corresponding supertype was identified for each allele (*50*, *51*) and the supertype-representative allele was used for NetMHC.

SNV annotation and neoantigen detection

Somatic mutations were annotated with VEP (Variant Effect Predictor) (*52*), and nonsynonymous mutations in protein coding regions were counted for TMB. All possible 9-mer peptide sequences with mutated amino acid inside for each nonsynonymous mutation locus were extracted, and binding affinities for patient HLA-A and HLA-B alleles were computed by using NetMHC (v3.4). The 9-mer peptide with the highest binding affinity with the HLA alleles from a nonsynonymous mutation locus was selected as the representative antigen for the mutation. Representative antigens with HLA-A or -B binding affinity of <50 nM were considered neoantigens.

Microsatellite instability (MSI) calling

MSI phenotype was detected by applying mSINGS on WES data from tumor samples (*22*). The stability of each mononucleotide microsatellite locus was evaluated, and the proportion of unstable microsatellite loci was determined as the MSI score. Samples with an MSI score of more than 20% were classified as MSI-high (MSI-H) positive. MSI was confirmed by PCR by using the Promega MSI analysis system, version 1.2.

Mutation signature analysis

Mutational signature analysis was performed by using the deconstructSigs package (v1.6.0) in R that selects the combination of known mutational signatures that can account for the observed mutational profile in each sample (*53*). Exome regions were defined by Agilent Sureselect V5 target region. Only somatic mutations in exome regions were considered, and trinucleotide counts were normalized by the number of times each trinucleotide context was observed in the exome region. Mutational signatures as defined by Alexandrov *et al.* (*54*) and named as *signatures. nature 2013* were the target signature set to be screened. The relationships of these various mutational signatures, including specific nucleotide changes, DNA repair, smoking, neoantigen, TP53, and APOBEC, with BOR and PFS were evaluated in patient samples in the pan-tumor cohort.

Allele-specific copy number and purity estimation

VarScan2 (*55*) output copy number ratio and SNP were input to Sequenza (*56*) to provide a maximum a posteriori estimation for cellularity and segmented allele-specific copy number for each sample.

Clonality

For each sample, MuTect-called somatic SNVs with variant allele frequency information, combined with Sequenza output allele-specific copy number and cellularity estimation, were input to PyClone to estimate cellular prevalence for all somatic SNVs. Mutational clonality was also inferred through the clustering process of PyClone (*57*).

PD-L1 expression

PD-L1 expression levels were evaluated in pretreatment samples by IHC staining by using the PD-L1 IHC 22C3 pharmDx kit (Agilent Technologies) in the pan-tumor and HNSCC cohorts (*39*); expression levels were reported as the CPS, defined as the number of PD-L1–positive cells (tumor cells, lymphocytes, macrophages) divided by the total number of tumor cells × 100. CPS was previously reported as a percentage and is now

reported as an equivalent unitless measure. This assay differs from the one used to determine PD-L1 positivity (≥1%, modified proportion score or interface pattern, QualTek IHC) for enrollment eligibility as described above for the pan-tumor and HNSCC clinical cohorts (*58*). For the melanoma cohort, PD-L1 levels were assessed by IHC by using the MEL score, and positivity was defined as a score of ≥2 membranous PD-L1 staining in at least 1% of tumor and tumor immune cells (*59*).

TCGA molecular data

Gene expression data for 9963 tumors and somatic alterations data for 6384 tumors were obtained through TCGA portal (*16*) as of September 2015.

Statistical methods

The retrospective, statistical analysis of clinical samples in this study was prespecified and performed in a blinded fashion, with genomic end points generated without access to clinical outcomes. Associations with BOR were tested by using logistic regression, and associations with PFS were examined by using Cox proportional hazards models. All models (logistic regression and Cox models) were adjusted for baseline Eastern Cooperative Oncology Group (ECOG) score performance. One-sided nominal P values were reported. Associations between continuous variables were assessed by using Spearman correlation, and associations between continuous variables and binary variables (e.g., BOR) were further assessed by using AUROC and rank sum P values. Statistical analyses and visualizations were performed with Matlab R2010 or with R3.4.1. TMB cutoffs for the pan-tumor and single-indication clinical cohorts were the Youden Index values derived in AUROC analysis. An additional, exploratory, pan-tumor TMB threshold was derived by using TMB and GEP data across each cohort, similar to a previously described method (*20*).

REFERENCES AND NOTES

1. S. J. Lee *et al*., Interferon regulatory factor-1 is prerequisite to the constitutive expression and IFN-gamma-induced upregulation of B7-H1 (CD274). *FEBS Lett.* **580**, 755–762 (2006). doi: 10.1016/j.febslet.2005.12.093; pmid: 16413538
2. C. Lu, P. S. Redd, J. R. Lee, N. Savage, K. Liu, The expression profiles and regulation of PD-L1 in tumor-induced myeloid-derived suppressor cells. *Oncoimmunology* **5**, e1247135 (2016). doi: 10.1080/2162402X.2016.1247135; pmid: 28123883
3. J. M. Taube *et al*., Association of PD-1, PD-1 ligands, and other features of the tumor immune microenvironment with response to anti-PD-1 therapy. *Clin. Cancer Res.* **20**, 5064–5074 (2014). doi: 10.1158/1078-0432.CCR-13-3271; pmid: 24714771
4. A. Snyder *et al*., Genetic basis for clinical response to CTLA-4 blockade in melanoma. *N. Engl. J. Med.* **371**, 2189–2199 (2014). doi: 10.1056/NEJMoa1406498; pmid: 25409260
5. E. M. Van Allen *et al*., Genomic correlates of response to CTLA-4 blockade in metastatic melanoma. *Science* **350**, 207–211 (2015). doi: 10.1126/science.aad0095; pmid: 26359337
6. W. Hugo *et al*., Genomic and transcriptomic features of response to anti-PD-1 therapy in metastatic melanoma. *Cell* **165**, 35–44 (2016). doi: 10.1016/j.cell.2016.02.065; pmid: 26997480
7. D. B. Johnson, M. H. Pollack, J. A. Sosman, Emerging targeted therapies for melanoma. *Expert Opin. Emerg. Drugs* **21**, 195–207 (2016). doi: 10.1080/14728214.2016.1184644; pmid: 27148822
8. N. A. Rizvi *et al*., Mutational landscape determines sensitivity to PD-1 blockade in non-small cell lung cancer. *Science* **348**, 124–128 (2015). doi: 10.1126/science.aaa1348; pmid: 25765070
9. M. Kowanetz *et al*., Tumor mutation load assessed by FoundationOne (FM1) is associated with improved efficacy of atezolizumab (atezo) in patients with advanced NSCLC. *Ann. Oncol.* **27**, 77P (2016). doi: 10.1093/annonc/mdw363.25
10. D. T. Le *et al*., Mismatch repair deficiency predicts response of solid tumors to PD-1 blockade. *Science* **357**, 409–413 (2017). doi: 10.1126/science.aan6733; pmid: 28596308
11. D. T. Le *et al*., PD-1 blockade in tumors with mismatch-repair deficiency. *N. Engl. J. Med.* **372**, 2509–2520 (2015). doi: 10.1056/NEJMoa1500596; pmid: 26028255
12. J. E. Rosenberg *et al*., Atezolizumab in patients with locally advanced and metastatic urothelial carcinoma who have progressed following treatment with platinum-based chemotherapy: A single-arm, multicentre, phase 2 trial. *Lancet* **387**, 1909–1920 (2016). doi: 10.1016/S0140-6736(16)00561-4; pmid: 26952546
13. M. S. Rooney, S. A. Shukla, C. J. Wu, G. Getz, N. Hacohen, Molecular and genetic properties of tumors associated with local immune cytolytic activity. *Cell* **160**, 48–61 (2015). doi: 10.1016/j.cell.2014.12.033; pmid: 25594174
14. P. C. Tumeh *et al*., PD-1 blockade induces responses by inhibiting adaptive immune resistance. *Nature* **515**, 568–571 (2014). doi: 10.1038/nature13954; pmid: 25428505
15. M. Ayers *et al*., IFN-γ-related mRNA profile predicts clinical response to PD-1 blockade. *J. Clin. Invest.* **127**, 2930–2940 (2017). doi: 10.1172/JCI91190; pmid: 28650338
16. The Cancer Genome Atlas, TCGA Data Portal; https://tcga-data.nci.nih.gov/docs/publications/tcga/.
17. L. B. Alexandrov, M. R. Stratton, Mutational signatures: The patterns of somatic mutations hidden in cancer genomes. *Curr. Opin. Genet. Dev.* **24**, 52–60 (2014). doi: 10.1016/j.gde.2013.11.014; pmid: 24657537
18. R. Haddad *et al*., Genomic determinants of response to pembrolizumab in head and neck squamous cell carcinoma. *J. Clin. Oncol.* **35** (Suppl.), 6009 (2017). doi: 10.1200/JCO.2017.35.15_suppl.6009
19. D. B. Johnson *et al*., Targeted next generation sequencing identifies markers of response to PD-1 blockade. *Cancer Immunol. Res.* **4**, 959–967 (2016). doi: 10.1158/2326-6066.CIR-16-0143; pmid: 27671167
20. A. Panda *et al*., Identifying a clinically applicable mutational burden threshold as a potential biomarker of response to immune checkpoint therapy in solid tumors. *JCO Precis. Oncol.* **2017**, 10.1200/PO.17.00146 (2017). pmid: 29951597
21. T. Y. Seiwert *et al*., "Biomarkers predictive of response to pembrolizumab in head and neck cancer (HNSCC)," in *Proceedings of the American Association for Cancer Research Annual Meeting 2018*, Chicago, IL, 14 to 18 April 2018 (American Association for Cancer Research, 2018), abstract no. LB-339.
22. S. J. Salipante, S. M. Scroggins, H. L. Hampel, E. H. Turner, C. C. Pritchard, Microsatellite instability detection by next generation sequencing. *Clin. Chem.* **60**, 1192–1199 (2014). doi: 10.1373/clinchem.2014.223677; pmid: 24987110
23. W. Roh *et al*., Integrated molecular analysis of tumor biopsies on sequential CTLA-4 and PD-1 blockade reveals markers of response and resistance. *Sci. Transl. Med.* **9**, eaah3560 (2017). doi: 10.1126/scitranslmed.aah3560; pmid: 28251903
24. N. McGranahan *et al*., Clonal neoantigens elicit T cell immunoreactivity and sensitivity to immune checkpoint blockade. *Science* **351**, 1463–1469 (2016). doi: 10.1126/science.aaf1490; pmid: 26940869
25. E. B. Garon *et al*., Pembrolizumab for the treatment of non-small-cell lung cancer. *N. Engl. J. Med.* **372**, 2018–2028 (2015). doi: 10.1056/NEJMoa1501824; pmid: 25891174
26. C. Robert *et al*., Pembrolizumab versus ipilimumab in advanced melanoma. *N. Engl. J. Med.* **372**, 2521–2532 (2015). doi: 10.1056/NEJMoa1503093; pmid: 25891173
27. M. R. Neagu, D. A. Reardon, An update on the role of immunotherapy and vaccine strategies for primary brain tumors. *Curr. Treat. Options Oncol.* **16**, 54 (2015). doi: 10.1007/s11864-015-0371-3; pmid: 26454859
28. U. Vaishampayan, Therapeutic options and multifaceted treatment paradigms in metastatic castrate-resistant prostate cancer. *Curr. Opin. Oncol.* **26**, 265–273 (2014). doi: 10.1097/CCO.0000000000000066; pmid: 24626129
29. D. Coppola *et al*., Unique ectopic lymph node-like structures present in human primary colorectal carcinoma are identified by immune gene array profiling. *Am. J. Pathol.* **179**, 37–45 (2011). doi: 10.1016/j.ajpath.2011.03.007; pmid: 21703392
30. J. Galon, H. K. Angell, D. Bedognetti, F. M. Marincola, The continuum of cancer immunosurveillance: Prognostic, predictive, and mechanistic signatures. *Immunity* **39**, 11–26 (2013). doi: 10.1016/j.immuni.2013.07.008; pmid: 23890060
31. M. Łuksza *et al*., A neoantigen fitness model predicts tumour response to checkpoint blockade immunotherapy. *Nature* **551**, 517–520 (2017). pmid: 29132144
32. H. Dai *et al*., A cell proliferation signature is a marker of extremely poor outcome in a subpopulation of breast cancer patients. *Cancer Res.* **65**, 4059–4066 (2005). doi: 10.1158/0008-5472.CAN-04-3953; pmid: 15899795
33. S. Spranger, R. Bao, T. F. Gajewski, Melanoma-intrinsic β-catenin signalling prevents anti-tumour immunity. *Nature* **523**, 231–235 (2015). doi: 10.1038/nature14404; pmid: 25970248
34. K. Yoshihara *et al*., Inferring tumour purity and stromal and immune cell admixture from expression data. *Nat. Commun.* **4**, 2612 (2013). doi: 10.1038/ncomms3612; pmid: 24113773
35. A. C. Huang *et al*., T-cell invigoration to tumour burden ratio associated with anti-PD-1 response. *Nature* **545**, 60–65 (2017). doi: 10.1038/nature22079; pmid: 28397821
36. P. F. Robbins *et al*., Mining exomic sequencing data to identify mutated antigens recognized by adoptively transferred tumor-reactive T cells. *Nat. Med.* **19**, 747–752 (2013). doi: 10.1038/nm.3161; pmid: 23644516
37. N. H. Segal *et al*., Epitope landscape in breast and colorectal cancer. *Cancer Res.* **68**, 889–892 (2008). doi: 10.1158/0008-5472.CAN-07-3095; pmid: 18245491
38. H. Matsushita *et al*., Cancer exome analysis reveals a T-cell-dependent mechanism of cancer immunoediting. *Nature* **482**, 400–404 (2012). doi: 10.1038/nature10755; pmid: 22318521
39. M. Dolled-Filhart *et al*., Development of a companion diagnostic for pembrolizumab in non–small cell lung cancer using immunohistochemistry for programmed death ligand-1. *Arch. Pathol. Lab. Med.* **140**, 1243–1249 (2016). doi: 10.5858/arpa.2015-0542-OA; pmid: 27552095
40. L. Q. Chow *et al*., A phase Ib study of pembrolizumab (pembro; MK-3475) in patients (pts) with human papillomavirus virus (HPV)-positive and negative head and neck cancer (HNC). *Ann. Oncol.* **25** (Suppl. 4), 14 (2014). doi: 10.1093/annonc/mdu438.32
41. T. Y. Seiwert *et al*., Safety and clinical activity of pembrolizumab for treatment of recurrent or metastatic squamous cell carcinoma of the head and neck (KEYNOTE-012): An open-label, multicentre, phase 1b trial. *Lancet Oncol.* **17**, 956–965 (2016). doi: 10.1016/S1470-2045(16)30066-3; pmid: 27247226
42. O. Hamid *et al*., Safety and tumor responses with lambrolizumab (anti-PD-1) in melanoma. *N. Engl. J. Med.* **369**, 134–144 (2013). doi: 10.1056/NEJMoa1305133; pmid: 23724846
43. H. Li, R. Durbin, Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009). doi: 10.1093/bioinformatics/btp324; pmid: 19451168
44. A. McKenna *et al*., The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **20**, 1297–1303 (2010). doi: 10.1101/gr.107524.110; pmid: 20644199
45. K. Cibulskis *et al*., Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. *Nat. Biotechnol.* **31**, 213–219 (2013). doi: 10.1038/nbt.2514; pmid: 23396013
46. S. T. Sherry, M. Ward, K. Sirotkin, dbSNP-database for single nucleotide polymorphisms and other classes of minor genetic variation. *Genome Res.* **9**, 677–679 (1999). pmid: 10447503
47. S. A. Forbes *et al*., COSMIC (the Catalogue of Somatic Mutations in Cancer): A resource to investigate acquired mutations in human cancer. *Nucleic Acids Res.* **38**, D652–D657 (2010). doi: 10.1093/nar/gkp995; pmid: 19906727
48. A. Szolek *et al*., OptiType: Precision HLA typing from next-generation sequencing data. *Bioinformatics* **30**, 3310–3316 (2014). doi: 10.1093/bioinformatics/btu548; pmid: 25143287
49. M. Nielsen *et al*., Reliable prediction of T-cell epitopes using neural networks with novel sequence representations. *Protein Sci.* **12**, 1007–1017 (2003). doi: 10.1110/ps.0239403; pmid: 12717023

50. O. Lund *et al*., Definition of supertypes for HLA molecules using clustering of specificity matrices. *Immunogenetics* **55**, 797–810 (2004). doi: 10.1007/s00251-004-0647-4; pmid: 14963618
51. J. Sidney, B. Peters, N. Frahm, C. Brander, A. Sette, HLA class I supertypes: A revised and updated classification. *BMC Immunol.* **9**, 1 (2008). doi: 10.1186/1471-2172-9-1; pmid: 18211710
52. W. McLaren *et al*., The Ensembl Variant Effect Predictor. *Genome Biol.* **17**, 122 (2016). doi: 10.1186/s13059-016-0974-4; pmid: 27268795
53. R. Rosenthal, N. McGranahan, J. Herrero, B. S. Taylor, C. Swanton, DeconstructSigs: Delineating mutational processes in single tumors distinguishes DNA repair deficiencies and patterns of carcinoma evolution. *Genome Biol.* **17**, 31 (2016). doi: 10.1186/s13059-016-0893-4; pmid: 26899170
54. L. B. Alexandrov *et al*., Signatures of mutational processes in human cancer. *Nature* **500**, 415–421 (2013). doi: 10.1038/nature12477; pmid: 23945592
55. D. C. Koboldt *et al*., VarScan 2: Somatic mutation and copy number alteration discovery in cancer by exome sequencing. *Genome Res.* **22**, 568–576 (2012). doi: 10.1101/gr.129684.111; pmid: 22300766
56. F. Favero *et al*., Sequenza: Allele-specific copy number and mutation profiles from tumor sequencing data. *Ann. Oncol.* **26**, 64–70 (2015). doi: 10.1093/annonc/mdu479; pmid: 25319062
57. A. Roth *et al*., PyClone: Statistical inference of clonal population structure in cancer. *Nat. Methods* **11**, 396–398 (2014). doi: 10.1038/nmeth.2883; pmid: 24633410
58. M. Dolled-Filhart *et al*., Development of a prototype immunohistochemistry assay to measure programmed death ligand-1 expression in tumor tissue. *Arch. Pathol. Lab. Med.* **140**, 1259–1266 (2016). doi: 10.5858/arpa.2015-0544-OA; pmid: 27788043
59. A. I. Daud *et al*., Programmed death-ligand 1 expression and response to the anti-programmed death 1 antibody pembrolizumab in melanoma. *J. Clin. Oncol.* **34**, 4102–4109 (2016). doi: 10.1200/JCO.2016.67.2477; pmid: 27863197

ACKNOWLEDGMENTS

We gratefully acknowledge D. Li, E. Rubin, and J.Yuan for critical review of the manuscript and valuable input and S. Erespe for editorial assistance, all of Merck, Kenilworth, NJ. **Funding:** This work was supported by Merck, Kenilworth, NJ. **Author contributions:** R.C., R.M., M.A., A.L., A.A., J.K.L., T.K.M., and D.K. conceived, designed, or planned the study. R.C., A.A., R.M., E.M., J.Y., X.S., E.R.P., P.A.O., X.Q.L., H.L., M.N., C.Z., J.C., J.K.L., N.I., A.L., M.A., A.J., A.L.W., T.Y.S., T.K.M., J.E.T., A.R., and D.K. contributed to the acquisition, analysis, or interpretation of the data. R.C., R.M., N.I., E.R.P., J.E.T., A.L., and D.K. drafted the manuscript. All authors critically reviewed or revised the manuscript for intellectual content and approved the final version. **Competing interests:** M.A., A.L., J.K.L., T.K.M., E.M., and M.N. are inventors on patent WO/2016/094377A1, submitted by Merck Sharp and Dohme, which covers "System and methods for deriving gene signature biomarkers of response to PD-1 antagonists." E.R.P. has served as a consultant or scientific advisor for AstraZeneca, Bristol-Myers Squibb, Clovis, Eli Lilly, Exelexis, Genentech, Horizon Pharma, Inovio, Novartis, Pfizer, and Roche and has received research funding from Agensys, AstraZeneca, Bristol-Myers Squibb, Merck, Peloton, Pfizer, and Genentech. P.A.O. has served as a consultant for Alexion Pharmaceuticals, Amgen, Bristol-Myers Squibb, Celldex, CytomX Therapeutics, Genentech, and Neon Therapeutics and has received research funding from ARMO BioSciences, AstraZeneca/MedImmune, Bristol-Myers Squibb, Celldex, and Merck. T.Y.S. has received honoraria from Amgen, AstraZeneca, Bayer/Onyx, Bristol-Myers Squibb, Merck, and Merck Serono and also research funding from Boehringer Ingelheim and Genentech/Roche. A.R. has served as a consultant-advisor for Merck, and his institution received research funding from Merck. R.C., R.M., M.A., A.A., E.M., J.Y., X.S., X.Q.L., H.L., M.N., C.Z., J.K.L., A.J., J.C., A.L.W., N.I., T.K.M., J.E.T., A.L., and D.K. are employees or former employees of Merck Sharp and Dohme, a subsidiary of Merck, Kenilworth, NJ, and may hold stock options in the company. **Data and materials availability:** Anonymized WES genomic data from tumor and normal specimens from patients in the KEYNOTE trials included in this study are available through the NCBI Database of Genotypes and Phenotyes (dbGaP) under accession number phs001572.v1.p1. Requests for access to patient-level clinical data from the KEYNOTE trials in this study can be submitted through the EngageZone site (http://engagezone.msd.com/ds_documentation.php) or via email (dataaccess@merck.com) per Merck's data sharing policy; those requests pertaining to the validation of work in the study will be reviewed in an expedited manner.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/362/6411/eaar3593/suppl/DC1
Figs. S1 to S6
Tables S1 to S7

10 November 2017; accepted 20 August 2018
10.1126/science.aar3593

RESEARCH ARTICLE SUMMARY

PROTEIN HOMEOSTASIS

Dimerization quality control ensures neuronal development and survival

Elijah L. Mena, Rachel A. S. Kjolby, Robert A. Saxton, Achim Werner, Brandon G. Lew, John M. Boyle, Richard Harland, Michael Rape*

INTRODUCTION: Protein complex formation is at the heart of all metazoan signal transduction networks. Facilitating cellular information flow, modular BTB domains, leucine zippers, or coiled coils have been reused in many proteins, where they often mediate crucial homodimerization events. While mutation of a single allele encoding homodimeric proteins might poison signaling complexes, aberrant heterodimerization between related modules can also inhibit or alter the output of signal transduction cascades. Whether cells detect and eliminate protein complexes of aberrant composition has remained unknown.

The globular BTB domain is found in ~220 human proteins that function as substrate adapters of CUL3 E3 ligases, transcription factors, or membrane channels. Proteins containing homodimeric BTB domains, such as KEAP1, KLHL3, KBTBD8, or BCL6, are essential for metazoan development, and their mutation or aberrant expression causes hypertension, cancer, or neurodegeneration. As it is not understood how organisms control the expression or activity of homodimeric BTB proteins, the BTB domain provides a physiologically important model to dissect the regulation of recurrent interaction modules.

RATIONALE: To identify regulatory mechanisms that impinge on modular interaction domains, we searched for shared binding partners of BTB proteins. Having found an E3 ligase that targets multiple BTB proteins for proteasomal degradation, we used biochemical reconstitution and protein complex engineering to dissect the underlying molecular control mechanism. Finally, we relied on *Xenopus laevis* embryos to study the organismal consequences of aberrant regulation of recurrent protein interaction modules.

RESULTS: Affinity purification and mass spectrometry experiments revealed that many BTB proteins heterodimerize, but also interacted with FBXL17, the substrate adapter of the SCF^{FBXL17} E3 ligase. SCF^{FBXL17} catalyzed the polyubiquitylation of BTB proteins to trigger their proteasomal degradation. SCF^{FBXL17} is therefore a rare example of an E3 ligase that targets a domain shared by many proteins, rather than a specific substrate.

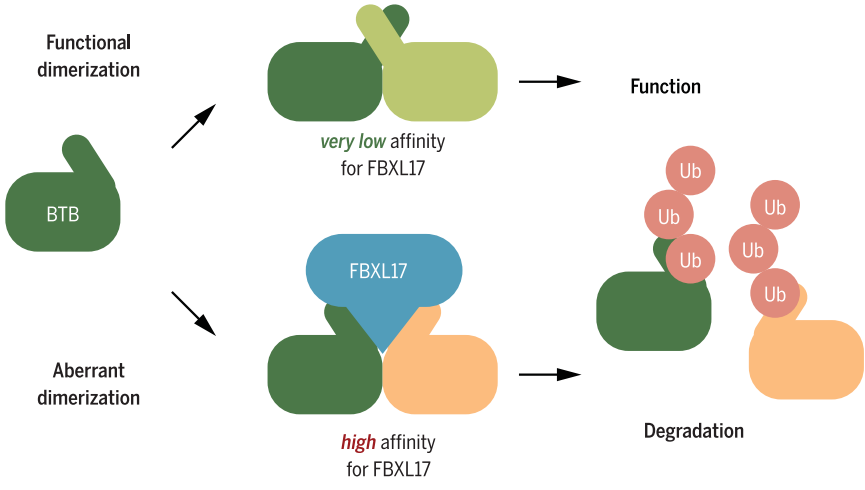
As shown by biochemical reconstitution and affinity purification from cells and animals, SCF^{FBXL17} is a quality control enzyme that detects and ubiquitylates inactive BTB heterodimers, yet ignores active homodimers of the same domains. Accordingly, the loss of FBXL17 increased heterodimerization of BTB proteins, yet at the same time reduced the ability of BTB proteins to engage their downstream targets. SCF^{FBXL17} therefore ensures that only functional BTB dimers are present in cells, an activity that we refer to as dimerization quality control (DQC).

ON OUR WEBSITE

Read the full article at http://dx.doi.org/10.1126/science.aap8236

Depletion of FBXL17 in differentiating human embryonic stem cells showed that DQC prevented heterodimerization of KBTBD8, a BTB protein that is an essential regulator of neural crest specification. In line with this observation, the loss of DQC from *Xenopus laevis* embryos interfered with the differentiation, function, and survival of cells of the central and peripheral nervous system, including the neural crest. By contrast, somitogenesis or general body plan formation were initially unaffected. Similar to other quality control networks, the loss of DQC thus caused specific neuronal phenotypes. However, in addition to the known consequence of muted quality control, i.e. premature neuronal death, the effects of aberrant DQC were already observed early during differentiation.

CONCLUSION: We discovered DQC as a surveillance pathway that detects protein complexes of aberrant composition, rather than protein misfolding. We speculate that other recurrent interaction modules, such as leucine zippers or coiled coils, are monitored by similar DQC networks that rely on distinct E3 ligases. The neuronal phenotypes caused by DQC inactivation point to an active role of quality control in fate decisions in the nervous system. During evolution, DQC appeared at the same time as BTB domains multiplied in the vertebrate genome, suggesting that the ability to eliminate inactive heterodimers formed by related BTB domains contributed to the widespread use of this domain as a dimerization module. ■



Dimerization quality control eliminates inactive heterodimers of a recurrent interaction module but leaves functional homodimers intact. SCF^{FBXL17} selectively ubiquitylates inactive BTB dimers, such as BTB heterodimers or dimers containing mutant BTB domains, which triggers their proteasomal degradation. Functional BTB homodimers escape detection by SCF^{FBXL17}.

The list of author affiliations is available in the full article online.
*Corresponding author. Email: mrape@berkeley.edu
Cite this article as E. L. Mena *et al.*, *Science* **362**, eaap8236 (2018). DOI: 10.1126/science.aap8236

RESEARCH ARTICLE

PROTEIN HOMEOSTASIS

Dimerization quality control ensures neuronal development and survival

Elijah L. Mena[1], Rachel A. S. Kjolby[1], Robert A. Saxton[1]*, Achim Werner[2], Brandon G. Lew[1], John M. Boyle[1], Richard Harland[1], Michael Rape[1,3]†

Aberrant complex formation by recurrent interaction modules, such as BTB domains, leucine zippers, or coiled coils, can disrupt signal transduction, yet whether cells detect and eliminate complexes of irregular composition is unknown. By searching for regulators of the BTB family, we discovered a quality control pathway that ensures functional dimerization [dimerization quality control (DQC)]. Key to this network is the E3 ligase SCF^{FBXL17}, which selectively binds and ubiquitylates BTB dimers of aberrant composition to trigger their clearance by proteasomal degradation. Underscoring the physiological importance of DQC, SCF^{FBXL17} is required for the differentiation, function, and survival of neural crest and neuronal cells. We conclude that metazoan organisms actively monitor BTB dimerization, and we predict that distinct E3 ligases similarly control complex formation by other recurrent domains.

As revealed by large-scale affinity-purification and mass spectrometry, human proteins associate on average with five partners (*1*), which frequently controls signal transduction by establishing the specificity of kinases, transcription factors, or E3 ligases (*2*, *3*). Modular domains, such as leucine zippers, coiled coils, or BTB domains, are found in functionally diverse proteins and often mediate dimerization. If homodimerization is required for function, mutation of a single allele can poison the complex and disrupt signaling, as seen with the tumor suppressor FBXW7 (*4*). Moreover, recurrent interaction modules may retain affinity for homologous domains in other proteins (*1*, *5*), which can result in heterooligomers with altered signaling output. Whether and how cells recognize aberrant dimers and prevent their formation is not known.

Providing a model to address this question, the BTB domain is a globular interaction fold that often mediates dimerization (*6*, *7*). The ~200 human BTB domain–containing proteins function as CUL3 E3 ligase subunits, transcription factors, or membrane channels (*8*–*12*). Members of this family, including the transcriptional regulators BCL6 and BACH1 or the E3 ligase subunits KLHL3 and KEAP1, have important roles in metazoan development (*13*–*18*), and their aberrant activity results in a wide range of diseases including hypertension, cancer, and neurodegeneration (*18*–*22*). Proteomic studies suggested that overexpressed BTB proteins heterodimerize with related family members (*1*, *5*), yet endogenous BTB proteins typically act as homodimers (*7*, *23*–*27*). These observations implied that cells prevent aberrant dimerization of BTB proteins, but the nature of this protective mechanism remains unknown.

SCF^{FBXL17} targets multiple BTB proteins

To identify regulators of the BTB family, we subjected 20 BTB proteins to heterologous expression, affinity purification, and mass spectrometry. As seen before (*1*, *5*, *25*), we found that several BTB proteins formed complexes with other members of this family (Fig. 1A), including KLHL12, KEAP1, and KBTBD8, which require homodimerization for function (*23*, *25*, *27*). By subjecting these immunoprecipitations to CompPASS analysis (*1*, *28*), we noticed that many BTB proteins bound the SCF E3 ligase adaptor F-box/LRR-repeat protein 17 (FBXL17) and the SCF core components SKP1 and CUL1, in line with recent high-throughput studies that independently pointed to potential interactions between FBXL17 and BTB proteins (*1*, *5*, *29*, *30*). Reciprocal immunoprecipitation of FBXL17 confirmed its association with BTB proteins, CUL1, and SKP1 (Fig. 1A). Notably, unsupervised clustering of these data revealed that FBXL17 preferentially recognized BTB proteins that engaged in heterodimerization (Fig. 1A).

To validate these interactions, we used CRISPR-Cas9 genome editing to append FLAG epitopes to all *FBXL17* alleles in 293T cells and analyzed FBXL17^{3xFLAG} immunoprecipitates by mass spectrometry or Western blotting. These experiments confirmed that endogenous FBXL17 interacted with many BTB proteins, as well as with CUL1 and SKP1 (Fig. 1B and fig. S1A). Expression of a dominant negative mutant of CUL1, which prevents ubiquitylation of SCF targets, stabilized the interaction between FBXL17 and BTB proteins (Fig. 1C and fig. S1B). By contrast, mutation of a conserved Cys residue of FBXL17 (FBXL17^{C627R}), a variant that was present in the COSMIC database of somatic breast cancer mutations, severely impaired the ability of FBXL17 to bind BTB proteins, but not SKP1 (Fig. 1C and fig. S1B). We conclude that SCF^{FBXL17} specifically binds and thus potentially regulates many BTB domain–containing proteins.

Overexpression of FBXL17 reduced the levels of many BTB proteins, which was rescued by dominant negative CUL1, proteasome inhibition, or mutation of Cys^{627} in FBXL17 (Fig. 1, C and E, and fig. S1C). Increased expression of FBXL17 also reduced the abundance of BTB proteins in complexes with endogenous CUL3 (fig. S1D), which together implied that SCF^{FBXL17} can ubiquitylate BTB proteins to trigger their proteasome-dependent degradation. Consistent with this, overexpressed FBXL17 ubiquitylated the BTB protein KLHL12, as shown by denaturing His-Ub (ubiquitin) pull-down experiments (fig. S1E), while depletion of FBXL17 or expression of dominant negative CUL1 reduced KLHL12 modification (Fig. 1D and fig. S1E). Moreover, deletion of *FBXL17* increased the stability of BTB proteins in cycloheximide chase experiments (Fig. 1F and fig. S2A), without strongly affecting mRNA levels (fig. S2B), and these effects were rescued by introduction of wild-type FBXL17, but not FBXL17^{C627R} (Fig. 1E and fig. S2C). However, overexpression of FBXL17 induced degradation of endogenous BTB proteins with a marked delay (fig. S2D), and deletion of *FBXL17* did not strongly affect the abundance of most endogenous BTB proteins (Fig. 1F). Thus, although SCF^{FBXL17} can induce the degradation of BTB proteins when overexpressed, it likely targets a specific subset of BTB proteins at the endogenous level.

SCF^{FBXL17} is a quality control enzyme for dimeric BTB complexes

To understand how SCF^{FBXL17} selects its substrates, we reconstituted the ubiquitylation of the BTB protein KLHL12 in vitro. After experiments using recombinant KLHL12 failed to produce modification by SCF^{FBXL17}, even if lysate was provided to modify substrate or enzyme, we asked whether SCF^{FBXL17} functions as a quality control enzyme to eliminate defective BTB proteins. As many quality control enzymes detect protein conformations that are transiently populated during folding (*31*), we tested whether SCF^{FBXL17} could recognize nascent BTB proteins that still expose surfaces buried in the mature state. Notably, if SCF^{FBXL17} was present as KLHL12 was being translated, it readily associated with KLHL12 and triggered the ubiquitylation and degradation of this substrate (Fig. 2, A to C). SCF^{FBXL17} also targeted nascent KLHL12 still bound to tRNA, owing to deletion of the stop codon (fig. S3, A to C). By contrast, if FBXL17 was added after synthesis of KLHL12 had been completed, substrate recognition was strongly impaired (Fig. 2, A to C), and even a large excess of purified BTB proteins failed to compete with recognition of nascent KLHL12 by SCF^{FBXL17} (fig. S3D). Time-of-addition experiments showed that wild-type KLHL12 was accessible

[1]Department of Molecular and Cell Biology, University of California, Berkeley, CA 94720, USA. [2]National Institute of Dental and Craniofacial Research (NIDCR), NIH, Bethesda, MD 20892, USA. [3]Howard Hughes Medical Institute, University of California, Berkeley, CA 94720, USA.
*Present address: Department of Molecular and Cellular Physiology, Stanford University School of Medicine, Stanford, CA 94305, USA.
†Corresponding author. Email: mrape@berkeley.edu

to FBXL17 only during a very brief time window during its synthesis, and KLHL12 ubiquitylation was prevented almost immediately after it became competent to bind CUL3 (fig. S3, E and F). SCF^{FBXL17} recognized the same BTB proteins in vitro that were substrates in cells (fig. S3G) and showed the strong preference for K48-linked ubiquitin chains that is expected of SCF ligases triggering proteasomal degradation (fig. S3, H and I). Thus, SCF^{FBXL17} recognizes wild-type BTB proteins only if these are incompletely folded, suggesting that it is a quality control enzyme specializing on BTB-domain containing proteins.

To identify which aspect of BTB structure or function is monitored, we searched for the substrate motif, or degron, recognized by SCF^{FBXL17}. After domain swap experiments showed that the BTB domain was required and sufficient to impose SCF^{FBXL17}-dependent degradation (fig. S4, A and B), we mutated ~50 residues in the BTB domain of KLHL12 and asked whether the resulting variants were protected from FBXL17 in cells. We found that His^{15}, Asp^{34}, and Ala^{60}, whose mutation in KLHL3 causes hypertension (22), were required for degradation of KLHL12 by FBXL17 (Fig. 2D and fig. S4C). The residue corresponding to Ala^{60}, Ser^{59}, was important for the SCF^{FBXL17}-dependent degradation of a BTB domain transcription factor, BCL6 (fig. S4D), while reintroduction of Ala^{60} converted the nontargeted

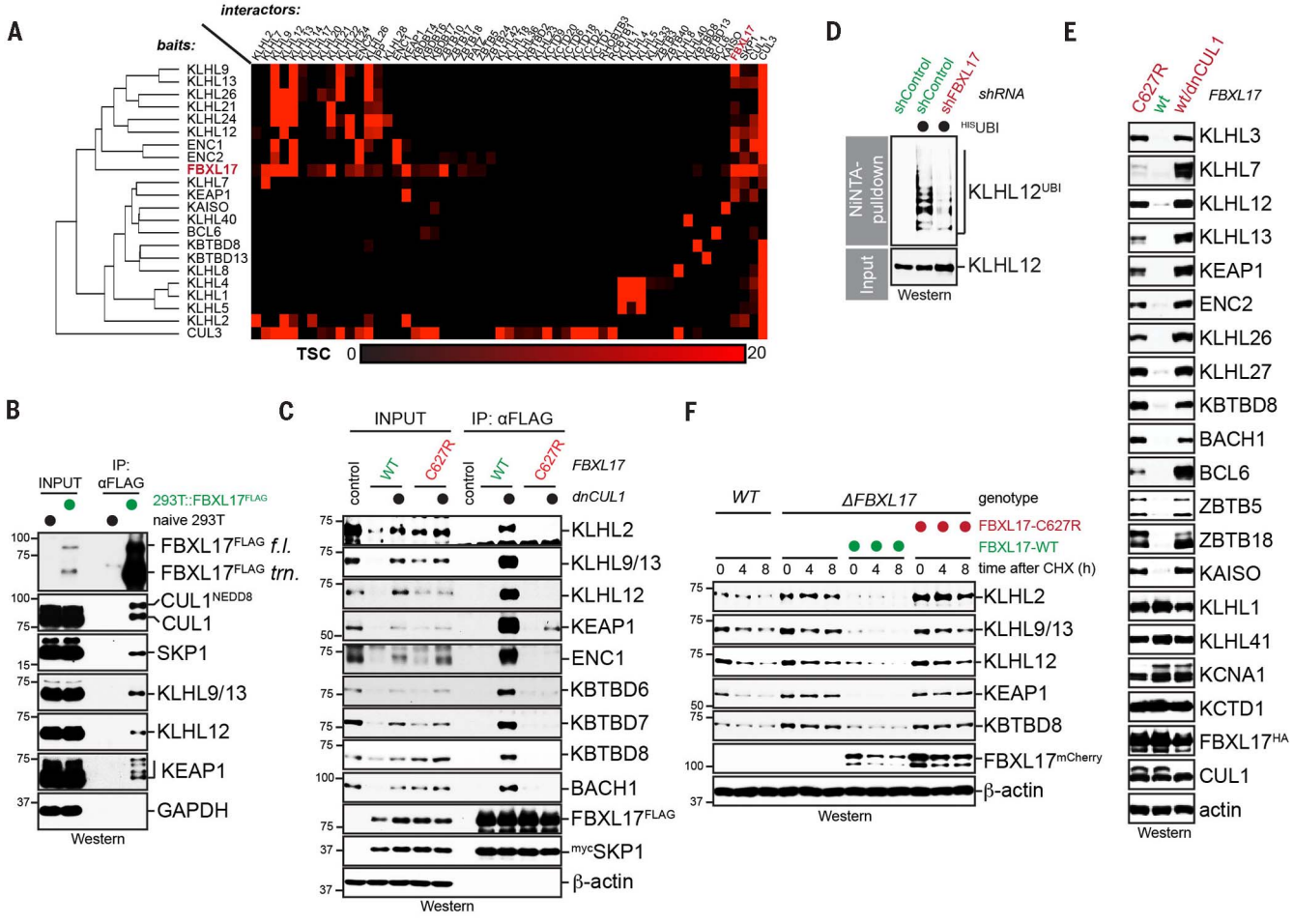


Fig. 1. SCF^{FBXL17} controls the stability and abundance of BTB proteins. (**A**) Proteomic analysis of binding partners of 20 BTB domain–containing proteins reveals FBXL17 as a common interactor. Twenty FLAG-tagged BTB proteins were affinity-purified from 293T cells and analyzed for binding partners by mass spectrometry. BTB proteins, CUL3, and components of SCF^{FBXL17} are shown. Shading corresponds to the abundance (total spectral counts, TSC) of interacting proteins. The reciprocal immunoprecipitation of $FBXL17^{FLAG}$ was from cells also expressing dominant-negative CUL1 to prevent substrate degradation. Bait proteins were ordered by unsupervised clustering. (**B**) Endogenous FBXL17 interacts with many BTB proteins. *FBXL17* loci were tagged in 293T cells with a 3xFLAG epitope by CRISPR-Cas9–dependent genome editing, and endogenous $FBXL17^{3xFLAG}$ was affinity-purified and analyzed by Western blotting for endogenous BTB proteins or components of the SCF machinery. (**C**) SCF^{FBXL17} specifically binds and regulates endogenous BTB proteins. Cells were transfected with $FBXL17^{FLAG}$, $FBXL17^{C627R/FLAG}$, and ^{6myc}SKP1. When shown, dominant negative CUL1 (dnCUL1) was coexpressed. FBXL17 variants were affinity-purified and analyzed for binding to endogenous BTB proteins by Western blotting. (**D**) FBXL17 promotes ubiquitylation of the BTB protein KLHL12 in cells. Cells stably transduced with control short hairpin RNAs (shRNAs) or shRNAs against FBXL17 were transfected with ^{HIS}ubiquitin and treated with proteasome inhibitor MG132. Ubiquitylated endogenous proteins were purified under denaturing conditions and analyzed for modified KLHL12 by Western blotting. (**E**) SCF^{FBXL17} induces the degradation of many, but not all, BTB proteins. Cells were transfected with distinct FLAG-tagged BTB proteins, $FBXL17^{HA}$ or $FBXL17^{C627R/HA}$, and ^{6myc}SKP1. As indicated, dominant negative CUL1 was coexpressed. The abundance of BTB proteins was determined by αFLAG Western blotting. Representative FBXL17, CUL1, and β-actin blots are shown. (**F**) SCF^{FBXL17} controls the stability of BTB proteins. 293T cells lacking functional FBXL17 were engineered by CRISPR-Cas9–dependent genome editing, and cell lines were verified after clonal selection by PCR. As indicated, cells were produced to stably express doxycycline-inducible FBXL17 or $FBXL17^{C627R}$. After 12 hours of induction, cycloheximide was added to measure the stability of endogenous proteins, using Western blotting and specific antibodies.

KLHL41 into an SCF^{FBXL17} substrate (Fig. 2E). In line with these results, mutation of Ala^{60} in KLHL12 or Ser^{59} in BCL6 interfered with the ubiquitylation or degradation of these SCF^{FBXL17} substrates in vitro (Fig. 2, A to C, and fig. S4E). The degron residues were conserved in most of 24 BTB proteins that were turned over by SCF^{FBXL17}, and at least one mutation was present in 20 out of 21 BTB proteins that were resistant to SCF^{FBXL17} (Fig. 2F).

When mapped onto the structure of a BTB domain, all SCF^{FBXL17} degron residues were in close proximity to the dimer interface, without involving residues required for dimerization (Fig. 2G, upper panel). Whereas Ala60 is surface exposed, His^{15} and Asp^{34} are buried in mature BTB dimers resistant to FBXL17 (Fig. 2G, lower panel). Thus, FBXL17 either targets monomeric BTB proteins, or it binds immature dimers that still present His^{15} and Asp^{34} on their surface. To distinguish between these possibilities, we mutated KLHL12 to prevent BTB dimerization (*25*) and found that monomeric KLHL12 was severely impaired in binding to and receiving ubiquitin chains by SCF^{FBXL17} (fig. S5, A and B). An excess of FBXL17 did not destabilize BTB dimers in vitro, which implied that FBXL17 does not interact with BTB monomers (fig. S5, C and D), and FBXL17 strongly bound KLHL12 dimers in cells (fig. S5E). Together, these findings indicated that SCF^{FBXL17} targets dimeric BTB proteins that have not yet matured into their final conformation.

SCF^{FBXL17} specifically targets inactive BTB dimers

To assess whether endogenous SCF^{FBXL17} recognizes nascent BTB dimers in vivo, we exposed cells to the mRNA translation inhibitor cycloheximide and determined binding partners of endogenous FBXL17 by mass spectrometry. FBXL17 interacted with the same BTB proteins if mRNA translation had been prevented (fig. S6A), showing that without overexpression, SCF^{FBXL17} engages most substrates posttranslationally. Other quality control enzymes, such as HSP70, similarly bind nascent wild-type clients cotranslationally, but posttranslationally detect proteins destabilized by mutations or stress (*32*). To test whether this was the case for SCF^{FBXL17}, we assessed whether unfolding or mutation of the BTB domain induced recognition by FBXL17. Whereas unfolding by heat, a condition that eliminates any structural information, did not enable BTB recognition (fig. S6B), specific mutations in the BTB domains of KLHL12 or KBTBD8 allowed for posttranslational binding and ubiquitylation by SCF^{FBXL17} (Fig. 3, A and B, and fig. S6, C and D). Time-of-addition experiments confirmed that these mutants remained accessible to SCF^{FBXL17}

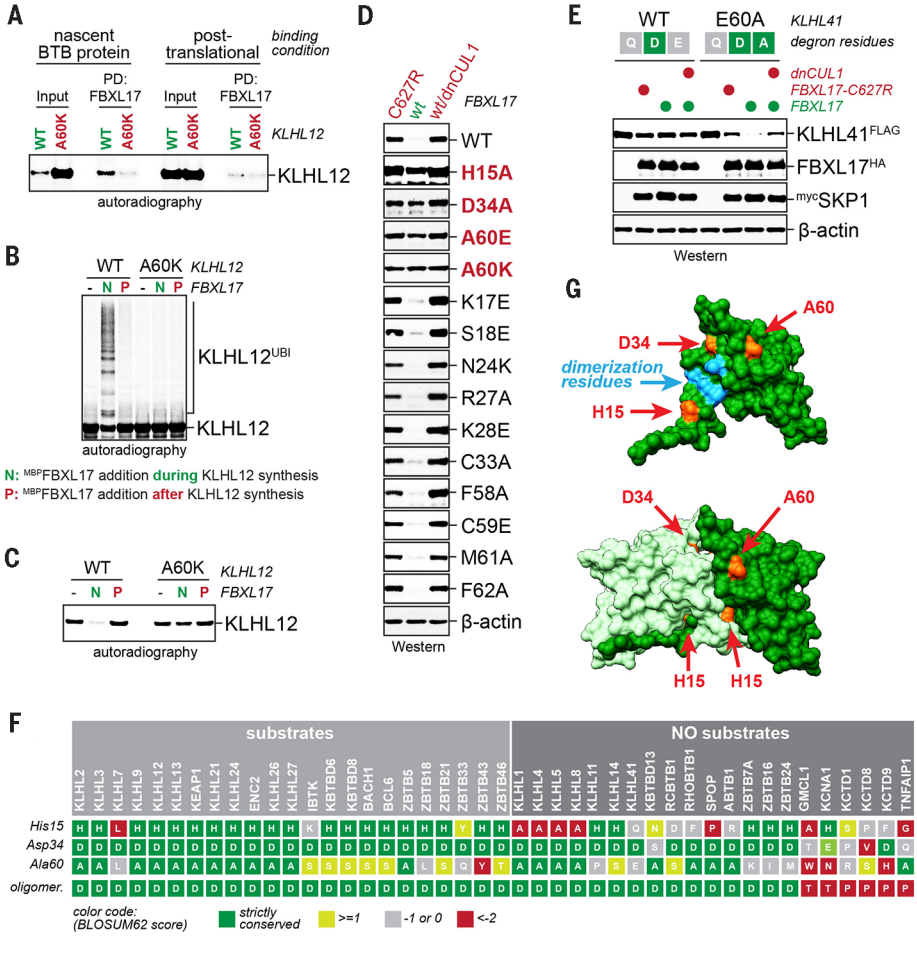


Fig. 2. SCF^{FBXL17} recognizes a conserved degron at the BTB dimer interface. (**A**) FBXL17 recognizes nascent wild-type BTB proteins. Recombinant ^{MBP}FBXL17 was added either during ("N"; nascent BTB protein binding) or after ("P"; posttranslational interaction) the in vitro synthesis of ^{35}S-labeled KLHL12 or KLHL12^{A60K}, immobilized, and analyzed for bound KLHL12 proteins by autoradiography. The input levels of KLHL12 or KLHL12^{A60K} are shown on the left. (**B**) SCF^{FBXL17} ubiquitylates KLHL12, but not KLHL12^{A60K}, cotranslationally in vitro. Recombinant FBXL17-SKP1 was added either during ("N") or after ("P") in vitro synthesis of KLHL12 or KLHL12^{A60K}. Ubiquitylated higher molecular weight KLHL12 species were detected after SDS gel electrophoresis by autoradiography. (**C**) FBXL17 induces cotranslational degradation of wild-type BTB proteins in vitro. ^{35}S-labeled KLHL12 or KLHL12^{A60K} were incubated either during ("N") or after ("P") in vitro synthesis with recombinant FBXL17 in rabbit reticulocyte lysate, which contains proteasomes. KLHL12 levels were analyzed by autoradiography. (**D**) Identification of an SCF^{FBXL17} degron in cells. The indicated KLHL12^{FLAG} variants were expressed in 293T cells together with either FBXL17^{HA} or FBXL17$^{C627R/HA}$. Dominant negative CUL1 (dnCUL1) was coexpressed as indicated. The levels of KLHL12 variants were analyzed by αFLAG-Western. (**E**) Key degron residues are required and sufficient for SCF^{FBXL17}-dependent degradation. KLHL41 or KLHL41^{E60A} were expressed in 293T cells with FBXL17, FBXL17^{C627R}, or FBXL17 and dnCUL1, as indicated. KLHL41 levels were analyzed by Western blotting. (**F**) Key degron residues are conserved in SCF^{FBXL17} targets, but not in BTB proteins that are resistant to FBXL17. Moreover, all FBXL17 targets are dimeric ("D") BTB proteins, whereas tetra- ("T") or pentameric ("P") family members are not recognized. The indicated BTB proteins were expressed in cells together with FBXL17, FBXL17^{C627R}, or FBXL17 and dnCUL1 and analyzed for FBXL17-dependent degradation by Western blotting. Substrates are marked with a green box, whereas BTB proteins resistant to FBXL17 degradation are marked with a red box. The conservation of key degron residues is shown, with the coloring of mutations according to BLOSUM62 matrix scores. (**G**) The SCF^{FBXL17} degron residues are in direct proximity to the dimerization interface and become buried upon completion of BTB dimer formation. Upper panel: Residues whose mutation caused significant stabilization of KLHL12 (red) were mapped onto the crystal structure of a KEAP1 BTB monomer (green; derived from Protein Data Bank: 4CXI). Residues required for dimerization (*25*) are shown in blue. Lower panel: Degron residues H15 and D34 (red) are buried in the folded KEAP1 BTB dimer (light and dark green). Single-letter abbreviations for the amino acid residues are as follows: A, Ala; D, Asp; F, Phe; 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.

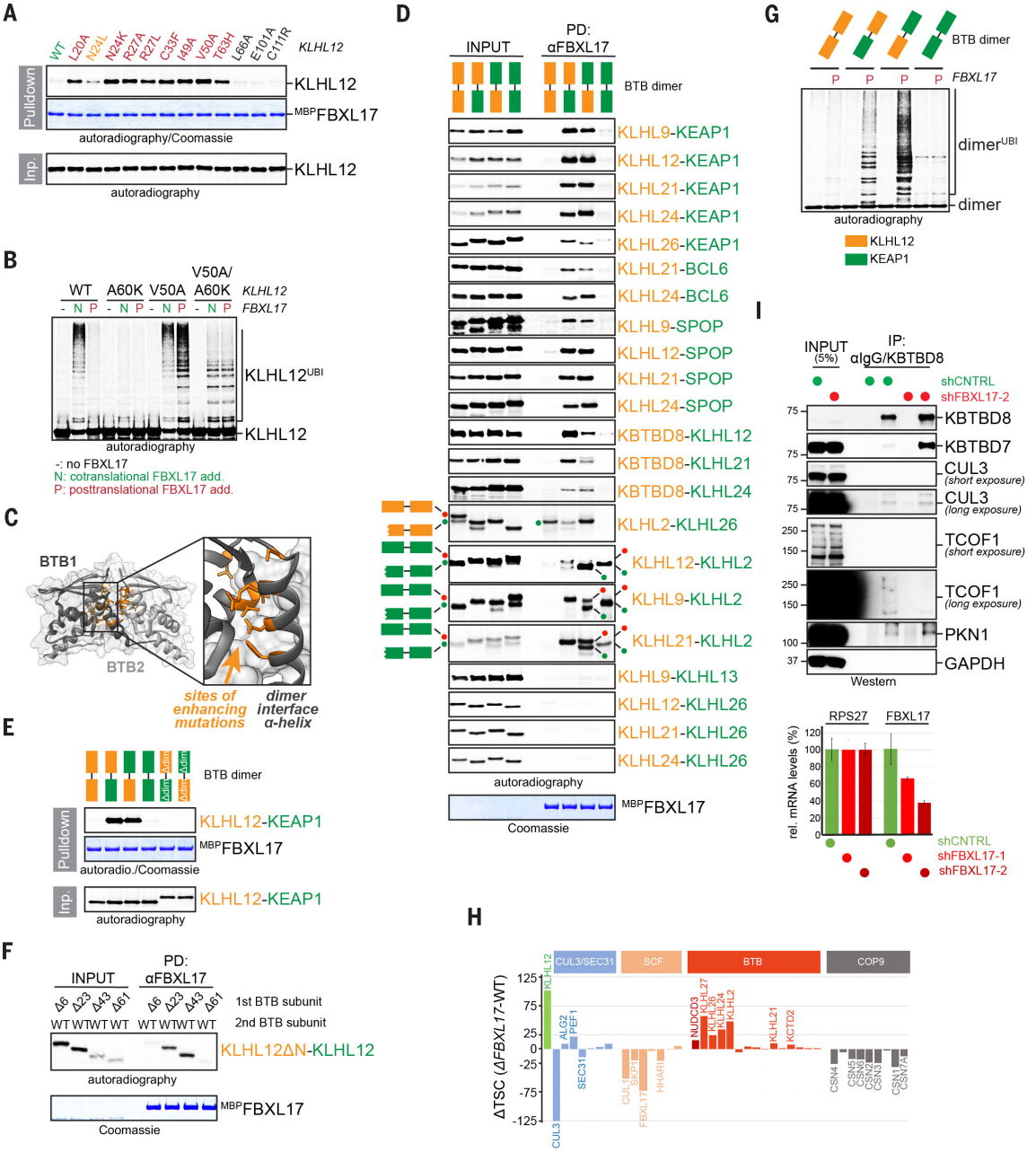


Fig. 3. SCF^{FBXL17} preferentially recognizes aberrant BTB dimers. (**A**) BTB mutations near the dimerization interface enhance substrate recognition by SCF^{FBXL17}. Variants of ^{35}S-labeled KLHL12 were tested for posttranslational binding to immobilized ^{MBP}FBXL17/SKP1 complexes. Binding was detected by autoradiography. (**B**) BTB mutations allow for posttranslational ubiquitylation by SCF^{FBXL17}. Recombinant FBXL17/SKP1 complexes were added either during ("N") or after ("P") synthesis of the indicated KLHL12 variants, and ubiquitylation was detected by SDS gel electrophoresis and autoradiography. (**C**) Structural representation of BTB mutations that allow for posttranslational substrate recognition by SCF^{FBXL17}. Sites of enhancing mutations (orange) were mapped onto the KEAP1-BTB homodimer (light and dark gray). The inset demonstrates that all enhancing mutations map to a central α helix at the interface of two BTB domains; note that an orange residue in the background binds the same α helix of the opposing BTB subunit. (**D**) FBXL17 preferentially binds BTB heterodimers. BTB domain fusions indicated on the right were produced as ^{35}S-labeled proteins and tested for posttranslational binding to ^{MBP}FBXL17-SKP1. Binding was monitored by SDS gel electrophoresis and autoradiography. The green dots mark heterodimers containing a truncated BTB domain of KLHL2 that lacks an amino-terminal region making stabilizing contacts with the opposing BTB subunit; the red dots mark the heterodimers with only full-length KLHL2 BTB domains. (**E**) Formation of a dimer interface between BTB domains is required for substrate binding by SCF^{FBXL17}. Mutations in residues required for BTB dimerization were introduced into heterodimers composed of KEAP1 and KLHL12 BTB domains, as indicated (Δdim). ^{35}S-labeled fusions were then tested for posttranslational binding to ^{MBP}FBXL17/SKP1, as described above. (**F**) Amino-terminal truncations that remove residues required for stabilization of the BTB dimer interface induce posttranslational recognition by SCF^{FBXL17}. Amino-terminal truncations were introduced into the first BTB domain of KLHL12-KLHL12 homodimers, as indicated. Posttranslational binding to ^{MBP}FBXL17-SKP1 was then analyzed as described above. (**G**) SCF^{FBXL17} preferentially ubiquitylates BTB heterodimers. Different combinations of the BTB domains of KLHL12 and KEAP1 were used to produce the forced dimers indicated above. BTB dimers were synthesized as ^{35}S-labeled proteins in vitro and incubated with recombinant FBXL17-SKP1 in reticulocyte lysate. Ubiquitylation was monitored by SDS gel electrophoresis and autoradiography. (**H**) Deletion of *FBXL17* stabilizes BTB heterodimers but reduces binding to substrate. All *KLHL12* alleles in 293T wild-type (WT) or Δ*FBXL17* cells were fused to FLAG-epitopes using CRISPR-Cas9-dependent genome editing. Endogenous KLHL12^{FLAG} was affinity-purified from both cell types and analyzed for binding partners by CompPASS mass spectrometry. The difference in total spectral counts (TSC) between Δ*FBXL17* and WT cells is shown. (**I**) Depletion of FBXL17 from differentiating hESCs induces heterodimer formation of KBTBD8. hESCs were stably transfected with control shRNAs or shRNAs depleting FBXL17. Cells were induced to undergo neural conversion, and at the time of neural crest specification (NC1), endogenous KBTBD8 was affinity-purified. Bound proteins were analyzed by Western blotting using specific antibodies. The efficiency of FBXL17 depletion was analyzed by quantitative reverse transcription (qRT)–PCR (below); RPS27 levels were tested as specificity control.

long after their synthesis had been completed (fig. S3, E and F). Although conservative mutations enhanced recognition by SCF^{FBXL17}, drastic substitutions that likely interfere with BTB folding, or unfolding by heat, obliterated this effect (fig. S6, B to D). Recognition and ubiquitylation of mutant BTB proteins was specific and relied on the SCF^{FBXL17} degron identified above (Fig. 3B and fig. S6E), and as shown for $KLHL12^{V50A}$, also resulted in more efficient detection by FBXL17 in cells (fig. S6F).

All mutations that enabled posttranslational recognition by SCF^{FBXL17} clustered around an α helix at the BTB dimer interface (Fig. 3C). As shown for $KLHL12^{V50A}$, these mutants formed stable dimers that did not noticeably dissociate within 24 hours (fig. S5, C and D), yet were strongly impaired in binding to substrate, i.e., SEC31 (fig. S6, G and H). Thus, SCF^{FBXL17} binds preferentially to BTB proteins that form inactive heterodimers (Fig. 1A), and conversely, mutations that inactivate BTB dimers promote posttranslational recognition by SCF^{FBXL17} (fig. S6, G and H). Together, these observations suggested that SCF^{FBXL17} might detect and eliminate aberrant BTB dimers, such as heterodimers, yet spare functional homodimers.

To test this hypothesis, we developed a system to control the composition of BTB dimers. Guided by structural analyses (*7*), we used a short linker to fuse the carboxy terminus of one BTB domain to the amino terminus of a second so that the high local BTB concentration competes off BTB proteins in trans and thus establishes defined BTB dimers. A fused homodimer of KLHL12 BTB domains was co-translationally recognized by SCF^{FBXL17} with the same efficiency as BTB domains that dimerize in trans (fig. S7A). Mutation of the degron residues Asp^{34} or Ala^{60} in one position of the fused dimer reduced binding by SCF^{FBXL17}, whereas loss of both degrons abolished its recognition (fig. S7, A and B). Notably, mutations that prevented dimerization of unfused BTBs also disrupted the recognition of fused BTB domains (fig. S5A), showing that even in the context of covalently linked BTB domains, formation of a dimer interface is required for substrate binding by SCF^{FBXL17}.

Having established this platform, we generated 55 distinct homo- or heterodimers of BTB domains and tested for posttranslational recognition by FBXL17. Notably, although FBXL17 readily bound many heterodimers composed of distinct BTB domains, it did not readily detect any homodimer (Fig. 3D). As revealed by mutation of dimerization residues in KEAP1-KLHL12 BTB heterodimers, formation of a dimer

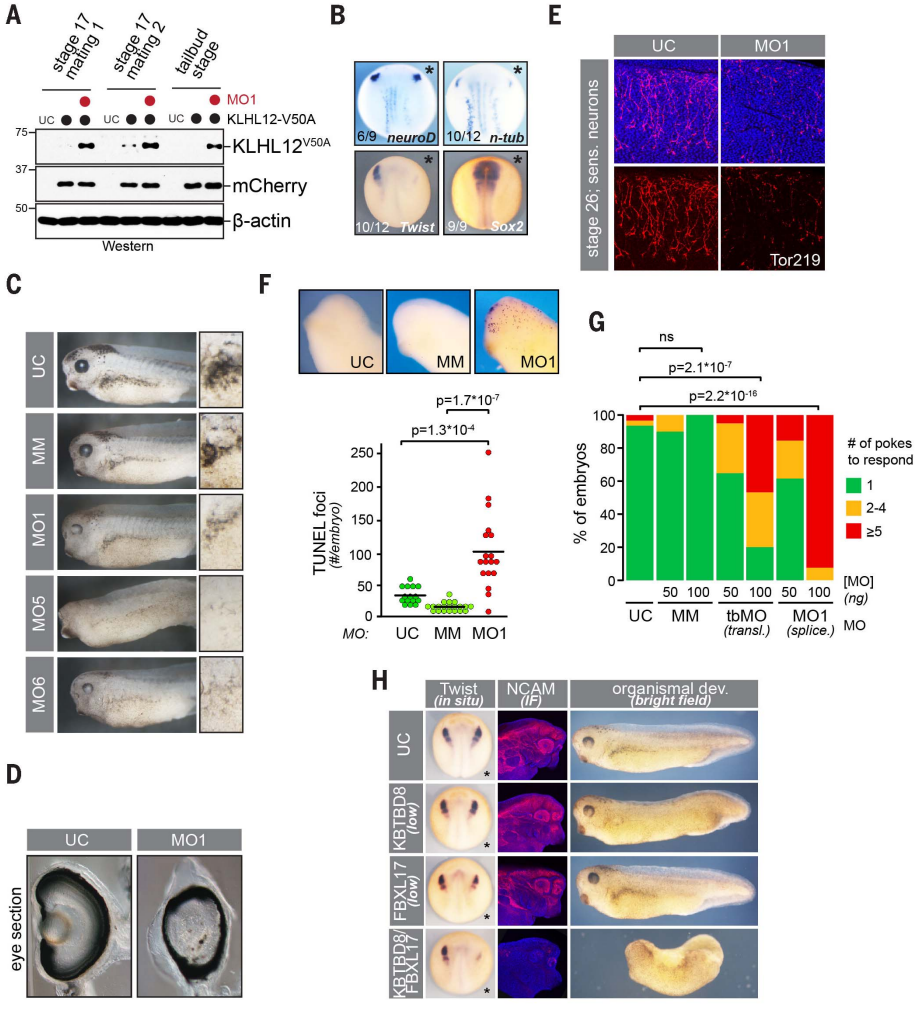


Fig. 4. FBXL17 is required for neuronal differentiation, survival, and function. (**A**) SCF^{FBXL17} functions in intact *X. laevis* embryos. Embryos derived from independent matings were injected with mRNA encoding $KLHL12^{V50A/3xFLAG}$, mCherry, and as indicated, morpholino (MO) 1 against *FBXL17*. Embryos were allowed to develop to stage 17 or the tailbud stage, lysed as pools of five embryos, and analyzed for levels of $KLHL12^{V50A}$ by Western blotting. UC: uninjected control. (**B**) Two-cell-stage *X. laevis* embryos were injected with morpholinos targeting FBXL17 (MO1: *n-tub; neuroD; Sox2*; MO5: *Twist*) on the side marked with an asterisk and allowed to develop to stage 18. In situ hybridization was used to monitor expression of developmental markers. (**C**) Depletion of FBXL17 impairs melanocyte differentiation. *X. laevis* embryos were injected with a mismatch control (MM) or three different morpholinos targeting FBXL17 and allowed to develop to the tailbud stage. Insets show the same embryonic area that is rich in melanocytes in uninjected (UC) or MM-injected embryos, but reduced in FBXL17-depleted embryos. (**D**) FBXL17 depletion impairs eye development. *X. laevis* embryos were uninjected (UC) or injected with MO1 targeting FBXL17, and eyes were sectioned and imaged. (**E**) FBXL17 is essential for neuronal development. Uninjected (UC) *X. laevis* embryos or those injected with MO1 against FBXL17 were stained for mechanosensory neurons with Tor219 antibodies [blue: DAPI (4′,6-diamidino-2-phenylindole); red: Tor219]. (**F**) Depletion of FBXL17 triggers apoptosis in heads of embryos. Apoptotic cells were detected by TUNEL staining in embryos left uninjected (UC) or those injected with either a mismatch (MM) control morpholino or an FBXL17-targeting morpholino (MO1). (**G**) FBXL17 is required for the escape response in *X. laevis* embryos. Embryos were left uninjected (UC) or were injected with a mismatch control morpholino (MM) or with two different morpholinos that either prevent translation (tbMO) or splicing (MO1) of FBXL17. Embryos were poked with a thin pipette tip to elicit a movement described as the escape response. Shown is the number of pokes required to elicit a response. Typically, 15 to 30 embryos were injected and measured per condition. (**H**) FBXL17 maintains KBTBD8 activity in developing *X. laevis* embryos. Morpholinos against FBXL17 and KBTBD8 were titrated so that only minor effects on neural crest development were observed, before both morpholinos were combined to test for synergistic effects of FBXL17 and KBTBD8 co-depletion. As seen by in situ hybridization against Twist, immunofluorescence analysis against NCAM, and bright-field microscopy of whole embryos, co-depletion of FBXL17 and KBTBD8 reveals their strong genetic interaction during neural crest development.

interface was essential for posttranslational heterodimer recognition by SCF^{FBXL17} (Fig. 3E). A single SCF^{FBXL17} degron was sufficient to mediate heterodimer recognition: Even though the BTB domain of SPOP lacks a key degron residue and is not recognized by itself (Fig. 2F), it induces binding by SCF^{FBXL17} when paired with BTB domains of KLHL9, KLHL12, KLHL21, or KLHL24 (Fig. 3D). SCF^{FBXL17} also detected heterodimers formed by distinct variants of the same BTB domain: In fusions of KLHL2 BTB domains, FBXL17 bound a truncated $KLHL2^{\Delta N}$-KLHL2 heterodimer, but not the more abundant full-length homodimer present in the same reaction (Fig. 3D); this specificity could be recapitulated by heterodimers between truncated and full-length BTB domains of a different protein, KLHL12 (Fig. 3F). Although SCF^{FBXL17} efficiently detects inactive BTB complexes, we found that it ignores functional heterodimers: Consistent with KLHL9 and KLHL13 binding each other to control mitosis (*33*), heterodimers composed of the KLHL9 and KLHL13 BTB domains escaped detection by SCF^{FBXL17} (Fig. 3D). We also noticed that KLHL26, which is present in many BTB affinity purifications (Fig. 1A), frequently formed heterodimers that were not recognized by FBXL17 (Fig. 3D), and we speculate that KLHL26 might establish BTB heterodimers with cellular activity. Consistent with these binding studies, SCF^{FBXL17} ubiquitylated BTB heterodimers, but not the respective homodimers (Fig. 3G), which required the SCF^{FBXL17} degron in at least one subunit (fig. S7, C and D). Together, these findings revealed that SCF^{FBXL17} indeed detects and ubiquitylates aberrant BTB dimers but ignores their functional counterparts.

To test whether SCF^{FBXL17} functions similarly in vivo, we used mass spectrometry to probe the dimerization status of FLAG-tagged endogenous KLHL12, an SCF^{FBXL17} substrate that needs to homodimerize to bind its substrate (*25*). Notably, *FBXL17* deletion increased the abundance of KLHL12 in heterodimers with KLHL27, KLHL26, KLHL24, and KLHL2 (Fig. 3H). At the same time, loss of FBXL17 reduced binding of KLHL12 to its substrate SEC31, suggesting that lack of SCF^{FBXL17} activity decreased overall KLHL12 activity. We made similar observations by affinity purification of endogenous KBTBD8, which formed heterodimers with KBTBD7 and ZBTB32 if *FBXL17* had been deleted, or KEAP1, which in the absence of FBXL17 formed heterodimers with KLHL2 (fig. S7, E and F). Similar to these experiments in 293T cells, SCF^{FBXL17} also counteracted heterodimerization in physiologically relevant settings: Depletion of FBXL17 strongly increased the heterodimerization of KBTBD8 with KBTBD7 in human embryonic stem cells instructed to differentiate into a neural crest fate (Fig. 3I), which is dependent upon KBTBD8 function (*15*). Depletion of FBXL17 also reduced the association of KBTBD8 with its key substrate, TCOF1, an interaction that requires KBTBD8 dimerization (*27*), whereas the dimerization-independent binding of KBTBD8 to CUL3 or PKN1 was unaffected. Thus, SCF^{FBXL17} eliminates BTB dimers of aberrant composition in cells. We propose to refer to the ability of SCF^{FBXL17} to detect and eliminate faulty dimers as dimerization quality control, or DQC.

SCF^{FBXL17} ensures nervous system development and function

We next wished to determine whether DQC, similar to other quality control networks, plays a role in organismal development or homeostasis and thus depleted FBXL17 from *Xenopus laevis* embryos (fig. S8A). We injected morpholino oligonucleotides that target either the start codon or four distinct exon-intron splice boundaries of FBXL17 into one cell of two-cell *X. laevis* embryos, thereby depleting FBXL17 by preventing mRNA translation or splicing. We confirmed the efficiency of splice site–directed morpholinos by quantitative polymerase chain reaction (qPCR) (fig. S8B) and analyzed uninjected embryos, untreated sides of injected embryos, or embryos injected with a mismatch morpholino as controls. All phenotypes described below were observed with multiple morpholinos depleting FBXL17, but not with ineffective morpholinos or any control approach.

Establishing that SCF^{FBXL17} prevents accumulation of aberrant BTB proteins in *X. laevis* embryos, down-regulation of FBXL17 resulted in a marked increase in the levels of coinjected $KLHL12^{V50A}$ (Fig. 4A), a prototypic SCF^{FBXL17} substrate in our cellular studies. Accumulation of $KLHL12^{V50A}$ was observed in early stage 17 embryos and the later tailbud stage. As in 293T cells or differentiating human embryonic stem cells (hESCs), depletion of FBXL17 also stabilized KBTBD7-KBTBD8 heterodimers in *X. laevis* embryos (fig. S8C). The inactivation of SCF^{FBXL17} had consequences for neuronal differentiation, as markers for neuronal identity (*n-Tubulin; Neuro D*) were strongly reduced, whereas those of early neural progenitors (*Sox2*) were unaffected (Fig. 4B and fig. S8D). The epithelial mesenchymal transition marker Twist was reduced; melanocyte specification and migration were diminished, indicative of impaired neural crest differentiation; and eye development was also impaired (Fig. 4, B to D). Analysis of Neurofilament, NCAM, or the Tor219 marker of sensory neurons revealed inefficient production of Rohon-Beard neurons and deficits at later stages of neuronal development (Fig. 4E and fig. S8, E and F), and TUNEL (terminal deoxynucleotidyl transferase–mediated deoxyuridine triphosphate nick end labeling) staining documented increased cell death specifically in the heads of embryos (Fig. 4F). In accordance with aberrant nervous system development, depletion of FBXL17 resulted in a strong and dose-dependent loss of the tadpole response to gentle touch (Fig. 4G and Movies 1 and 2). Loss of SCF^{FBXL17}, however, did not affect somitogenesis (fig. S8G), and FBXL17-depleted embryos initially did not show gross morphological changes (fig. S8H). Consistent with FBXL17 being expressed in the neuro-ectoderm of *X. laevis* embryos (fig. S8I), these observations thus revealed that DQC is particularly important for the differentiation, function, and survival of cells in the peripheral and central nervous system.

To assess whether phenotypes of aberrant DQC were caused by misregulation of specific BTB proteins, we tested for a genetic interaction between FBXL17 and KBTBD8 in neural crest specification. KBTBD8 is a prototypic DQC substrate: Although SCF^{FBXL17} posttranslationally bound heterodimers containing the KBTBD8 BTB domain, it did not detect KBTBD8 homodimers (Fig. 3D); KBTBD8 showed increased heterodimerization in 293T cells, differentiating hESCs, or *X. laevis* embryos lacking FBXL17 (Fig. 3I and fig. S7, E and F); and the heterodimerization in hESCs lacking FBXL17 was accompanied by a decrease of KBTBD8-substrate binding (Fig. 3I). We first titrated morpholinos against KBTBD8 or FBXL17 to cause only very mild neural crest phenotypes and then combined both morpholinos: Additive effects of co-depletion would indicate independent functions of KBTBD8 and FBXL17; rescue of KBTBD8 phenotypes by FBXL17 co-depletion would suggest that FBXL17 degrades and inactivates KBTBD8; yet, synthetic lethality

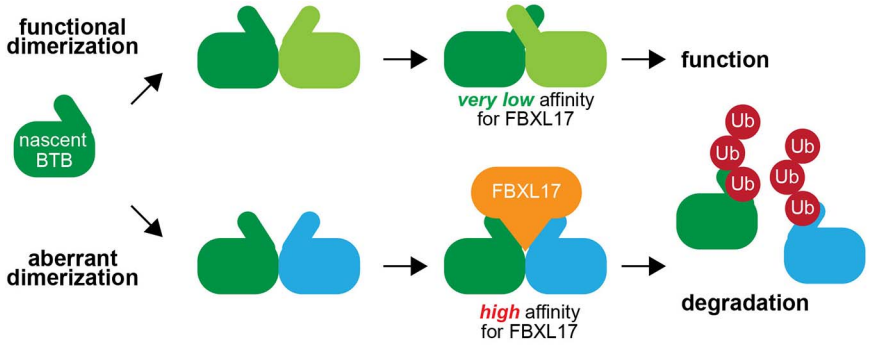


Fig. 5. Model of dimerization quality control (DQC). BTB domain–containing proteins rapidly dimerize upon synthesis. Wild-type BTB homodimers or functional BTB heterodimers mature quickly to bury essential SCF^{FBXL17}-degron residues. Mature and active dimers therefore escape recognition by SCF^{FBXL17}. By contrast, the maturation of mutant BTB homodimers or aberrant heterodimers into a closed conformation is delayed, and such inactive dimers therefore remain accessible to SCF^{FBXL17}, leading to their ubiquitylation and proteasomal degradation. Thus, DQC detects and eliminates complexes of aberrant composition, rather than targeting misfolded and aggregation-prone proteins.

Movie 1. *Xenopus laevis* embryos respond to touch. Embryos were poked with a pipette tip to evoke a movement referred to as escape response.

Movie 2. *Xenopus laevis* embryos lacking FBXL17 fail to respond to touch. Embryos treated with morpholino oligonucleotides targeting FBXL17 were analyzed for their escape response, as described above.

would be anticipated if quality control by SCF^{FBXL17} ensured KBTBD8 activity. Notably, in situ hybridization against Twist, immunofluorescence against NCAM, and bright-field microscopy all revealed strong synthetic lethality of KBTBD8 and FBXL17 depletion (Fig. 4H). This demonstrates that FBXL17 sustains neural crest differentiation by maintaining the cellular pool of active KBTBD8, which highlights the physiological importance of DQC for specific BTB proteins in embryos.

Discussion

Dimerization quality control monitors a recurrent interaction module

We have identified a quality control system, DQC, that ensures dimerization and function of a widespread protein interaction module, the BTB domain. By monitoring complex composition, rather than folding, DQC complements quality control networks that provide defense against unfolded and aggregation-prone proteins (*31*, *32*, *34*). DQC relies on SCF^{FBXL17}, which selectively ubiquitylates aberrant BTB dimers that arise as a consequence of heterodimerization or BTB mutation (Fig. 5). Our results suggest that SCF^{FBXL17} accesses residues at the interface of BTB subunits that are transiently exposed in nascent BTB dimers or persistently presented by mutant homo- or aberrant heterodimers, yet buried upon successful dimerization, as shown in structures of mature KEAP1 and KLHL3 (*35*, *36*). The mechanism of substrate detection by SCF^{FBXL17} allows for the possibility that cellular conditions, such as oxidative stress (*30*), induce BTB recognition—for example, by post-translational modifications that expose degron residues. Conversely, localized mRNA translation or cotranslational BTB homodimerization might limit heterodimerization and reduce the burden on DQC. Our model of DQC is consistent with the finding that overexpressed FBXL17 degrades wild-type BTB proteins (Fig. 1C), as overexpressed FBXL17 could capture nascent BTB dimers before these have reached their mature conformation.

Specific roles of dimerization quality control in the nervous system

The ability of SCF^{FBXL17} to eliminate compromised BTB dimers ensures that only functional BTB complexes are present in cells, and genetic interaction studies demonstrated the importance of DQC for maintaining active KBTBD8 for neural crest specification. The compromised neuronal differentiation and survival that are caused by aberrant DQC may result from a combinatorial effect of dysregulating multiple BTB proteins at a time. We also do not wish to exclude the possibility that SCF^{FBXL17} has targets in addition to inactive BTB dimers, whose misregulation contributes to the aberrant neuronal function observed in the absence of this E3 ligase. Notably, phenotypes similar to those of FBXL17 depletion have been observed in neurodegenerative diseases, and mutation of the F-box of *FBXL17* has been reported in Parkinson's disease (*37*). However, whether and how *FBXL17* mutations affect dopaminergic neurons remains to be determined.

Implications for the evolution of protein interaction modules

Basal eukaryotes lack a recognizable FBXL17 and accordingly contain few BTB proteins that could heterodimerize; *Drosophila melanogaster* also does not possess a FBXL17 ortholog but uses alternative splicing to link a single BTB domain to multiple proteins for homodimerization (*38*). Conversely, FBXL17 is present in basal metazoans and conserved in most metazoans, which correlates with the rapid expansion of dimeric BTB proteins early in the metazoan lineage (*39*). We speculate that FBXL17-dependent DQC allowed expansion of BTB proteins by ensuring homodimerization despite a steep increase in the number of proteins that are prone to aberrant heterodimerization. In addition to BTB domains, recurrent dimerization modules include leucine zippers, coiled coils, or zinc fingers. We predict that the interaction status of these domains is surveilled by additional branches of the DQC pathway that depend on E3 ligases distinct from SCF^{FBXL17}. Analyzing the coevolution of E3 ligases with dimerization modules could allow us to identify such DQC E3 ligases that we expect to be similarly important for metazoan development as SCF^{FBXL17}.

Materials and methods

Detailed materials and methods can be found in the supplementary materials. In vitro binding and ubiquitylation assays were performed with in vitro–transcribed and –translated substrates, as described (*40*). Cellular experiments used 293T cells, as well as differentiating hESCs subjected to neural conversion (*15*). Animal experiments were performed with embryos of *X. laevis*.

REFERENCES AND NOTES

1. E. L. Huttlin *et al.*, Architecture of the human interactome defines protein communities and disease networks. *Nature* **545**, 505–509 (2017). doi: 10.1038/nature22366; pmid: 28514442
2. N. J. Marianayagam, M. Sunde, J. M. Matthews, The power of two: Protein dimerization in biology. *Trends Biochem. Sci.* **29**, 618–625 (2004). doi: 10.1016/j.tibs.2004.09.006; pmid: 15501681
3. M. A. Lemmon, J. Schlessinger, K. M. Ferguson, The EGFR family: Not so prototypical receptor tyrosine kinases. *Cold Spring Harb. Perspect. Biol.* **6**, a020768 (2014). doi: 10.1101/cshperspect.a020768; pmid: 24691965
4. M. Welcker, B. E. Clurman, FBW7 ubiquitin ligase: A tumour suppressor at the crossroads of cell division, growth and differentiation. *Nat. Rev. Cancer* **8**, 83–93 (2008). doi: 10.1038/nrc2290; pmid: 18094723
5. E. L. Huttlin *et al.*, The BioPlex Network: A Systematic Exploration of the Human Interactome. *Cell* **162**, 425–440 (2015). doi: 10.1016/j.cell.2015.06.043; pmid: 26186194
6. V. J. Bardwell, R. Treisman, The POZ domain: A conserved protein-protein interaction motif. *Genes Dev.* **8**, 1664–1677 (1994). doi: 10.1101/gad.8.14.1664; pmid: 7958847

7. M. Zhuang *et al*., Structures of SPOP-substrate complexes: Insights into molecular architectures of BTB-Cul3 ubiquitin ligases. *Mol. Cell* **36**, 39–50 (2009). doi: 10.1016/j.molcel.2009.09.022; pmid: 19818708
8. L. Aravind, E. V. Koonin, Fold prediction and evolutionary analysis of the POZ domain: Structural and evolutionary relationship with the potassium channel tetramerization domain. *J. Mol. Biol.* **285**, 1353–1361 (1999). doi: 10.1006/jmbi.1998.2394; pmid: 9917379
9. L. Pintard *et al*., The BTB protein MEL-26 is a substrate-specific adaptor of the CUL-3 ubiquitin-ligase. *Nature* **425**, 311–316 (2003). doi: 10.1038/nature01959; pmid: 13679921
10. R. Geyer, S. Wee, S. Anderson, J. Yates III, D. A. Wolf, BTB/POZ domain proteins are putative substrate adaptors for cullin 3 ubiquitin ligases. *Mol. Cell* **12**, 783–790 (2003). doi: 10.1016/S1097-2765(03)00341-1; pmid: 14527422
11. L. Xu *et al*., BTB proteins are substrate-specific adaptors in an SCF-like modular ubiquitin ligase containing CUL-3. *Nature* **425**, 316–321 (2003). doi: 10.1038/nature01985; pmid: 13679922
12. M. Furukawa, Y. J. He, C. Borchers, Y. Xiong, Targeting of protein ubiquitination by BTB-Cullin 3-Roc1 ubiquitin ligases. *Nat. Cell Biol.* **5**, 1001–1007 (2003). doi: 10.1038/ncb1056; pmid: 14528312
13. W. Ci, J. M. Polo, A. Melnick, B-cell lymphoma 6 and the molecular pathogenesis of diffuse large B-cell lymphoma. *Curr. Opin. Hematol.* **15**, 381–390 (2008). doi: 10.1097/MOH.0b013e328302c7df; pmid: 18536578
14. L. Jin *et al*., Ubiquitin-dependent regulation of COPII coat size and function. *Nature* **482**, 495–500 (2012). doi: 10.1038/nature10822; pmid: 22358839
15. A. Werner *et al*., Cell-fate determination by ubiquitin-dependent regulation of translation. *Nature* **525**, 523–527 (2015). doi: 10.1038/nature14978; pmid: 26399832
16. H. Sies, C. Berndt, D. P. Jones, Oxidative Stress. *Annu. Rev. Biochem.* **86**, 715–748 (2017). doi: 10.1146/annurev-biochem-061516-045037; pmid: 28441057
17. H. Cai, A. Liu, Spop promotes skeletal development and homeostasis by positively regulating Ihh signaling. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 14751–14756 (2016). doi: 10.1073/pnas.1612520114; pmid: 27930311
18. K. Igarashi, T. Kurosaki, R. Roychoudhuri, BACH transcription factors in innate and adaptive immunity. *Nat. Rev. Immunol.* **17**, 437–450 (2017). doi: 10.1038/nri.2017.26; pmid: 28461702
19. N. K. Hayward *et al*., Whole-genome landscapes of major melanoma subtypes. *Nature* **545**, 175–180 (2017). doi: 10.1038/nature22071; pmid: 28467829
20. The Cancer Genome Atlas Research Network, The Molecular Taxonomy of Primary Prostate Cancer, *Cell* **163**, 1011–1025 (2015). doi: 10.1016/j.cell.2015.10.025; pmid: 26544944
21. S. Aharoni *et al*., Novel homozygous missense mutation in GAN associated with Charcot-Marie-Tooth disease type 2 in a large consanguineous family from Israel. *BMC Med. Genet.* **17**, 82 (2016). doi: 10.1186/s12881-016-0343-x; pmid: 27852232
22. L. M. Boyden *et al*., Mutations in kelch-like 3 and cullin 3 cause hypertension and electrolyte abnormalities. *Nature* **482**, 98–102 (2012). doi: 10.1038/nature10814; pmid: 22266938
23. K. I. Tong, A. Kobayashi, F. Katsuoka, M. Yamamoto, Two-site substrate recognition model for the Keap1-Nrf2 system: A hinge and latch mechanism. *Biol. Chem.* **387**, 1311–1320 (2006). doi: 10.1515/BC.2006.164; pmid: 17081101
24. L. M. Zipper, R. T. Mulcahy, The Keap1 BTB/POZ dimerization function is required to sequester Nrf2 in cytoplasm. *J. Biol. Chem.* **277**, 36544–36552 (2002). doi: 10.1074/jbc.M206530200; pmid: 12145307
25. C. A. McGourty *et al*., Regulation of the CUL3 Ubiquitin Ligase by a Calcium-Dependent Co-adaptor. *Cell* **167**, 525–538.e14 (2016). doi: 10.1016/j.cell.2016.09.026; pmid: 27716508
26. K. Hatzi *et al*., A hybrid mechanism of action for BCL6 in B cells defined by formation of functionally distinct complexes at enhancers and promoters. *Cell Reports* **4**, 578–588 (2013). doi: 10.1016/j.celrep.2013.06.016; pmid: 23911289
27. A. Werner, R. Baur, N. Teerikorpi, D. U. Kaya, M. Rape, Multisite dependency of an E3 ligase controls monoubiquitylation-dependent cell fate decisions. *eLife* **7**, e35407 (2018). doi: 10.7554/eLife.35407; pmid: 29999490
28. M. E. Sowa, E. J. Bennett, S. P. Gygi, J. W. Harper, Defining the human deubiquitinating enzyme interaction landscape. *Cell* **138**, 389–403 (2009). doi: 10.1016/j.cell.2009.04.042; pmid: 19615732
29. M. Raducu *et al*., SCF (Fbxl17) ubiquitylation of Sufu regulates Hedgehog signaling and medulloblastoma development. *EMBO J.* **35**, 1400–1416 (2016). doi: 10.15252/embj.201593374; pmid: 27234298
30. M. K. Tan, H. J. Lim, E. J. Bennett, Y. Shi, J. W. Harper, Parallel SCF adaptor capture proteomics reveals a role for SCFFBXL17 in NRF2 activation via BACH1 repressor turnover. *Mol. Cell* **52**, 9–24 (2013). doi: 10.1016/j.molcel.2013.08.018; pmid: 24035498
31. F. U. Hartl, A. Bracher, M. Hayer-Hartl, Molecular chaperones in protein folding and proteostasis. *Nature* **475**, 324–332 (2011). doi: 10.1038/nature10317; pmid: 21776078
32. D. Balchin, M. Hayer-Hartl, F. U. Hartl, In vivo aspects of protein folding and quality control. *Science* **353**, aac4354 (2016). doi: 10.1126/science.aac4354; pmid: 27365453
33. I. Sumara *et al*., A Cul3-based E3 ligase removes Aurora B from mitotic chromosomes, regulating mitotic progression and completion of cytokinesis in human cells. *Dev. Cell* **12**, 887–900 (2007). doi: 10.1016/j.devcel.2007.03.019; pmid: 17543862
34. K. Yanagitani, S. Juszkiewicz, R. S. Hegde, UBE2O is a quality control factor for orphans of multiprotein complexes. *Science* **357**, 472–475 (2017). doi: 10.1126/science.aan0178; pmid: 28774922
35. A. X. Ji, G. G. Privé, Crystal structure of KLHL3 in complex with Cullin3. *PLOS ONE* **8**, e60445 (2013). doi: 10.1371/journal.pone.0060445; pmid: 23573258
36. A. Cleasby *et al*., Structure of the BTB domain of Keap1 and its interaction with the triterpenoid antagonist CDDO. *PLOS ONE* **9**, e98896 (2014). doi: 10.1371/journal.pone.0098896; pmid: 24896564
37. C. Kun-Rodrigues *et al*., A systematic screening to identify de novo mutations causing sporadic early-onset Parkinson's disease. *Hum. Mol. Genet.* **24**, 6711–6720 (2015). doi: 10.1093/hmg/ddv376; pmid: 26362251
38. S. Goeke *et al*., Alternative splicing of lola generates 19 transcription factors controlling axon guidance in Drosophila. *Nat. Neurosci.* **6**, 917–924 (2003). doi: 10.1038/nn1105; pmid: 12897787
39. B. S. Dhanoa, T. Cogliati, A. G. Satish, E. A. Bruford, J. S. Friedman, Update on the Kelch-like (KLHL) gene family. *Hum. Genomics* **7**, 13 (2013). doi: 10.1186/1479-7364-7-13; pmid: 23676014
40. K. E. Wickliffe, S. Lorenz, D. E. Wemmer, J. Kuriyan, M. Rape, The mechanism of linkage-specific ubiquitin chain elongation by a single-subunit E2. *Cell* **144**, 769–781 (2011). doi: 10.1016/j.cell.2011.01.035; pmid: 21376237

ACKNOWLEDGMENTS

We thank all members of M.R.'s and R.H.'s lab, as well as H. Malik, for discussions and advice; A. Harris and F. Lorbeer for help with in vitro work and genome editing; and J. Schaletzky and R. Zoncu for comments on the manuscript. **Funding:** E.M. was supported by an NSF predoctoral fellowship. R.H. was funded by NIH GM42341. M.R. is an Investigator of the Howard Hughes Medical Institute. **Competing interests:** M.R. is a founder and consultant of Nurix, a biotech company working in the ubiquitin space. All other authors declare no competing interests. **Data and materials availability**: All data are available in the manuscript or the supplementary materials.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/362/6411/eaap8236/suppl/DC1
Materials and Methods
Figs. S1 to S8
Table S1
References (*41*–*44*)

30 August 2017; resubmitted 26 July 2018
Accepted 14 August 2018
Published online 6 September 2018
10.1126/science.aap8236

RESEARCH ARTICLE SUMMARY

STRUCTURAL BIOLOGY

Structure and dynamics of the yeast SWR1-nucleosome complex

Oliver Willhoft*, Mohamed Ghoneim*, Chia-Liang Lin*, Eugene Y. D. Chua, Martin Wilkinson, Yuriy Chaban, Rafael Ayala, Elizabeth A. McCormack, Lorraine Ocloo, David S. Rueda†, Dale B. Wigley†

INTRODUCTION: Canonical nucleosomes contain two copies of each of four histone proteins: H2A, H2B, H3, and H4. However, variants of these histones can be inserted by adenosine triphosphate (ATP)–dependent chromatin-remodeling machines. The yeast SWR1 chromatin-remodeling complex, a member of the INO80 remodeler family, catalyzes the exchange of H2A-H2B dimers for dimers containing Htz1 (H2A.Z in human) in an ATP-dependent manner. However, the mechanism by which SWR1 exchanges histones is poorly understood. Despite having a DNA translocase subunit similar to that in the INO80 complex that slides nucleosomes, no net translocation of nucleosomes has been reported for SWR1. Consequently, the function of the ATPase activity, which is required for histone exchange in SWR1, has remained enigmatic.

RATIONALE: To obtain sufficient quantities for structural analysis, we generated the complete 14-subunit yeast SWR1 complex in insect cells. Binding of nucleosomes to SWR1 is stabilized in the presence of an ATP analog (ADP•BeF_3), which we used to prepare a complex with a canonical yeast H2A-containing nucleosome. Structural analysis was undertaken by cryo–electron microscopy (cryo-EM). We also used single-molecule FRET (smFRET) techniques to probe the dynamics of nucleosomes bound to SWR1. Fluorescent probes were positioned on the H2A histones and the end of the DNA to monitor changes in nucleosome dynamics upon binding of SWR1 and ATP (or ATP analogs).

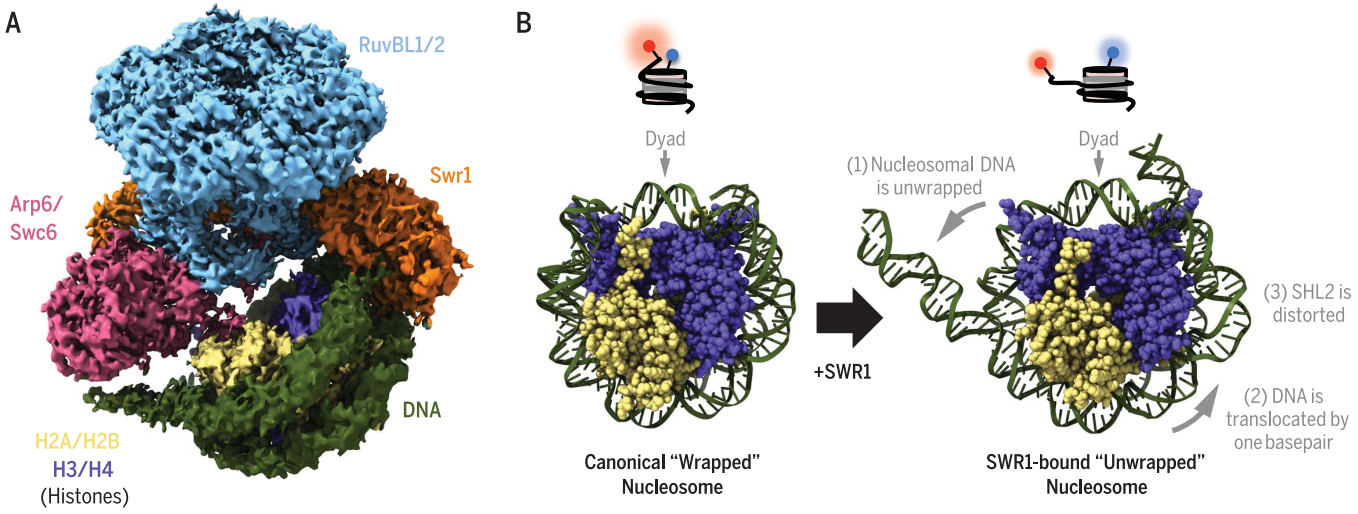
RESULTS: We determined the cryo-EM structure of the SWR1-nucleosome complex at 3.6-Å resolution. The architecture of the complex shows how the SWR1 complex is assembled around a heterohexameric core of the RuvBL1 and RuvBL2 subunits. The Swr1 motor subunit binds at superhelical location 2 (SHL2), a position it shares in common with other remodelers but not with its most closely related complex, INO80, which binds at SHL6-SHL7. Binding of ATP or ADP•BeF_3 to the SWR1-nucleosome complex induces substantial unwrapping of the DNA wrap. Conformational changes in the motor domains of the Swr1 subunit drive a single–base pair translocation of the DNA wrap from the DNA entry site. The single–base pair DNA translocation accompanies conformational changes in the histone core that begin to destabilize the histone dimer interface. Using smFRET methods, we further probed these conformational changes to show how an increase in the dynamics of the SWR1-bound nucleosomes is dependent on binding of ATP but not hydrolysis.

ON OUR WEBSITE

Read the full article at http://dx.doi.org/10.1126/science.aat7716

CONCLUSION: The cryo-EM structure of the SWR1 complex bound to a nucleosome reveals details of the intricate interactions between components of the SWR1 complex and its nucleosome substrate. Interactions between the Swr1 motor domains and the DNA wrap at SHL2 distort the DNA, causing a bulge with concomitant translocation of the DNA by one base pair, coupled to conformational changes of the histone core that likely destabilize the dimer interface. Furthermore, partial unwrapping of the DNA from the histone core takes place upon binding of nucleosomes to the SWR1 complex. Single-molecule data monitor this unwrapping and show how the dynamics are altered by ATP binding prior to hydrolysis. ■

The list of author affiliations is available in the full article online.
*These authors contributed equally to this work.
†Corresponding author. Email: david.rueda@imperial.ac.uk (D.S.R.); d.wigley@imperial.ac.uk (D.B.W.)
Cite this article as O. Willhoft *et al.*, *Science* **362**, eaat7716 (2018). DOI: 10.1126/science.aat7716



Structure of the SWR1-nucleosome complex. (**A**) 3.6-Å SWR1-nucleosome map. (**B**) Binding of SWR1-ADP•BeF_3 to the nucleosome induces multiple changes: (i) The DNA wrap is peeled away by ~2.5 turns, (ii) DNA is translocated by one base pair, and (iii) SHL2 is distorted as a consequence of motor domain closure. These distortions are a precursor to histone exchange and can be monitored by smFRET.

RESEARCH ARTICLE

STRUCTURAL BIOLOGY

Structure and dynamics of the yeast SWR1-nucleosome complex

Oliver Willhoft[1]*, Mohamed Ghoneim[2,3]*, Chia-Liang Lin[1]*†, Eugene Y. D. Chua[1], Martin Wilkinson[1], Yuriy Chaban[1]‡, Rafael Ayala[1], Elizabeth A. McCormack[1], Lorraine Ocloo[1], David S. Rueda[2,3]§, Dale B. Wigley[1]§

The yeast SWR1 complex exchanges histone H2A in nucleosomes with Htz1 (H2A.Z in humans). The cryo–electron microscopy structure of the SWR1 complex bound to a nucleosome at 3.6-angstrom resolution reveals details of the intricate interactions between components of the SWR1 complex and its nucleosome substrate. Interactions between the Swr1 motor domains and the DNA wrap at superhelical location 2 distort the DNA, causing a bulge with concomitant translocation of the DNA by one base pair, coupled to conformational changes of the histone core. Furthermore, partial unwrapping of the DNA from the histone core takes place upon binding of nucleosomes to SWR1 complex. The unwrapping, as monitored by single-molecule data, is stabilized and has its dynamics altered by adenosine triphosphate binding but does not require hydrolysis.

The specific incorporation of histone variants into nucleosomes can trigger changes in chromatin structure (*1*) and the recruitment of cellular mediators (*2*). Histone variants can be inserted by histone chaperones during chromatin deposition (*3*), or at a later stage, by adenosine triphosphate (ATP)–dependent chromatin remodeling machines (*4*). Several variants exist for canonical H2A and H3 histones, such as the replication-independent yeast H3 variant H3.3, which is present at actively transcribed genes (*5*), or human H2A.X and H2A.Z, both of which have essential roles in DNA repair (*6*).

The yeast SWR1 chromatin-remodeling complex is a member of the INO80 remodeler family (*7*). SWR1 drives the exchange of H2A/H2B dimers for Htz1/H2B in an ATP-dependent manner (*8*). In humans, the equivalent histone exchange reaction is catalyzed by the SRCAP (*9*) and TIP60 (*10*) complexes, although the latter complex also catalyzes histone acetylation (*10*). However, the mechanism by which all of these complexes exchange histones has remained enigmatic. Despite having a superfamily 2 (SF2) helicase/translocase (Swr1 subunit) at its heart, and despite the dependence of histone exchange on the ATPase activity of this subunit (*8*), no net translocation of nucleosomes has been reported. By contrast, the related yeast INO80 remodeler [reported, albeit controversially, to catalyze the reverse exchange reaction by replacing Htz1/H2A.Z with H2A (*11*–*13*)] appears to use a mechanism of limited translocation at the nucleosomal entry/exit region in order to exchange a histone dimer (*14*). Recent structures of the INO80-nucleosome complex place the ATPase motor at superhelical location SHL6-SHL7 (*15*, *16*), where it could potentially expose the H2A(Z)-H2B dimer for histone exchange by translocating DNA toward the nucleosome dyad. By contrast, the SWR1 motor subunit (Swr1) binds at SHL2 (*17*), a position it shares in common with other remodelers (*18*–*20*).

Structure of a complex of SWR1 with a bound nucleosome

Biochemical data have shown that SWR1 prefers nucleosome substrates with overhangs on both sides, with at least one of these needing to be long [>60 base pairs (bp)] (*17*, *21*). We therefore used a canonical H2A-containing yeast nucleosome with overhangs on both sides (113N25), in the presence of the ATP analog ADP•BeF_3, to prepare a complex suitable for structure determination by cryo-EM (Fig. 1 and figs. S1 to S3).

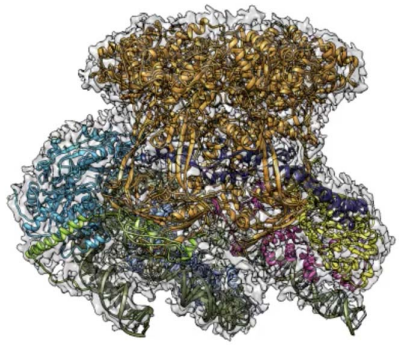
As shown previously (*21*, *22*), and in common with the related INO80 complex (*23*), the SWR1 complex contains a single heterohexamer of the RuvBL proteins (Fig. 1B) that serves as a scaffold around which many of the other subunits are assembled (Movie 1). Each ATPase site is occupied by an adenosine diphosphate (ADP) molecule. If there are any additional functions of the RuvBL hexamer, these do not require ATP hydrolysis because catalytically dead ATPase subunits show wild-type histone exchange activity (fig. S4).

The motor subunit of INO80-family proteins is characterized by a large insertion in the second ATPase domain (*7*) (Fig. 1A). As with the INO80 motor subunit (*23*), the insert in Swr1 forms a large extended structure that makes intimate contacts with the RuvBL hexamer (Fig. 1, B to E). However, the structures of this insert are quite different, although largely planar in both cases. One notable difference between the insert structures is a two-helix extension that protrudes between two RuvBL subunits in SWR1 and makes contact with the Arp6 subunit, in a region that is an insert in the actin fold (Fig. 1E). A function for this interaction, other than architectural, is unknown.

As with the INO80-nucleosome complex (*15*, *16*), the N-terminal region of the SWR1 complex [termed Subcomplex 1 (SC1), comprising the N-terminal region of Swr1 and the actin, Arp4, Swc4, Swc7, Yaf9, and Bdf1 subunits] is present but disordered in our structure.

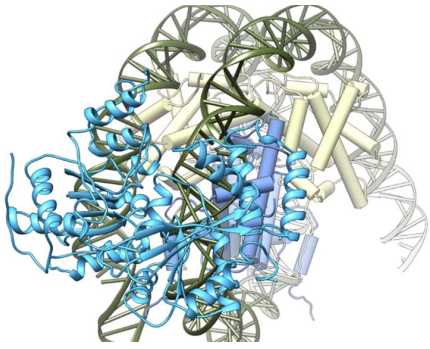
Interactions between the Swr1 ATPase domains and the nucleosome

The nucleosome is bound in a cleft between two lobes that extend from the RuvBL core of the complex. Multiple contacts are made with the histone core and DNA wrap. The Swr1 subunit ATPase domains are located at the canonical SHL2 position seen in the structures of most other remodelers (*19*, *20*) and consistent with SWR1-nucleosome footprinting studies (*17*). There is clear density for a bound nucleotide (ADP•BeF_3) at the ATP-binding site (fig. S3) and, as a consequence, the domains are in the "closed" conformation observed in the ATP-bound state of other SF2 translocases (*24*) such as Chd1 (*19*) (fig. S5), placing the motor domains on the "tracking" strand consistent with 3′-5′ translocation toward the nucleosome dyad axis (Fig. 2, A and B). Although located at SHL2, the binding site is displaced by 1 bp toward the dyad relative to Chd1 and Snf2, resulting in a ~35° rotation of the complex around the bound DNA duplex toward the other gyre of the DNA wrap (Movie 2). The DNA of the bound nucleosome at the SHL2 site contacts the N-terminal ATPase domain (HD1), as observed in other remodelers (*19*, *20*), but as a result of the 1-bp shift, it now makes much more



Movie 1. Rotating overview of SWR1-nucleosome complex coordinates fitted into 3.6-Å cryo-EM envelope. Subunits are colored as in Fig. 1B.

[1]Section of Structural Biology, Department of Medicine, Imperial College London, London SW7 2AZ, UK. [2]Single Molecule Imaging Group, MRC London Institute of Medical Sciences, London W12 0NN, UK. [3]Molecular Virology, Department of Medicine, Imperial College London, London W12 0NN, UK.
*These authors contributed equally to this work. †Present address: Department of Molecular and Cell Biology, University of Leicester, Leicester LE1 9HN, UK. ‡Present address: Electron Bio-Imaging Centre (eBIC), Diamond Light Source, Harwell Science and Innovation Campus, Didcot, Oxfordshire OX11 0DE, UK.
§Corresponding author. Email: david.rueda@imperial.ac.uk (D.S.R.); d.wigley@imperial.ac.uk (D.B.W.)



Movie 2. Rotation of Chd1 motor domain relative to Swr1 motor. Chd1 (pink, PDB ID 5O9G) is shown moving from a superposition with the Swr1 motor domain to its position in the deposited Chd1-nucleosome structure. This shows that Swr1 is rotated by 35° about the DNA axis.

extensive contacts with the other gyre, suggesting a much tighter interaction on the DNA (Fig. 2C). Furthermore, across the second ATPase domain (HD2), the bound DNA shows substantial distortion (Fig. 2, D and E, and fig. S5). Unlike DNA binding across the motor domains of Chd1, the DNA duplex makes a sharp kink, the result of which is lifting of the DNA from the histone surface to make a bulge in the wrap (fig. S5). The kink is a consequence of two α helices being rammed against the duplex (Fig. 2E). In order to create this bulge, the DNA has also been translocated by 1 bp (Movie 3). This suggests that the power stroke is ATP-binding, causing a translocation of 1 bp together with a distortion of the bound DNA. A step size of one base per ATP has been deduced for other SF1 and SF2 translocases (*15*, *25*–*27*).

The location of the Swr1 ATPase domains at the SHL2 position (Fig. 2, A and B) is quite different from that of the INO80 complex, which binds at SHL6-SHL7 (*14*–*16*). However, SWR1 differs from INO80 in a number of important ways. Notably, a single complex of SWR1 is required for activity (*21*), whereas INO80 functions as a dimer (*28*). An INO80 dimer is able to slide nucleosomes along DNA, whereas a SWR1 monomer instead swaps histone dimers (*21*) and is unable to slide nucleosomes (*17*). Although yeast INO80 has been reported to also catalyze histone exchange (*11*, *14*), this activity is controversial (*12*, *13*) and it is not known whether exchange activity requires INO80 dimers. Furthermore, although SWR1 is unable to slide nucleosomes, the ATPase activity of the "motor" domains is nonetheless essential for activity (*8*). Limited

Movie 3. Morph between canonical nucleosome (PDB ID 1ID3) and Swr1-bound nucleosome, which highlights the translocation event induced by SWR1 binding as well as the distortion of the histone octamer from the canonical positions.

sliding of DNA within the confines of the DNA wrap has recently been demonstrated for SWR1 (*29*). The tighter interactions between the N-terminal motor domain (HD1) and the other DNA gyre may act as a block to prevent net translocation by SWR1 but appears to allow

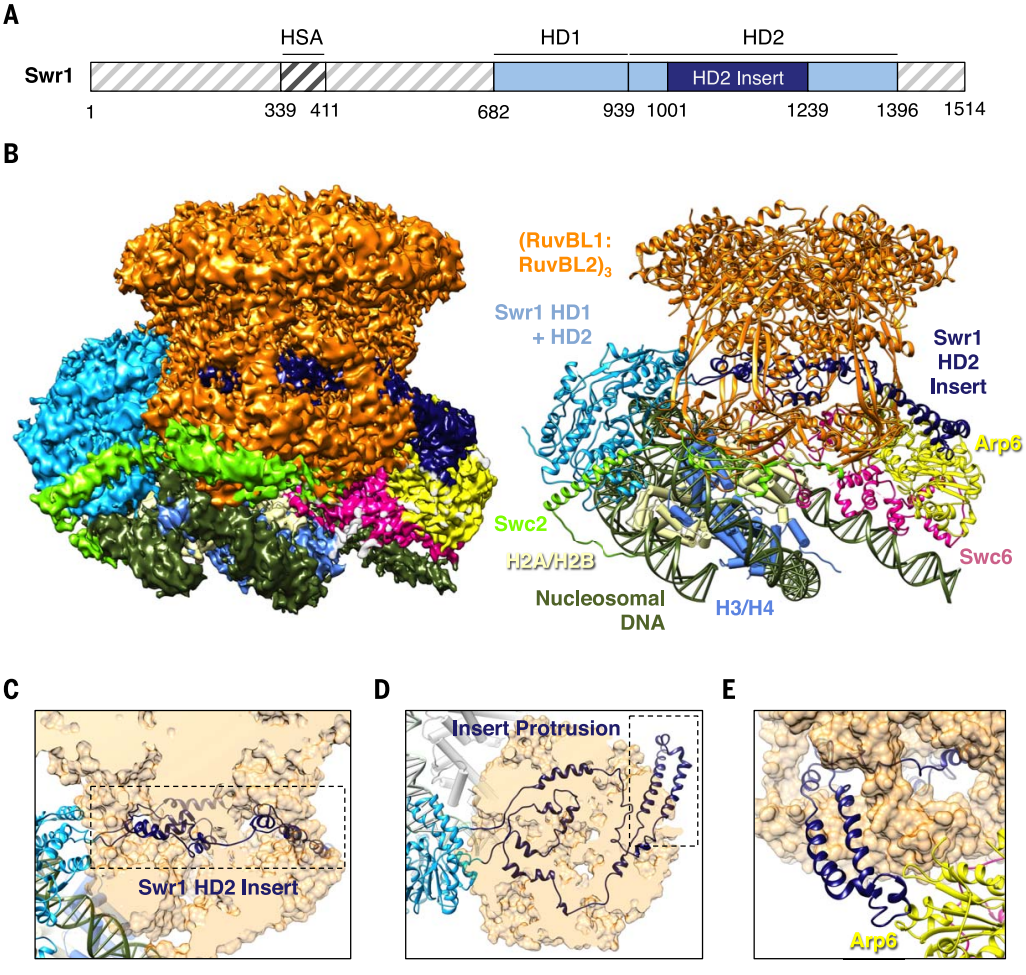


Fig. 1. Overview of the SWR1-nucleosome complex architecture. (**A**) Linearized cartoon of Swr1 subunit showing domain boundaries. HD1 and HD2 are the two ATPase domains; HSA is the helicase/SANT-associated domain. (**B**) Overview of SWR1-nucleosome complex at 3.6 Å (left) and built coordinates (right). (**C**) View of Swr1 insert (ribbon) inside RuvBL1/RuvBL2 ring structure (surface). (**D**) Rotated view of SWR1, showing Swr1 insert following the inner contour of the RuvBL1/RuvBL2 ring, with the Swr1 insert protrusion emerging from the ring. (**E**) Close-up of Swr1 protrusion interacting with Arp6.

limited local translocation of the DNA wrap leading up to that block.

Other contacts with the nucleosome

Arp6 and Swc6 form a tight heterodimer with Swc6 entwined around the actin fold of Arp6 that forms extensive contacts with two adjacent OB-folds of the RuvBL hexamer, with density for an ADP•BeF_3 moiety at its center (fig. S3). This module provides a further tether to the nucleosome via contacts made with nucleosomal H2A and linker DNA on the opposite side to the Swr1 motor (Fig. 3, A to D). At this site, the DNA is unwrapped by ~65° from the canonical nucleosome path, exposing 2.5 turns of the wrap (Fig. 3B). This open conformation is stabilized via specific contacts with Swc6 (Fig. 3C) and may be further stabilized by Swc3 (fig. S6). In addition, a two-pronged interface of Swc6 with H2A, involving a hydrophobic core centered around the αC region (Fig. 3D), could contribute to the specificity of SWR1 for H2A-containing nucleosomes (*30*). Although the resolution of the interface between Swc6 and the nucleosome is insufficient to identify the Swc6 sequence, the region of contact with H2A involves residues 86 to 94 and from residue 109 to the C terminus—regions with differences between H2A and Htz1 that may contribute to the specificity. Interactions with Asp^{91} and Glu^{93} of the H2A acidic patch anchor this module firmly on the face of the octamer. The acidic patch has previously been shown to be important for SWR1 histone

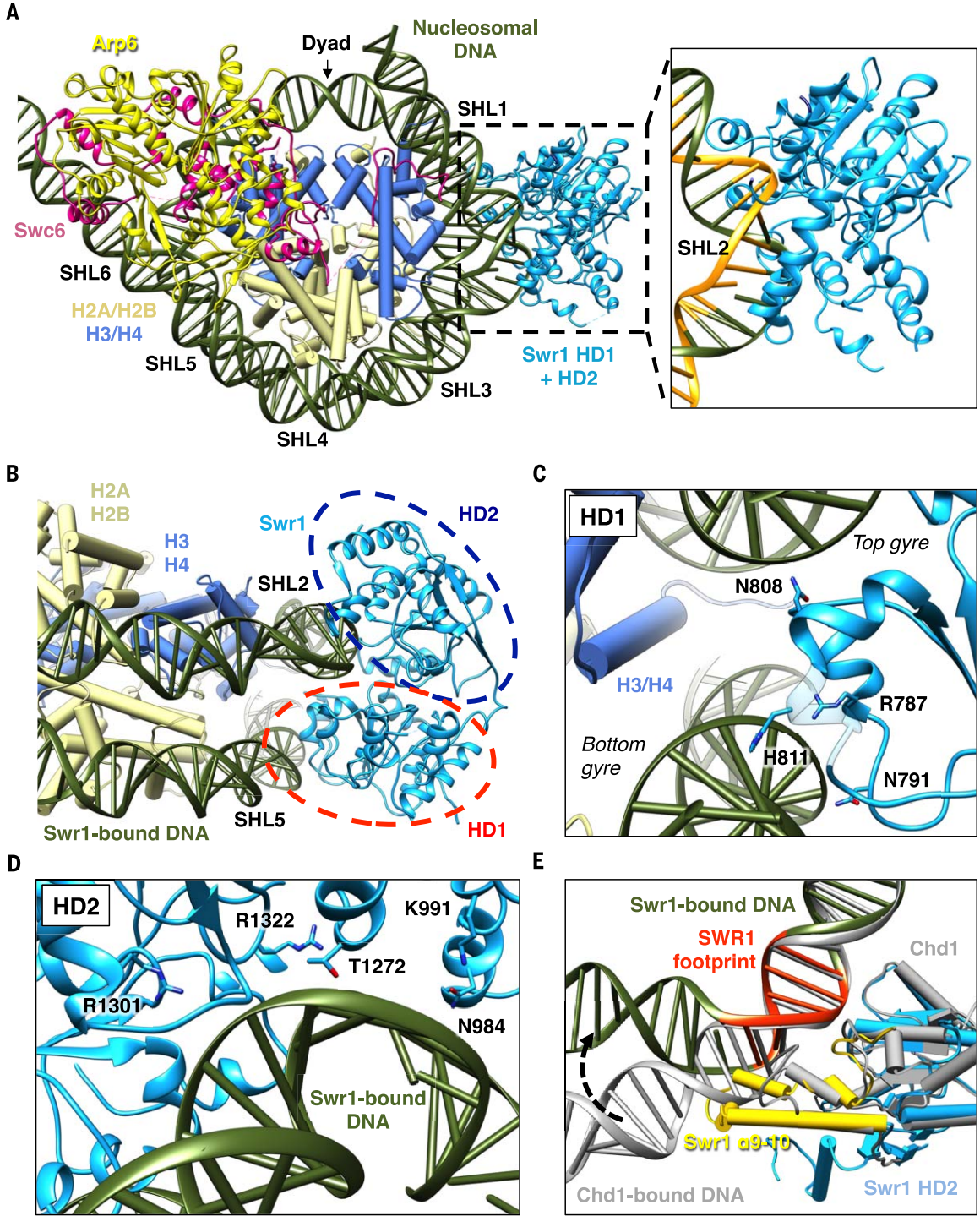


Fig. 2. Details of the Swr1 motor and nucleosomal DNA distortions at SHL2. (**A**) Top-down view onto the nucleosome showing the position of the Swr1 motor domain and the Arp6/Swc6 module. Inset: Swr1 motor domain bound to the tracking strand (gold). (**B**) Nucleosomal position of Swr1 motor domain at SHL2, highlighting approximate delineations of HD1 and HD2 motor domain lobes. (**C**) Details of the Swr1 HD1 interactions with nucleosomal DNA. HD1 wedges between the top and bottom gyres. (**D**) Details of the Swr1 HD2-DNA interactions. The DNA is distorted in the SWR1-nucleosome complex as it lifts onto the HD2 motor domain. (**E**) Comparison of Swr1 and Chd1 nucleosomal DNA trajectories, showing Swr1 motor domain helices α9 and α10 (yellow) pushing on DNA. Structures are aligned on the motor domains of Swr1 and Chd1. The SWR1 footprint (*17*) coincides well with the Swr1 binding site.

exchange (17). Similarly, we observe a loss in histone exchange activity when Arp6/Swc6 is deleted (Fig. 3E and fig. S6). Contacts made with Arp6/Swc6 are mirrored by Arp5/Ies6 in the INO80 complex (15, 16), which have been proposed to facilitate a ratchet-like mechanism for nucleosome sliding by INO80.

A region of extended density, which we assign as Swc2, forms a large interface crossing the main body of the complex, interacting with two consecutive subunits of the RuvBL hexamer before traversing the motor domains, crossing the DNA wrap and then making contacts with the distal face of the nucleosome on the acidic patch (fig. S6). A strikingly similar region of density is observed in INO80 corresponding to the Ies2 subunit (16). Although these subunits do not share recognizable sequence similarity, both their structures and locations share much in common. Ies2 plays an important role in regulating ATPase and sliding activities (31–33). However, although there are similarities in the interaction (such as interaction with the acidic patch), the actual binding site is quite different, so it is unclear whether the role of this interaction is similar or different in the two complexes.

SWR1 causes unwrapping of the DNA at the nucleosome entry and exit sites

By creating nucleosomes with two overhangs (113N25), we satisfied the requirement of the SWR1 complex for at least a 60-bp overhang on at least one side for optimal activity, although a substrate with overhangs on both sides appears to bind even better (34). The structure shows interactions with both DNA overhangs (Fig. 3 and fig. S7). The DNA wrap at one end, corresponding to the SHL6-SHL7 region, is peeled back from the histone surface where it interacts with Arp6/Swc6 (Fig. 3). DNA unwrapping is seen in many other remodeling systems (INO80, Chd1), so it may be a common way to reduce friction as the DNA wrap slides around histones (15, 16, 19). At the other DNA overhang, we also see interactions with the SWR1 complex (figs. S6 and S7). However, this interaction is rather variable and several different conformational classes could be distinguished in the data set, suggesting a high degree of mobility in this region (figs. S1 and S7).

Single-molecule assay to monitor SWR1-nucleosome unwrapping

To characterize SWR1-induced nucleosome unwrapping observed in the cryo-EM structure, we developed a single-molecule Förster resonance energy transfer (smFRET) assay to monitor nucleosome conformation upon SWR1 binding (Fig. 4A) (35, 36). Surface-immobilized fluorescently labeled nucleosomes were imaged by FRET in the absence and presence of SWR1. Nucleosomes alone exhibit three main peaks (Fig. 4B), corresponding to bilabeled (~0.7 FRET), proximal-only (~0.9 FRET), and distal-only (~0.5 FRET) labeling, based on their photobleaching pattern (fig. S8). Counting the fraction of donor-acceptor labeled nucleosomes as a function of time (Fig. 4C) shows that SWR1 remains as active on surface-immobilized nucleosomes as in bulk (21). In the presence of the slowly hydrolyzable ATP analog ATPγS, exchange was not observed (Fig. 4C). This single-molecule histone exchange assay enabled us to quantify whether either of the H2A/H2B dimers (distal or proximal) was exchanged preferentially (fig. S8). The fraction of proximal-only fluorescent nucleosomes (D_p) remained constant over time (Fig. 4D).

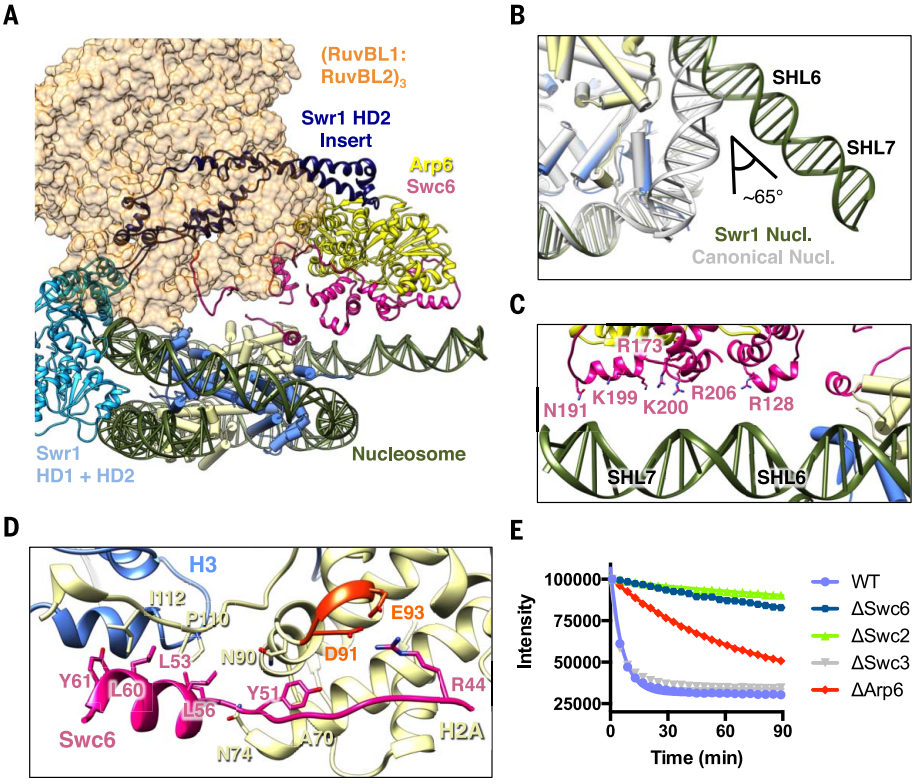


Fig. 3. Contributions of nonmotor subunits to structural changes in the nucleosome. (**A**) Overview of SWR1-nucleosome structure showing the location of the Arp6-Swc6 module relative to Swr1 motor domains. (**B**) Close-up of SHL6 region bound by Arp6-Swc6. The canonical nucleosome (PDB ID 1AOI; gray) is superimposed onto a Swr1-bound nucleosome. The change in angle corresponds to approximately a ~65° rotation away from the nucleosomal wrap, exposing ~2.5 turns of the canonical wrap. (**C**) Swc6 interactions pinning DNA in the SHL6-SHL7 region. (**D**) Swc6 interactions with nucleosomal H2A showing specific interactions with the C-terminal tail and acidic patch (red). (**E**) The effect of subunit deletions on histone exchange activity of SWR1. Amino acid abbreviations: A, Ala; D, Asp; E, Glu; H, His; I, Ile; K, Lys; L, Leu; N, Asn; P, Pro; R, Arg; S, Ser; Y, Tyr. WT, wild type.

SWR1 unwrapping is dynamic and stepwise

The single-molecule FRET trajectories reveal a large increase in conformational dynamics in the presence of SWR1 and ATP. In the absence of SWR1, almost all observed trajectories appear static within our time resolution (250 ms; Fig. 4E and fig. S8), in agreement with previous studies (37, 38). In the presence of SWR1 and ATP, the fraction of dynamic traces increases to almost half (Fig. 4E). Dynamic FRET trajectories (Fig. 5A and fig. S9) show rapid excursions to lower FRET states, consistent with SWR1-induced unwrapping of nucleosomal DNA as observed in the structure (Fig. 3 and fig. S7). In the absence of ATP, the fraction of static traces is comparable to that of nucleosomes alone (Fig. 4E), indicating that ATP binding or hydrolysis is required for unwrapping.

A zoom into these excursions (Fig. 5B) reveals the presence of various unwrapped states with FRET ratios ranging from 0.5 to 0.8. To analyze the observed transitions, and to identify the various observed states (Fig. 5A and fig. S9), we used hidden Markov modeling (38). To simplify the analysis, we focused exclusively on bona fide proximal-only fluorescent nucleosomes, although distal-only fluorescent nucleosome trajectories exhibited similar behavior (fig. S10). A transition density plot (TDP; Fig. 5C) summarizes the observed FRET transitions in all trajectories analyzed. SWR1 induces multiple, partially unwrapped states with the most frequent transition being between the fully wrapped state (W, ~0.9 FRET) and the second unwrapped state (U2, ~0.7 FRET). Additional transitions to ~0.8, ~0.6, and ~0.5 FRET indicate that unwrapping is a stepwise process involving multiple intermediates. However, most observed transitions occur directly without stopping at each intermediate state,

likely indicating that stepwise transitions between intermediate states are faster than our 250-ms time resolution. The symmetric shape of the TDP about the diagonal shows that these transitions are reversible. Lowering the ATP concentration to 100 μM results in similar but less frequent excursions (fig. S11), indicating that SWR1 can unwrap the DNA multiple times per encounter.

SWR1 unwrapping requires ATP binding but not hydrolysis

In the presence of a slowly hydrolyzable ATP analog (ATPγS) that does not support histone exchange over the time scale of these experiments (Fig. 4C), the fraction of dynamic traces remains high (Fig. 4E). The corresponding smFRET trajectories (Fig. 5D and fig. S12) exhibit similar rapid excursions into lower FRET states, indicating that nucleosome unwrapping requires ATP binding but not hydrolysis. These excursions (Fig. 5E) involve similar transitions into various unwrapped states; however, steps into multiple intermediate states are now apparent even at 250-ms time resolution, likely due to slower transitions in the presence of ATPγS. Likewise, the resulting TDP (Fig. 5F) reveals additional transitions to intermediate states (Fig. 5, C and

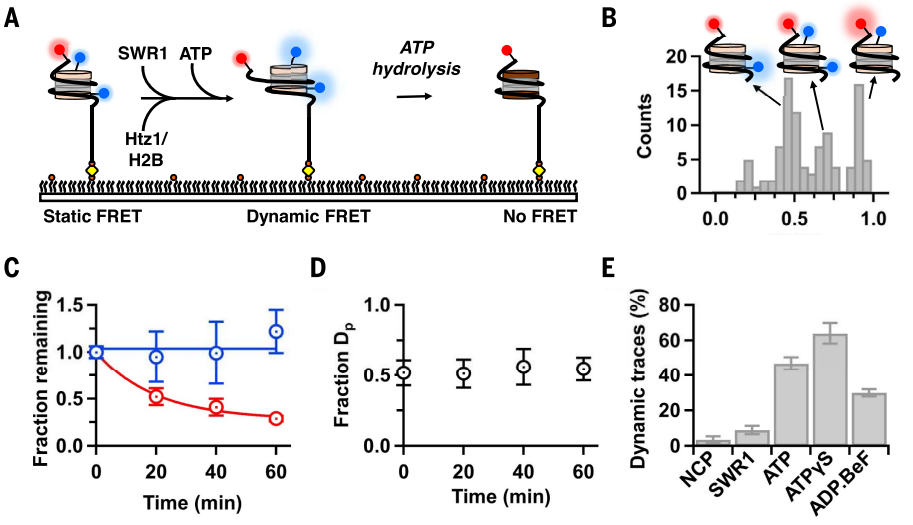


Fig. 4. Single-molecule nucleosome dynamics induced by SWR1. (**A**) Schematic diagram of the experimental setup: Bilabeled nucleosomes (AF555 on each H2A, blue), bound to biotinylated (orange) and labeled (AF647, red) 257-bp DNA, are surface-immobilized on neutravidin-coated (yellow) biotin-PEG slides. Nucleosomes comprise canonical H2A/H2B (beige discs) and H3/H4 dimers (gray discs). SWR1-nucleosome interactions are monitored via FRET between the donors and the acceptor. (**B**) FRET histogram of nucleosomes (in the absence of SWR1) shows three major populations: bilabeled (~0.7 FRET), proximal-only (~0.9 FRET), and distal-only (~0.5 FRET) fluorescent nucleosomes (N = 103). (**C**) Single-molecule histone exchange assay shows that SWR1 is active on immobilized nucleosomes at levels comparable to bulk experiments (*21*) in the presence of ATP (red, observed rate constant k_{obs} = 0.05 ± 0.01 min^{-1}) but not ATPγS (blue). Histone exchange removes fluorescently labeled H2A; therefore, activity is determined as the fraction of remaining fluorescent nucleosomes as a function of time (N = 290 at t = 0 min in ATP or ATPγS). (**D**) The fraction of proximal-only (D_p) fluorescent nucleosomes remains constant with time, indicating that proximal and distal H2A/H2B dimers are exchanged without preference (N = 74, 79, 36, and 40). (**E**) Percentage of dynamic traces in the presence of nucleosome alone (NCP, N = 55); nucleosome and SWR1 complex (SWR1, N = 67); nucleosome, SWR1, and ATP (N = 85); ATPγS (N = 121); or ADP•BeF_3 (N = 102). Error bars in (C) to (E) denote SE.

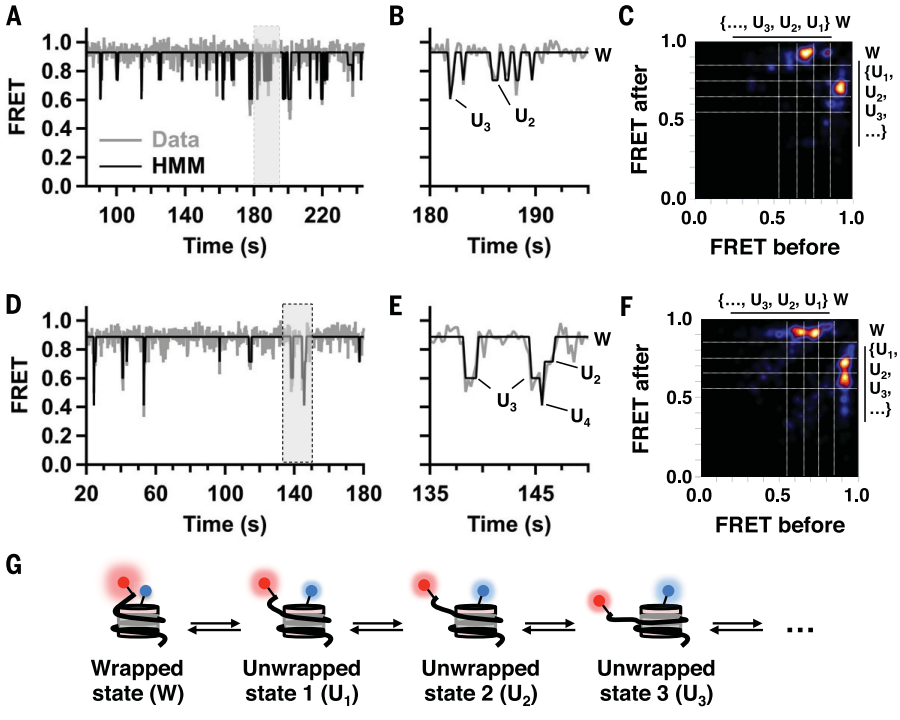


Fig. 5. Analysis of unwrapped states of nucleosomal DNA. (**A**) FRET time trajectory (gray) of a proximal-only (D_p) fluorescent nucleosome in the presence of SWR1 and ATP, with resulting hidden Markov model fit (HMM, black). (**B**) Zoom of gray box in (A) showing direct and reversible transient excursions into multiple mid-FRET unwrapped states ({U_i}). (**C**) Calculated transition density plot (TDP) of unwrapping of nucleosomal DNA in the presence of ATP (2000 transitions from 60 trajectories). TDP confirms that multiple partially unwrapped states are accessed from the high-FRET wrapped state (W). (**D**) FRET trajectory (gray) of a proximal-only fluorescent nucleosome in the presence of SWR1 complex and ATPγS with resulting HMM fit (black). Fewer and slower transitions are observed than in (A). (**E**) Zoom of the gray box in (D) showing direct and reversible transient excursions into multiple mid-FRET unwrapped states ({U_i}). (**F**) Calculated TDP of unwrapping nucleosomal DNA in presence of ATPγS (1404 transitions from 65 trajectories). (**G**) Schematic of various nucleosomal unwrapping states.

G). Taken together, these data show that ATP binding (but not hydrolysis) is required for SWR1 unwrapping on the pathway to H2A/H2B histone exchange. Similar results are observed with ADP•BeF_3 (Fig. 4E and fig. S13).

To correlate the observed FRET values to the number of unwrapped base pairs, we determined an empirical distance calibration linking the distance between the end of the DNA and the labeled H2A histones (fig. S14). On the basis of this calibration, which we regard as qualitative rather than quantitative, we estimate that SWR1 unwraps 10 to 12 bp from the nucleosome, which is consistent with the degree of unwrapping observed in the structure (Fig. 3 and fig. S7).

Exchange of histone dimers is distributive with no preference for linker-proximal or -distal dimers

Although dimer exchange by SWR1 is sequential (*30*), it is not known whether this process is distributive or processive, nor is it known whether there is any preference for which dimer is exchanged first. We looked at histone dimer exchange under catalytic conditions (1:10 ratio of SWR1 to nucleosomes, i.e., 20 exchange reactions) at various ratios of dimer to nucleosome (fig. S15). At a 1:1 ratio, we observed around 31% doubly exchanged nucleosomes, although the majority (69%) have a single dimer exchanged, indicating a distributive rather than processive mechanism. At lower dimer/nucleosome ratios, the proportion of single exchanges is even higher (fig. S15). Furthermore, the proportion of dye-proximal and -distal dimer exchanges was approximately equal, suggesting little preference for which dimer is exchanged first. Together, these data imply that SWR1 exchanges dimers one at a time, most likely binding in two orientations (probably at each SHL2 site) to achieve this process. The exchange reaction, therefore, is better considered in terms of which dimer is proximal or distal to the SWR1 motor domain binding site rather than whether it is linker-distal or -proximal, because in vivo the nucleosomes will be flanked by linker DNA on both sides. As discussed above, the structure strongly suggests that the dimer that is exchanged is the one proximal to the SHL2 site at which SWR1 is bound, independent of which side any linker DNA might be attached.

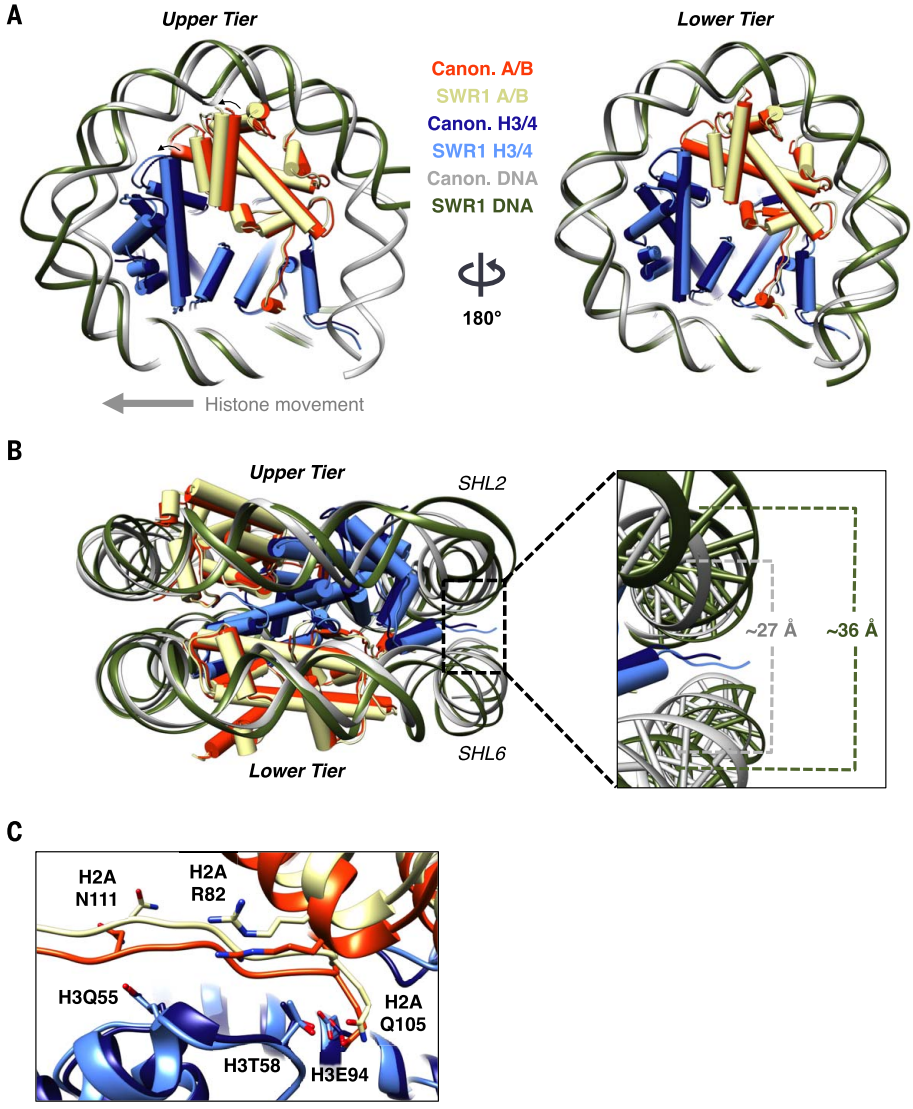


Fig. 6. Distortions of core histones as a consequence of DNA translocation. (**A**) Summary of distortions in the histone core. Histones are shown as cylinders, with coloring to distinguish between canonical and SWR1-bound nucleosome positions. Upper- and lower-tier histones were defined on the basis of the views shown here. Canonical yeast nucleosome (PDB ID 1ID3) and SWR1-bound nucleosome were superimposed on residues 67 to 74 of H4, 108 to 126 of H2B, and 56 to 98 of H3. The general direction of histone movement is toward the motor-bound side of the nucleosome. (**B**) Edge-on view of the nucleosome, highlighting upper and lower histone tiers as well as upper and lower DNA gyres. Coloring is as in (A). Inset: Close-up of the motor-bound nucleosome region, showing a widening of the gap between the two gyres and distortion of the lower gyre in the presence of SWR1. (**C**) Change in the upper-tier H2A-H3 interface. As a consequence of SWR1-nucleosome interactions, the upper H2A-H3 interface decreases by approximately 180 $Å^2$, as determined in a comparison with a canonical nucleosome (PDB ID 1ID3) (470 $Å^2$ versus 650 $Å^2$). By contrast, the lower-tier interface remains largely the same (660 $Å^2$ versus 650 $Å^2$). These changes are further detailed in fig. S16.

Implications for histone exchange

The mechanism for histone exchange by SWR1 is still poorly understood. Very early studies showed that not only was the Swr1 subunit a member of the superfamily 2 translocases, like other chromatin remodelers that slide nucleosomes, but also that the ATPase activity of this subunit was essential for activity (*8*); these findings suggested that DNA translocation at some level might be a component of the histone exchange mechanism. However, no group has been able to demonstrate net nucleosome sliding by SWR1, although limited sliding within the nucleosome has recently been reported (*29*).

Experiments with nucleosomes containing two-base gaps on the tracking strand within the SHL2 sites on either side of the nucleosome dyad have interesting effects on histone exchange (*17*). A two-base gap on one strand, at positions located between base pairs ~17 to 22 from the nucleosome dyad, permits exchange of one dimer but prevents exchange of the other. Furthermore, a nucleosome in which gaps are placed at both SHL2 regions (in both cases on the tracking strand) prevents exchange of both histone dimers, even though binding to this nucleosome is unaffected. The region of SHL2 that is sensitive to the presence of a gap (bases 17 to 22 from the dyad) corresponds very nicely with the contact regions of the motor domains with the DNA wrap (Fig. 2E). However, placing a gap on either side of this contact region, but on the same DNA strand, has only a modest effect on the rate of exchange and both dimers can be exchanged. This latter observation suggests that, at best, there can be only limited translocation of DNA by the motor domains; similar experiments with bona fide

nucleosome sliders show that a single-base gap ahead of the motor domain binding site will block translocation, but only when placed in the tracking strand (*18*). By contrast, for SWR1, a gap in the tracking strand close to the SHL2 site is sufficient to prevent histone exchange, which suggests a requirement for some form of ATP-driven conformational changes of the DNA at the binding site. Although footprinting shows that there is no net sliding of nucleosome position as a consequence of exchange even at base pair resolution (*17*), limited transient translocation (i.e., a few base pairs) of part of the DNA wrap has been proposed (*29*).

In our structure, we do indeed observe limited translocation of the DNA wrap from the entry site to the Swr1 binding site by a single base pair to create a bulge. We also observe some distortion of the histone core (Fig. 6, fig. S16, and Movie 3). The histone core "flexes" and the upper tier, comprising a histone dimer and part of the H3/H4 tetramer, twists relative to the lower tier. These conformational changes result in changes at the interface between the H2A/H2B dimer and H3/H4 (Fig. 6 and fig. S16). The changes in the nucleosome core likely also explain the increased dynamics of the DNA tails, as observed in our single-molecule studies (Figs. 4 and 5).

Consequently, our structural and single-molecule data suggest that the role of the Swr1 ATPase activity is to induce local distortion in the DNA at SHL2 that results from limited translocation against the HD1 domain of Swr1, and that this begins to destabilize the DNA wrap and histone dimer contacts within the nucleosome. Such a distortion of the dimer interface with H4 as a consequence of limited, and transient, DNA translocation has been suggested previously on the basis of biochemical data (*17*, *29*), and our structure now provides mechanistic insights into that process. Our structure likely represents an initial stage in the reaction cycle, and the distortions we observe in both the DNA and the histone core could be amplified by additional translocation steps. Although ADP•BeF_3 induces conformational changes similar to those likely accompanied by ATP binding, we note that non-hydrolyzable ATP analogs do not support histone exchange; thus, at least one, possibly several, rounds of ATP binding and hydrolysis may be required. As a consequence of histone exchange, there need not be any net translocation of the DNA wrap beyond the Swr1 motor binding site at SHL2, if any translocation of the DNA up to that point were transient until histone dimer release and/or exchange.

Materials and methods

Briefly, nucleosomes and SWR1 complex were prepared as described (*21*). The SWR1-nucleosome complex was assembled by incubating SWR1-ADP•BeF_3:Htz1/H2B dimer complex with nucleosomes at a 1:1 molar ratio. Cryo-EM data acquisition, image acquisition, and structure reconstruction followed a similar procedure as described (*15*). Data processing and refinement statistics for the two cryo-EM structures are summarized in figs. S1 to S3 and table S1. See the supplementary materials for further details, including details of single-molecule experiments.

REFERENCES AND NOTES

1. P. B. Talbert, S. Henikoff, Histone variants on the move: Substrates for chromatin dynamics. *Nat. Rev. Mol. Cell Biol.* **18**, 115–126 (2017). doi: 10.1038/nrm.2016.148; pmid: 27924075
2. M. Yun, J. Wu, J. L. Workman, B. Li, Readers of histone modifications. *Cell Res.* **21**, 564–578 (2011). doi: 10.1038/cr.2011.42; pmid: 21423274
3. S. J. Elsässer, S. D'Arcy, Towards a mechanism for histone chaperones. *Biochim. Biophys. Acta* **1819**, 211–221 (2013). doi: 10.1016/j.bbagrm.2011.07.007; pmid: 24459723
4. B. Bartholomew, Regulating the chromatin landscape: Structural and mechanistic perspectives. *Annu. Rev. Biochem.* **83**, 671–696 (2014). doi: 10.1146/annurev-biochem-051810-093157; pmid: 24606138
5. A. Jamai, R. M. Imoberdorf, M. Strubin, Continuous histone H2B and transcription-dependent histone H3 exchange in yeast cells outside of replication. *Mol. Cell* **25**, 345–355 (2007). doi: 10.1016/j.molcel.2007.01.019; pmid: 17289583
6. O. Gursoy-Yuzugullu, N. House, B. D. Price, Patching Broken DNA: Nucleosome Dynamics and the Repair of DNA Breaks. *J. Mol. Biol.* **428**, 1846–1860 (2016). doi: 10.1016/j.jmb.2015.11.021; pmid: 26625977
7. C. R. Clapier, B. R. Cairns, The biology of chromatin remodeling complexes. *Annu. Rev. Biochem.* **78**, 273–304 (2009). doi: 10.1146/annurev.biochem.77.062706.153223; pmid: 19355820
8. G. Mizuguchi *et al.*, ATP-driven exchange of histone H2AZ variant catalyzed by SWR1 chromatin remodeling complex. *Science* **303**, 343–348 (2004). doi: 10.1126/science.1090701; pmid: 14645854
9. D. D. Ruhl *et al.*, Purification of a human SRCAP complex that remodels chromatin by incorporating the histone variant H2A.Z into nucleosomes. *Biochemistry* **45**, 5671–5677 (2006). doi: 10.1021/bi060043d; pmid: 16634648
10. T. Kusch *et al.*, Acetylation by Tip60 is required for selective histone variant exchange at DNA lesions. *Science* **306**, 2084–2087 (2004). doi: 10.1126/science.1103455; pmid: 15528408
11. M. Papamichos-Chronakis, S. Watanabe, O. J. Rando, C. L. Peterson, Global regulation of H2A.Z localization by the INO80 chromatin-remodeling enzyme is essential for genome integrity. *Cell* **144**, 200–213 (2011). doi: 10.1016/j.cell.2010.12.021; pmid: 21241891
12. F. Wang, A. Ranjan, D. Wei, C. Wu, Comment on "A histone acetylation switch regulates H2A.Z deposition by the SWR-C remodeling enzyme". *Science* **353**, 358 (2016). doi: 10.1126/science.aad5921; pmid: 27463665
13. S. Watanabe, C. L. Peterson, Response to Comment on "A histone acetylation switch regulates H2A.Z deposition by the SWR-C remodeling enzyme". *Science* **353**, 358 (2016). doi: 10.1126/science.aad6398; pmid: 27463666
14. S. Brahma *et al.*, INO80 exchanges H2A.Z for H2A by translocating on DNA proximal to histone dimers. *Nat. Commun.* **8**, 15616 (2017). doi: 10.1038/ncomms15616; pmid: 28604691
15. R. Ayala *et al.*, Structure and regulation of the human INO80-nucleosome complex. *Nature* **556**, 391–395 (2018). doi: 10.1038/s41586-018-0021-6; pmid: 29643506
16. S. Eustermann *et al.*, Structural basis for ATP-dependent chromatin remodelling by the INO80 complex. *Nature* **556**, 386–390 (2018). doi: 10.1038/s41586-018-0029-y; pmid: 29643509
17. A. Ranjan *et al.*, H2A histone-fold and DNA elements in nucleosome activate SWR1-mediated H2A.Z replacement in budding yeast. *eLife* **4**, e06845 (2015). doi: 10.7554/eLife.06845; pmid: 26116819
18. A. Saha, J. Wittmeyer, B. R. Cairns, Chromatin remodeling through directional DNA translocation from an internal nucleosomal site. *Nat. Struct. Mol. Biol.* **12**, 747–755 (2005). doi: 10.1038/nsmb973; pmid: 16086025
19. L. Farnung, S. M. Vos, C. Wigge, P. Cramer, Nucleosome-Chd1 structure and implications for chromatin remodelling. *Nature* **550**, 539–542 (2017). doi: 10.1038/nature24046; pmid: 29019976
20. X. Liu, M. Li, X. Xia, X. Li, Z. Chen, Mechanism of chromatin remodelling revealed by the Snf2-nucleosome structure. *Nature* **544**, 440–445 (2017). doi: 10.1038/nature22036; pmid: 28424519
21. C. L. Lin *et al.*, Functional characterization and architecture of recombinant yeast SWR1 histone exchange complex. *Nucleic Acids Res.* **45**, 7249–7260 (2017). doi: 10.1093/nar/gkx414; pmid: 28499038
22. V. Q. Nguyen *et al.*, Molecular architecture of the ATP-dependent chromatin-remodeling complex SWR1. *Cell* **154**, 1220–1231 (2013). doi: 10.1016/j.cell.2013.08.018; pmid: 24034246
23. R. J. Aramayo *et al.*, Cryo-EM structures of the human INO80 chromatin-remodeling complex. *Nat. Struct. Mol. Biol.* **25**, 37–44 (2018). doi: 10.1038/s41594-017-0003-7; pmid: 29323271
24. M. R. Singleton, M. S. Dillingham, D. B. Wigley, Structure and mechanism of helicases and nucleic acid translocases. *Annu. Rev. Biochem.* **76**, 23–50 (2007). doi: 10.1146/annurev.biochem.76.052305.115300; pmid: 17506634
25. S. S. Velankar, P. Soultanas, M. S. Dillingham, H. S. Subramanya, D. B. Wigley, Crystal structures of complexes of PcrA DNA helicase with a DNA substrate indicate an inchworm mechanism. *Cell* **97**, 75–84 (1999). doi: 10.1016/S0092-8674(00)80716-3; pmid: 10199404
26. M. S. Dillingham, D. B. Wigley, M. R. Webb, Demonstration of unidirectional single-stranded DNA translocation by PcrA helicase: Measurement of step size and translocation speed. *Biochemistry* **39**, 205–212 (2000). doi: 10.1021/bi992105o; pmid: 10625495
27. D. B. Wigley, G. D. Bowman, A glimpse into chromatin remodeling. *Nat. Struct. Mol. Biol.* **24**, 498–500 (2017). doi: 10.1038/nsmb.3415; pmid: 28586327
28. O. Willhoft *et al.*, Crosstalk within a functional INO80 complex dimer regulates nucleosome sliding. *eLife* **6**, e25782 (2017). doi: 10.7554/eLife.25782; pmid: 28585918
29. R. K. Singh, S. Watanabe, O. Bilsel, C. L. Peterson, Transient kinetic analysis of SWR1C-catalyzed H2A.Z deposition unravels the impact of nucleosome dynamics and the asymmetry of stepwise histone exchange. BioRxiv [preprint]. 19 April 2018. pmid: 304998
30. E. Luk *et al.*, Stepwise histone replacement by SWR1 requires dual activation with histone H2A.Z and canonical nucleosome. *Cell* **143**, 725–736 (2010). doi: 10.1016/j.cell.2010.10.019; pmid: 21111233
31. L. Chen, R. C. Conaway, J. W. Conaway, Multiple modes of regulation of the human Ino80 SNF2 ATPase by subunits of the INO80 chromatin-remodeling complex. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 20497–20502 (2013). doi: 10.1073/pnas.1317092110; pmid: 24297934
32. W. Yao *et al.*, Assembly of the Arp5 (Actin-related Protein) Subunit Involved in Distinct INO80 Chromatin Remodeling Activities. *J. Biol. Chem.* **290**, 25700–25709 (2015). doi: 10.1074/jbc.M115.674887; pmid: 26306040
33. O. Willhoft, R. Bythell-Douglas, E. A. McCormack, D. B. Wigley, Synergy and antagonism in regulation of recombinant human INO80 chromatin remodeling complex. *Nucleic Acids Res.* **44**, 8179–8188 (2016). doi: 10.1093/nar/gkw509; pmid: 27257055
34. A. Ranjan *et al.*, Nucleosome-free region dominates histone acetylation in targeting SWR1 to promoters for H2A.Z replacement. *Cell* **154**, 1232–1245 (2013). doi: 10.1093/nar/gkw509; pmid: 27257055
35. T. R. Blosser, J. G. Yang, M. D. Stone, G. J. Narlikar, X. Zhuang, Dynamics of nucleosome remodelling by individual ACF complexes. *Nature* **462**, 1022–1027 (2009). doi: 10.1038/nature08627; pmid: 20033040
36. S. Deindl *et al.*, ISWI remodelers slide nucleosomes with coordinated multi-base-pair entry steps and single-base-pair exit steps. *Cell* **152**, 442–452 (2013). doi: 10.1016/j.cell.2012.12.040; pmid: 23374341
37. G. Li, M. Levitus, C. Bustamante, J. Widom, Rapid spontaneous accessibility of nucleosomal DNA. *Nat. Struct. Mol. Biol.* **12**, 46–53 (2005). doi: 10.1038/nsmb869; pmid: 15580276
38. S. A. McKinney, C. Joo, T. Ha, Analysis of single-molecule FRET trajectories using hidden Markov modeling. *Biophys. J.* **91**, 1941–1951 (2006). doi: 10.1529/biophysj.106.082487; pmid: 16766620

ACKNOWLEDGMENTS

We thank Diamond for access and support of the Cryo-EM facilities at the UK national electron bio-imaging center (eBIC), funded by the Wellcome Trust, MRC, and BBSRC. **Funding:** Supported by the Wellcome Trust [095519/Z/11/Z

and 209327/Z/17/Z (D.B.W.), Cancer Research UK (C6913/A21608 (D.B.W.)], the Medical Research Council [MR/N009258/1 and MR/R009023/1 (D.B.W.)], a core grant from the MRC London Institute of Medical Sciences (D.S.R.), and start-up funds from Imperial College London (D.S.R.). **Author contributions:** O.W., M.G., D.S.R., and D.B.W. designed the studies; O.W., E.Y.D.C., M.W., and R.A. performed the cryo-EM analysis; C.-L.L. and Y.C. conducted initial preliminary studies; O.W., E.A.M., and L.O. prepared the samples; M.G. conducted and analyzed the single-molecule experiments; C.-L.L. and O.W. conducted the biochemical experiments; and D.B.W. and D.S.R. analyzed the data and wrote the manuscript with input from all the authors. **Competing interests:** Authors declare no competing interests. **Data and materials availability:** Density maps are deposited at the Electron Microscopy Database (accession code EMD-4395) and protein coordinates are deposited at the Protein Data Bank (PDB ID codes 6GEJ and 6GEN).

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/362/6411/eaat7716/suppl/DC1
Materials and Methods
Figs. S1 to S16
Table S1
References (*39*–*48*)

27 April 2018; accepted 8 August 2018
10.1126/science.aat7716

RESEARCH ARTICLE SUMMARY

NEURODEVELOPMENT

Thyroid hormone signaling specifies cone subtypes in human retinal organoids

Kiara C. Eldred, Sarah E. Hadyniak, Katarzyna A. Hussey, Boris Brenerman, Ping-Wu Zhang, Xitiz Chamling, Valentin M. Sluch, Derek S. Welsbie, Samer Hattar, James Taylor, Karl Wahlin, Donald J. Zack, Robert J. Johnston Jr.*

INTRODUCTION: Cone photoreceptors in the human retina enable daytime, color, and high-acuity vision. The three subtypes of human cones are defined by the visual pigment that they express: blue-opsin (short wavelength; S), green-opsin (medium wavelength; M), or red-opsin (long wavelength; L). Mutations that affect opsin expression or function cause various forms of color blindness and retinal degeneration.

RATIONALE: Our current understanding of the vertebrate eye has been derived primarily from the study of model organisms. We studied the human retina to understand the developmental mechanisms that generate the mosaic of mutually exclusive cone subtypes. Specification of human cones occurs in a two-step process. First, a decision occurs between S versus L/M cone fates. If the L/M fate is chosen, a subsequent choice is made between expression of L- or M-opsin. To determine the mechanism that controls the first decision between S and L/M cone fates, we studied human retinal organoids derived from stem cells.

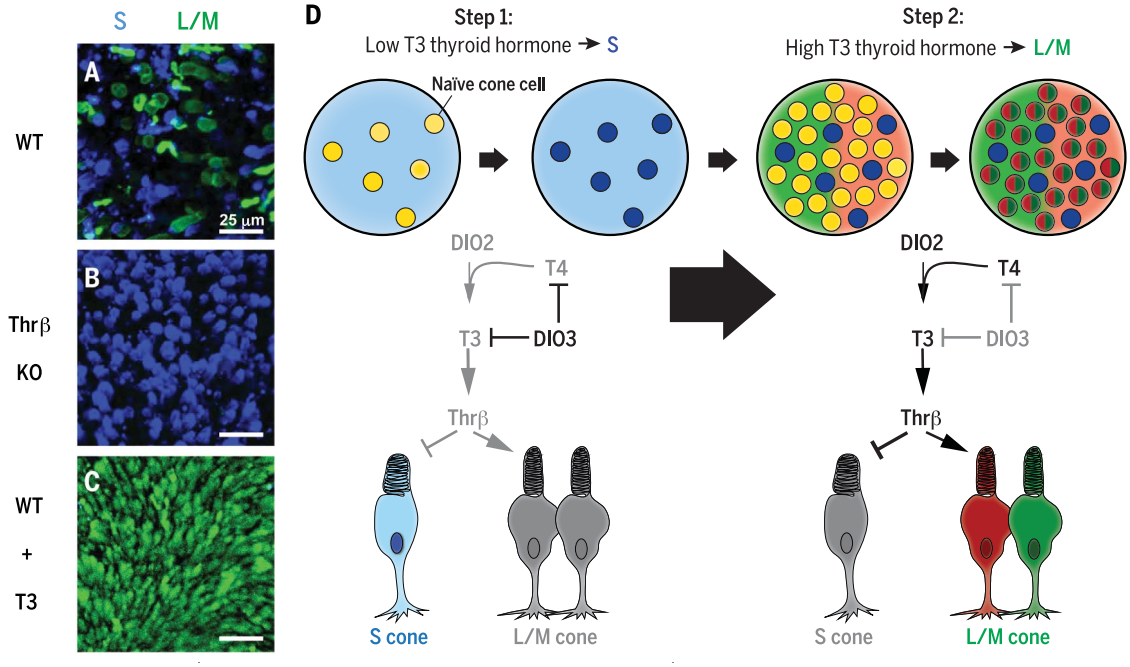
RESULTS: We found that human organoids and retinas have similar distributions, gene expression profiles, and morphologies of cone subtypes. During development, S cones are specified first, followed by L/M cones. This temporal switch from specification of S cones to generation of L/M cones is controlled by thyroid hormone (TH) signaling. In retinal organoids that lacked thyroid hormone receptor β (*Thrβ*), all cones developed into the S subtype. *Thrβ* binds with high affinity to triiodothyronine (T3), the more active form of TH, to regulate gene expression. We observed that addition of T3 early during development induced L/M fate in nearly all cones. Thus, TH signaling through *Thrβ* is necessary and sufficient to induce L/M cone fate and suppress S fate. TH exists largely in two states: thyroxine (T4), the most abundant circulating form of TH, and T3, which binds TH receptors with high affinity. We hypothesized that the retina itself could modulate TH levels to control subtype fates. We found that deiodinase 3 (*DIO3*), an enzyme that degrades both T3 and T4, was expressed early in organoid and retina development. Conversely, deiodinase 2 (*DIO2*), an enzyme that converts T4 to active T3, as well as TH carriers and transporters, were expressed later in development. Temporally dynamic expression of TH-degrading and -activating proteins supports a model in which the retina itself controls TH levels, ensuring low TH signaling early to specify S cones and high TH signaling later in development to produce L/M cones.

CONCLUSION: Studies of model organisms and human epidemiology often generate hypotheses about human biology that cannot be studied in humans. Organoids provide a system to determine the mechanisms of human development, enabling direct testing of hypotheses in developing human tissue. Our studies identify temporal regulation of TH signaling as a mechanism that controls cone subtype specification in humans. Consistent with our findings, preterm human infants with low T3 and T4 have an increased incidence of color vision defects. Moreover, our identification of a mechanism that generates one cone subtype while suppressing the other, coupled with successful transplantation and incorporation of stem cell–derived photoreceptors in mice, suggests that the promise of therapies to treat human diseases such as color blindness, retinitis pigmentosa, and macular degeneration will be achieved in the near future. ■

ON OUR WEBSITE

Read the full article at http://dx.doi.org/10.1126/science.aau6348

The list of author affiliations is available in the full article online.
*Corresponding author. Email: robertjohnston@jhu.edu
Cite this article as K. C. Eldred *et al.*, *Science* **362**, eaau6348 (2018). DOI: 10.1126/science.aau6348



Temporally regulated TH signaling specifies cone subtypes. (**A**) Embryonic stem cell–derived human retinal organoids [wild type (WT)] generate S and L/M cones. Blue, S-opsin; green, L/M-opsin. (**B**) Organoids that lack thyroid hormone receptor β (*Thrβ* KO) generate all S cones. (**C**) Early activation of TH signaling (WT + T3) specifies nearly all L/M cones. (**D**) TH-degrading enzymes (such as DIO3) expressed early in development lower TH and promote S fate, whereas TH-activating regulators (such as DIO2) expressed later promote L/M fate.

RESEARCH ARTICLE

NEURODEVELOPMENT

Thyroid hormone signaling specifies cone subtypes in human retinal organoids

Kiara C. Eldred[1], Sarah E. Hadyniak[1], Katarzyna A. Hussey[1], Boris Brenerman[1], Ping-Wu Zhang[2], Xitiz Chamling[2], Valentin M. Sluch[2], Derek S. Welsbie[3], Samer Hattar[4], James Taylor[1,5], Karl Wahlin[3], Donald J. Zack[2,6,7,8], Robert J. Johnston Jr.[1]*

The mechanisms underlying specification of neuronal subtypes within the human nervous system are largely unknown. The blue (S), green (M), and red (L) cones of the retina enable high-acuity daytime and color vision. To determine the mechanism that controls S versus L/M fates, we studied the differentiation of human retinal organoids. Organoids and retinas have similar distributions, expression profiles, and morphologies of cone subtypes. S cones are specified first, followed by L/M cones, and thyroid hormone signaling controls this temporal switch. Dynamic expression of thyroid hormone–degrading and –activating proteins within the retina ensures low signaling early to specify S cones and high signaling late to produce L/M cones. This work establishes organoids as a model for determining mechanisms of human development with promising utility for therapeutics and vision repair.

Cone photoreceptors in the human retina enable daytime, color, and high-acuity vision (*1*). The three subtypes of human cones are defined by the visual pigment that they express: blue-opsin (short wavelength; S), green-opsin (medium wavelength; M), or red-opsin (long wavelength; L) (*2*). Specification of human cones occurs in a two-step process. First, a decision occurs between S versus L/M cone fates (Fig. 1A). If the L/M fate is chosen, a subsequent choice is made between expression of L- or M-opsins (*3*–*6*). Mutations that affect opsin expression or function cause various forms of color blindness and retinal degeneration (*7*–*9*). Great progress has been made in our understanding of the vertebrate eye through the study of model organisms. However, little is known about the developmental mechanisms that generate the mosaic of mutually exclusive cone subtypes in the human retina. We studied the specification of human cone subtypes using human retinal organoids differentiated from stem cells (Fig. 1, D to K).

Human retinal organoids generate photoreceptors that respond to light (*10*–*14*). We found that human organoids recapitulate the specification of cone subtypes observed in the human retina, including the temporal generation of S cones followed by L and M cones. Moreover, we found that this regulation is controlled by thyroid hormone signaling, which is necessary and sufficient to control cone subtype fates through the nuclear hormone receptor thyroid hormone receptor β (Thr*β*). Expression of thyroid hormone–regulating genes suggests that retina-intrinsic temporal control of thyroid hormone levels and activity governs cone subtype specification. Whereas retinal organoids have largely been studied for their promise of therapeutic applications (*15*), our work demonstrates that human organoids can also be used to reveal fundamental mechanisms of human development.

Specification of cone cells in organoids recapitulates development in the human retina

We compared features of cone subtypes in human organoids with those of adult retinal tissue. Adult human retinas and organoids at day 200 of differentiation displayed similar ratios of S to L/M cones as indicated by expression of S- or L/M-opsins (adult, S = 13%, L/M = 87%; organoid, S = 29%, L/M = 71%) (Fig. 1, B and C, and fig. S1A). The difference in the ratios is likely due to the immaturity of the organoid at ~6 months compared with the terminally differentiated adult retina. We examined L/M cones with an antibody that recognizes both L- and M-opsin proteins because of their extremely high similarity. Both S and L/M cones expressed the cone-rod-homeobox transcription factor (CRX), a critical transcription factor for photoreceptor differentiation (Fig. 2, A and E) (*16*–*18*), indicating proper fate specification in organoids. Additionally, cones in organoids and retinas displayed similar morphologies, with L/M cones that had longer outer segments and wider inner segments than those of S cones (Fig. 2, B to D and F to H) (*19*). The outer segments of cones were shorter in organoids than in adult retinas, which is consistent with postnatal maturation (Fig. 2, D and H) (*20*). Thus, cone subtypes in human retinal organoids displayed distributions, gene expression patterns, and morphologies similar to those of cones of the human retina.

We next examined the developmental dynamics of cone subtype specification in organoids. In the human retina, S cones are generated during fetal weeks 11 to 34 (days 77 to 238), whereas L/M cones are specified later, during fetal weeks 14 to 37 (days 98 to 259) (*21*, *22*). We tracked the ratios and densities of S and L/M cones in organoids by means of antibody staining over 360 days of differentiation. Cones expressing S-opsin were first observed at day 150 (Fig. 2, I, L, and M). The density of S cones leveled off at day 170 (Fig. 2M), at the time point when cones expressing L/M-opsin began to be observed (Fig. 2, J to M). The population of L/M cones increased dramatically until day 300 (Fig. 2, K to M), when they reached a steady-state density. The 20-day difference between S- and L/M-opsin expression onset in retinal organoids is similar to the 20-day difference observed in the appearance of S and L/M cones in the fetal retina (*21*). These observations show a temporal switch from S cone specification to L/M cone specification during retinal development.

We next conducted RNA sequencing (RNA-seq) through 250 days of induced pluripotent stem cell (iPSC)–derived organoid development. We found that *S-opsin* RNA was expressed first at day 111 and leveled off at day 160, whereas *L/M-opsin* RNA was expressed at day 160 and remained steady after day 180, which is consistent with the timeline of photoreceptor maturation in organoids and fetal retinas (Fig. 2N and fig. S1B). Moreover, *CRX* RNA and CRX protein were expressed before opsins in organoids, which is similar to human development (Fig. 2N and fig. S1, B to G) (*23*). Thus, human organoids recapitulate many aspects of the developmental timeline of cone subtype specification observed in human retinas, providing a model system with which to uncover the mechanisms of these developmental changes.

Thyroid hormone signaling and the temporal switch between S and L/M fate specification

Seminal work in mice identified Thr*β*2 as a critical regulator of cone subtype specification: *Thrβ2* mutants display a complete loss of M-opsin expression and a complete gain of S-opsin expression in cone photoreceptors (*24*–*26*). Similar roles for *Thrβ2* have been characterized in other organisms with highly divergent cone patterning (*27*–*29*). Additionally, rare human mutations in *Thrβ2* are reported to alter color perception, which is indicative of a change in the S-to-L/M cone ratio (*30*). To directly test the role of *Thrβ2* in human cone subtype specification, we used CRISPR/Cas9 in human embryonic stem cells (ESCs) to generate a homozygous mutation that resulted in early translational termination in the

[1]Department of Biology, Johns Hopkins University, 3400 N. Charles Street, Baltimore, MD 21218, USA. [2]Wilmer Eye Institute, Johns Hopkins University School of Medicine, Baltimore, MD 21287, USA. [3]Shiley Eye Institute, University of California, San Diego, La Jolla, CA 92093, USA. [4]National Institute of Mental Health, National Institutes of Health, Bethesda, MD 20892, USA. [5]Department of Computer Science, Johns Hopkins University, 3400 N. Charles Street, Baltimore, MD 21218, USA. [6]Department of Molecular Biology and Genetics, Johns Hopkins University School of Medicine, Baltimore, MD 21287, USA. [7]Department of Neuroscience, Johns Hopkins University School of Medicine, Baltimore, MD 21287, USA. [8]Institute of Genetic Medicine, Johns Hopkins University School of Medicine, Baltimore, MD 21287, USA.
*Corresponding author. Email: robertjohnston@jhu.edu

first exon of *Thrβ2* (fig. S2A). Surprisingly, organoids derived from these mutant stem cells displayed no differences in cone subtype ratio from genotypically wild-type organoids [wild type, S = 62%, L/M = 38%; *Thrβ2* knockout (KO), S = 59%, L/M = 41%; $P = 0.83$]. The S-to-L/M ratio is high for both wild-type controls and *Thrβ2* KO organoids, likely owing to variability in organoid differentiation. Thus, unlike previous suggestions based on other species, *Thrβ2* is dispensable for cone subtype specification in humans (Fig. 3, A to C).

Because *Thrβ2* alone is not required for human cone subtype specification, we reexamined data from Weiss *et. al* (*30*) and found that missense mutations in exons 9 and 10 affected both *Thrβ2* and another isoform of the human *Thrβ* gene, *Thrβ1* (fig. S2A). Thus, we asked whether *Thrβ1* and *Thrβ2* together are required for cone subtype specification in humans. To completely ablate Thrβ function (Thrβ1 and Thrβ2), we used CRISPR/Cas9 in human ESCs to delete a shared exon that codes for part of the DNA binding domain of *Thrβ* (fig. S2A). *Thrβ* null mutant retinal organoids displayed a complete conversion of all cones to the S subtype (wild type, S = 27%, L/M = 73%; *Thrβ* KO, S = 100%, L/M = 0%; $P < 0.0001$) (Fig. 3, D to E and H). In these mutants, all cones expressed S-opsin and had the S cone morphology (Fig. 3, I and J). Thus, *Thrβ* is required to activate L/M and to repress S cone fates in the human retina.

Thrβ binds with high affinity to triiodothyronine (T3), the more active form of thyroid hormone, to regulate gene expression (*31*). Depletion or addition of T3 alters the ratios of S to M cones in rodents (*25*, *32*, *33*). Because L/M cones differentiate after S cones, we hypothesized that T3 acts through Thrβ late in retinal development to induce L/M cone fate and repress S cone fate. One prediction of this hypothesis is that addition of T3 early in development will induce L/M fate and repress S fate. To test this model, we added 20 nM T3 to ESC- and iPSC-derived organoids starting from days 20 to 50 and continued until day 200 of differentiation. We observed a dramatic conversion of cone cells to L/M fate (wild type, S = 27%, L/M = 73%; wild type + T3, S = 4%, L/M = 96%; $P < 0.01$) (Fig. 3, F and H, and fig. S2B). Thus, early addition of T3 is sufficient to induce L/M fate and suppress S fate.

To test whether T3 acts specifically through Thrβ to control cone subtype specification, we differentiated *Thrβ* mutant organoids with early T3 addition. *Thrβ* mutation completely suppressed the effects of T3, generating organoids with only S cones (wild type + T3, S = 4%, L/M = 96%; *Thrβ* KO + T3, S = 100%, L/M = 0%; $P < 0.0001$) (Fig. 3, F to H). We conclude that T3 acts though Thrβ to promote L/M cone fate and suppress S cone fate.

We confirmed the regulation of L/M-opsin expression through thyroid hormone signaling in a retinoblastoma cell line, which expresses L/M-opsin when treated with T3 (fig. S2, C and D) (*34*). T3-induced activation of *L/M-opsin* expression was suppressed upon RNA interference knockdown of *Thrβ* (fig. S2, E and F), which is similar to the suppression observed in human organoids.

In organoids, early T3 addition not only converted cone cells to L/M fate but also dramatically increased cone density (Fig. 3, F and K). Moreover, T3 acts specifically through Thrβ to control cone density (Fig. 3, G and K). Early T3 addition may increase cone density by advancing and extending the temporal window of L/M cone generation.

Together, these results demonstrate that T3 signals though Thrβ to promote L/M cone fate and repress S cone fate in developing human retinal tissue.

Dynamic expression of thyroid hormone–regulating genes during development

Our data suggest that temporal control of thyroid hormone signaling determines the S-versus-L/M cone fate decision, in which low signaling early induces S fate and high signaling late induces L/M fate. Thyroid hormone exists largely in two states: thyroxine (T4), the most abundant circulating form of thyroid hormone, and T3, which binds thyroid hormone receptors with high affinity (*31*, *35*). Because the culture medium contains low amounts of T3 and T4, we hypothesized that the retina itself could modulate and/or generate thyroid hormone to control subtype fates.

Conversion of T4 to T3 occurs locally in target tissues to induce gene expression responses (*36*, *37*). Deiodinases—enzymes that modulate the levels of T3 and T4—are expressed in the retinas of mice, fish, and chickens (*29*, *38*–*42*). Therefore, we predicted that T3- and T4-degrading enzymes would be expressed during early human eye development to reduce thyroid hormone signaling and specify S cones, whereas T3-producing enzymes, carriers,

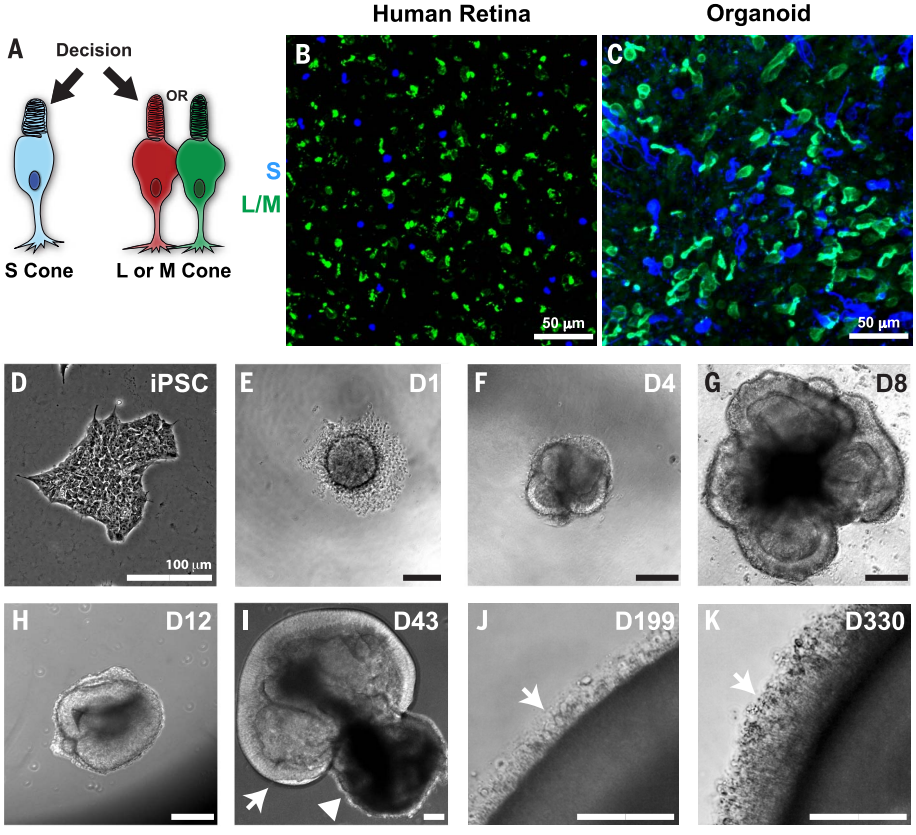


Fig. 1. S and L/M cone generation in human retinal organoids. (**A**) Decision between S and L/M cone subtype fate. (**B** and **C**) S-opsin (blue) and L/M-opsin (green). (B) Human adult retina age 53. (C) iPSC-derived organoid, day 200 of differentiation. (**D** to **K**) Bright-field images of organoids derived from iPSCs. (D) Undifferentiated iPSCs. (E) Day 1, aggregation. (F) Day 4, formation of neuronal vesicles. (G) Day 8, differentiation of retinal vesicles. (H) Day 12, manual isolation of retinal organoid. (I) Day 43, arrow indicates developing retinal tissue, and arrowhead indicates developing retinal pigment epithelium. (J) Day 199, arrow indicates outer segments. (K) Day 330, arrow indicates outer segments.

and transporters would be expressed later in human eye development to increase signaling and generate L/M cones.

To test these predictions, we examined gene expression across 250 days of organoid development. The expression patterns of thyroid hormone–regulating genes were grouped into three classes: changing expression (Fig. 4A), consistent expression (Fig. 4B), or no expression (Fig. 4C). Deiodinase 3 (*DIO3*), an enzyme that degrades T3 and T4 (*36*), was expressed at high levels early in organoid development but at low levels later (Fig. 4A). Conversely, deiodinase 2 (*DIO2*), an enzyme that converts T4 to active T3 (*36*), was expressed at low levels early but then dramatically increased over time (Fig. 4A). We examined RNA-seq data from Hoshino *et. al* (*23*) and found that developing human retinas display similar temporal changes in expression of *DIO3* and *DIO2* (fig. S3A). Deiodinase 1 (*DIO1*), which regulates T3 and T4 predominantly in the liver and kidney (*43*), was not expressed in organoids or retinas (Fig. 4C and fig. S3C). Thus, the dynamic expression of *Dio3* and *Dio2* supports low thyroid hormone signaling early in development to generate S cones and high thyroid hormone signaling late to produce L/M cones.

Consistent with a role for high thyroid hormone signaling in the generation of L/M cones later in development, expression of transthyretin (*TTR*), a thyroid hormone carrier protein, increased during organoid and retinal development (Fig. 4A and fig. S3A) (*23*). By contrast, albumin (*ALB*) and thyroxine-binding globulin (*SERPINA7*), other carrier proteins of T3 and T4, were not expressed in organoids or retinas (Fig. 4C and fig. S3C) (*23*).

T3 and T4 are transported into cells via membrane transport proteins (*44*). The T3/T4 transporters *SLC7A5* and *SLC7A8* increased in expression during organoid differentiation (Fig. 4A). Additionally, two T3/T4 transporters, *SLC3A2* and *SLC16A2*, were expressed at high and consistent levels throughout organoid development (Fig. 4B). Other T3/T4 transporters (*SLC16A10*, *SLCO1C1*, and *SLC5A5*) were not expressed in organoids (Fig. 4C), suggesting tissue-specific regulation of T3/T4 uptake. We observed similar expression patterns of T3/T4 transporters in human retinas (fig. S3, A to C) (*23*).

We next examined expression of transcriptional activators and repressors that mediate the response to thyroid hormone. Consistent with *Thrβ* expression in human cones (*45*), expression of *Thrβ* in organoids increased with time as cone cells were specified (Fig. 4A). Expression of thyroid hormone receptor α (*Thrα*) similarly increased with time (Fig. 4A). Thyroid hormone receptor cofactors, corepressor *NCoR2* and coactivator *MED1*, were expressed at steady levels during organoid differentiation (Fig. 4B). Similar temporal expression patterns were observed in human retinas (fig. S3, A and B) (*23*). Thus, our data suggest that expression of Thrβ and other transcriptional regulators enables gene regulatory responses to differential thyroid hormone levels.

A complex pathway controls production of thyroid hormone. Thyrotropin-releasing hormone (TRH) is produced by the hypothalamus and other neural tissue. TRH stimulates release of thyroid-stimulating hormone α (CGA) and thyroid-stimulating hormone β (TSHβ) from the pituitary gland. CGA and TSHβ bind the thyroid-stimulating hormone receptor (TSHR) in the thyroid gland. T3 and T4 production requires thyroglobulin (TG), the substrate for T3/T4 synthesis, and thyroid peroxidase (TPO), an enzyme that iodinates tyrosine residues in TG (*46*). *TRH* was expressed in organoids and retinas, but the other players were not (Fig. 4, A to C, and fig. S3, A to C) (*23*, *47*, *48*), suggesting that the retina itself does not generate thyroid hormone; rather, it modulates the relative levels of T3 and T4 and expresses TRH to signal for thyroid hormone production in other tissues.

Therefore, the temporal expression of thyroid hormone signaling regulators supports our model that the retina intrinsically controls T3 and T4 levels, ensuring low thyroid hormone signaling early to promote S fate and high thyroid hormone signaling late to specify L/M fate (Fig. 4D).

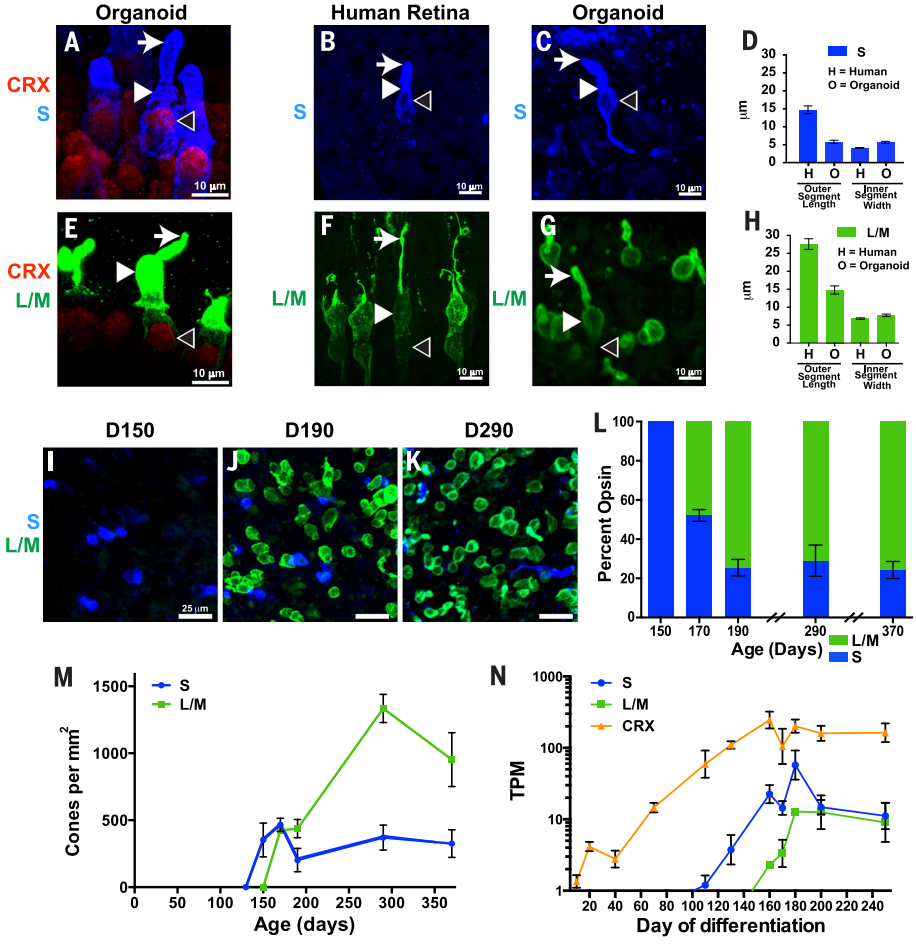


Fig. 2. Human cone subtype specification is recapitulated in organoids. (**A** to **K**) S-opsin (blue) and L/M-opsin (green) were examined in human iPSC-derived organoids [(A), (C) to (E), and (G) to (M)] and human retinas [(B), (D), (F), and (H)]. [(A) to (C) and (E) to (G)] Arrows indicate outer segments, solid arrowheads indicate inner segments, and open arrowheads indicate nuclei. [(A) and (E)] CRX (a general marker of photoreceptors) is expressed in S cones and L/M cones. [(B) to (D)] S cones display short outer segments and thin inner segments in both human retinas and organoids. [(F) to (H)] L/M cones display long outer segments and wide inner segments in both human retinas and organoids. [(D) and (H)] Quantification of outer segment lengths and inner segment widths (adult retina, L/M, n = 13 cones, S, n = 10 cones; organoid, L/M, n = 35 cones, S, n = 42 cones). [(I) to (N)] S cones are generated before L/M cones in organoids. (**L**) Ratio of S:L/M cones during organoid development. (**M**) Density of S and L/M cones during organoid development. (**N**) *S-opsin* expression precedes *L/M-opsin* expression in human iPSC-derived organoids. *CRX* expression starts before opsin expression. TPM, transcripts per kilobase million.

Organoids provide a powerful system with which to determine the mechanisms of human development. Model organism and epidemiological studies generate important hypotheses about human biology that are often experimentally intractable. This work shows that organoids enable direct testing of hypotheses in developing human tissue.

Our studies identify temporal regulation of thyroid hormone signaling as a mechanism that controls cone subtype specification in humans. Consistent with our findings, preterm human infants with low T3/T4 have an increased incidence of color vision defects (*49*–*52*). Moreover, our identification of a mechanism that generates one cone subtype while suppressing the other, coupled with successful transplantation and incorporation of stem cell–derived photoreceptors in mice (*53*–*56*), suggests that the promise of therapies to treat human diseases such as color blindness, retinitis pigmentosa, and macular degeneration will be achieved in the near future.

Materials and methods summary

Cell lines

H7 ESC (WA07, WiCell) and episomal-derived EP1.1 iPSC lines were used for differentiation. WERI-Rb1 retinoblastoma cells were obtained from ATCC. Cell maintenance and organoid differentiation protocols are described in the supplementary materials.

CRISPR mutations

All mutations were generated in H7 ESCs. Cells were modified to express an inducible Cas9 element. Plasmids for guide RNA (gRNA) transfection were generated by using the pSpCas9(BB)-P2A-Puro plasmid modified from the pX459_V2.0 plasmid (62988, Addgene) by replacing T2A with a P2A sequence. Mutations were confirmed with polymerase chain reaction sequencing. Gene diagrams of deletions are displayed in fig. S2A. Detailed transfection procedures, gRNA sequences, and homology arm sequences are included in the supplementary materials.

Immunohistochemistry

Primary antibodies were used at the following dilutions: goat anti-SW-opsin (1:200 for organoids, 1:500 for human retinas) (Santa Cruz Biotechnology), rabbit anti-LW/MW-opsins (1:200 for organoids, 1:500 for human retinas) (Millipore), mouse anti-CRX (1:500) (Abnova), and mouse anti-Rhodopsin (1:500) (GeneTex). All secondary antibodies were Alexa Fluor–conjugated (1:400) and made in donkey (Molecular Probes). Detailed methods for fixation, microscopy, and image processing of organoids, retinas, and WERI-Rb1 cells are included in the supplementary materials.

Organoid age

Opsin expression time course

EP1 iPSC–derived organoids for time course experiments were binned into 10-day increments for analysis. Organoids were binned into day 130 [actual day 129 (n = 3 organoids)], day 150 [actual day 152 (n = 4 organoids)], day 170 [actual day 173 (n = 2 organoids)], day 200 [actual days 194 to 199 (n = 7 organoids)], day 290 [actual day 291 (n = 3 organoids)], and day 360 [actual day 361 (n = 3 organoids)]. Quantifications of outer-segment lengths and inner-segment widths were measured in day 361 organoids (n = 3 organoids).

Opsin expression in different conditions

iCas9 H7 ESC–derived organoids for *Thrβ2* KOs and controls were analyzed at day 200. Organoids for *Thrβ* KO, control, and wild-type + T3 were analyzed at two time points: two organoids were

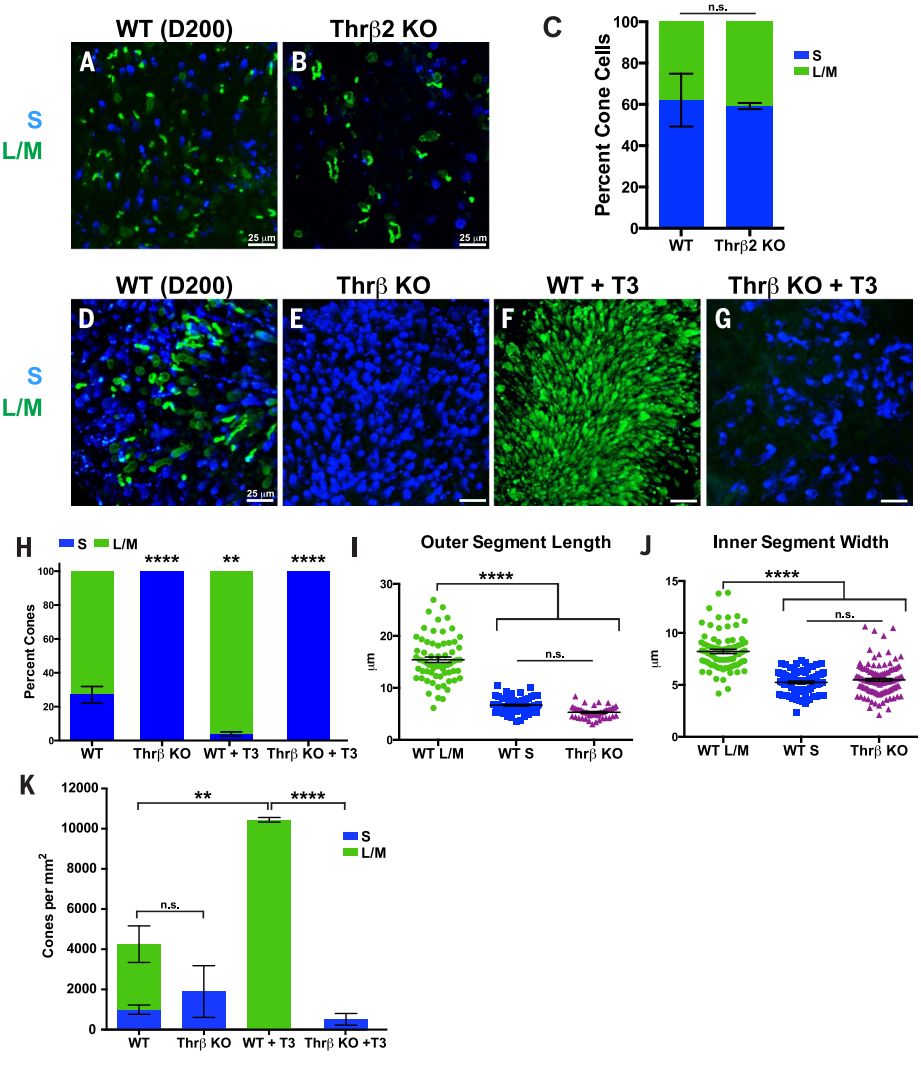


Fig. 3. Thyroid hormone signaling is necessary and sufficient for the temporal switch between S and L/M fate specification. (**A** to **K**) S-opsin (blue) and L/M-opsin (green) were examined in human ESC-derived organoids. (A) Wild-type (WT). (B) *Thrβ2* early termination mutant (*Thrβ2* KO). (C) Quantification of (A) and (B) (WT, n = 3 organoids; *Thrβ2 KO*, n = 3 organoids). (D) WT. (E) *Thrβ* KO. (F) WT treated with 20 nM T3 (WT + T3). (G) *Thrβ* KO treated with 20 nM T3 (*Thrβ* KO + T3). (H) Quantification of (D) to (G) (WT, n = 9 organoids; *Thrβ* KO, n = 3 organoids; WT + T3, n = 6 organoids; *Thrβ* KO + T3, n = 3 organoids. Tukey's multiple comparisons test: WT versus *Thrβ* KO, $P < 0.0001$; WT versus WT + T3, $P < 0.01$; WT + T3 versus *Thrβ* KO + T3, $P < 0.0001$). (I) Length of outer segments. WT, L/M n = 66 cells; WT, S n = 66 cells; *Thrβ* KO, n = 50 cells (Tukey's multiple comparisons test, WT L/M versus WT S, $P < 0.0001$; WT L/M versus *Thrβ* KO, $P < 0.0001$; WT S versus *Thrβ* KO, not significantly different). (J) Width of inner segments. WT, L/M n = 78 cells; WT, S n = 78 cells; *Thrβ* KO, n = 118 cells (Tukey's multiple comparisons test, WT L/M versus WT S, $P < 0.0001$; WT L/M versus *Thrβ* KO, $P < 0.0001$; WT S versus *Thrβ* KO, not significantly different). (K) T3 acts through *Thrβ* to increase total cone number. Quantification of density of S and L/M cones; WT, n = 6 organoids; *Thrβ* KO, n = 3 organoids; WT + T3, n = 3 organoids; *Thrβ* KO + T3, n = 3 organoids (Tukey's multiple comparisons test between total cone numbers, WT versus *Thrβ* KO, not significantly different; WT versus WT + T3, $P < 0.01$; WT + T3 versus *Thrβ* KO + T3, $P < 0.0001$).

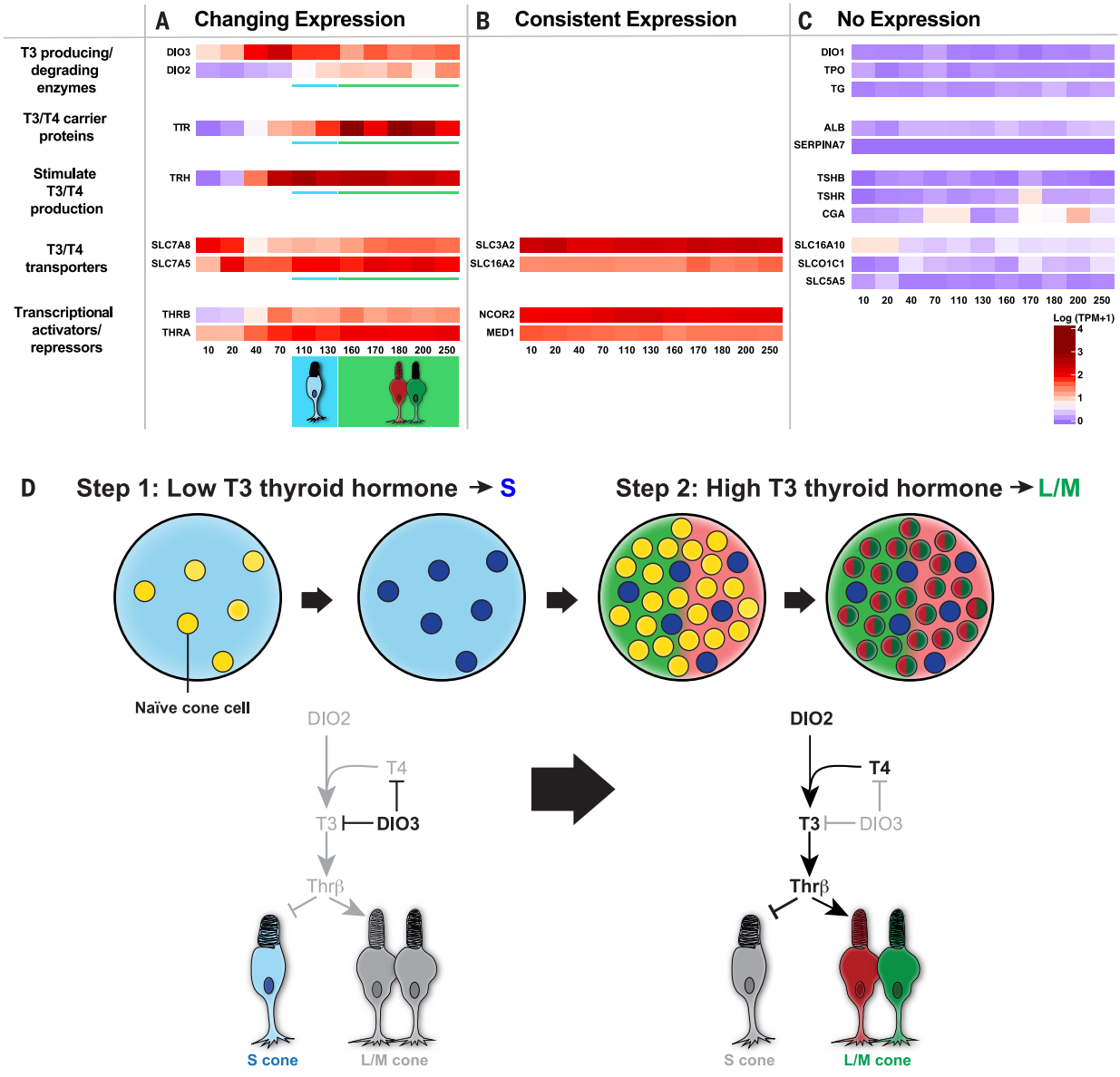


Fig. 4. Dynamic expression of thyroid hormone signaling regulators during development. (**A** to **C**) Heat maps of log(TPM + 1) values for genes with (A) changing expression, (B) consistent expression, and (C) no expression. Numbers at the bottom of heat maps indicate organoid age in days. (**D**) Model of the temporal mechanism of cone subtype specification in humans. For simplicity, only the roles of DIO3 and DIO2 are illustrated. In step 1, expression of DIO3 degrades T3 and T4, leading to S cone specification. In step 2, expression of DIO2 converts T4 to T3 to signal Thrβ to repress S and induce L/M cone fate.

taken at day 199 for each group, and one was taken at day 277 for each group. T3-treated organoids were taken at time points between day 195 and day 200 for different differentiations. For each treatment group and genotype, organoids were compared with control organoids grown in parallel.

RNA-seq time course

EP1 iPSC–derived organoids were analyzed at time points ranging from day 10 to day 250 of differentiation. We took samples at day 10 (n = 3 organoids), day 20 (n = 2 organoids), day 35 (n = 3 organoids), day 69 (n = 3 organoids), day 111 (n = 3 organoids), day 128 (n = 3 organoids), day 158 (n = 2 organoids), day 173 (n = 3 organoids), day 181 (n = 3 organoids), day 200 (n = 3 organoids), and day 250 (n = 3 organoids). RNA from individual organoids was extracted by using the Zymo Direct-zol RNA Microprep Kit (Zymo Research) according to manufacturer's instructions. Libraries were prepared by using the Illumina TruSeq stranded mRNA kit and sequenced on an Illumina NextSeq 500 with single 200–base pair reads.

RNA-seq time course analysis

Expression levels were quantified by using Kallisto (version 0.34.1) with the following parameters: "-b 100 -l 200 -s 10 -t 20–single". The Gencode release 28 comprehensive annotation was used as the reference transcriptome (*57*). Transcripts per million (TPM) values (table S1) were then used to generate graphs in Prism and heatmaps in R by using ggplot2. The distributions of transcripts were plotted so as to identify the best low TPM cutoff (fig. S5A). The threshold was determined to be 0.7 log(TPM + 1)—5 TPM—and this value was used as an inflection point for the heatmaps. Heatmaps for fig. S3, A to C, were made similarly, by using CPM values from Hoshino *et. al* (fig. S5B) (*23*).

Measurements and quantification

Measurements of retinal area and cell morphology were done by using ImageJ software. Quantifications and statistics (except for RNA-seq data)

were done in GraphPad Prism, with a significance cutoff of 0.01. Statistical tests are listed in figure legends. All error bars represent the SEM.

REFERENCES AND NOTES

1. K. Viets, K. Eldred, R. J. Johnston Jr., Mechanisms of photoreceptor patterning in vertebrates and invertebrates. *Trends Genet.* **32**, 638–659 (2016). doi: 10.1016/j.tig.2016.07.004; pmid: 27615122
2. J. Nathans, D. Thomas, D. S. Hogness, Molecular genetics of human color vision: The genes encoding blue, green, and red pigments. *Science* **232**, 193–202 (1986). doi: 10.1126/science.2937147; pmid: 2937147
3. D. Vollrath, J. Nathans, R. W. Davis, Tandem array of human visual pigment genes at Xq28. *Science* **240**, 1669–1672 (1988). doi: 10.1126/science.2837827; pmid: 2837827
4. Y. Wang *et al*., A locus control region adjacent to the human red and green visual pigment genes. *Neuron* **9**, 429–440 (1992). doi: 10.1016/0896-6273(92)90181-C; pmid: 1524826
5. P. M. Smallwood, Y. Wang, J. Nathans, Role of a locus control region in the mutually exclusive expression of human red and green cone pigment genes. *Proc. Natl. Acad. Sci. U.S.A.* **99**, 1008–1011 (2002). doi: 10.1073/pnas.022629799; pmid: 11773636
6. Y. Wang *et al*., Mutually exclusive expression of human red and green visual pigment-reporter transgenes occurs at high frequency in murine cone photoreceptors. *Proc. Natl. Acad. Sci. U.S.A.* **96**, 5251–5256 (1999). doi: 10.1073/pnas.96.9.5251; pmid: 10220452
7. J. Nathans *et al*., Molecular genetics of human blue cone monochromacy. *Science* **245**, 831–838 (1989). doi: 10.1126/science.2788922; pmid: 2788922
8. A. S. Ladekjaer-Mikkelsen, T. Rosenberg, A. L. Jørgensen, A new mechanism in blue cone monochromatism. *Hum. Genet.* **98**, 403–408 (1996). doi: 10.1007/s004390050229; pmid: 8792812
9. E. J. Patterson *et al*., Cone photoreceptor structure in patients with X-linked cone dysfunction and red-green color vision deficiency. *Invest. Ophthalmol. Vis. Sci.* **57**, 3853–3863 (2016). doi: 10.1167/iovs.16-19608; pmid: 27447086
10. T. Nakano *et al*., Self-formation of optic cups and storable stratified neural retina from human ESCs. *Cell Stem Cell* **10**, 771–785 (2012). doi: 10.1016/j.stem.2012.05.009; pmid: 22704518
11. X. Zhong *et al*., Generation of three-dimensional retinal tissue with functional photoreceptors from human iPSCs. *Nat. Commun.* **5**, 4047 (2014). doi: 10.1038/ncomms5047; pmid: 24915161
12. K. J. Wahlin *et al*., Photoreceptor outer segment-like structures in long-term 3D retinas from human pluripotent stem cells. *Sci. Rep.* **7**, 766 (2017). doi: 10.1038/s41598-017-00774-9; pmid: 28396597
13. R. Kaewkhaw *et al*., Transcriptome dynamics of developing photoreceptors in three-dimensional retina cultures recapitulates temporal sequence of human cone and rod differentiation revealing cell surface markers and gene networks. *Stem Cells* **33**, 3504–3518 (2015). doi: 10.1002/stem.2122; pmid: 26235913
14. M. J. Phillips *et al*., A novel approach to single cell RNA-sequence analysis facilitates in silico gene reporting of human pluripotent stem cell-derived retinal cell types. *Stem Cells* **36**, 313–324 (2018). doi: 10.1002/stem.2755; pmid: 29230913
15. A. Artero Castro, D. Lukovic, P. Jendelova, S. Erceg, Concise review: Human induced pluripotent stem cell models of retinitis pigmentosa. *Stem Cells* **36**, 474–481 (2018). doi: 10.1002/stem.2783; pmid: 29345014
16. T. Furukawa, E. M. Morrow, C. L. Cepko, Crx, a novel otx-like homeobox gene, shows photoreceptor-specific expression and regulates photoreceptor differentiation. *Cell* **91**, 531–541 (1997). doi: 10.1016/S0092-8674(00)80439-0; pmid: 9390562
17. C. L. Freund *et al*., Cone-rod dystrophy due to mutations in a novel photoreceptor-specific homeobox gene (CRX) essential for maintenance of the photoreceptor. *Cell* **91**, 543–553 (1997). doi: 10.1016/S0092-8674(00)80440-7; pmid: 9390563
18. S. Chen *et al*., Crx, a novel Otx-like paired-homeodomain protein, binds to and transactivates photoreceptor cell-specific genes. *Neuron* **19**, 1017–1030 (1997). doi: 10.1016/S0896-6273(00)80394-3; pmid: 9390516
19. C. A. Curcio *et al*., Distribution and morphology of human cone photoreceptors stained with anti-blue opsin. *J. Comp. Neurol.* **312**, 610–624 (1991). doi: 10.1002/cne.903120411; pmid: 1722224
20. A. Hendrickson, D. Drucker, The development of parafoveal and mid-peripheral human retina. *Behav. Brain Res.* **49**, 21–31 (1992). doi: 10.1016/S0166-4328(05)80191-3; pmid: 1388798
21. M. Xiao, A. Hendrickson, Spatial and temporal expression of short, long/medium, or both opsins in human fetal cones. *J. Comp. Neurol.* **425**, 545–559 (2000). doi: 10.1002/1096-9861(20001002)425:4<545::AID-CNE6>3.0.CO;2-3; pmid: 10975879
22. C. A. Curcio, K. R. Sloan, R. E. Kalina, A. E. Hendrickson, Human photoreceptor topography. *J. Comp. Neurol.* **292**, 497–523 (1990). doi: 10.1002/cne.902920402; pmid: 2324310
23. A. Hoshino *et al*., Molecular anatomy of the developing human retina. *Dev. Cell* **43**, 763–779.e4 (2017). doi: 10.1016/j.devcel.2017.10.029; pmid: 29233477
24. L. Ng *et al*., A thyroid hormone receptor that is required for the development of green cone photoreceptors. *Nat. Genet.* **27**, 94–98 (2001). doi: 10.1038/83829; pmid: 11138006
25. M. R. Roberts, M. Srinivas, D. Forrest, G. Morreale de Escobar, T. A. Reh, Making the gradient: Thyroid hormone regulates cone opsin expression in the developing mouse retina. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 6218–6223 (2006). doi: 10.1073/pnas.0509981103; pmid: 16606843
26. M. L. Applebury *et al*., Transient expression of thyroid hormone nuclear receptor TRbeta2 sets S opsin patterning during cone photoreceptor genesis. *Dev. Dyn.* **236**, 1203–1212 (2007). doi: 10.1002/dvdy.21155; pmid: 17436273
27. S. C. Suzuki *et al*., Cone photoreceptor types in zebrafish are generated by symmetric terminal divisions of dedicated precursors. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 15109–15114 (2013). doi: 10.1073/pnas.1303551110; pmid: 23980162
28. M. Sjöberg, B. Vennström, D. Forrest, Thyroid hormone receptors in chick retinal development: Differential expression of mRNAs for alpha and N-terminal variant beta receptors. *Development* **114**, 39–47 (1992). pmid: 1576965
29. J. M. Trimarchi, S. Harpavat, N. A. Billings, C. L. Cepko, Thyroid hormone components are expressed in three sequential waves during development of the chick retina. *BMC Dev. Biol.* **8**, 101 (2008). doi: 10.1186/1471-213X-8-101; pmid: 18854032
30. A. H. Weiss, J. P. Kelly, D. Bisset, S. S. Deeb, Reduced L- and M- and increased S-cone functions in an infant with thyroid hormone resistance due to mutations in the THRβ2 gene. *Ophthalmic Genet.* **33**, 187–195 (2012). doi: 10.3109/13816810.2012.681096; pmid: 22551329
31. H. H. Samuels, J. S. Tsai, J. Casanova, F. Stanley, Thyroid hormone action: In vitro characterization of solubilized nuclear receptors from rat liver and cultured GH1 cells. *J. Clin. Invest.* **54**, 853–865 (1974). doi: 10.1172/JCI107825; pmid: 4372251
32. A. Glaschke, M. Glösmann, L. Peichl, Developmental changes of cone opsin expression but not retinal morphology in the hypothyroid Pax8 knockout mouse. *Invest. Ophthalmol. Vis. Sci.* **51**, 1719–1727 (2010). doi: 10.1167/iovs.09-3592; pmid: 19834026
33. A. Glaschke *et al*., Thyroid hormone controls cone opsin expression in the retina of adult rodents. *J. Neurosci.* **31**, 4844–4851 (2011). doi: 10.1523/JNEUROSCI.6181-10.2011; pmid: 21451022
34. Y. Liu, L. Fu, D. G. Chen, S. S. Deeb, Identification of novel retinal target genes of thyroid hormone in the human WERI cells by expression microarray analysis. *Vision Res.* **47**, 2314–2326 (2007). doi: 10.1016/j.visres.2007.04.023; pmid: 17655910
35. A. Schroeder, R. Jimenez, B. Young, M. L. Privalsky, The ability of thyroid hormone receptors to sense t4 as an agonist depends on receptor isoform and on cellular cofactors. *Mol. Endocrinol.* **28**, 745–757 (2014). doi: 10.1210/me.2013-1335; pmid: 24673558
36. M. Dentice, A. Marsili, A. Zavacki, P. R. Larsen, D. Salvatore, The deiodinases and the control of intracellular thyroid hormone signaling during cellular differentiation. *Biochim. Biophys. Acta* **1830**, 3937–3945 (2013). doi: 10.1016/j.bbagen.2012.05.007; pmid: 22634734
37. V. M. Darras, A. M. Houbrechts, S. L. Van Herck, Intracellular thyroid hormone metabolism as a local regulator of nuclear thyroid hormone receptor-mediated impact on vertebrate development. *Biochim. Biophys. Acta* **1849**, 130–141 (2015). doi: 10.1016/j.bbagrm.2014.05.004; pmid: 24844179
38. L. Ng *et al*., Type 3 deiodinase, a thyroid-hormone-inactivating enzyme, controls survival and maturation of cone photoreceptors. *J. Neurosci.* **30**, 3347–3357 (2010). doi: 10.1523/JNEUROSCI.5267-09.2010; pmid: 20203194
39. P. J. Bonezzi, M. E. Stabio, J. M. Renna, The development of mid-wavelength photoresponsivity in the mouse retina. *Curr. Eye Res.* **43**, 666–673 (2018). doi: 10.1080/02713683.2018.1433859; pmid: 29447486
40. E. Bagci *et al*., Deiodinase knockdown during early zebrafish development affects growth, development, energy metabolism, motility and phototransduction. *PLOS ONE* **10**, e0123285 (2015). doi: 10.1371/journal.pone.0123285; pmid: 25855985
41. C. Guo *et al*., Intrinsic expression of a multiexon type 3 deiodinase gene controls zebrafish embryo size. *Endocrinology* **155**, 4069–4080 (2014). doi: 10.1210/en.2013-2029; pmid: 25004091
42. S. L. Bruhn, C. L. Cepko, Development of the pattern of photoreceptors in the chick retina. *J. Neurosci.* **16**, 1430–1439 (1996). doi: 10.1523/JNEUROSCI.16-04-01430.1996; pmid: 8778294
43. A. C. Bianco, D. Salvatore, B. Gereben, M. J. Berry, P. R. Larsen, Biochemistry, cellular and molecular biology, and physiological roles of the iodothyronine selenodeiodinases. *Endocr. Rev.* **23**, 38–89 (2002). doi: 10.1210/edrv.23.1.0455; pmid: 11844744
44. D. S. Sharlin, T. J. Visser, D. Forrest, Developmental and cell-specific expression of thyroid hormone transporters in the mouse cochlea. *Endocrinology* **152**, 5053–5064 (2011). doi: 10.1210/en.2011-1372; pmid: 21878515
45. T. C. Lee, D. Almeida, N. Claros, D. H. Abramson, D. Cobrinik, Cell cycle-specific and cell type-specific expression of Rb in the developing human retina. *Invest. Ophthalmol. Vis. Sci.* **47**, 5590–5598 (2006). doi: 10.1167/iovs.06-0063; pmid: 17122153
46. E. J. Barrett, in *Medical Physiology, 2e Updated Edition*, W. B. E. Boulpaep, Ed. (Elsevier, Inc., Philadelphia, PA, 2012), chap. 49.
47. S. R. Dubovy *et al*., Expression of hypothalamic neurohormones and their receptors in the human eye. *Oncotarget* **8**, 66796–66814 (2017). doi: 10.18632/oncotarget.18358; pmid: 28977997
48. E. Martino *et al*., Thyrotropin-releasing hormone-like material in human retina. *J. Endocrinol. Invest.* **3**, 267–271 (1980). doi: 10.1007/BF03348274; pmid: 6776180
49. J. Rovet, N. Simic, The role of transient hypothyroxinemia of prematurity in development of visual abilities. *Semin. Perinatol.* **32**, 431–437 (2008). doi: 10.1053/j.semperi.2008.09.009; pmid: 19007682
50. N. Simic, C. Westall, E. V. Astzalos, J. Rovet, Visual abilities at 6 months in preterm infants: Impact of thyroid hormone deficiency and neonatal medical morbidity. *Thyroid* **20**, 309–315 (2010). doi: 10.1089/thy.2009.0128; pmid: 20144040
51. S. A. Yassin, A. J. Al-Dawood, W. M. Al-Zamil, M. A. Al-Ghamdi, Z. N. Al-Khudairy, Comparative study of visual dysfunctions in 6-10-year-old very preterm- and full-term-born children. *Int. Ophthalmol.* (2018). doi: 10.1007/s10792-018-0959-2; pmid: 29916121
52. H. J. Dowdeswell, A. M. Slater, J. Broomhall, J. Tripp, Visual deficits in children born at less than 32 weeks' gestation with and without major ocular pathology and cerebral damage. *Br. J. Ophthalmol.* **79**, 447–452 (1995). doi: 10.1136/bjo.79.5.447; pmid: 7612557
53. R. A. Pearson *et al*., Restoration of vision after transplantation of photoreceptors. *Nature* **485**, 99–103 (2012). doi: 10.1038/nature10997; pmid: 22522934
54. A. O. Barnea-Cramer *et al*., Function of human pluripotent stem cell-derived photoreceptor progenitors in blind mice. *Sci. Rep.* **6**, 29784 (2016). doi: 10.1038/srep29784; pmid: 27405580
55. D. A. Lamba, J. Gust, T. A. Reh, Transplantation of human embryonic stem cell-derived photoreceptors restores some visual function in Crx-deficient mice. *Cell Stem Cell* **4**, 73–79 (2009). doi: 10.1016/j.stem.2008.10.015; pmid: 19128794
56. B. A. Tucker *et al*., Transplantation of adult mouse iPS cell-derived photoreceptor precursors restores retinal structure and function in degenerative mice. *PLOS ONE* **6**, e18992 (2011). doi: 10.1371/journal.pone.0018992; pmid: 21559507
57. J. Harrow *et al*., GENCODE: The reference human genome annotation for The ENCODE Project. *Genome Res.* **22**, 1760–1774 (2012). doi: 10.1101/gr.135350.111; pmid: 22955987

ACKNOWLEDGMENTS

We thank A. Kolodkin, J. Nathans, and members of the Johnston laboratory for helpful comments on the manuscript. **Funding:** K.C.E. was a Howard Hughes Medical Institute Gilliam Fellow and was supported by the National Science Foundation Graduate Research Fellowship Program under grant 1746891. R.J.J. was

supported by the Pew Scholar Award 00027373. **Author Contributions:** K.C.E.: Conception, data acquisition, new reagent contribution, data analysis, and data interpretation; drafted and revised manuscript. S.E.H.: Data acquisition and data interpretation. K.A.H.: Data acquisition, data analysis, and data interpretation. B.B.: Data analysis and data interpretation. P.-W.Z.: New reagent contribution. X.C.: New reagent contribution. V.M.S.: New reagent contribution. D.S.W.: New reagent contribution. S.H.: Data interpretation. J.T.: Data analysis and data interpretation. K.W.: Data acquisition and new reagent contribution. D.J.Z.: Data acquisition and new reagent contribution. R.J.J.: Conception and data interpretation; drafted and revised manuscript. **Competing interests:** None. **Data and materials availability:** RNA-seq data are available on Gene Expression Omnibus, accession no. GSE119320. All other data and methods are in the supplementary materials. H7 stem cells are available from WiCell under a materials transfer agreement with WiCell.

SUPPLEMENTARY MATERIALS

30 June 2018; accepted 30 August 2018
10.1126/science.aau6348

RESEARCH ARTICLE

SUPERNOVAE

A hot and fast ultra-stripped supernova that likely formed a compact neutron star binary

K. De[1]*, M. M. Kasliwal[1], E. O. Ofek[2], T. J. Moriya[3], J. Burke[4,5], Y. Cao[6], S. B. Cenko[7,8], G. B. Doran[9], G. E. Duggan[1], R. P. Fender[10], C. Fransson[11], A. Gal-Yam[2], A. Horesh[12], S. R. Kulkarni[1], R. R. Laher[13], R. Lunnan[11], I. Manulis[2], F. Masci[13], P. A. Mazzali[14,15], P. E. Nugent[16,17], D. A. Perley[14], T. Petrushevska[18,19], A. L. Piro[20], C. Rumsey[21], J. Sollerman[11], M. Sullivan[22], F. Taddia[11]

Compact neutron star binary systems are produced from binary massive stars through stellar evolution involving up to two supernova explosions. The final stages in the formation of these systems have not been directly observed. We report the discovery of iPTF 14gqr (SN 2014ft), a type Ic supernova with a fast-evolving light curve indicating an extremely low ejecta mass (≈0.2 solar masses) and low kinetic energy (≈2 × 10^{50} ergs). Early photometry and spectroscopy reveal evidence of shock cooling of an extended helium-rich envelope, likely ejected in an intense pre-explosion mass-loss episode of the progenitor. Taken together, we interpret iPTF 14gqr as evidence for ultra-stripped supernovae that form neutron stars in compact binary systems.

Core-collapse supernovae (SNe) are the violent deaths of massive stars when they run out of nuclear fuel in their cores and collapse, forming a neutron star (NS) or black hole (BH) (*1*). For massive stars that have lost some or all of their outer hydrogen (H) and helium (He) envelope, the resulting collapse produces a stripped-envelope supernova (SN) (*2*). The amount of material stripped from the star is a sensitive function of the initial mass of the star and its environment; if the star was born in a binary system, it also depends on the orbital properties of the system and the nature of the companion (*2*, *3*).

Because most massive stars are born in close binary systems (*4*), stripping via binary interactions likely plays a large role in producing the observed diversity of stripped-envelope SNe (*5*, *6*). For the most compact companions in close orbits, the stripping of massive stars may be extensive enough to completely remove their outer layers, leaving behind a naked metal core close to the minimum mass required for the core to collapse (the Chandrasekhar mass). If massive enough, the highly stripped core eventually collapses to produce a faint and fast-evolving SN explosion that ejects a small amount of material (*7*, *8*). Although it has been difficult to securely identify these explosions, such "ultra-stripped" SNe have been suggested to lead to the formation of a variety of compact NS binary systems [i.e., a NS in orbit around another NS, white dwarf (WD), or BH] (*7*, *9*).

Discovery and follow-up of iPTF 14gqr

iPTF 14gqr (SN 2014ft) was discovered by the intermediate Palomar Transient Factory (iPTF) (*10*, *11*) on 14.18 October 2014 UTC (universal time coordinated) at a *g*-band optical magnitude of ≈20.2 mag. The source was not detected in the previous observation on 13.32 October 2014 UTC (0.86 days before discovery), with a limiting magnitude of $g \geq 21.5$ mag. The transient was found in the outskirts (at a projected offset of ≈29 kpc from the center) of a tidally interacting spiral galaxy (IV Zw 155) at redshift $z = 0.063$ and luminosity distance $D = 284.5$ megaparsecs (Fig. 1). We obtained rapid ultraviolet (UV), optical, and near-infrared follow-up observations of the source, including a sequence of four spectra within 24 hours of the first detection (*12*).

We also obtained multiepoch x-ray and radio observations and found that the source remained undetected at these wavelengths (*12*). These upper limits rule out luminous nonthermal emission, such as that typically seen in relativistic and gamma-ray burst–associated SNe, but are not stringent enough to constrain the environment of the progenitor (figs. S11 and S12).

Our photometric follow-up indicated that the source rapidly faded within a day of detection, followed by rebrightening to a second peak on a longer time scale (rising over ≈7 days) (*12*) (Fig. 2). The early decline was detected in all optical and UV photometric bands and was characterized by a blackbody spectrum that cooled rapidly from a temperature $T > 32{,}000$ K near first detection to $T \sim 10{,}000$ K at 1 day after discovery (Figs. 3 and 4). Our early spectra also exhibit blackbody continua with temperatures consistent with those inferred from the photometry, superimposed with intermediate-width emission lines of He II, C III, and C IV. Such high ionization lines, which are typically associated with elevated pre-explosion mass-loss episodes in massive stars, have not been seen in early spectra of previously observed hydrogen-poor SNe. Although similar features are present in the early spectra of some hydrogen-rich core-collapse SNe (*13*–*15*) (fig. S6), the relatively large widths of the lines (full width at half maximum ~ 2000 to 4000 km s^{-1}) and the rapid evolution of the 4686-Å emission feature (Fig. 3) are not.

Spectra obtained near the second peak are dominated by emission from the expanding photosphere and exhibit relatively blue continua, with broad absorption features reminiscent of normal stripped-envelope SNe of type Ic that do not exhibit absorption lines of H or He in the spectra (*16*) (fig. S7). We find associated absorption velocities of ~10,000 km s^{-1} (*12*). The photometric properties of the second peak are broadly consistent with a number of previously observed fast type Ic events (figs. S3 and S4), but the rapidly declining first peak and the fast rise time to the second

[1]Cahill Centre for Astrophysics, California Institute of Technology, 1200 East California Boulevard, Pasadena, CA 91125, USA. [2]Department of Particle Physics and Astrophysics, Faculty of Physics, The Weizmann Institute of Science, Rehovot 76100, Israel. [3]Division of Theoretical Astronomy, National Astronomical Observatory of Japan, National Institutes of Natural Sciences, 2-21-1 Osawa, Mitaka, Tokyo 181-8588, Japan. [4]Las Cumbres Observatory, 6740 Cortona Drive, Suite 102, Goleta, CA 93117, USA. [5]Department of Physics, University of California, Santa Barbara, CA 93106, USA. [6]Department of Astronomy, University of Washington, Box 351580, Seattle, WA 98195, USA. [7]Astrophysics Science Division, NASA Goddard Space Flight Center, Mail Code 661, Greenbelt, MD 20771, USA. [8]Joint Space-Science Institute, University of Maryland, College Park, MD 20742, USA. [9]Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA 91109, USA. [10]Department of Physics, Astrophysics, University of Oxford, Denys Wilkinson Building, Oxford OX1 3RH, UK. [11]Oskar Klein Centre, Department of Astronomy, Stockholm University, 106 91 Stockholm, Sweden. [12]Racah Institute of Physics, The Hebrew University of Jerusalem, Jerusalem 91904, Israel. [13]Infrared Processing and Analysis Center, California Institute of Technology, MS 100-22, Pasadena, CA 91125, USA. [14]Astrophysics Research Institute, Liverpool John Moores University, Liverpool L3 5RF, UK. [15]Max-Planck-Institut für Astrophysik, Karl-Schwarzschild-Str. 1, D-85748 Garching bei München, Germany. [16]Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA. [17]Department of Astronomy, University of California, Berkeley, CA, 94720, USA. [18]Oskar Klein Centre, Department of Physics, Stockholm University, 106 91 Stockholm, Sweden. [19]Centre for Astrophysics and Cosmology, University of Nova Gorica, Vipavska 11c, 5270 Ajdovščina, Slovenia. [20]The Observatories of the Carnegie Institution for Science, 813 Santa Barbara Street, Pasadena, CA 91101, USA. [21]Astrophysics Group, Cavendish Laboratory, 19 J J Thomson Avenue, Cambridge CB3 0HE, UK. [22]Department of Physics and Astronomy, University of Southampton, Southampton SO17 1BJ, UK.
*Corresponding author. Email: kde@astro.caltech.edu

peak are unlike previously observed events. The source quickly faded after the second peak, declining at a rate of 0.21 mag day^{-1} in the g band (*12*). Our final spectrum taken at ≈34 days after explosion shows that the source exhibited an early transition to the nebular phase on a time scale faster than that for previously observed core-collapse SNe. The nebular phase spectrum exhibits prominent [Ca II] emission similar to several other type Ic SNe (fig. S8).

Multicolor photometry at multiple epochs allow us to trace the evolution of the optical/UV spectral energy distribution (SED), which we use to construct bolometric light curves that contain flux integrated over all wavelengths (Figs. 3 and 4) (*12*). We fit the pseudo-bolometric light curve of iPTF 14gqr with a simple Arnett model (*17*) to estimate the explosion

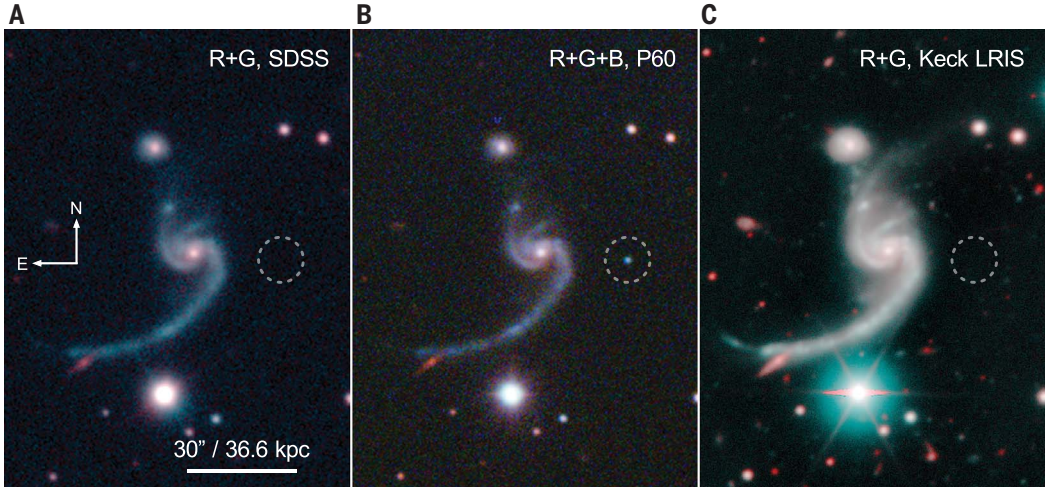


Fig. 1. Discovery field and host galaxy of iPTF 14gqr. (**A**) Optical image of the field from the Sloan Digital Sky Survey (SDSS) (*51*); R and G filter images have been used for red and cyan colors, respectively. (**B**) Composite RGB image (R, G, and B filter images have been used for red, green, and blue colors, respectively) of the iPTF 14gqr field from images taken near the second peak (19 October 2014) with the Palomar 60-inch telescope (P60), showing a blue transient inside the white dashed circle at the discovery location. (**C**) Late-time composite R+G image (R and G filter images have been used for red and cyan colors, respectively) of the host galaxy taken with the Low Resolution Imaging Spectrograph on the Keck I telescope.

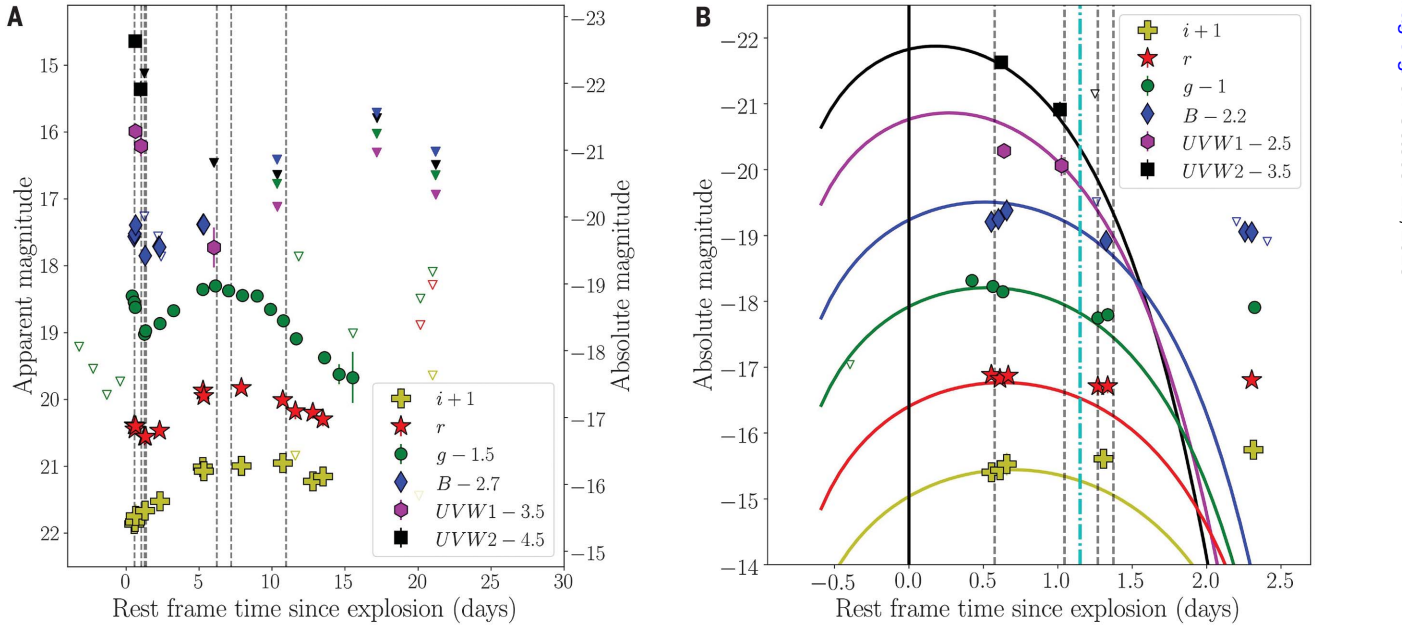


Fig. 2. Multicolor photometric observations of iPTF 14gqr. (**A**) Multicolor light curves of iPTF 14gqr from our photometric follow-up observations (magnitudes are corrected for galactic extinction and offset vertically as indicated in the legend). Inverted triangles denote 5σ upper limits, whereas other symbols denote detections. Hollow inverted triangles are upper limits from P48/P60 imaging, and the filled inverted triangles are upper limits from Swift observations (filled green triangles are *V* band limits from Swift). Epochs when spectra were obtained are marked in both panels by vertical black dashed lines. (**B**) Zoomed-in view of the early evolution of the light curve. The black solid line shows the assumed explosion epoch. The colored solid lines show the best-fitting shock-cooling model for extended progenitors (*25*). Only photometric data before the cyan dot-dashed vertical line were used in the fitting (*12*).

parameters. Allowing the explosion time to vary as a free parameter, we estimate an ejecta mass $M_{ej} \approx 0.15$ to 0.30 solar masses ($M_\odot$), an explosion kinetic energy $E_K \approx (1.0 \text{ to } 1.9) \times 10^{50}$ ergs, and a synthesized Ni mass $M_{Ni} \approx 0.05\ M_\odot$ (*12*) (Fig. 4 and fig. S9). The inferred ejecta mass is lower than known core-collapse type Ic SNe (*18*–*20*), which have ejecta masses in the higher range of ~0.7 to 15 $M_\odot$, with a mean of 2 to 3 $M_\odot$ over a sample of ≈20 SNe. However, the parameters of iPTF 14gqr are similar to those inferred for SN 2005ek (*21*) and 2010X (*22*), rapidly evolving type I SNe whose physical origins remain a matter of debate.

The rapid decline of the first peak observed in iPTF 14gqr is reminiscent of shock-cooling emission from the outer layers of a progenitor after the core-collapse SN shock breaks out (*23*, *24*) (fig. S5). We consider alternative explanations (*12*) and find them to be inconsistent with the data. In particular, the observed double-peaked light curve in the redder optical bands requires the presence of an extended low-mass envelope around the progenitor (*24*, *25*). To constrain the properties of such an envelope, we use models (*25*) to construct multicolor light curves for a range of masses and radii of the envelope (M_e and R_e, respectively). We find a best-fitting model of $M_e \sim 8 \times 10^{-3}\ M_\odot$ and $R_e \sim 3 \times 10^{13}$ cm [~450 solar radii ($R_\odot$)] (*12*) (Fig. 2 and fig. S10). Even though the model considered here is simplified (e.g., it ignores the density structure of the envelope), we expect the estimated parameters to be accurate within an order of magnitude (*26*), leading us to conclude that the progenitor was surrounded by an extended envelope with a mass of ~0.01 $M_\odot$ at a radius of ~500 $R_\odot$.

We use the early spectra to constrain the composition of the outer envelope. The emission lines observed in the early spectra of iPTF 14gqr can be understood as arising from recombination in the outer regions of the extended circumstellar material (CSM), which was ionized by the high-energy radiation produced in the shock breakout (*13*, *15*) (fig. S6). We estimate the location and mass of the emitting He II from the luminosity of the early 4686-Å line, assuming a CSM density profile that varies with radius r as $\propto r^{-2}$ (*15*). We find the emitting region to be located at $r \sim 6 \times 10^{14}\tau^{-2}$ cm and to contain a helium mass $M_{He} \sim 0.01\tau^{-3} M_\odot$, where τ is the optical depth of the region (*12*). The absence of prominent Lorentzian scattering profiles in the lines suggests that the optical depth is small, and assuming $\tau \approx 1$, we find $r \sim 6 \times 10^{14}$ cm ($8 \times 10^3\ R_\odot$) and $M_{He} \sim 0.01\ M_\odot$. Because our calculations are based on fitting a simple two-component Gaussian profile to the 4686-Å emission line (to estimate the unknown contamination of C III at 4650 Å), these estimates are uncertain by a factor of a few (2 to 4).

Using the C IV 5801-Å lines and similar methods as above, we estimate a CSM carbon mass of $\sim 4 \times 10^{-3}\ M_\odot$, whereas the hydrogen mass is constrained to be $< 10^{-3} M_\odot$. Additional constraints based on light travel time arguments

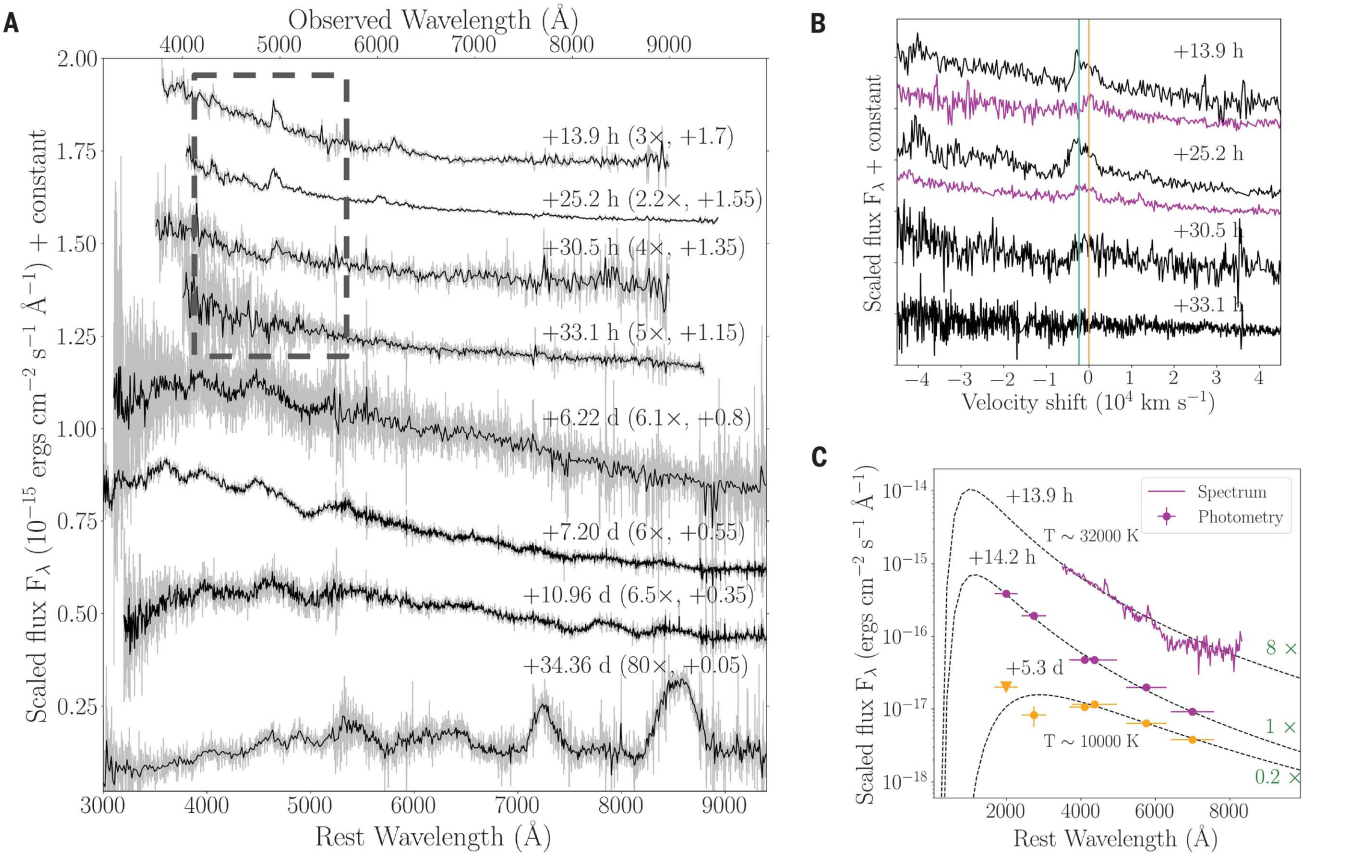


Fig. 3. Spectroscopic evolution of iPTF 14gqr. (**A**) Observed spectra before (gray) and after (black) binning. The epochs of the spectra, along with the scaling and vertical shifts used, are indicated next to each spectrum. (**B**) Zoomed-in view of the early spectra, indicated by the black dashed box in (A), showing rapid evolution of the λ4686 feature within 24 hours of discovery. The *x* axis indicates the velocity shift from the He II λ4686 line. The orange and cyan lines mark the locations of the λ4686 line and the C III λ4650 line, respectively. For the +13.9-hour and +25.2-hour spectra, additional magenta lines show the profiles of the C IV λ5801 and the C III λ5696 features, respectively, at the same epochs. (**C**) Scaled optical/UV SEDs of the photometry and spectra obtained within the first light curve peak (see Fig. 2) in magenta, along with photometry near the second peak in orange. The circles indicate observed photometric fluxes, whereas the triangle is a 5σ upper limit. The dashed black lines indicate the best-fitting blackbody SEDs, including all optical/UV data points for the first peak and only the optical data points for the second peak (*12*).

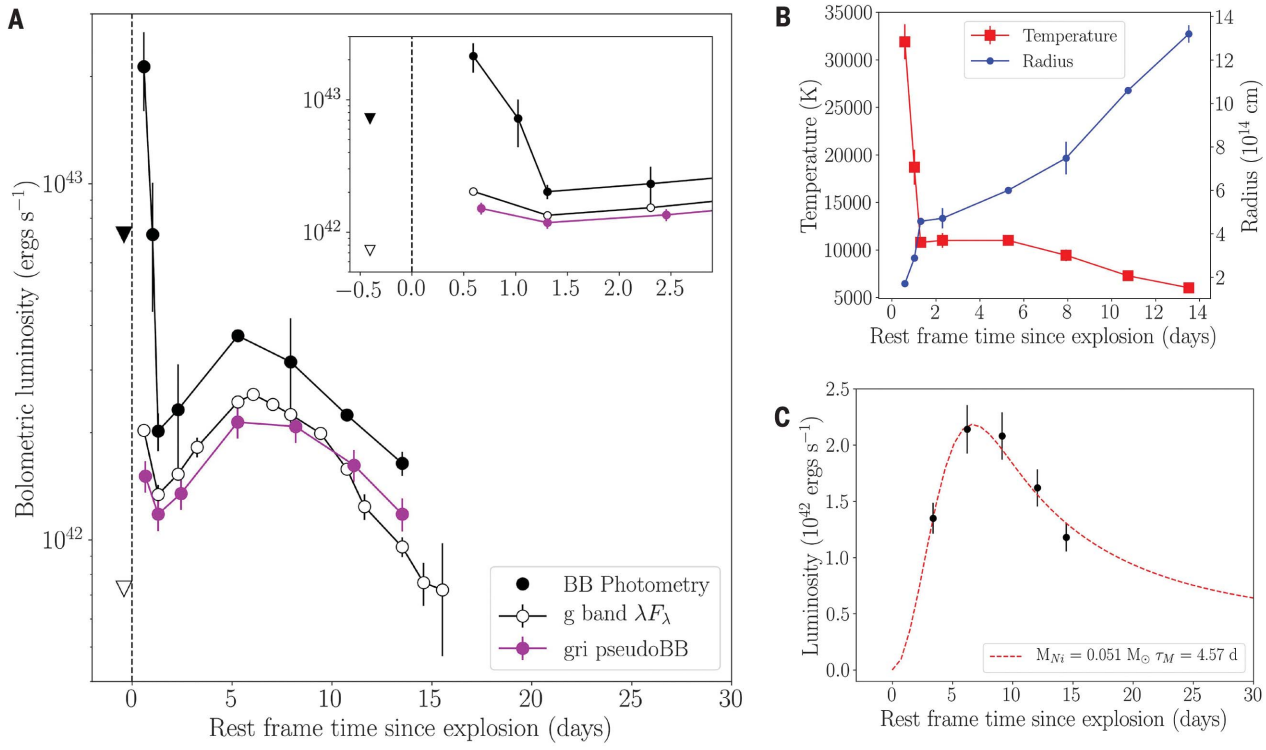


Fig. 4. Bolometric light curve and Arnett modeling of iPTF 14gqr. (**A**) Bolometric light curve of iPTF 14gqr. The filled black points indicate blackbody (BB) luminosities obtained from fitting multicolor photometry, whereas the magenta points correspond to pseudo-bolometric luminosities (*12*). The empty black circles indicate *g*-band luminosities obtained by multiplying the *g*-band flux F_λ with the wavelength λ of the filter. The inverted triangles denote estimated predetection 5σ upper limits on the respective luminosities (*12*). The inset shows the bolometric light curves zoomed into the region of the first peak. (**B**) Radius and temperature evolution of the fitted blackbody functions. (**C**) Best-fitting Arnett model of the pseudo-bolometric light curve of the main (second) peak of iPTF 14gqr. The ^{56}Ni mass M_{Ni} and diffusion time scale τ_M corresponding to the model are indicated in the legend (*12*).

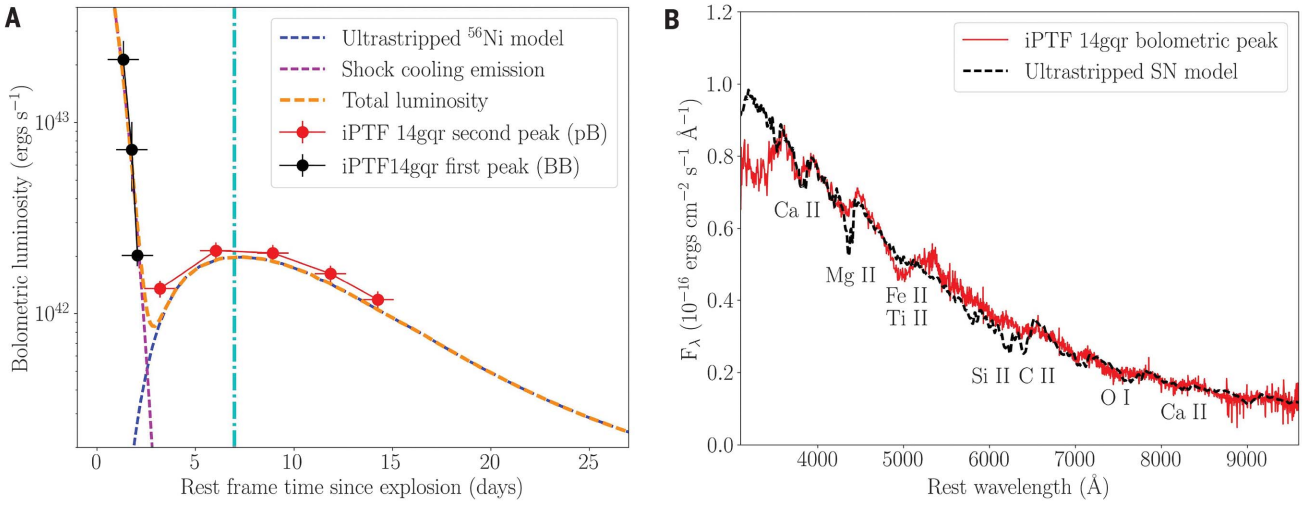


Fig. 5. Comparison of iPTF 14gqr to theoretical models of ultra-stripped SNe. (**A**) Bolometric light curve of iPTF 14gqr shown with a composite light curve consisting of ultra-stripped type Ic SN models (*28*) and early shock-cooling emission (*25*). The blue dashed line corresponds to the ^{56}Ni powered peak in the ultra-stripped SN models for $M_{ej} = 0.2\ M_\odot$, $M_{Ni} = 0.05\ M_\odot$, and $E_K = 2 \times 10^{50}$ ergs; the magenta line corresponds to the early shock-cooling emission; and the orange line represents the total luminosity from the sum of the two components. Blackbody (BB) luminosities represent the early emission, whereas pseudo-bolometric (pB) luminosities are used for the second peak (*12*). (**B**) Comparison of the peak photospheric spectra of iPTF 14gqr [the epoch is indicated by the cyan dashed line in (A)] to that of the model in (A). The overall continuum shape, as well as absorption features of O I, Ca II, Fe II, and Mg II, are reproduced (*12*).

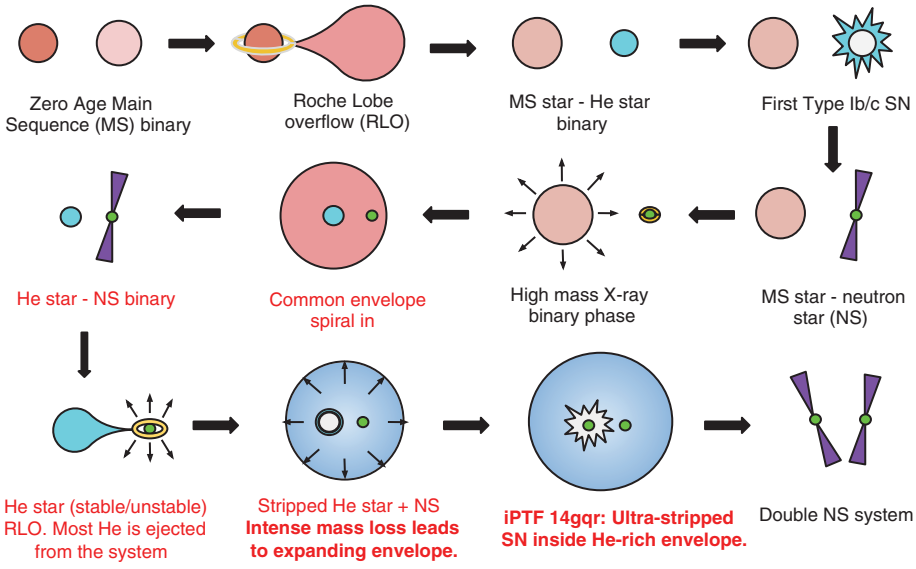


Fig. 6. Stellar evolutionary sequence leading from a binary system of massive stars (starting from the top left) to a NS-NS system. NS-BH systems are expected to arise from binaries in which the first formed compact object is a BH. NS-WD systems follow a similar evolutionary sequence starting from the HMXB (high-mass x-ray binary) stage (where the NS is replaced by the WD) but require additional mass transfer in the earlier stages (*52*). The material composition of the stars is indicated by their colors: red, H-rich material; cyan/blue, He-rich material; gray, CO-rich material; green, degenerate matter (in NS). The specific phase of the evolution is indicated by the text under each diagram, with black text indicating previously observed phases, red text denoting phases that have not been previously observed, and bold red text indicating phases observed in this work. [Adapted from (*9*)]

also suggest that the envelope was located at $r \leq 6 \times 10^{15}$ cm from the progenitor (*12*). The flash-ionized emission lines exhibit complex asymmetric profiles (Fig. 3) that we attribute to light travel time effects, given the large size of the envelope and the high inferred wind velocities (*12*, *27*).

An ultra-stripped progenitor

The low ejecta mass and explosion energy, as well as the presence of an extended He-rich envelope, indicate an unusual progenitor channel for iPTF 14gqr. The detection of the early shock-cooling emission indicates a core-collapse origin of the explosion, whereas the bright radioactivity powered emission suggests that this explosion is associated with the class of iron core-collapse explosions. The low ejecta mass, together with the small remaining amount of He in the progenitor, rules out models of single star evolution as well as a nondegenerate massive star companion for the progenitor of iPTF 14gqr (*12*), leaving only the most compact companions (such as a NS, WD, or BH) as possible explanations of the highly stripped (or ultra-stripped) progenitor.

Ultra-stripped explosions have been modeled in the case of He star–NS binaries, in which stripping of the He star by a NS in a close orbit leads to the subsequent collapse of an ultra-stripped He star (*7*, *8*, *28*). Hence, we compare theoretical bolometric light curves for ultra-stripped explosions (*28*) to those of iPTF 14gqr in Fig. 5 for a model with $M_{ej} = 0.2\,M_{\odot}$, $M_{Ni} = 0.05\,M_{\odot}$, and $E_K = 2 \times 10^{50}$ ergs. To account for the early declining emission, we also add a component corresponding to shock cooling of an extended envelope for $M_e = 0.01\,M_{\odot}$ and $R_e = 6 \times 10^{13}$ cm. The two-component light curve matches the light curve data. We also compare the spectroscopic properties of iPTF 14gqr to those of ultra-stripped SN models in Fig. 5. The models (*28*) assumed fully mixed ejecta that led to the production of strong line blanketing features below 4000 Å, unlike this source. Thus, we recalculated the models for ejecta with no mixing (as with the light curve calculations) and were able to match to the spectra of iPTF 14gqr near the second peak (Fig. 5 and fig. S13).

Our observations indicate the presence of an extended He-rich envelope around the progenitor at the time of collapse, thus providing insight into the terminal evolution of the progenitors of ultra-stripped SNe and, more broadly, the lowest-mass progenitors of core-collapse SNe. By using the line widths in our early spectra, we estimate that the emitting envelope was expanding with a velocity of ~1000 to 2000 km s^{-1} at the time of collapse, consistent with the escape velocity from a compact He star (*12*). When considered with the inferred size of the envelope (at least $\sim 500\,R_{\odot}$), the velocities suggest that the envelope was ejected ~8 to 20 days before the explosion.

The temporal coincidence of the ejection with the final SN suggests that the envelope was likely associated with an intense pre-SN mass-loss episode of the progenitor (*12*). Despite the close stripping, ultra-stripped progenitors are expected to retain a small amount of He ($\sim 0.01\,M_{\odot}$) in their outer layers. The prominent He and C lines in the early spectra are consistent with eruptive mass loss when considering the expected surface compositions of ultra-stripped progenitors (*8*). The time scale of the ejection is similar to that expected for silicon flashes (~2 weeks before explosion) in the terminal evolution of low-mass metal cores (*29*) that have been suggested to lead to elevated mass-loss episodes before the explosion. Such mass-loss episodes are relevant to ultra-stripped progenitors as well (*28*–*30*).

iPTF 14gqr exhibits a projected offset of ~15 kpc from the nearest spiral arms of its star-forming host galaxy (*12*), which is puzzling when compared to the expected locations of ultra-stripped SNe (*8*). Although we do not find evidence of an underlying stellar association or of galaxy emission features in late-time imaging and spectroscopy, the limits are not sensitive enough to rule out the presence of a dwarf galaxy or a star-forming H II region (characterized by its H α emission) at or near the transient location (*12*). Nonetheless, the tidally interacting environment of the host galaxy suggests that outlying star formation in collisional debris is likely in this system (*12*, *31*), which could harbor young stellar systems (with ages of ~5 to 100 million years) in the faint tidal tails (fig. S14). Hence, the discovery of a core-collapse SN in these outskirts is consistent with our interpretation.

Although a number of previously observed fast type Ic SNe [e.g., SN 2005ek (*21*) and SN 2010X (*22*)] were suggested to be members of the ultra-stripped SN class, it has been difficult to confirm a core-collapse origin for these explosions because these events were discovered only near maximum of the radioactively powered peak. Specifically, without early photometry and spectroscopy that can reveal the presence of a shock-cooling component, these fast transients are also consistent with variants of models involving thermonuclear detonations on WDs (*32*–*34*). The early discovery and prompt follow-up of iPTF 14gqr establish the presence of a shock-cooling emission component that requires an extended progenitor consistent with a core-collapse explosion. In the probable scenario that iPTF 14gqr formed a NS in the explosion [we find a BH remnant to be unlikely given the observed properties of the SN (*12*)], the low ejecta mass in the system suggests that the SN results in the formation of a bound and compact NS binary system (*12*).

Implications for formation of compact NS binaries

Our interpretation of iPTF 14gqr as an ultra-stripped SN has implications in the wider context of stellar evolution. Compact NS binary systems evolve from binary massive stars that undergo several phases of mass transfer over their lifetime (Fig. 6). The initial phases of such evolution, in which two massive stars evolve into interacting binaries consisting of a compact object in orbit around a massive star (x-ray binaries), have been

observed in several systems in the local Universe (*35*, *36*). However, the subsequent phases that lead to the formation of compact NS binary systems have not been observed. This is due to the low occurrence rates of such systems, the short lifetimes (~10^6 years) of the final stages, and observational selection effects disfavoring their detection (*8*, *37*, *38*).

Binary evolution models suggest that the subsequent evolution proceeds via a common envelope phase, during which the loss of angular momentum via dynamical friction leads to the formation of a close He star–compact object binary (*9*, *39*, *40*). An additional phase of close gravitational stripping by the compact companion then leads to the formation of an ultra-stripped SN progenitor (*9*), with properties that can be inferred from our observations of iPTF 14gqr. The measured orbital properties of known double NS systems suggest that the second NSs were created in weak and low–ejecta mass explosions that impart a small natal kick to the newborn NS (*41*, *42*).

The presence of the extended He-rich envelope in iPTF 14gqr, along with the lack of He in the low mass of ejecta, suggests that the progenitor was highly stripped by a compact companion, such that only a thin He layer was retained on its surface. This He layer was then ejected in an intense pre-SN mass-loss episode, as shown by the high velocity of the envelope. Taken together, these observations provide evidence of the terminal evolution of a post–common envelope He star–compact object binary leading to the formation of a compact NS binary system (Fig. 6).

Although wide binaries containing a NS and another compact object may be formed in non-interacting systems of binary massive stars, ultra-stripped SNe have been suggested to precede the formation of almost all compact NS binary systems (*8*). Thus, these explosions likely represent the only channel to forming NS-NS and NS-BH systems that are sufficiently compact to merge within the age of the Universe and produce observable merger signals for joint gravitational wave (*43*) and electromagnetic (*44*–*46*) observations (*8*, *47*, *48*). Given that only a fraction of the systems produced by these explosions will merge within that time, the rates of ultra-stripped explosions must be higher than the rates of their mergers.

REFERENCES AND NOTES

1. S. E. Woosley, A. Heger, T. A. Weaver, *Rev. Mod. Phys.* **74**, 1015–1071 (2002).
2. S. J. Smartt, *Annu. Rev. Astron. Astrophys.* **47**, 63–106 (2009).
3. N. Langer, *Annu. Rev. Astron. Astrophys.* **50**, 107–164 (2012).
4. H. Sana *et al*., *Science* **337**, 444–446 (2012).
5. S.-C. Yoon, S. E. Woosley, N. Langer, *Astrophys. J.* **725**, 940–954 (2010).
6. N. Smith, W. Li, A. V. Filippenko, R. Chornock, *Mon. Not. R. Astron. Soc.* **412**, 1522–1538 (2011).
7. T. M. Tauris *et al*., *Astrophys. J.* **778**, L23 (2013).
8. T. M. Tauris, N. Langer, P. Podsiadlowski, *Mon. Not. R. Astron. Soc.* **451**, 2123–2144 (2015).
9. T. M. Tauris *et al*., *Astrophys. J.* **846**, 170 (2017).
10. Y. Cao, P. E. Nugent, M. M. Kasliwal, *Publ. Astron. Soc. Pac.* **128**, 114502 (2016).
11. F. J. Masci *et al*., *Publ. Astron. Soc. Pac.* **129**, 014002 (2017).
12. See supplementary materials.
13. A. Gal-Yam *et al*., *Nature* **509**, 471–474 (2014).
14. D. Khazov *et al*., *Astrophys. J.* **818**, 3 (2016).
15. O. Yaron *et al*., *Nat. Phys.* **13**, 510–517 (2017).
16. A. Gal-Yam, in *Handbook of Supernovae*, A. W. Alsabti, P. Murdin, Eds. (Springer, 2017), pp. 1–43.
17. W. D. Arnett, *Astrophys. J.* **253**, 785 (1982).
18. M. R. Drout *et al*., *Astrophys. J.* **741**, 97 (2011).
19. J. D. Lyman *et al*., *Mon. Not. R. Astron. Soc.* **457**, 328–350 (2016).
20. F. Taddia *et al*., *Astron. Astrophys.* **609**, A136 (2018).
21. M. R. Drout *et al*., *Astrophys. J.* **774**, 58 (2013).
22. M. M. Kasliwal *et al*., *Astrophys. J.* **723**, L98–L102 (2010).
23. E. Nakar, A. L. Piro, *Astrophys. J.* **788**, 193 (2014).
24. N. Sapir, E. Waxman, *Astrophys. J.* **838**, 130 (2017).
25. A. L. Piro, *Astrophys. J.* **808**, L51 (2015).
26. A. L. Piro *et al*., *Astrophys. J.* **846**, 94 (2017).
27. G. Gräfener, J. S. Vink, *Mon. Not. R. Astron. Soc.* **455**, 112–126 (2016).
28. T. J. Moriya *et al*., *Mon. Not. R. Astron. Soc.* **466**, 2085–2098 (2017).
29. S. E. Woosley, A. Heger, *Astrophys. J.* **810**, 34 (2015).
30. B. Müller *et al*., *Mon. Not. R. Astron. Soc.* **479**, 3675–3689 (2018).
31. M. Boquien *et al*., *Astron. J.* **137**, 4561–4576 (2009).
32. K. J. Shen, D. Kasen, N. N. Weinberg, L. Bildsten, E. Scannapieco, *Astrophys. J.* **715**, 767–774 (2010).
33. B. D. Metzger, *Mon. Not. R. Astron. Soc.* **419**, 827–840 (2012).
34. S. Darbha *et al*., *Mon. Not. R. Astron. Soc.* **409**, 846–854 (2010).
35. O. G. Benvenuto, M. C. Bersten, in *Handbook of Supernovae*, A. W. Alsabti, P. Murdin, Eds. (Springer, 2017), pp. 1–22.
36. R. Walter, A. A. Lutovinov, E. Bozzo, S. S. Tsygankov, *Astron. Astrophys. Rev.* **23**, 2 (2015).
37. Y. Götberg, S. E. de Mink, J. H. Groh, *Astron. Astrophys.* **608**, A11 (2017).
38. E. Zapartas *et al*., *Astron. Astrophys.* **601**, A29 (2017).
39. D. Bhattacharya, E. P. J. van den Heuvel, *Phys. Rep.* **203**, 1–124 (1991).
40. T. M. Tauris, E. P. J. van den Heuvel, in *Compact Stellar X-ray Sources*, W. Lewin, M. van der Klis (Cambridge Astrophysics Series, Cambridge Univ. Press, 2006), pp. 623–665.
41. R. D. Ferdman *et al*., *Astrophys. J.* **767**, 85 (2013).
42. P. Beniamini, T. Piran, *Mon. Not. R. Astron. Soc.* **456**, 4089–4099 (2016).
43. B. P. Abbott *et al*., *Astrophys. J.* **848**, L13 (2017).
44. B. P. Abbott *et al*., *Astrophys. J.* **848**, L12 (2017).
45. E. Pian *et al*., *Nature* **551**, 67–70 (2017).
46. D. Kasen, B. Metzger, J. Barnes, E. Quataert, E. Ramirez-Ruiz, *Nature* **551**, 80–84 (2017).
47. R. Voss, T. M. Tauris, *Mon. Not. R. Astron. Soc.* **342**, 1169–1184 (2003).
48. Compact systems may also form by dynamical capture in dense stellar environments (*49*).
49. J. Grindlay, S. Portegies Zwart, S. McMillan, *Nat. Phys.* **2**, 116–119 (2006).
50. P. A. Mazzali, L. B. Lucy, *Astron. Astrophys.* **279**, 447 (1993).
51. B. Abolfathi *et al*., *Astrophys. J. Suppl. Ser.* **235**, 42 (2018).
52. T. M. Tauris, T. Sennels, *Astron. Astrophys.* **355**, 236 (2000).

ACKNOWLEDGMENTS

We thank the anonymous referees for a careful reading of the manuscript, which helped improve the quality of the paper. We thank C. Steidel, N. Stone, D. Stern, P. Hopkins, S. de Mink, Y. Suwa, A. Heger, and T. M. Tauris for valuable discussions. M.M.K. thanks J. Fuller, E. S. Phinney, L. Bildsten, and E. Quataert for stimulating discussions at the Skyhouse during a PTF-TN meeting. We also thank T. Staley and G. Anderson for help with scheduling of the AMI observations. Additional facility acknowledgments are provided in the supplementary materials. **Funding:** The Intermediate Palomar Transient Factory project is a scientific collaboration among the California Institute of Technology; Los Alamos National Laboratory; the University of Wisconsin, Milwaukee; the Oskar Klein Center; the Weizmann Institute of Science; the TANGO Program of the University System of Taiwan; and the Kavli Institute for the Physics and Mathematics of the Universe. This work was supported by the GROWTH (Global Relay of Observatories Watching Transients Happen) project funded by the National Science Foundation under PIRE Grant 1545949. GROWTH is a collaborative project among California Institute of Technology (USA); University of Maryland, College Park (USA); University of Wisconsin, Milwaukee (USA); Texas Tech University (USA); San Diego State University (USA); Los Alamos National Laboratory (USA); Tokyo Institute of Technology (Japan); National Central University (Taiwan); Indian Institute of Astrophysics (India); Indian Institute of Technology Bombay (India); Weizmann Institute of Science (Israel); The Oskar Klein Centre at Stockholm University (Sweden); Humboldt University (Germany); and Liverpool John Moores University (UK). A.H. acknowledges support by the I-Core Program of the Planning and Budgeting Committee and the Israel Science Foundation. A.G.-Y. is supported by the EU via ERC grant 725161, the Quantum Universe I-Core program, the ISF, the BSF Transformative program, and a Kimmel award. E.O.O. is grateful for support by grants from the Willner Family Leadership Institute Ilan Gluzman (Secaucus, NJ), Israel Science Foundation, Minerva, BSF, BSF-transformative, and the I-Core program by the Israeli Committee for Planning and Budgeting and the Israel Science Foundation (ISF). F.T. and J.S. gratefully acknowledge support from the Knut and Alice Wallenberg Foundation. The Oskar Klein Centre is funded by the Swedish Research Council. This research used resources of the National Energy Research Scientific Computing Center, a DOE Office of Science User Facility supported by the Office of Science of the U.S. Department of Energy under contract DE-AC02-05CH11231. M.S. acknowledges support from EU/FP7 ERC grant 615929. P.E.N. acknowledges support from the DOE through DE-FOA-0001088, Analytical Modeling for Extreme-Scale Computing Environments. T.J.M. is supported by Grants-in-Aid for Scientific Research of the Japan Society for the Promotion of Science (16H07413 and 17H02864). Numerical computations were, in part, carried out on the PC cluster at the Center for Computational Astrophysics, National Astronomical Observatory of Japan. Part of this research was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration. **Author contributions:** K.D. and M.M.K. initiated the study, conducted analysis, and wrote the manuscript. I.M. initiated the follow-up of the young transient. D.A.P., G.E.D., and Y.C. conducted Keck and Palomar observations and contributed to data reduction and manuscript preparation. S.B.C. conducted Keck and Swift observations and contributed to data reduction and manuscript preparation. M.S. conducted the WHT observations and data reduction. F.T. and J.S. conducted NOT observations and data analysis and contributed to manuscript preparation. J.B. conducted the LCO observations and data reduction. T.P. conducted Gemini observations and data analysis. C.R. and R.P.F. conducted the AMI observations and data reduction. A.H. conducted the VLA observations and data reduction. S.R.K. is iPTF PI and contributed to manuscript preparation. T.J.M. and P.A.M. prepared the ultra-stripped SN models presented in the paper. E.O.O., C.F., A.G.-Y., R.L., P.E.N., and A.L.P. contributed to manuscript preparation. G.B.D., R.R.L., and F.M. contributed to the machine learning codes used to search for young transients. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** All photometric data used in this paper are provided in the supplementary materials (tables S1 and S2), and all observed spectra are available via the WISeREP repository at https://wiserep.weizmann.ac.il/ (under source name iPTF 14gqr). The software used for the ultra-stripped SN modeling is presented in (*50*), and the synthetic model spectra are available in data S1.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/362/6411/201/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S15
Tables S1 to S7
References (*53*–*181*)
Data S1

31 December 2017; accepted 26 July 2018
10.1126/science.aas8693

VIROLOGY

Genome hypermobility by lateral transduction

John Chen[1]*†, Nuria Quiles-Puchalt[2]*, Yin Ning Chiang[1]*, Rodrigo Bacigalupe[3], Alfred Fillol-Salom[2], Melissa Su Juan Chee[1], J. Ross Fitzgerald[3], José R. Penadés[2,4,5]†

Genetic transduction is a major evolutionary force that underlies bacterial adaptation. Here we report that the temperate bacteriophages of *Staphylococcus aureus* engage in a distinct form of transduction we term lateral transduction. Staphylococcal prophages do not follow the previously described excision-replication-packaging pathway but instead excise late in their lytic program. Here, DNA packaging initiates in situ from integrated prophages, and large metameric spans including several hundred kilobases of the *S. aureus* genome are packaged in phage heads at very high frequency. In situ replication before DNA packaging creates multiple prophage genomes so that lateral-transducing particles form during normal phage maturation, transforming parts of the *S. aureus* chromosome into hypermobile regions of gene transfer.

Bacteriophages are the most abundant gene-transfer particles, and phage transduction is generally regarded as the most important mechanism of horizontal gene transfer (HGT) between bacterial cells. HGT is of considerable importance in medicine because it is the major route by which bacteria acquire virulence factors and antibiotic resistance.

Prophages are phage genomes that are integrated into bacterial chromosomes and replicate passively along with the host genome. Mature phages are produced in the lytic cycle during host cell infection or lysogenic induction, when rapid viral DNA replication and capsid assembly lead to formation of infectious particles that are released after cell lysis (fig. S1). Transducing particles are also produced during the phage lytic cycle, when bacterial DNA can also become packaged into newly formed procapsids. The acquisition of host DNA by transducing particles depends on the packaging mechanism. Most packaging begins with the cleavage of concatemeric DNA, generated by rolling-circle replication of the phage genome. A phage-specific packaging site (*pac* or *cos*) is recognized by the phage small terminase (TerS), which forms heterooligomers with the phage large terminase (TerL) to process DNA into procapsids (*1*). To complete DNA packaging, *pac*-type terminases make a nonspecific sequence cut when capsid "headful" capacity (i.e., slightly longer than a genome unit length) has been reached. In an alternative mechanism, *cos*-type terminases require a second *cos* site for terminal cleavage and thus package precise genome monomers (*2*).

Phage-mediated HGT is known to occur by either generalized or specialized transduction

[1]Department of Microbiology and Immunology, Yong Loo Lin School of Medicine, National University of Singapore, 5 Science Drive 2, Singapore. [2]Institute of Infection, Immunity, and Inflammation, College of Medical, Veterinary, and Life Sciences, University of Glasgow, Glasgow G12 8TA, UK. [3]The Roslin Institute, University of Edinburgh, Easter Bush Campus, Edinburgh EH25 9RG, UK. [4]Departamento de Ciencias Biomédicas, Universidad CEU Cardenal Herrera, 46113 Moncada, Spain. [5]MRC–University of Glasgow Centre for Virus Research, Glasgow G61 1QH, UK.
*These authors contributed equally to this work.
†Corresponding author. Email: miccjy@nus.edu.sg (J.C.); joser.penades@glasgow.ac.uk (J.R.P.)

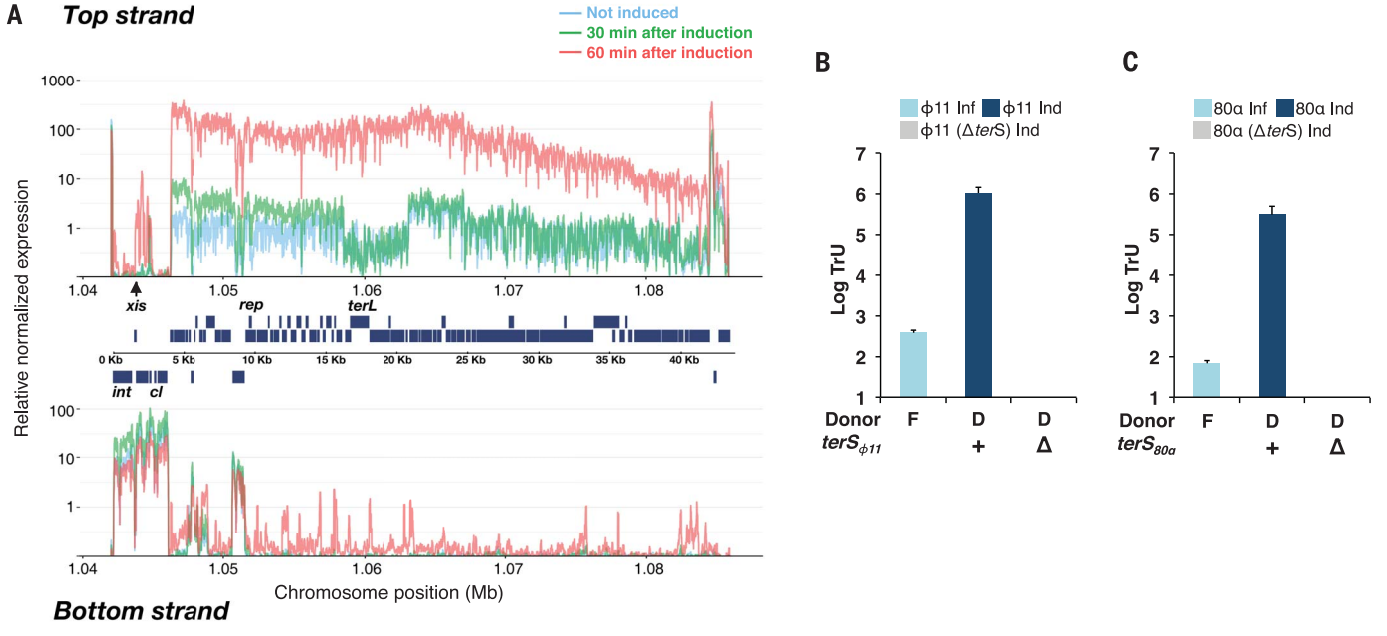


Fig. 1. DNA packaging initiates from the *pac* site within integrated prophage genomes. (**A**) Transcriptomic analysis of the early and late genes from phage 80α for the positive (top) and the negative (bottom) DNA strands. An 80α lysogen was treated with mitomycin C, and samples were analyzed without induction (light blue) or at 30 min (early genes, green) and 60 min (late genes, red) after induction. Expression was normalized by the number of aligned reads. (**B** and **C**) Transfer of Cd^R markers downstream of *attB* sites for (B) ϕ11 and a Cd^R_1 5 kb downstream of the $attB_{ϕ11}$ and (C) 80α and a Cd^R 4 kb downstream of the $attB_{80α}$. Nonlysogenic strains (light blue) were infected (Inf or F) and the lysogenic WT (+, dark blue) and *terS* deletion (Δ, gray) strains were induced (Ind or D) with mitomycin C, and the lysates were tested for transduction into *S. aureus*. Transduction units (TrU) per milliliter were normalized by plaque forming units (PFU) per milliliter and represented as the log TrU of an average phage titer (1×10^9 PFU). TrU per milliliter amounts for Δ*terS* were <10. Values are means ($n = 3$ independent samples). Error bars indicate standard deviation.

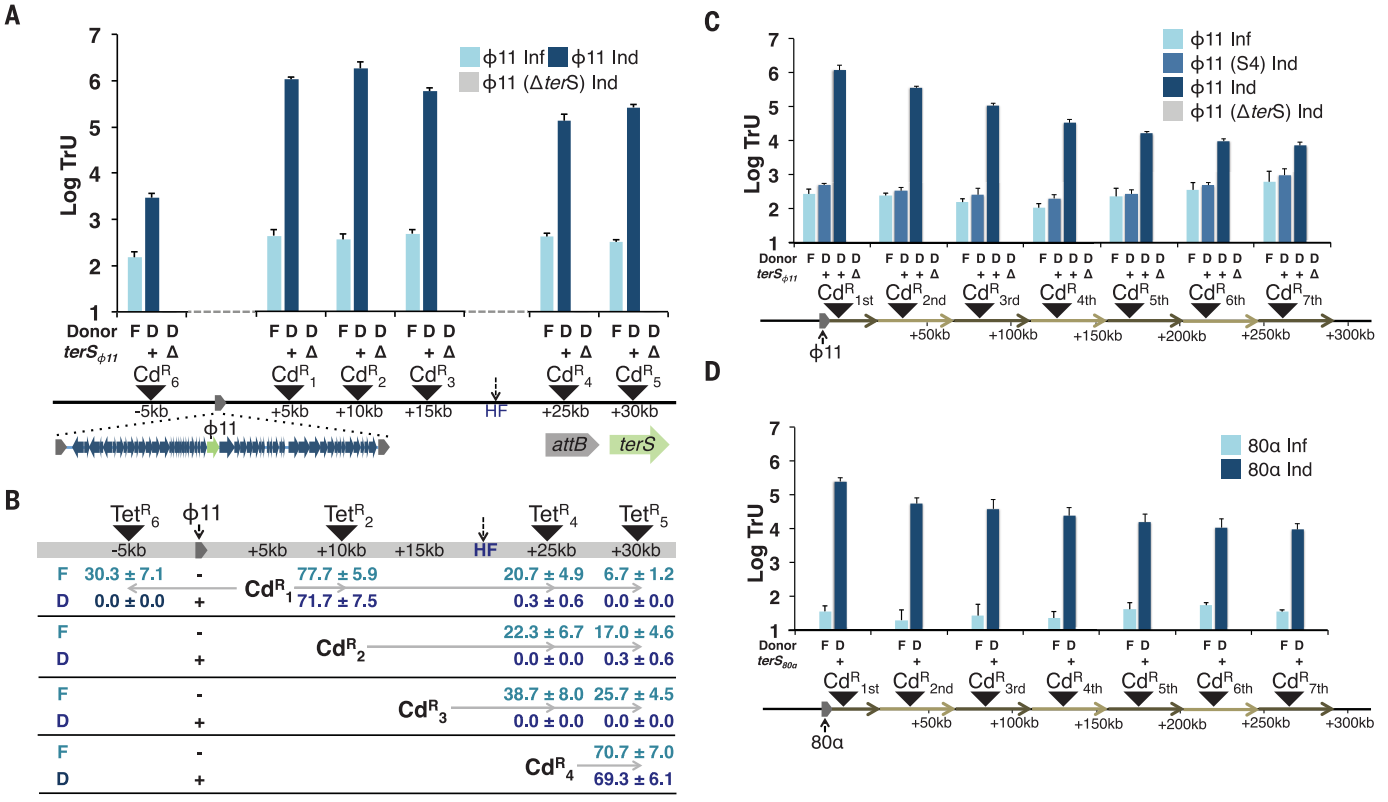


Fig. 2. Lateral transduction transfers large metameric spans of the bacterial chromosome at high frequencies by the headful packaging mechanism. (**A**) Phage ϕ11 was tested for transfer of Cd^R markers upstream (Cd^R_6) of the $attB_{ϕ11}$ and downstream within (Cd^R_1 to Cd^R_3) and beyond (Cd^R_4 and Cd^R_5) a capsid headful (HF) capacity. (**B**) Cotransduction frequencies for strains containing both a Cd^R and a Tet^R marker at varying distances apart. One hundred Cd^R transductants from a ϕ11 infection (light blue) or lysogen induction (dark blue) were tested for Tet^R, and the frequency is represented as a percentage, calculated as (Tet^R/Cd^R) × 100. The HF limit indicated does not account for marker expansion. +, ϕ11 lysogen; −, nonlysogenic strain. (**C** and **D**) Transfer of Cd^R markers in seven successive HFs for (C) ϕ11 and (D) 80α. For (A), (C), and (D), nonlysogenic strains (light blue) were infected (Inf or F) and the lysogenic WT (+, dark blue), ϕ11 (S4, blue), and *terS* deletion (Δ, gray) strains were induced (Ind or D) with mitomycin C, and the lysates were tested for transduction into *S. aureus*. Strain ϕ11 (S4) has ϕ11 integrated at the SaPI 4 *attB* instead of the natural $attB_{ϕ11}$. TrU per milliliter were normalized by PFU per milliliter and represented as the log TrU of an average phage titer (1×10^9 PFU). TrU per milliliter amounts for Δ*terS* were <10. Error bars indicate standard deviation. For all panels, values are means (n = 3 independent samples).

(GT or ST, respectively) (*3*–*5*). GT is the process by which *pac*-type phages can package any bacterial DNA and transfer it to another bacterium, whereas ST is limited to the transfer of specific sets of genes. GT results from the recognition of *pac*-site homologs (also called pseudo-*pac* sites) (*6*, *7*) in host chromosomal or plasmid DNA by the *pac*-type headful mechanism (*2*). The *cos*-type phages typically are not involved in GT because the probability of two *cos*-site homologs being found in the host DNA at an optimal length apart is exceedingly rare (*2*).

The formation of ST particles is more complicated than that of the GT mechanism, and our current understanding is based on the classical λ phage model, in which aberrant excision events join part of the prophage to bacterial genes adjacent to their attachment site (*attB*) in the excised DNA (*5*). Because ST is limited to restricted sets of genes, it is assumed that most phage-mediated HGT events are governed by the GT mechanism. However, aberrant prophage excision is not the only means of ST particle formation, and other mechanisms involving the initiation of DNA packaging from the unexcised viral genome (in situ packaging) have been proposed (*8*). Early studies showed that artificially generated mutants of *cos*- and *pac*-type phages (λ and P22, respectively) could package headfuls of viral DNA still connected to adjacent bacterial DNA. These particles were not mature and required in vitro nuclease treatment (in λ) and the attachment of purified phage tails to be fully infective (*8*–*10*). Other studies focusing on headful packaging showed that completed ST particles could be formed in vivo from "locked-in" prophage hybrids and mutants that were unable to excise; however, these experimental systems also did not produce viable phage (*11*–*13*). Therefore, most integrated phages are probably capable of ST, but the role of in situ packaging, if any, remains uncertain for normal phage production.

Prophages typically excise and circularize early after induction, and the ensuing replication forms head-to-tail concatemers that are packaged by the terminase machinery. The order of this sequence of events is important because DNA packaging before excision and replication would split the viral genome and render the phage nonviable. Accordingly, most phages, including the λ and P22 phages, follow this temporal program (*14*, *15*).

We have discovered that the resident prophages of *Staphylococcus aureus* have atypical lytic programs, because they do not excise until late in their life cycles. The potentially detrimental effects of delayed excision are offset by in situ bidirectional replication, which creates multiple integrated genomes so that both in situ DNA packaging and phage maturation can proceed in parallel. As a result, staphylococcal phages naturally generate high titers of transducing particles in the process of wild-type phage production.

Delayed prophage excision results in DNA packaging from integrated viral genomes

In a previous study on phage transcriptional activators, using tiling microarray analysis on a *S. aureus* strain lysogenic for phage 80α, we showed that transcription of the excisionase gene (*xis*) was not activated until late (30 to 60 min) after induction (*16*). This result indicated that the 80α prophage may delay excision but is not defective for phage production. Therefore,

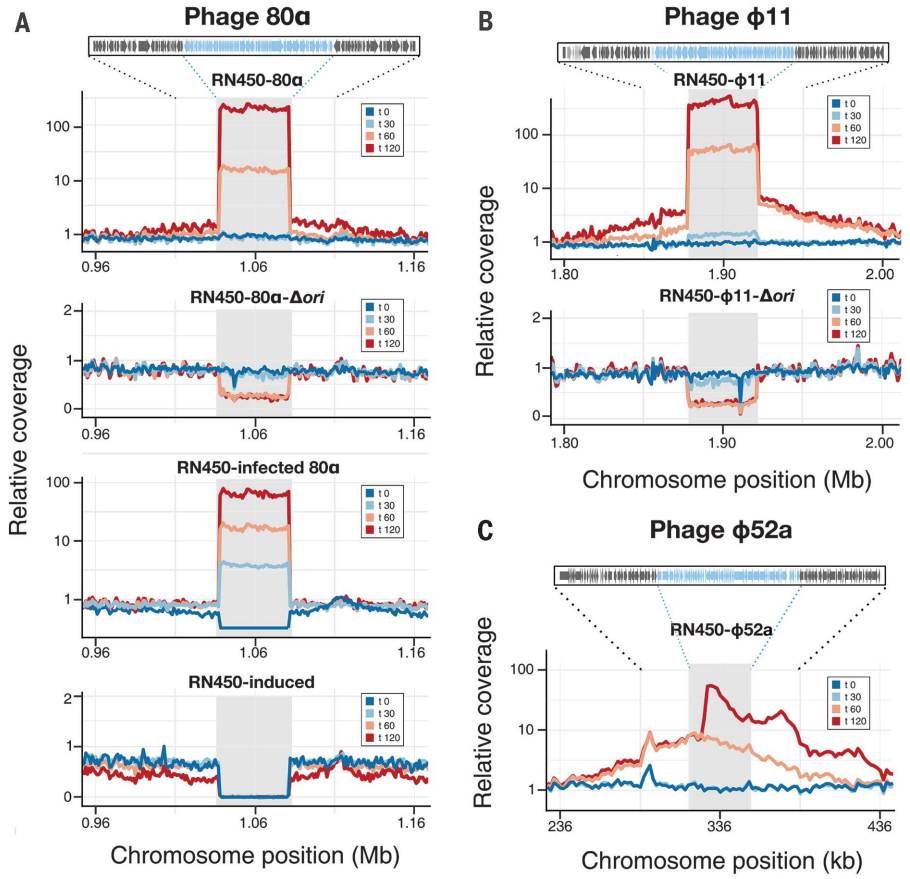


Fig. 3. Staphylococcal phages replicate in situ before excision. (**A** to **C**) Relative abundance of phage genomic DNA and the chromosomal regions proximal to where they integrate for (A) 80α, (B) ϕ11, and (C) ϕ52a. Samples were analyzed at 0 (blue), 30 (light blue), 60 (orange), and 120 min (red) after induction with mitomycin C. Relative coverage is the DNA relative to the average bacterial genomic coverage (excluding phages). Shaded gray rectangles represent the location of the prophages in the *S. aureus* chromosome. The following strains were analyzed: 80α lysogen (RN450-80α), ϕ11 lysogen (RN450-ϕ11), and ϕ52a lysogen (RN450-ϕ52a); derivatives of these strains carrying mutations in both the *ori* site and *rep* genes [80α *ori* deletion (RN450-80α-Δ*ori*) or ϕ11 *ori* deletion (RN450-ϕ11-Δ*ori*)]; a nonlysogenic strain infected with 80α (RN450-infected 80α); and the nonlysogenic strain (RN450-induced). For the 80α and ϕ52a lysogens, the *t* = 0 min samples obscure the *t* = 30 min samples.

we performed transcriptional profiling using 80α as a model and compared it with other staphylococcal phages, including ϕ11, ϕNM1, and ϕNM2. Lysogenic derivatives were treated with mitomycin C to elicit the SOS response, which activates the resident prophages, and total RNA for RNA-sequencing analysis was isolated before prophage induction and at 30 and 60 min afterward. Consistent with our previous findings, transcriptional activation of *xis* occurred in all the phages between 30 and 60 min after prophage induction (Fig. 1 and fig. S2). To correlate the onset of *xis* expression with prophage excision, the 80α and ϕ11 lysogens were induced under the same conditions, but instead of RNA, we isolated total chromosomal DNA for whole-genome sequencing. At each time point, we identified the sequencing reads corresponding to empty *attB* sites (which gave a measure for excised prophage) and the reads covering *att*L sites (i.e., the left end of the integrated prophage) and represented the results as the percentage of integrated prophage. The percentage of integrated 80α steadily declined after 30 min (fig. S3), matching the timing of *xis* transcriptional activation and confirming that excision was delayed. Interestingly, after 60 min, the decrease in percentage of 80α integration began to slow and that of ϕ11 began to increase (see below).

We reasoned that DNA packaging could initiate from the integrated genome because of the delay in prophage excision and because expression of the *xis* gene overlaps with the phage DNA-packaging module (late operon) (Fig. 1 and fig. S2). To test this, we first identified the *pac* sites of packaging initiation for 80α and ϕ11. Phage *pac* sites are often embedded in structural *terS* genes and direct unidirectional packaging toward the 3′ end of the gene (*1*), and likewise for the staphylococcal phages (fig. S4). Because the *terS* genes are located near the center of the 80α and ϕ11 prophage genomes, unidirectional packaging initiated in situ can only reach headful capacity (~105% of a phage genome or ~46 kb) by including the adjacent host DNA (fig. S5). Thus, to test for in situ prophage and host-DNA packaging, a cadmium-resistance cassette (Cd^R) was inserted in the *S. aureus* chromosome 5 kb downstream of the ϕ11 *attB* site ($attB_{\phi11}$), in a phage-free or ϕ11 lysogen, for both intact wild-type (WT) and *terS* deleted mutants (Δ*terS*) (fig. S5). A distance of 5 kb was chosen because it is well within a headful capacity for ϕ11 (~32 kb from the $terS_{\phi11}$) and also provides sufficient flanking DNA for homologous recombination in the nonlysogenic recipient host strain. Because the viral genome is extrachromosomal after infection, a nonlysogenic host was infected with ϕ11 to measure GT. To measure in situ packaging, the ϕ11 lysogenic derivatives were induced with mitomycin C, and the resulting lysates were tested as donors of cadmium resistance to *S. aureus*.

The lysates resulting from ϕ11 infection of the nonlysogen transferred a Cd^R marker (Cd^R_1) at a frequency of 100 to 1000 transductants ml^{-1} (Fig. 1B), consistent with GT frequencies of other chromosomal markers in *S. aureus* (*17*). By contrast, ϕ11 prophage induction transferred the Cd^R_1 at very high frequencies, three orders of magnitude greater than those observed for GT, in a $terS_{\phi11}$-dependent manner (Fig. 1B). Because expected low levels of GT were observed, it was unlikely that an exceptionally strong pseudo-*pac* site was directing high-frequency transfer of the Cd^R_1 marker. Other explanations include ST by an unusually efficient aberrant excision mechanism, superinfection by phages released early in the lytic cycle, or, as the transcriptomic analyses indicates, the prophage initiated packaging before excision.

For excision to be efficient, one possibility is that cryptic repeated sequences in both the ϕ11 genome and in the adjacent bacterial chromosome result in excision that generates ST particles. Changing the phage and/or the *attB* site should abolish this activity. First, we inserted a Cd^R_1 marker 5 kb downstream of the $attB_{80\alpha}$ in an 80α lysogen and tested for transfer of cadmium resistance. Induction of the 80α lysogen resulted in high transfer frequencies like those from ϕ11 induction (Fig. 1C). Other phages—including ϕNM1 and ϕNM2, which use the same *attB* as ϕ11 and 80α—also showed high-frequency transduction of the Cd^R_1 markers after prophage induction (fig. S6). Changing the phage and *attB* location did not abolish excision, and so it seemed unlikely that an aberrant excision mechanism was responsible for the high-frequency transfer. To rule out recombination, we engineered a *recA* Asn^{303}→Asp [*recA* (N303D)] mutant that is defective for recombination but not for LexA cleavage or the SOS response (*18*). Lysates from ϕ11 induction in a WT or *recA* (N303D) background transferred the Cd^R_1 marker at comparable frequencies (fig. S7), showing that recombination is not involved in Cd^R_1 packaging. Moreover, polymerase chain reaction analysis of 100 cadmium-resistant colonies confirmed that the transductants of 80α

prophage induction did not have integrated defective prophages linked to the transferred marker. Together, these results show that aberrant excision is unlikely to be involved.

Next, we determined if phages released early after SOS induction could superinfect the remaining cells to initiate packaging from the resident prophage genome. A previous study showed that *Salmonella typhimurium* carrying deleted P22 prophages that lost lysogenic immunity could be superinfected and that the WT phage could initiate packaging from the unexcised genomes (*12*). Therefore, we induced ϕ11 in the presence of sodium citrate at 100 and 200 mM, which was sufficient to block all phage adsorption. This experiment showed that the packaging and transfer frequencies of the Cd^R_1 marker were unaffected (fig. S8). These results rule out a mechanism by which early released phages superinfect the remaining cells for in situ packaging. Hence, neither GT nor ST are involved, and we propose here the existence of a distinct mechanism of natural transduction that we term lateral transduction.

Lateral transduction of large spans of the bacterial genome by the headful packaging mechanism

The high frequency of Cd^R transfer by the lateral transfer mechanism indicated that headful packaging occurs off the integrated prophage after induction of the lysogen. To test this possibility, we inserted two additional downstream markers (Cd^R_2 and Cd^R_3) within a headful capacity, two markers (Cd^R_4 and Cd^R_5) beyond a headful capacity, and one marker (Cd^R_6) upstream of the $attB_{\phi11}$ in nonlysogenic strains or ϕ11 lysogens. Next, lysates resulting from ϕ11 infection of a nonlysogenic strain or from prophage induction were tested for marker transfer. As expected, in lysates from ϕ11 infection, we observed low levels of transfer typical of GT (Fig. 2A). By contrast, lysates from ϕ11 induction resulted in high frequencies of transduction for all markers, located in the directionality of packaging, that is, Cd^R_2 to Cd^R_5 (Fig. 2A). Only minor reductions in the transfer frequencies of the Cd^R_4 and Cd^R_5 markers were observed, indicating either that packaging did not occur by the classical headful mechanism or that initiation on the next successive headful was highly efficient.

To test whether packaging occurred by the headful mechanism, we inserted tetracycline-resistance (Tet^R) cassettes upstream and downstream of the $attB_{\phi11}$ and paired them with the previous Cd^R markers—that is, with Cd^R_6 or downstream Cd^R_1 to Cd^R_4, respectively—in nonlysogenic strains or ϕ11 lysogens (Fig. 2B). Then we tested the lysates generated from ϕ11 infection or by prophage induction (GT or lateral transduction, respectively) for cotransduction of the two markers by selecting for cadmium resistance and scoring for Tet^R. As expected for GT, we found that for lysates generated by ϕ11 infection, all cotransduction frequencies were inversely proportional to their distance apart, and the shortest distances exhibited the highest percentage of cotransduction (Fig. 2B). However, when we tested lysates generated by ϕ11 induction, cotransduction was observed for markers within a headful, but even very close markers were completely unlinked when they were separated by the predicted headful limit (Fig. 2B). Next, we determined if *cos* phages could also mediate lateral transduction by testing for the transfer of Cd^R markers downstream of the ϕ12 or phage DI *attB* sites in nonlysogenic or lysogenic strains. As expected, lysates of ϕ12 or DI produced by infection or lysogenic induction did not transfer cadmium resistance, showing that *cos* phages do not mediate lateral transduction. These results confirm our model that headful packaging initiates from the $terS_{\phi11}$ gene and efficiently initiates the next headful.

Because the frequency of lateral transduction was so high from the first ϕ11 headful and because packaging initiated so efficiently for the second headful (Fig. 2A), we reasoned that the packaging machinery could continue for many headfuls before diminishing into the low levels of GT. To test this, we used previous markers (Cd^R_2 and Cd^R_5) for the first two headfuls and inserted five additional markers 10 kb into each successive headful—that is, seven in total—into nonlysogenic strains or ϕ11 lysogens. As an additional control to measure GT by prophage induction, a strain was generated in which the $attB_{\phi11}$ was deleted so that ϕ11 could be lysogenized at a new $attB_{\phi11}$ inserted at the *S. aureus* pathogenicity island 4 (SaPI 4) *attB* (strain S4); in this strain, the ϕ11 prophage is not linked to the Cd^R markers. We found that for lysates generated by ϕ11 induction, lateral transduction transferred up to seven headfuls of markers at levels that were substantially higher than the frequencies observed for lysates generated from ϕ11 infection or induction from ϕ11 (strain S4) (Fig. 2C). Results obtained with phage 80α showed that lateral transduction–mediated Cd^R marker transfer was much greater than GT-mediated transfer for at least seven headfuls (Fig. 2D). These results show that lateral transduction can mediate high-frequency HGT of bacterial host DNA for several hundred kilobases before the frequencies diminish and smooth out into the basal levels of GT.

To simulate a more natural test for HGT, we assayed for lateral transduction resulting from spontaneous lysogenic induction. To test this, we mixed intact cells carrying ϕ11 lysogenic derivatives and containing Cd^R_2 with a streptomycin-resistant host recipient and plated the mixture for cadmium resistance and streptomycin selection. As shown in fig. S9, spontaneous induction of ϕ11 resulted in a steady increase in lateral transductants from 4, 8, and 24 hours (>4000 transductants ml^{-1}). By contrast, spontaneous GT by the strain carrying the ϕ11 prophage integrated into the SaPI 4 *attB* (strain S4) was just slightly more than that of the ϕ11 (Δ*terS*) negative control and of spontaneous streptomycin resistance of the donor strain. These results show that lateral transduction is a powerful mode of HGT that promotes considerable levels of genetic exchange, even in natural conditions of rare spontaneous lysogenic induction.

In situ bidirectional replication enables phage maturation

Phage production requires early genome excision, which is at odds with the in situ lateral transduction-packaging mechanism. In situ replication, understood as the ability some prophages have to initiate replication before excision, could create sufficient genomic redundancy to enable both lateral transduction and phage maturation to proceed in parallel. In support of this hypothesis, λ and P22 mutants defective for excision have been observed to replicate in situ (*10*, *19*). Moreover, transcriptional analysis of the staphylococcal phages shows that the genes required for phage replication are expressed early, before *xis* transcription (Fig. 1 and fig. S2).

To test for in situ replication, we first determined whether the staphylococcal prophages exhibit escape replication—a phenomenon whereby the bacterial genome adjacent to occupied

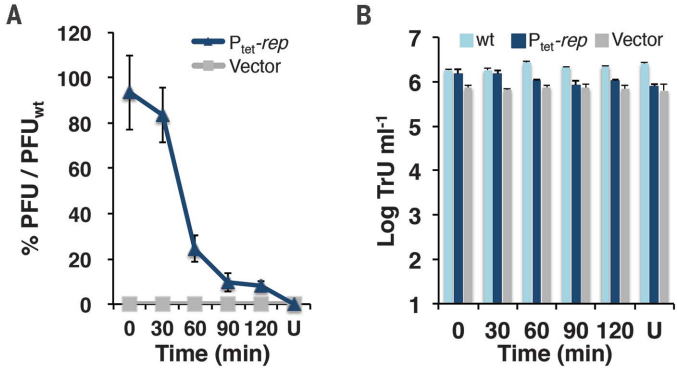


Fig. 4. In situ replication is required for phage reproduction but not for lateral transduction. (**A**) Lysogens of ϕ52a and ϕ52a-fs ($rep_{\phi52a}$ frameshift) were induced with mitomycin C. Anhydrotetracycline was added at 0, 30, 60, 90, and 120 min for complementation by P_{tet}-$rep_{\phi52a}$ or not added [uninduced (U)]. The cells were mechanically lysed 2 hours after the addition of anhydrotetracycline, and the lysates were tested for PFU per milliliter on a *S. aureus* host that constitutively expresses $rep_{\phi52a}$. The results are represented as a percentage, calculated as (PFU/PFU_{WT}) × 100. (**B**) The same lysates in (A) were tested for lateral transduction of the Cd^R marker 5 kb downstream of the $attB_{\phi52a}$ site to *S. aureus*. The results are represented as the log TrU per milliliter. Values are means (n = 3 independent samples). Error bars indicate standard deviation.

attB sites is amplified, owing to the initiation of bidirectional (theta) replication before prophage excision. We infected nonlysogenic strains or mitomycin C–induced lysogenic derivatives of 80α (carrying the WT or the *rep-ori* mutant prophage and thus incapable of replicating) in *S. aureus* and collected the total chromosomal DNA for whole-genome sequencing. At 0, 30, 60, and 120 min, we quantified the reads corresponding to 80α and the host DNA adjacent to the $attB_{80\alpha}$ site and measured coverage relative to the average of the entire genome. Induction of the 80α lysogen showed strong amplification of 80α DNA (Fig. 3). Phage replication started before 60 min and was robust by 60 to 120 min after induction. Interestingly, host DNA flanking the 80α lysogen also showed considerable amplification by 60 to 120 min, confirming that the phage was still integrated and that escape replication had indeed amplified these regions. This phenomenon also explains the earlier observation (fig. S3) that the percentage of integrated ϕ11 began to increase 60 min after induction. Host DNA amplification was distinct from phage replication, because it decreased linearly away from the phage origin of replication, similarly to that observed for chromosomal replication (fig. S10). By contrast, 80α infection of nonlysogenic strains showed strong amplification of phage DNA but not of host DNA. Similar results were observed for ϕ11, ϕ52a, and several Newman phages (Fig. 3, B and C, and fig. S11). Hence, many staphylococcal prophages can initiate replication in situ.

To determine the role of in situ replication in lateral transduction and phage production, we designed a system in which the replication of ϕ52a could be tightly regulated and inducible. To do this, we constructed a frameshift mutant (ϕ52a-fs) of the ϕ52a *rep* gene (that controls bidirectional replication) so that the embedded origin of replication remained intact and the mutant could be complemented in trans with $rep_{\phi52a}$ under the control of a tetracycline-inducible promoter. To test for lateral transduction, a Cd^R marker was inserted 5 kb downstream of the $attB_{\phi52a}$ site. Derivatives of these ϕ52a lysogenic strains were induced with mitomycin C [time (t) = 0 min], and anhydrotetracycline was added at 0, 30, 60, 90, and 120 min for replicase expression. Because ϕ52a replicase mutants are unable to lyse their host cells, owing to the lack of viral DNA replication, we used mechanical disruption to release all intracellular particles 2 hours after the addition of inducer. To assay for lateral transduction, the resulting phage lysates were tested as donors of cadmium resistance to *S. aureus*. For phage production, the lysates were tested for plaque formation on a recipient *S. aureus* that constitutively expressed $rep_{\phi52a}$. The WT and ϕ52a-fs mutants were capable of high-frequency lateral transduction of the Cd^R marker (Fig. 4), indicating that in situ replication is not required for lateral transduction. However, the ϕ52a-fs mutants were completely unable to produce plaque-forming units without complementation with $rep_{\phi52a}$. More importantly, complementation of the ϕ52a-fs mutants only resulted in normal levels of phage production if $rep_{\phi52a}$ was provided within 90 min of mitomycin C induction. This timing indicates that the ϕ52a-fs mutants were incapable of being complemented at later time points, presumably because in situ packaging had compromised the integrity of the genomes. These results are consistent with the model (fig. S12) that early in situ replication provides genomic redundancy so that both in situ DNA packaging and excision (followed by phage maturation) can proceed in parallel.

Lateral transduction drives genome organization and evolution

On the basis of the high frequencies of lateral transduction, we speculated that the regions adjacent to phage *attB* sites in the direction of packaging could serve as platforms for high-frequency mobility for any DNA element. There are 10 phage and 5 SaPI *attB* sites scattered throughout the *S. aureus* chromosome (*20*, *21*). Further analysis of the regions flanking the phage *attB* sites revealed that nearly all of the SaPIs and the three staphylococcal chromosomal islands (νSaα, νSaβ, and νSaγ) were positioned such that they could be highly transferred by lateral transduction (fig. S13). Notably, the localization of the phage *attB* sites and the directionality of the phage packaging suggest that most of the bacterial chromosome could also be mobilized by lateral transduction (fig. S13).

Although the SaPIs are well-characterized, highly mobile parasites of helper phages (*22*), specific mechanisms of transfer have not been identified for most of the chromosomal islands of all bacterial species, including νSaα, νSaβ, and νSaγ. Lateral transduction could provide a mechanism to mobilize these islands, too. Indeed, we have already shown lateral transduction–mediated transfer of νSaγ by 80α in the cadmium-resistance transfer experiments, because the Cd^R_2 marker in the second headful was inserted within this island (Fig. 2D). We have also directly tested for lateral transduction–mediated transfer of νSaα by prophage ϕ52a and found that a Cd^R marker 83 kb from the ϕ52a *attB* site (Sa6) was transferred at frequencies three orders of magnitude greater than those observed for GT (fig. S14). These results indicate that lateral transduction is a general mechanism for the high-frequency transfer of mobile genetic elements and pathogenicity islands in *S. aureus*.

To investigate the impact of lateral transduction on genome structure, gene content, and genetic variability, we compared 140-kb regions upstream and downstream of the Sa6 phage *attB* site from 100 complete *S. aureus* genomes. We found that the upstream regions were more highly conserved than the downstream regions, in terms of both gene synteny and gene similarity (fig. S15). This was primarily due to the presence of other phages and SaPIs downstream of the *attB* site. In addition, the number of predicted recombinant fragments in the conserved genes downstream of the *attB* site was significantly higher compared with those upstream ($P = 1.066 \times 10^{-7}$ for paired Wilcoxon nonparametric hypothesis test for comparison of matched samples, and $P = 0.0316$ for unpaired t test for length of recombinant fragments) (fig. S16). By contrast, the nearby regions of the ϕ12 *attB* site (Sa2) showed little gene variability and retained high levels of conservation, both upstream and downstream (fig. S17). However, in these regions, recombination is significantly much greater distal to one headful downstream of the *attB* site (fig. S18; paired Wilcoxon test $P = 1.308 \times 10^{-8}$, unpaired Wilcoxon test for recombinant nucleotides $P = 5.514 \times 10^{-7}$). This region can still be classified as a recombination hotspot and encompasses the gene encoding the surface giant protein Ebh, which is involved in adhesion, bacterial envelope stability, and pathogenesis of staphylococcal infections (*23*, *24*). Variants of Ebh are associated with increased sensitivity to certain antibiotics and reduced virulence (*24*). In a genome-wide analysis of recombination in *S. aureus*, Everitt *et al.* found hotspots of recombination at insertion sites for mobile genetic elements (*25*). Moon *et al.* also observed phage-mediated transfer of virulence-associated genes located in the νSaβ island flanking the phage ϕSaBov insertion site, without elucidating the mechanism (*26*). We found that HGT induced by lateral transduction has a major measurable impact on *S. aureus* genome structure and evolution: First, by promoting gene mobilization, lateral transduction leads to gain and loss of new functions, and second, it provides source material on which homologous recombination can act to generate genetic variability.

Discussion

Of the three modes of bacterial gene transfer (i.e., transformation, conjugation, and transduction), phage transduction is often regarded as the primary driving force of microbial evolution. Here we have identified and characterized not just an additional mode of natural phage transduction but potentially the most powerful one. In our model (fig. S12), bidirectional replication creates multiple integrated prophages so that lateral transduction and normal phage maturation can proceed in parallel. Interestingly, a similar model was proposed long ago for λ phage, where the induction of *doc*L mutants, unable to excise, resulted in the production of noninfectious particles carrying bacterial DNA located to the right of the λ *cos* site (*8*–*10*). Although the model was similar to lateral transduction, the result was quite different, because *doc*L mutants were unable to produce viable phage and the transducing particles required deoxyribonuclease treatment and the addition of purified tails to be infectious. By contrast, the staphylococcal phages naturally generate high titers of infectious transducing particles in the process of wild-type phage production.

Although late prophage excision is the first step in lateral transduction, the timing of *xis* expression has only been investigated in a few phages. In most phage genomes, the integrase

(*int*) and *xis* genes are located in tandem and transcribed together, but for most staphylococcal phages, these genes are opposed, so that they are transcribed from different promoters. Because both integrase and excisionase are generally required for efficient excision, phages that have opposed *int* and *xis* genes or that differentially regulate these genes could be candidates for lateral transduction.

We have demonstrated here that the headful mechanism is essential for lateral transduction. Because this mechanism is not exclusive of the *S. aureus* phages but is widespread in nature, we anticipate lateral transduction will be a universal mechanism of gene transfer. This is currently being studied.

Because lateral transduction can promote the efficient transfer of several hundred kilobases, these spans essentially become large platforms of high-frequency gene transfer for any DNA element located within their boundaries. Considering that bacterial chromosomes often contain multiple prophages, this mode of transduction can transmit a large portion of the bacterial genome at exceptionally high frequencies in a single lytic event. Thus, lateral transduction creates high-volume channels of genetic exchange among hosts, which, in return, provides selection to keep prophages intact and functional. We believe this mechanism sheds light on many genomic conundrums that have gone unexplained for decades in microbiology, such as the mosaicism of the packaging modules of phages, why phage-carried toxins and virulence factors are typically found on certain ends of phage genomes, how some chromosomal islands can be highly transferred, and the occurrence of inexplicable hotspots for recombination in bacterial genomes, just to name a few. Thus, our results indicate that phage-mediated lateral transduction is an extremely powerful force driving both bacterial and phage evolution.

REFERENCES AND NOTES

1. V. B. Rao, M. Feiss, *Annu. Rev. Virol.* **2**, 351–378 (2015).
2. J. R. Penadés, J. Chen, N. Quiles-Puchalt, N. Carpena, R. P. Novick, *Curr. Opin. Microbiol.* **23**, 171–178 (2015).
3. E. S. Lennox, *Virology* **1**, 190–206 (1955).
4. N. D. Zinder, J. Lederberg, *J. Bacteriol.* **64**, 679–699 (1952).
5. M. L. Morse, E. M. Lederberg, J. Lederberg, *Genetics* **41**, 142–156 (1956).
6. C. A. Chelala, P. Margolin, *Genet. Res.* **27**, 315–322 (1976).
7. H. Schmieger, *Mol. Gen. Genet.* **187**, 516–518 (1982).
8. D. Y. Kwoh, J. Kemper, *J. Virol.* **27**, 519–534 (1978).
9. J. W. Little, M. E. Gottesman, in *The Bacteriophage Lambda*, A. D. Hershey, Ed. (Cold Spring Harbor Laboratory, 1971), pp. 371–394.
10. N. Sternberg, R. Weisberg, *Nature* **256**, 97–103 (1975).
11. J. P. Gratia, *Genetics* **95**, 525–544 (1980).
12. B. Kufer, H. Backhaus, H. Schmieger, *Mol. Gen. Genet.* **187**, 510–515 (1982).
13. P. Youderian, P. Sugiono, K. L. Brewer, N. P. Higgins, T. Elliott, *Genetics* **118**, 581–592 (1988).
14. D. L. Court, A. B. Oppenheim, S. L. Adhya, *J. Bacteriol.* **189**, 298–304 (2007).
15. M. M. Susskind, D. Botstein, *Microbiol. Rev.* **42**, 385–413 (1978).
16. N. Quiles-Puchalt *et al.*, *Nucleic Acids Res.* **41**, 7260–7275 (2013).
17. E. Maiques *et al.*, *J. Bacteriol.* **189**, 5608–5616 (2007).
18. A. K. Adikesavan *et al.*, *PLOS Genet.* **7**, e1002244 (2011).
19. S. Weaver, M. Levine, *J. Mol. Biol.* **118**, 389–411 (1978).
20. G. Xia, C. Wolz, *Infect. Genet. Evol.* **21**, 593–601 (2014).
21. A. Subedi, C. Ubeda, R. P. Adhikari, J. R. Penadés, R. P. Novick, *Microbiology* **153**, 3235–3245 (2007).
22. J. R. Penadés, G. E. Christie, *Annu. Rev. Virol.* **2**, 181–201 (2015).
23. M. Kuroda *et al.*, *Biochem. Biophys. Res. Commun.* **374**, 237–241 (2008).
24. A. G. Cheng, D. Missiakas, O. Schneewind, *J. Bacteriol.* **196**, 971–981 (2014).
25. R. G. Everitt *et al.*, *Nat. Commun.* **5**, 3956 (2014).
26. B. Y. Moon *et al.*, *Sci. Rep.* **5**, 9784 (2015).

ACKNOWLEDGMENTS

Funding: This work was supported in part by the Singapore Ministry of Education (T1-2015Sep-13), National Medical Research Council (BNIG15may013), and National University of Singapore (start-up funds) to J.C.; by a project grant (BB/I013873/1) and institute strategic grant funding (ISP2: BBS/E/D/20002173) from the Biotechnology and Biological Sciences Research Council (UK) to J.R.F.; by grant MR/M003876/1 from the Medical Research Council (UK), BB/N002873/1 from the Biotechnology and Biological Sciences Research Council (BBSRC, UK), and ERC-ADG-2014 Proposal n° 670932 Dut-signal (from EU) to J.R.P.; and by Wellcome Trust 201531/Z/16/Z to J.R.F. and J.R.P.
Author contributions: J.C. and J.R.P. conceived the study. J.C., N.Q.-P., Y.N.C., M.S.J.C., and A.F.-S. conducted the experiments. R.B. and J.R.F. performed the genomic and transcriptomic analyses. J.C. and J.R.P. wrote the manuscript. **Competing interests:** The authors declare no competing interests. **Data and materials availability**: All data and code to understand and assess the conclusions of this research are available in the main text, supplementary materials, and the European Bioinformatics Institute repository via EBI accession number PRJEB27527.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/362/6411/207/suppl/DC1
Materials and Methods
Figs. S1 to S18
Tables S1 to S4
References (*27*–*49*)

15 March 2018; accepted 14 August 2018
10.1126/science.aat5867

REPORT

NANOMATERIALS

Entropy-driven stability of chiral single-walled carbon nanotubes

Yann Magnin[1]*, Hakim Amara[2], François Ducastelle[2], Annick Loiseau[2], Christophe Bichara[1]†

Single-walled carbon nanotubes are hollow cylinders that can grow centimeters long via carbon incorporation at the interface with a catalyst. They display semiconducting or metallic characteristics, depending on their helicity, which is determined during their growth. To support the quest for a selective synthesis, we develop a thermodynamic model that relates the tube-catalyst interfacial energies, temperature, and the resulting tube chirality. We show that nanotubes can grow chiral because of the configurational entropy of their nanometer-sized edge, thus explaining experimentally observed temperature evolutions of chiral distributions. Taking the chemical nature of the catalyst into account through interfacial energies, we derive structural maps and phase diagrams that will guide a rational choice of a catalyst and growth parameters toward a better selectivity.

Electronic properties of single-walled carbon nanotubes (SWNTs) depend on their chirality—i.e., the way the SWNTs are rolled along their axis— which is characterized by two indices (n, m). Controlling chirality during the tube's synthesis would enable us to avoid costly sorting and trigger the implementation of promising applications [such as the use of SWNT yarns as strong, light, and conductive wires (*1*) or the development of SWNT-based electronics (*2*)], with the ultimate goal of overcoming the limitations of silicon. Notable breakthroughs have been reported (*3*, *4*), and progress toward carbon nanotube computers (*5*, *6*) has been very rapid. However, selective synthesis still appears to be the weak link, though new studies using solid-state catalysts (*7*–*9*) have reported a chiral-specific growth of SWNTs. Detailed mechanisms underlying this selective growth are still being debated, thus underlining the need for realistic growth models explicitly including the role of the catalyst. Existing models focus on kinetics (*10*), neglecting the role of the catalyst (*11*, *12*), but fail to calculate chiral distributions in line with experiments. Atomistic computer simulations emphasize chemical accuracy (*13*, *14*) but need to be complemented with a model so as to provide a global understanding of the process. In this study, we developed a thermodynamic modeling of the interface between the tube and the catalyst to relate its properties to the resulting chiral distribution obtained during chemical vapor deposition (CVD) synthesis experiments.

Vapor-liquid-solid and vapor-solid-solid CVD processes have both been used to grow SWNTs (*8*), the latter leading to a (n, m) selectivity. Growth can proceed through tangential or perpendicular modes (*15*), and ways to control these modes have been proposed recently (*16*). For specific catalysts and growth conditions favoring the perpendicular mode, a pronounced near-armchair selectivity can be observed (*16*). In such a mode, the interface between the tube and the catalyst nanoparticle (NP) is limited to a line, and a simple model describing the thermodynamic stability of the tube-NP system can be developed. We thus considered an ensemble of configurations of a catalyst NP, possibly a metal or a carbide, in perpendicular contact with a (n, m) SWNT, as in Fig. 1. The total numbers of carbon and catalyst atoms are constant. Configurations differ by the structure of the NP-tube interface, defined by (n, m), for which we have ($n + m$) SWNT-NP bonds, with typically $10 < n + m < 50$. On the tube edge, $2m$ among the bonds are armchair, and ($n - m$) are zigzag (*17*). In a first approximation, the atomic structure of the NP is neglected, and the catalyst appears as a smooth flat surface, in a jellium-like approximation. The interface is then a simple closed loop with two kinds of species: armchair and zigzag contact atoms. Under these conditions, the total energy of the system can be separated into three terms

$$E(n, m) = E_0 + E_{\mathrm{Curv}}(n, m) + E_{\mathrm{Int}}(n, m) \quad (1)$$

where E_0 includes all terms independent of (n, m), such as the energy of the threefold coordinated carbon atoms in the tube wall and the atoms forming the NP. The surface energy of the NP and the very weak surface energy of the tube are also included in E_0, because these surfaces are kept constant. Additionally, E_{Curv} is the curvature energy, and E_{Int} is the interfacial energy. Note that this model could possibly also apply in tangential mode, if the lateral tube-catalyst interaction does not depend on (n, m).

The (n, m)-dependent energy terms concern the tube curvature and its interface with the NP. Using density functional theory (DFT) calculations, Gülseren *et al.* (*18*) evaluated the curvature energy of the isolated tube as $E_{\mathrm{Curv}} = 4\,\alpha\,D_{\mathrm{CNT}}^{-2}$, where D_{CNT} is the tube diameter and $\alpha = 2.14$ eV·Å^2 per C atom. We assume that the interfacial energy for a (n, m) tube in contact with the NP surface depends only on the number of its $2m$ armchair and (n, m) zigzag contacts

$$E_{\mathrm{Int}}^{(n,m)} = 2mE_{\mathrm{Int}}^{\mathrm{A}} + (n - m)E_{\mathrm{Int}}^{\mathrm{Z}} \quad (2)$$

where the armchair ($E_{\mathrm{Int}}^{\mathrm{A}}$) and zigzag ($E_{\mathrm{Int}}^{\mathrm{Z}}$) interfacial energies are given by $E_{\mathrm{Int}}^{\mathrm{X}} = \gamma_{\mathrm{G}}^{\mathrm{X}} + E_{\mathrm{Adh}}^{\mathrm{X}}$, with X standing for A or Z. The edge energy per dangling bond, $\gamma_{\mathrm{G}}^{\mathrm{X}}$, is positive because it is the energy cost of cutting a tube or a graphene ribbon and depends on the type of edge created. The adhesion energy of the tube in contact with the NP, $E_{\mathrm{Adh}}^{\mathrm{X}}$, is negative because energy is gained by reconnecting a cut tube to the NP. $E_{\mathrm{Int}}^{\mathrm{X}}$, the sum of these two terms, has to be positive to create a driving force for SWNT formation. DFT calculations of the edge energies of different edge configurations of a (8, 4) tube (Fig. 2) show no preferential ordering. We thus assume that all tube-catalyst interfaces with the same number of armchair and zigzag contacts have the same energy.

This leads us to introduce the edge configurational entropy as a central piece of the model. We assume that the tube is cut almost perpendicular to its axis, forming the shortest possible interface, for a given (n, m). We neglect vibrational entropy contributions, which are essentially the same for all tubes, except for radial breathing modes. Armchair twofold coordinated C atoms always come as a pair; thus, this entropy (S) that relates the number of ways of putting ($n - m$) zigzag C atoms and m pairs of armchair atoms on n sites (degeneracy) is

$$\frac{S(n, m)}{k_{\mathrm{B}}} = \ln\frac{n\,!}{m!(n - m)!} \quad (3)$$

where k_{B} is Boltzmann's constant. Interfacial energies can be evaluated using DFT calculations, described in the materials and methods. In agreement with previous studies (*17*, *19*), we find $\gamma_{\mathrm{G}}^{\mathrm{A}} = 2.06$ eV per bond and $\gamma_{\mathrm{G}}^{\mathrm{Z}} = 3.17$ eV per bond for graphene, and 1.99 and 3.12 eV

[1]Aix Marseille Université, CNRS, Centre Interdisciplinaire de Nanoscience de Marseille, Campus de Luminy, Case 913, F-13288 Marseille, France. [2]Laboratoire d'Etude des Microstructures, ONERA-CNRS, UMR104, Université Paris-Saclay, BP 72, 92322 Châtillon Cedex, France.
*Present address: MultiScale Material Science for Energy and Environment, MIT-CNRS Joint Laboratory at MIT, Cambridge, MA 02139, USA.
†Corresponding author. Email: bichara@cinam.univ-mrs.fr

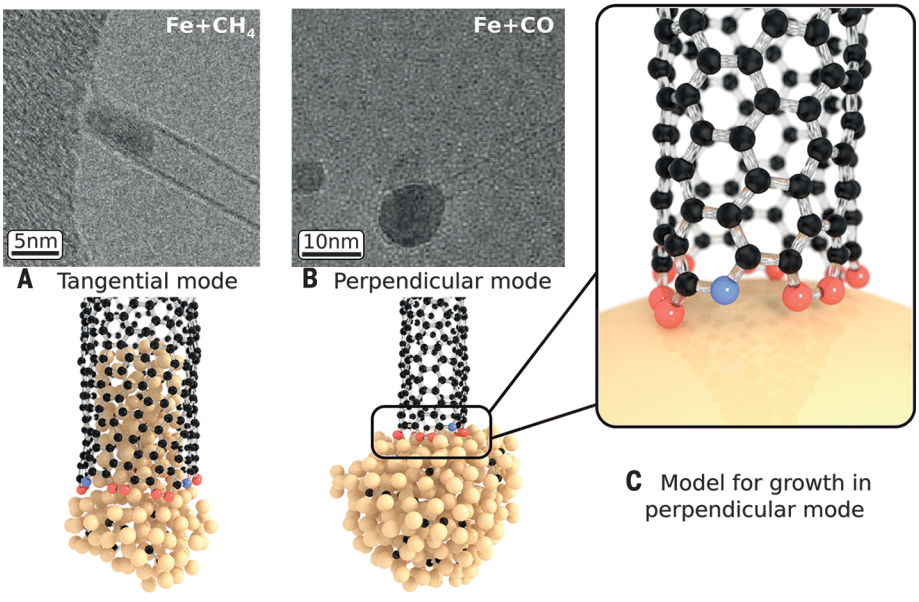


Fig. 1. From experiments to a model. (**A** and **B**) (Top) Postsynthesis transmission electron microscopy (TEM) images of a SWNT attached to the NP from which it grew at 1073 K, using either CH_4 (A) or CO (B) feedstocks, leading to a tangential or perpendicular growth mode, illustrated at the atomic scale (bottom). TEM images are reproduced from (*16*) with permission from the Royal Society of Chemistry. The experiments and our analysis of the growth modes are described in (*16*). (**C**) Sketch of the model, with a SWNT in perpendicular contact with a structureless catalyst. Armchair edge atoms are in red, zigzag ones in blue.

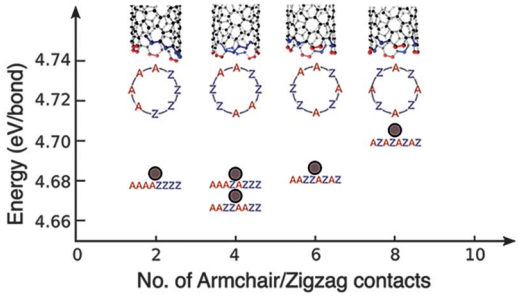


Fig. 2. Key elements of the model. (**Top**) Different ways of cutting a (8, 4) tube, leading to the formation of zigzag (blue) and armchair (red) undercoordinated atoms. For a (8, 4) tube, there are 70 different edge configurations with almost the same energy. (**Bottom**) Formation energies of all possible (8, 4) edges, from DFT calculations described in the materials and methods. The energy levels lie within 25 meV per bond and can thus be considered degenerate.

per bond, respectively, for cutting (6, 6) and (12, 0) tubes. The lower value of γ_G^A is due to the relaxation (shortening) of the C–C bonds of the armchair edge that stabilizes it. Adhesion energies of (10, 0) and (5, 5) tubes on icosahedral clusters of various metals, including Fe, Co, Ni, Cu, Pd, and Au, were calculated in (*20*, *21*). Thus, orders of magnitude for interface energies, E_{Int}^X, of armchair and zigzag terminations in contact with typical catalysts can be estimated: They lie between 0.0 and 0.5 eV per bond, with $E_{Int}^A < E_{Int}^Z$ for these metals. An example of free energy and corresponding probability distribution is plotted as a function of (n, m) in fig. S1.

Instead of focusing on a specific catalytic system, it is more relevant at this stage to study the general properties of the model that links the (n, m) indexes of a SWNT to three parameters characterizing its CVD growth conditions—namely, temperature and the interfacial energies of armchair (E_{Int}^A) and zigzag (E_{Int}^Z) tube-catalyst contacts. For each set of parameters, a free energy can be calculated, and its minimization yields the stable (n, m) value. This model displays similarities with a simple alloy model on a linear chair, but the curvature term, dominant for small diameters, and the small and discrete values of n and m prevent it from being analytically solvable, except for ground states (i.e., stable structures at zero kelvin), for which a solution is provided in the materials and methods. We thus define a three-dimensional (3D) space of stable configurations in the (T, E_{Int}^A, E_{Int}^Z) coordinates.

Setting T and, hence, the entropy contribution to zero, the ground states are readily calculated and displayed in Fig. 3A. Only armchair or zigzag tubes are found to be stable, separated by a line $E_{Int}^Z = \frac{4}{3} E_{Int}^A$. With increasing temperature, they become unstable, and a transition toward chiral tubes takes place. Figure 3B is a contour plot of the surface defined by the transition temperatures. Above this surface, for each set of (T, E_{Int}^A, E_{Int}^Z) parameters, a chiral (n, m) tube is found stable, defining "volumes" of stability for each chirality. To explore it, we can cut slices at constant temperature to obtain an isothermal stability map (Fig. 3C). In such maps, only the most stable (n, m) tube structures are shown, whereas the model yields a distribution of chiralities for each (T, E_{Int}^A, E_{Int}^Z) point. Within a (n, m) domain, this distribution is not constant, especially close to the boundaries, which are calculated by searching for points where the free energies and hence the probabilities of two competing structures are equal. As illustrated in fig. S1B, around the chirality that displays a maximal probability set to 1, neighboring chiralities have non-negligible contributions that depend on (T, E_{Int}^A, E_{Int}^Z). We can also fix either E_{Int}^A or E_{Int}^Z to obtain temperature-dependent phase diagrams, as in Fig. 4A (for $E_{Int}^A = 0.15$ eV per bond) and Fig. 4B (for $E_{Int}^Z = 0.25$ eV per bond). As an example, we can follow the temperature stability of a (6, 6) tube. Figure 4A shows a large stability range with a maximal stability temperature rising from 200 to 800 K by increasing E_{Int}^Z from 0.20 to 0.30 eV per bond, whereas the second map, orthogonal to the first one in the 3D configuration space, shows an upper temperature limit varying from 500 to 700 K, within a narrower E_{Int}^A range. Above the armchair tubes, chiral (n, $n - m$) tubes become stable starting with (n, $n - 1$) and then with increasing ($n - m$) values such as (6, 5), (7, 5), etc. Chiral tubes—i.e., tubes different from armchair or zigzag tubes—are stabilized at finite temperature by the configurational entropy of the tube edge.

An isothermal map calculated at 1000 K is plotted in Fig. 3C. Chiral tubes are spread along the $E_{Int}^Z = \frac{4}{3} E_{Int}^A$ diagonal, between armchair and zigzag ones. Small-diameter tubes are stabilized for larger values of (E_{Int}^A, E_{Int}^Z) and hence for weaker adhesion energies of the tube on the catalyst. Larger-diameter tubes are obtained for small values of (E_{Int}^A, E_{Int}^Z) because the entropy cannot counterbalance the energy cost of the interface, proportional to $n + m$. A comparison of maps at 1000 and 1400 K is provided in fig. S3. As shown in movie S1, the effect of increasing temperature is to expand and shift the stability domain of chiral tubes along and on both sides of the $E_{Int}^Z = \frac{4}{3} E_{Int}^A$ diagonal, with a larger spread on the armchair side. The stability domain of chiralities between central ($2n$, n) and near-armchair (n, $n - 1$) expands substantially at high temperature. However, the free-energy differences become smaller, leading to broader chiral distributions and thus explaining the lack of selectivity reported for tubes grown at very high temperature by electric arc or laser ablation methods (*22*).

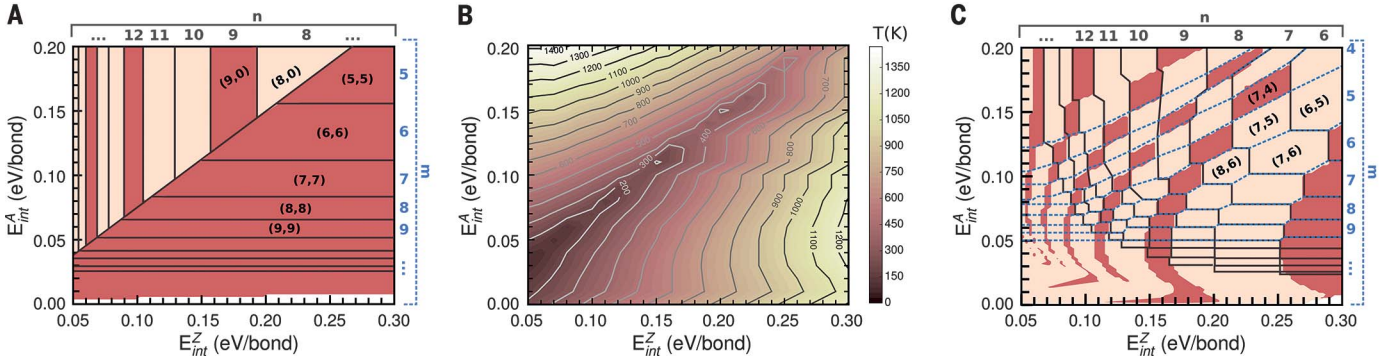


Fig. 3. Structural maps. (**A**) Map of the ground states, with armchair tubes in the lower right corner and zigzag ones in the upper left corner, separated by a line $E^{Z}_{\mathrm{Int}} = \frac{4}{3} E^{A}_{\mathrm{Int}}$. Small-diameter tubes [e.g., (5, 5) and (8, 0)] are obtained for large values of the interfacial energies E_{Int}, whereas stability domains of large-diameter tubes are narrower, with a width decaying as $\frac{1}{n(n+1)}$, and are obtained for small values of E_{Int}. (**B**) Contour plot of the highest temperatures of stability of the ground-state structures, armchair or zigzag. Chiral tubes are found only above this surface, stabilized by the configurational entropy of the tube's edge. Armchair and zigzag tubes can remain stable at high temperatures, in the bottom right and upper left corners, respectively. (**C**) Chirality map at 1000 K. Iso-n (iso-m) values are delimited by solid black (dashed blue) lines. Metallic tubes, for which $(n - m)$ is a multiple of 3, are shown in red, and semiconducting ones are flesh colored. The parameter space for armchair (metallic) and $(n, n - 1)$ and $(n, n - 2)$ (semiconducting) tubes is larger than for other chiralities.

This very simple model displays a fair agreement with literature data, as illustrated in the following examples. Figure 3B suggests a way to grow either zigzag or armchair tubes, the latter being metallic for any diameter. For both, growth kinetics is slow, because each new ring of carbon atoms has to nucleate once the previous one has been completed (*11*, *12*). To overcome this nucleation barrier, one should seek regions in the map where such tubes remain stable at high temperature. For armchair species, this corresponds to the lower right corner of the map in Fig. 3B, where the adhesion energy of armchair edges is strong and that of zigzag ones is weak, and the opposite is true for the interfacial energies. Such requirements have possibly been met in high-temperature (1473 K) CVD experiments (*23*) that also used thiophene in the feedstock. Those experiments might indicate that the presence of sulfur at the interface could modify the relative interaction strength of zigzag and armchair edges with the Fe NP.

The temperature dependence of the chiral distributions, measured by photoluminescence (*24*–*26*) or Raman and transmission electron spectroscopies (*27*) in previous studies, seems more robust. The maps presented in Fig. 4, A and B, are consistent with these experiments, showing that armchair or near-armchair chiralities [(6, 6) and (6, 5)] are grown at low temperature (873 K) and that the chiral distribution gradually shifts toward larger chiral angles [(7, 5), (7, 6), and (8, 4)...] at higher temperatures. Referring to our model, this suggests that Co- and Fe-based catalysts used in these experiments correspond to interfacial energy values around $E^{A}_{\mathrm{Int}} = 0.15$ eV per bond and $E^{Z}_{\mathrm{Int}} = 0.24$ eV per bond, as indicated by the dashed boxes in the maps. A quantitative comparison with four different sets of experimental data is provided in fig. S2, showing a slight tendency to overestimate the width of the distributions. This overestimation partly results from the fact that we use a two-parameter thermodynamic model to account for experiments that include the variability in the catalyst size and chemical composition and the growth kinetics. Our results also confirm that overlooking metallic tubes in photoluminescence experiments introduces a serious bias in the resulting chiral distribution. Further, the dependence of the quantum yield on chiral angles of semiconducting tubes may also contribute to underestimating the width of the experimental distributions.

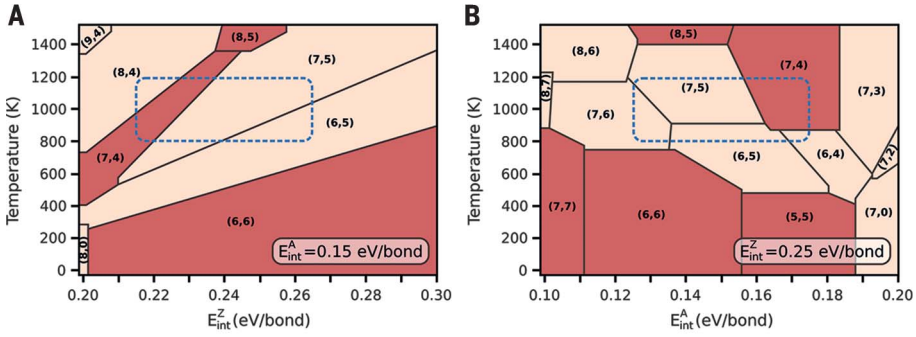


Fig. 4. Chirality phase diagrams. Phase diagrams calculated for constant values of E^{A}_{Int} (**A**) and E^{Z}_{Int} (**B**). These diagrams would be orthogonal in a 3D plot. The blue dashed boxes indicate possible parameter ranges corresponding to the analysis of growth products by He *et al*. (*26*), based on a photoluminescence assignment of tubes grown using a FeCu catalyst. (6, 5) tubes are reported stable up to 1023 K, (7, 5) and (8, 4) become dominant at 1023 K, and (7, 6) at 1073 K.

The present model thus sets a framework for understanding why a number of experiments, using metallic catalysts in perpendicular growth conditions as discussed in (*16*), report a near-armchair selectivity. For such catalysts, E^{A}_{Int} is generally lower than E^{Z}_{Int} (*20*). At low temperature, zigzag or armchair tubes are thermodynamically favored but may not always be obtained, owing to kinetic reasons. On the armchair side, our model indicates that near-armchair helicities are then favored by a temperature increase, because their stability domain is large and they are less kinetically impaired (*11*). At even higher temperatures, tube chiralities tending toward $(2n, n)$ indexes should be stabilized by their larger edge configurational entropy, but their stability domains turn out to be narrower in the present model. Taking the atomic structure of the catalyst into account in our model could rule out some neighboring structures and contribute to open up these domains.

Concerning the practical use of these maps, a first issue is to select the appropriate location for a catalyst in the $(T, E^{A}_{\mathrm{Int}}, E^{Z}_{\mathrm{Int}})$ coordinates, so as to favor the desired tube helicity. Looking at Fig. 3C, one can see that the largest and most interesting parameter ranges correspond to either metallic armchair tubes or to

(n, $n-1$) and (n, $n-2$) semiconducting tubes. A second, more difficult issue is to design a catalyst that would display appropriate $E_{\mathrm{Int}}^{\mathrm{A}}$ and $E_{\mathrm{Int}}^{\mathrm{Z}}$ values. DFT-based calculations, in the same spirit as those used in other studies (*7*, *9*, *20*, *21*, *28*), should probably be helpful. However, the evidence of the important role of the edge configurational entropy calls into question the possibility of explaining the high selectivity reported in (*7*, *9*) on the basis of a structural or symmetry matching. The intrinsic disorder at the edge could be taken into account by averaging over various atomic configurations and using molecular dynamics at finite temperature.

The present model reevaluates the role of thermodynamics in the understanding of SWNT growth mechanisms. It accounts for experimental evidence, such as the near-armchair preferential selectivity, hitherto attributed to kinetics (*11*), and the temperature-dependent trends in chiralities. It also provides a guide to design better, more selective catalysts. However, one must also consider the importance of kinetics in a global understanding of the SWNT growth process. An attempt to combine thermodynamic and kinetic aspects of the growth has been proposed in (*12*), but, overlooking the role of the edge configurational entropy, it led to unrealistic chiral distributions. Those resulting from the present thermodynamic analysis are slightly broader than the experimental distributions (fig. S2) but should be narrower if the reported chirality dependence of the growth kinetics (*11*) is taken into account. Owing to the high synthesis temperatures and the very small size of the interface, there may be SWNT growth regimes where the atomic mobility and the residence time of atoms close to the interface are large enough to achieve a local thermodynamic equilibrium.

REFERENCES AND NOTES

1. N. Behabtu *et al.*, *Science* **339**, 182–186 (2013).
2. A. D. Franklin, *Nature* **498**, 443–444 (2013).
3. Q. Cao *et al.*, *Science* **350**, 68–72 (2015).
4. D. Zhong *et al.*, *Nat. Electron.* **1**, 40–45 (2018).
5. M. M. Shulaker *et al.*, *Nature* **501**, 526–530 (2013).
6. M. M. Shulaker *et al.*, *Nature* **547**, 74–78 (2017).
7. F. Yang *et al.*, *Nature* **510**, 522–524 (2014).
8. M. Li *et al.*, *Top. Curr. Chem.* **375**, 29 (2017).
9. S. Zhang *et al.*, *Nature* **543**, 234–238 (2017).
10. A. A. Puretzky, D. B. Geohegan, S. Jesse, I. N. Ivanov, G. Eres, *Appl. Phys. A* **81**, 223–240 (2005).
11. F. Ding, A. R. Harutyunyan, B. I. Yakobson, *Proc. Natl. Acad. Sci. U.S.A.* **106**, 2506–2509 (2009).
12. V. I. Artyukhov, E. S. Penev, B. I. Yakobson, *Nat. Commun.* **5**, 4892 (2014).
13. A. J. Page, F. Ding, S. Irle, K. Morokuma, *Rep. Prog. Phys.* **78**, 036501 (2015).
14. H. Amara, C. Bichara, *Top. Curr. Chem.* **375**, 55 (2017).
15. M.-F. C. Fiawoo *et al.*, *Phys. Rev. Lett.* **108**, 195503 (2012).
16. M. He *et al.*, *Nanoscale* **10**, 6744–6750 (2018).
17. Y. Liu, A. Dobrinsky, B. I. Yakobson, *Phys. Rev. Lett.* **105**, 235502 (2010).
18. O. Gülseren, T. Yildirim, S. Ciraci, *Phys. Rev. B* **65**, 153405 (2002).
19. T. Wassmann, A. Seitsonen, A. M. Saitta, M. Lazzeri, F. Mauri, *Phys. Rev. Lett.* **101**, 096402 (2008).
20. F. Ding *et al.*, *Nano Lett.* **8**, 463–468 (2008).
21. A. Börjesson, K. Bolton, *J. Phys. Chem. C* **114**, 18045–18050 (2010).
22. L. Henrard, A. Loiseau, C. Journet, P. Bernier, *Synth. Met.* **103**, 2533–2536 (1999).
23. R. M. Sundaram, K. K. K. Koziol, A. H. Windle, *Adv. Mater.* **23**, 5064–5068 (2011).
24. X. Li *et al.*, *J. Am. Chem. Soc.* **129**, 15770–15771 (2007).
25. H. Wang *et al.*, *J. Am. Chem. Soc.* **132**, 16747–16749 (2010).
26. M. He *et al.*, *J. Am. Chem. Soc.* **132**, 13994–13996 (2010).
27. M. Fouquet *et al.*, *Phys. Rev. B* **85**, 235411 (2012).
28. F. Silvearv, P. Larsson, S. Jones, R. Ahuja, J. A. Larsson, *J. Mater. Chem. C* **3**, 3422–3427 (2015).

ACKNOWLEDGMENTS

Funding: Support from the French research funding agency (ANR) under grant 13-BS10-0015-01 (SYNAPSE) is gratefully acknowledged. C.B. thanks P. Müller for stimulating discussions. **Author contributions:** C.B. designed the project; Y.M., H.A., F.D., and C.B. developed the model; and all authors contributed to the data analysis and manuscript preparation. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** All data are available in the main text or the supplementary materials.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/362/6411/212/suppl/DC1
Materials and Methods
Figs. S1 to S3
References (*29*–*35*)
Movie S1
Data S1 and S2

21 March 2018; accepted 8 August 2018
10.1126/science.aat6228

ORGANIC CHEMISTRY

Confined acids catalyze asymmetric single aldolizations of acetaldehyde enolates

Lucas Schreyer, Philip S. J. Kaib, Vijay N. Wakchaure, Carla Obradors, Roberta Properzi, Sunggi Lee, Benjamin List*

Reactions that form a product with the same reactive functionality as that of one of the starting compounds frequently end in oligomerization. As a salient example, selective aldol coupling of the smallest, though arguably most useful, enolizable aldehyde, acetaldehyde, with just one partner substrate has proven to be extremely challenging. Here, we report a highly enantioselective Mukaiyama aldol reaction with the simple triethylsilyl (TES) and *tert*-butyldimethylsilyl (TBS) enolates of acetaldehyde and various aliphatic and aromatic acceptor aldehydes. The reaction is catalyzed by recently developed, strongly acidic imidodiphosphorimidates (IDPi), which, like enzymes, display a confined active site but, like small-molecule catalysts, have a broad substrate scope. The process is scalable, fast, efficient (0.5 to 1.5 mole % catalyst loading), and greatly simplifies access to highly valuable silylated acetaldehyde aldols.

The aldol addition is one of the most fundamental C–C bond–forming reactions in chemical synthesis (*1*). The resulting β-hydroxy carbonyl compounds are versatile building blocks to numerous synthetic targets, in particular oligoketides (*2*, *3*), with antibiotic, antiproliferating, antifungal, and cholesterol-lowering properties (*4*). A fundamental challenge is the synthesis of β-hydroxy aldehydes, which can be envisioned either in a direct cross-aldol reaction between two aldehydes or in an indirect aldol reaction, in which the donor aldehyde is used as its corresponding preformed enolate (*5*). Both approaches frequently suffer from oligomerization or polymerization that results from the preservation of the reacting functional group in the product (Fig. 1A) (*6*).

Traditional approaches have circumvented this problem by installing less reactive surrogates for the critical aldehyde moiety, e.g., olefins (*7*) or esters (*8*), and therefore require additional transformations to access the desired aldehyde species. In recent years, researchers have turned their attention to the development of enantioselective aldol reactions to afford the targeted β-hydroxy aldehydes in a single transformation. The first report came from Denmark and Ghosh in 2001 and comprised an indirect, chiral Lewis base–catalyzed aldol reaction of aldehyde-derived trichlorosilyl enolates with aldehydes (*9*). Subsequent reports were based on chiral amines as Lewis basic catalysts for the direct cross-aldol reaction between nonequivalent aldehydes (*1*, *6*, *10*–*12*). Despite all previous efforts, the smallest enolizable donor aldehyde, acetaldehyde, has remained a challenge in this transformation. The few reports in which this particular donor aldehyde is employed have encountered severe limitations in the scope of substrates to nonenolizable, typically electron-poor acceptor aldehydes (*1*, *13*, *14*). Another drawback is the intrinsic lability (*15*) of the products obtained by these methods, which frequently require protection via *O*-silylation of the β-hydroxy group after direct cross-aldol reactions (*16*), or in situ derivatization of the valuable aldehyde moiety by acetal formation (*14*) or reduction (*1*, *13*). By contrast, the single addition products of the Mukaiyama aldol reaction are appreciably more stable and can be readily isolated by conventional purification methods when silyl groups larger than trimethylsilyl are employed. However, simple (e.g., triethyl-) silyl enolates of acetaldehyde mostly afford oligomers when the powerful small molecule triflimide ($HNTf_2$) is used as catalyst in this transformation (Fig. 1B). In 2006, Boxer and Yamamoto successfully realized a non-enantioselective variant, in which high selectivities toward single and double additions were controlled by the exceptionally bulky tris(trimethylsilyl)silyl group (Fig. 1C) (*3*). Yet, an enantioselective version has remained elusive. In light of the challenging transformations enabled by our recently developed, highly acidic and confined imidodiphosphorimidate (IDPi) catalysts (*17*–*22*), we envisioned an enantioselective Mukaiyama aldol reaction with simple enolsilanes of acetaldehyde. This transformation would potentially benefit from the confined space of the binding site in IDPi catalysts, thereby effecting selective

Max-Planck-Institut für Kohlenforschung, Kaiser-Wilhelm-Platz 1, D-45470 Mülheim an der Ruhr, Germany.
***Corresponding author. Email: list@kofo.mpg.de**

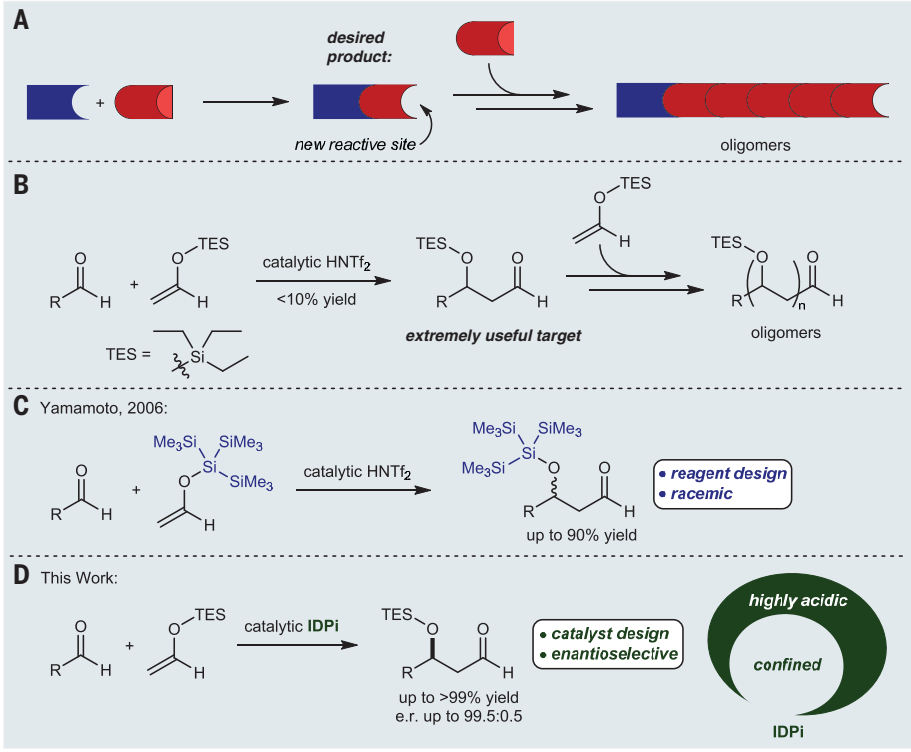


Fig. 1. Indirect aldol reactions with enolsilanes of aldehydes. (**A**) The principal selectivity problem. (**B**) A Mukaiyama aldol reaction with the simple TES enolate of acetaldehyde: no control over multiple additions with triflimide ($HNTf_2$) as catalyst. (**C**) A reagent-controlled, non-enantioselective Mukaiyama aldol reaction with the tris(trimethylsilyl)silyl enolate of acetaldehyde (*3*). (**D**) This work: a catalyst-controlled, highly enantioselective Mukaiyama aldol reaction with simple enolsilanes of acetaldehyde.

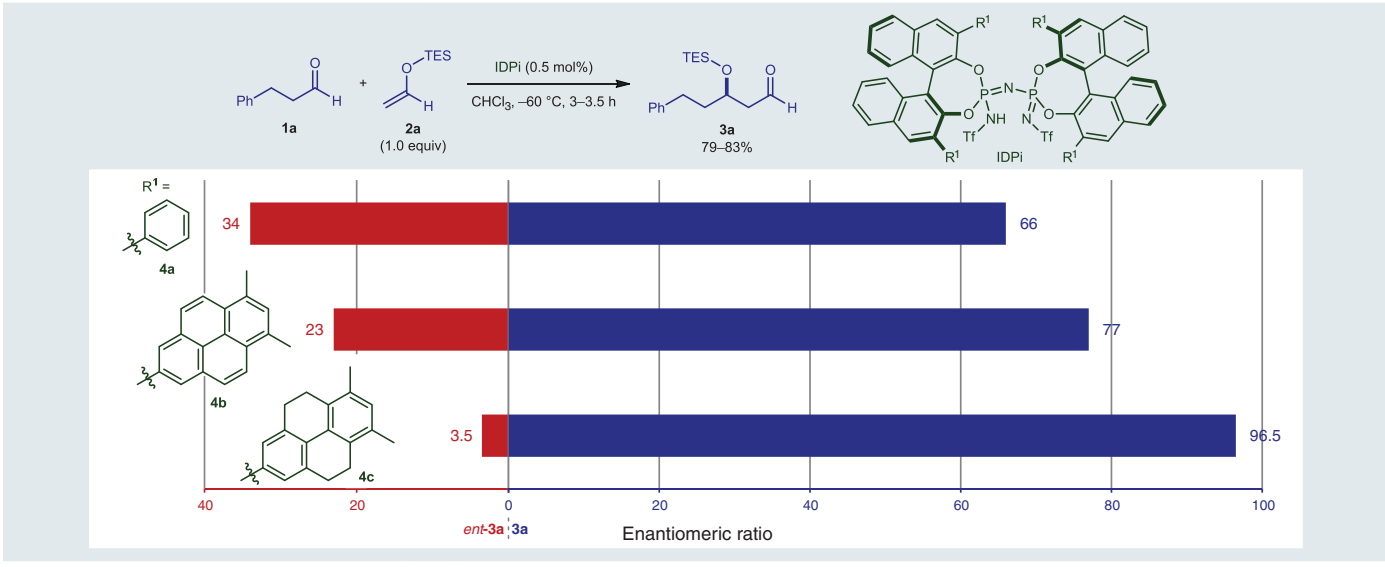


Fig. 2. Catalyst optimization. All yields were determined by ^{1}H NMR spectroscopy, with Ph_3CH as internal standard. For the determination of the e.r., aldol **3a** was reduced and desilylated to the corresponding 1,3-diol **S1a** with $NaBH_4$ and tetra(*n*-butyl)ammonium fluoride (TBAF), respectively (for details, see supplementary materials). We determined e.r. by means of high-performance liquid chromatography (HPLC) with a chiral stationary phase. The absolute configuration was determined by comparison of the optical rotation of the corresponding 1,3-diol **S1a** with a value reported in the literature. Tf, trifluoromethanesulfonyl.

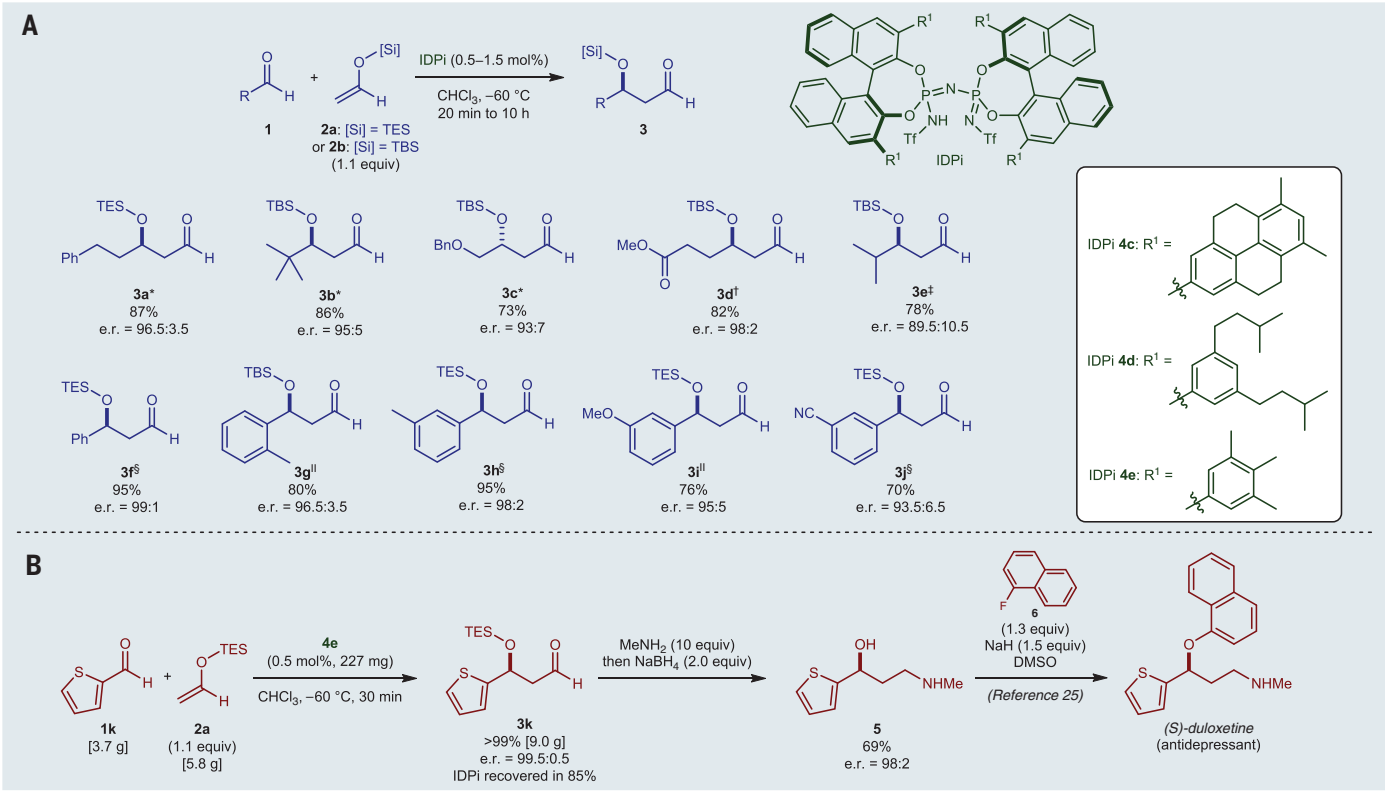


Fig. 3. IDPi-catalyzed Mukaiyama aldol reaction with enolsilanes of acetaldehyde. (**A**) Substrate scope. Reactions were performed with 0.5 mmol of aldehydes **1**. *Using IDPi **4c**. †Using IDPi **4d** and 1.4 equivalents of enolsilane **2b**. ‡Using IDPi **4d** and 1.2 equivalents of enolsilane **2b**. §Using IDPi **4e**. ||Reactions were performed in $CHCl_3$/*n*-hexane (5:4) at –78°C using IDPi **4e**. (**B**) Scaled-up aldol reaction and application to the formal synthesis of (*S*)-duloxetine. All yields are those of isolated materials. For the determination of the e.r., aldols **3** were reduced and desilylated to the corresponding 1,3-diols **S1** with $NaBH_4$ and TBAF, respectively (the 1,3-diols derived from aldols **3b** and **3e** were further derivatized to acetonides **S2b** and **S2e**, respectively; for details, see supplementary materials). We determined e.r. by means of either HPLC or gas chromatography (GC) with chiral stationary phases. Absolute configurations were determined by comparison of optical rotations of the aldol products and/or the corresponding 1,3-diols **S1** with values reported in the literature. The absolute configuration of aldol **3c** was additionally determined by HPLC analysis of the corresponding 1,3-diol **S1c** derived from commercial, enantiomerically enriched (*R*)-1,2,4-butantriol (**S1n**; for details, see supplementary materials).

single additions of enolsilane due to steric differentiation between the substrate and product aldehydes (Fig. 1D).

Although several different classes of previously developed chiral acids failed in this endeavor, owing to lack of any catalytic activity [e.g., chiral phosphoric acids and imidodiphosphates (*23*)] or unselective multiple additions [disulfonimides (*24*)], we were intrigued to observe that phenyl-substituted IDPi **4a** cleanly converted 3-phenylpropanal (**1a**) and enolsilane **2a** to afford aldol **3a** in 79% yield and a promising enantiomeric ratio (e.r.) of 66:34 (Fig. 2).

Inspired by this initial result, we aimed to enhance the enantioselectivity through fine-tuning of our IDPi's substituents R^1. Whereas our previously developed dimethylpyrenyl-substituted IDPi **4b**—the most powerful catalyst in the allylation of aliphatic aldehydes (*17*)—performed with only moderate enantioselectivity, its partially saturated analog **4c** gave access to the desired aldol of 3-phenylpropanal (**3a**) in excellent yield and enantioselectivity (83%, e.r. = 96.5:3.5).

Other aliphatic aldehydes such as pivalaldehyde (**1b**) and benzyloxyacetaldehyde (**1c**) were

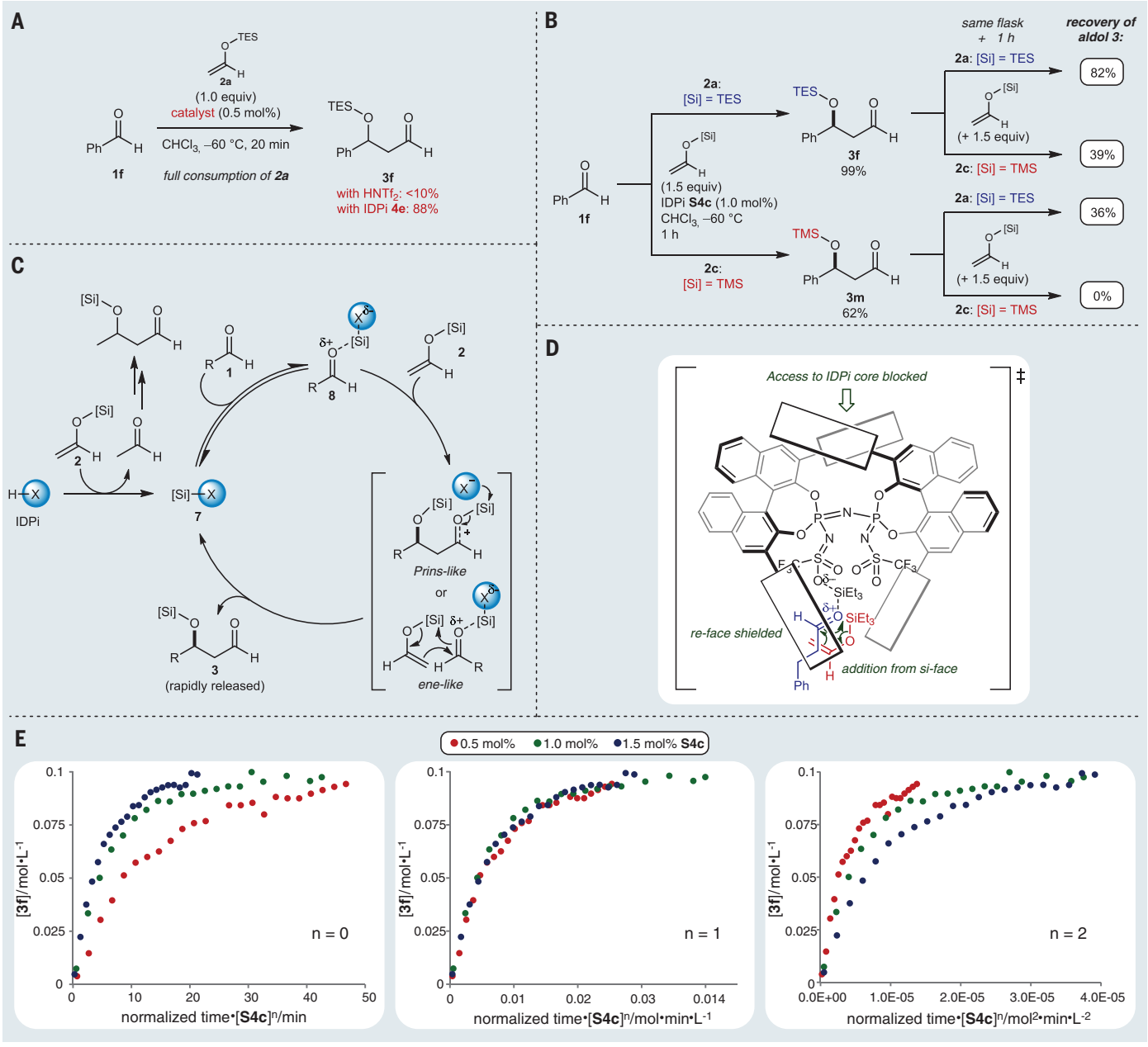


Fig. 4. Mechanistic investigations. (**A**, **B**) Influence of the catalyst and the silyl group, respectively, on the selectivity toward single addition of enolsilane **2**. IDPi **S4c**: see fig. S1. (**C**) Proposed catalytic cycle, based on reaction progress kinetic analysis and NMR spectroscopic studies. (**D**) Proposed transition state for the aldol reaction of 3-phenylpropanal (**1a**) and enolsilane **2a** to afford (*R*)-configured aldol **3a**, following the ene-like mechanism. The superscript double dagger symbol denotes a transition state. (**E**) The method of Burés revealed a first-order dependence on the catalyst concentration. In this method, the substrate (or product) concentration is plotted against a normalized time scale, t•[cat]^n, where "t" is the normalized time (for details, see supplementary materials) and [cat] is the (total) catalyst concentration. The exponent n equals the order of the catalyst where the plots of several reactions with different catalyst concentrations overlap, which in this case is $n = 1$. Yields were determined either by ^{1}H NMR spectroscopy, with Ph_3CH as internal standard, or GC, using *n*-octane as internal standard.

transformed into aldols **3b** and **3c** in similarly high yields and enantioselectivities with the use of *tert*-butyldimethylsilyl (TBS) enolate **2b**, yet in the case of aldol **3c** with an unexpected inversion of the facial selectivity of the nucleophilic addition (Fig. 3A).

For the aldol reactions of methyl 4-oxobutanoate (**1d**) and isobutyraldehyde (**1e**), the best performances were obtained with 3,5-dialkylphenyl substituted IDPi **4d**, which afforded aldol **3d** in high yield and an excellent e.r. of 98:2, and aldol **3e** in 78% yield and an e.r. of 89.5:10.5.

We subsequently turned our attention to aromatic acceptor aldehydes, for which a systematic methylation of the core phenyl substituents R^1 in IDPi **4a** revealed trimethylphenyl analog **4e** to be the optimal catalyst, affording benzaldehyde-derived aldol **3f** in 95% yield and a remarkable e.r. of 99:1 (for details, see supplementary materials). Tolualdehyde-derived aldols **3g** and **3h** were obtained in high yields and enantioselectivities (Fig. 3A). *meta*-Anisaldehyde (**1i**) and *meta*-cyanobenzaldehyde (**1j**) also were well-tolerated substrates, rendering the corresponding aldols **3i** and **3j** in good yields and enantioselectivities.

To illustrate the utility of the obtained aldol products, we envisioned accessing the antidepressant (*S*)-duloxetine in a concise synthesis from aldol **3k** via reductive amination and the reported nucleophilic aromatic substitution with 1-fluoronaphthalene (**6**; Fig. 3B) (*25*). A gram-scale aldol synthesis was therefore performed, in which 0.5 mol % (227 mg) of IDPi **4e** furnished 9.0 g (>99% yield) of thiophene-2-carbaldehyde–derived aldol **3k** in a superb e.r. of 99.5:0.5. The subsequent reductive amination using methylamine and sodium borohydride, and a concomitant desilylation under the reaction conditions, afforded the desired, known (*25*) amino alcohol **5**, which can be converted into (*S*)-duloxetine in a single step.

The rationale for the differentiation between substrate and product aldehyde by the developed catalyst system is a pivotal question. The distinctive properties of our IDPi catalysts in this regard are best illustrated by a direct comparison with triflimide as catalyst. Namely, whereas IDPi **4e** afforded aldol **3f** in 88% yield, under otherwise identical conditions, the yield with triflimide was below 10% because of oligomerization (Fig. 4A).

In addition to the catalyst, the solvent and the silyl group had notable effects on the reaction profile. Typically, the highest yields of aldols **3** were obtained when the reaction was conducted in chloroform (*26*). The steric bulk of the silyl group greatly affected the extent to which aldols **3** were consumed in the course of the reaction (Fig. 4B). Under otherwise identical conditions, with 1.5 equivalents of the TES (**2a**) and TMS (**2c**) enolates of acetaldehyde, the aldol products of benzaldehyde (**3f**, **3m**) were formed in 99 and 62% yield, respectively. Upon addition of a further 1.5 equivalents of either nucleophile, both aldols **3** were consumed in varying amounts depending on the silyl group combinations, lowering the yields of aldol **3f** to 82% (TES/TES) and 39% (TES/TMS) and that of **3m** to 36% (TMS/TES). When TMS enolate **2c** was used at both stages, aldol **3m** was fully consumed (for details, see supplementary materials). These results suggest that the discrimination between substrate and product aldehyde are of steric nature, with the sterically demanding β-silyloxy groups in the aldol products, as well as the size of the nucleophile, as contributing factors. On the basis of reaction progress kinetic analysis (*27*) and nuclear magnetic resonance (NMR) spectroscopic studies (*28*), we propose the reaction to commence with in situ silylation of the IDPi by enolsilane **2** (Fig. 4C). Aldehyde **1** then reversibly coordinates to the silylated catalyst **7**, giving rise to intermediate **8**, in which the formyl group of the substrate is sufficiently activated for the irreversible addition of enolsilane **2**. As the incorporation of the silyl group of nucleophile **2** into the aldol product **3** is preferred to the incorporation of the silyl group of a differently pre-silylated catalyst **7**, we propose two competing mechanisms for the C–C bond–forming step. One involves a Prins-like intermediate as part of a stepwise process, in which the silyl group of the catalyst results in aldol **3**, whereas the other entails a concerted ene-like transition state, in which the silyl group of nucleophile **2** results in the aldol product. In either case, aldol **3** is rapidly released from the catalyst, reestablishing the silylated species **7**. On the basis of x-ray crystallographic data of related catalysts (*17*), we propose that the stereoselectivity of the C–C bond–forming step primarily arises from the shielding of the *re*-face of the catalyst-bound substrate by one of the 3,3′-substituents on the binaphthyl backbones. This allows the addition of the enolsilane (following either mechanism) only from the *si*-face, as illustrated for the aldol reaction of 3-phenylpropanal (**1a**) with enolsilane **2a** (Fig. 4D). We currently aim to confirm this model with high-level computational methods. First-order dependence on the catalyst concentration confirmed the involvement of a single catalyst molecule in the rate-determining step of the reaction, as identified by the graphical method of Burés (Fig. 4E) (*29*).

We believe that our method strongly simplifies the enantioselective access to a class of broadly applicable compounds, of which the synthesis to this point typically involved sequences of several distinct steps (*30*).

REFERENCES AND NOTES

1. Y. Hayashi, T. Itoh, S. Aratake, H. Ishikawa, *Angew. Chem. Int. Ed.* **47**, 2082–2084 (2008).
2. J. Matsuo, M. Murakami, *Angew. Chem. Int. Ed.* **52**, 9109–9118 (2013).
3. M. B. Boxer, H. Yamamoto, *J. Am. Chem. Soc.* **128**, 48–49 (2006).
4. A. M. P. Koskinen, K. Karisalmi, *Chem. Soc. Rev.* **34**, 677–690 (2005).
5. S. Mukherjee, J. W. Yang, S. Hoffmann, B. List, *Chem. Rev.* **107**, 5471–5569 (2007).
6. A. B. Northrup, D. W. C. MacMillan, *J. Am. Chem. Soc.* **124**, 6798–6799 (2002).
7. E. J. Corey, C.-M. Yu, S. S. Kim, *J. Am. Chem. Soc.* **111**, 5495–5496 (1989).
8. E. M. Carreira, R. A. Singer, W. Lee, *J. Am. Chem. Soc.* **116**, 8837–8838 (1994).
9. S. E. Denmark, S. K. Ghosh, *Angew. Chem. Int. Ed.* **40**, 4759–4762 (2001).
10. N. Mase, F. Tanaka, C. F. Barbas 3rd, *Angew. Chem. Int. Ed.* **43**, 2420–2423 (2004).
11. T. Kano, Y. Yamaguchi, Y. Tanaka, K. Maruoka, *Angew. Chem. Int. Ed.* **46**, 1738–1740 (2007).
12. M. Markert, U. Scheffler, R. Mahrwald, *J. Am. Chem. Soc.* **131**, 16642–16643 (2009).
13. S. Hu, L. Zhang, J. Li, S. Luo, J.-P. Cheng, *Eur. J. Org. Chem.* **2011**, 3347–3352 (2011).
14. S. E. Denmark, T. Bui, *J. Org. Chem.* **70**, 10190–10193 (2005).
15. S. K. Mandal, A. Sarkar, P. Chakraborty, A. Chattopadhyay, *Synlett* **29**, 75–78 (2018).
16. P. M. Pihko, A. Erkkilä, *Tetrahedron Lett.* **44**, 7607–7609 (2003).
17. P. S. J. Kaib, L. Schreyer, S. Lee, R. Properzi, B. List, *Angew. Chem. Int. Ed.* **55**, 13200–13203 (2016).
18. Y. Xie *et al.*, *J. Am. Chem. Soc.* **138**, 14538–14541 (2016).
19. S. Lee, P. S. J. Kaib, B. List, *J. Am. Chem. Soc.* **139**, 2156–2159 (2017).
20. L. Liu *et al.*, *J. Am. Chem. Soc.* **139**, 13656–13659 (2017).
21. T. Gatzenmeier, P. S. J. Kaib, J. B. Lingnau, R. Goddard, B. List, *Angew. Chem. Int. Ed.* **57**, 2464–2468 (2018).
22. N. Tsuji *et al.*, *Science* **359**, 1501–1505 (2018).
23. I. Čorić, B. List, *Nature* **483**, 315–319 (2012).
24. P. García-García, F. Lay, P. García-García, C. Rabalakos, B. List, *Angew. Chem. Int. Ed.* **48**, 4363–4366 (2009).
25. A. Träff, R. Lihammar, J.-E. Bäckvall, *J. Org. Chem.* **76**, 3917–3921 (2011).
26. Contaminations of chloroform with HCl could be excluded as source for the observed phenomenon, because neither yield nor enantioselectivity was affected when neutralized chloroform was applied.
27. D. G. Blackmond, *Angew. Chem. Int. Ed.* **44**, 4302–4320 (2005).
28. Z. Zhang *et al.*, *Nat. Commun.* **7**, 12478 (2016).
29. J. Burés, *Angew. Chem. Int. Ed.* **55**, 2028–2031 (2016).
30. C. Bressy, F. Allais, J. Cossy, *Synlett* 3455–3456 (2006).

ACKNOWLEDGMENTS

We thank the HPLC, GC, MS, and NMR departments, especially J. Lingnau, J. Rosentreter, and H. Hinrichs, of the Max-Planck-Institut (MPI) für Kohlenforschung for analytics. We thank M. Klußmann for helpful discussions about the mechanistic studies. **Funding:** Generous support from the Max Planck Society, the Deutsche Forschungsgemeinschaft (Leibniz Award to B.L.) and Cluster of Excellence Ruhr Explores Solvation (RESOLV, EXC 1069), and the European Research Council (Advanced Grant "C–H Acids for Organic Synthesis, CHAOS") is gratefully acknowledged. C.O. acknowledges Alexander von Humboldt Foundation and Bayer Science & Education Foundation for the Humboldt Bayer Fellowship for Postdoctoral Researchers. **Author contributions:** B.L. and L.S. jointly conceived and directed the project. S.L. first identified the reactivity discussed in this publication. L.S., with the assistance of P.S.J.K. and R.P., developed the optimized catalysts. L.S. and V.N.W. conducted screenings of catalysts, substrates, and reaction conditions. P.S.J.K., R.P., and S.L. prepared the racemates. C.O. conducted the mechanistic studies. L.S. and B.L. prepared the manuscript. **Competing interests:** B.L., P.S.J.K., L.S., R.P., and S.L. are inventors on patent WO2017037141 (A1) filed by the MPI für Kohlenforschung covering the IDPi catalyst class and its applications in asymmetric synthesis. **Data and materials availability:** All data are available in the main text or the supplementary materials.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/362/6411/216/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S40
Tables S1 to S32
References (*31*–*68*)

4 May 2018; accepted 7 August 2018
10.1126/science.aau0817

POLYMERS

Key-and-lock commodity self-healing copolymers

Marek W. Urban[1,2,3]*, Dmitriy Davydovich[1,3], Ying Yang[1,3], Tugba Demir[1,3], Yunzhi Zhang[2], Leah Casabianca[2]

Self-healing materials are notable for their ability to recover from physical or chemical damage. We report that commodity copolymers, such as poly(methyl methacrylate)/n-butyl acrylate [p(MMA/nBA)] and their derivatives, can self-heal upon mechanical damage. This behavior occurs in a narrow compositional range for copolymer topologies that are preferentially alternating with a random component (alternating/random) and is attributed to favorable interchain van der Waals forces forming key-and-lock interchain junctions. The use of van der Waals forces instead of supramolecular or covalent rebonding or encapsulated reactants eliminates chemical and physical alterations and enables multiple recovery upon mechanical damage without external intervention. Unlike other self-healing approaches, perturbation of ubiquitous van der Waals forces upon mechanical damage is energetically unfavorable for interdigitated alternating/random copolymer motifs that facilitate self-healing under ambient conditions.

Advances in the last two decades in materials capable of self-healing focused primarily on incorporating physical and chemical mechanisms into polymer networks. These mechanisms can be conveniently classified into the following categories: embedding reactive encapsulated fluids that burst open upon damage to fill and repair damaged areas (*1*); incorporating covalent (*2*–*6*) or supramolecular (*7*–*12*) dynamic bonds that, upon cleavage, reform polymer networks; physically dispersing nanomaterials that enable repair in response to magnetic or electromagnetic fields (*13*, *14*); introducing phase-separated morphologies that facilitate damage closure (*11*, *15*); and incorporating living organisms capable of remending damaged structures (*16*). Polymers—and in particular copolymers, if designed properly—can encode molecular features by placement of repeating units that interact with each other (*17*). However, for self-repair to occur, synchronized chemical and physical events (*18*, *19*), potentially driven by van der Waals (vdW) interactions, must take place.

We synthesized a series of copolymers using atom transfer radical polymerization (ATRP), statistical free radical polymerization, and colloidal polymerization. The methyl methacrylate/n-butyl acrylate (MMA/nBA) molar ratios were varied from 30/70 to 70/30, while maintaining similar molecular weights for all compositions for each synthesis method (ATPR: ~25 kD; statistical: ~60 kD; colloidal: ~700 kD). Copolymer synthesis and properties are summarized in table S1. Figure 1A illustrates selected optical images of p(MMA/nBA) copolymer films produced by ATRP in the 40/60 to 55/45 compositional range that were damaged (0 hours) and allowed to self-heal (~14 hours). Self-repair occurs without external intervention only within narrow 45/55 to 50/50 MMA/nBA compositional ranges (movie S1) Outside this range, self-repair does not take place even days after damage, even though the glass transition temperature (T_g) for nBA-richer 40/60 copolymers is below ambient conditions (25°C, relative humidity = 50%). For undamaged copolymer films, when MMA/nBA molar ratios increase, Young's moduli (E) also increase (Fig. 1B). However, ~14 hours after damage, only 45/55 to 50/50 p(MMA/nBA) copolymer compositions recover 90 to 100% (±5%) of their original tensile strains, respectively (Fig. 1B, B5 and C5). The 45/55 self-healing copolymer exhibits moderate toughness with tensile strain of ~550% and stress values of ~8.6 MPa after self-healing (~600% and 10 MPa before damage). By contrast, the copolymer films outside this range exhibit ~55 and 10% recovery (Fig. 1B, A5 and D5), respectively. Similar behavior, although with longer self-healing times (~86 hours), are observed for copolymers produced by colloidal radical polymerization (fig. S1 and table S1-B).

It is reasonable to hypothesize that for copolymers with 45/55 to 50/50 MMA/nBA molar ratios, the neighboring MMA and nBA copolymer units and their distribution may play some role in self-healing as these compositions are expected to form random and/or alternating chain topologies. To test this hypothesis, MMA and nBA monomers were copolymerized to obtain number average molecular weight M_n = ~20- to 30-kDa pMMA-b-pnBA block copolymers with controlled bock sizes and the number of blocks ranging from two to six (tables S2 and S3). These block copolymers do not exhibit self-healing under the same conditions.

To experimentally assess molecular events associated with self-healing or lack thereof, we used internal reflection infrared imaging (IRIRI), proton nuclear magnetic resonance (^{1}H NMR), and electron spin resonance (ESR), along with stress-strain and dynamic mechanical analysis (DMA). The results of these experiments show that reversible spectroscopic changes are only observed for self-healable copolymer compositions. In IR analysis (figs. S2 and S3), they are manifested by the intensity changes of the C=O (1728 cm^{-1}) and C-O-C (1158 cm^{-1}) normal vibrations due to conformational changes of MMA and nBA repeating units (*20*). In ^{1}H NMR, the key features are the changes in the methyl group shielding-deshielding during the damage-repair cycle for self-healing copolymer compositions (figs. S4 to S6 and tables S4 and S5) (*21*, *22*). Upon mechanical damage, the resonances at 0.98 parts per million (ppm) (a) and 0.96 ppm (b) increase (deshielded) at the expense of diminishing 0.93-ppm (c) and 0.90-ppm (d) peaks (shielded), suggesting a closer chain packing. If strong vdW forces contribute to interchain cohesiveness, mechanical damage will alter the distribution of shielded and deshielded methyl groups along the polymer backbone. Because mechanical damage may also lead to the formation of free radicals, ESR analysis of damaged copolymers showed that, regardless of the copolymer composition, the concentrations of free radicals are in the 4.5 to 8 × 10^{-7} mol/liter range (fig. S7) and appear to have no relation to self-healing. Junction densities (ν_j) due to chain entanglements or adjacent chain interactions were obtained from the measurements of viscoelastic length transitions (VLTs) in dynamic mechanical analysis (DMA) as a function of copolymer composition (fig. S8 and table S6) (*23*). For MMA and nBA homopolymers, the ν_j values are 93.1 and 60 mol/m^3, respectively, but an increase up to 123.6 mol/m^3 is observed for self-healable compositions.

Molecular dynamics (MD) simulations were employed under isothermal (NVT) and isoenergetic equilibration (NVE) conditions as a function of copolymer composition to determine copolymer conformations, end-to-end distances (r), and cohesive energy densities (CEDs). These results are plotted in Fig. 2A (table S7A), and further experimental details along with the results of MD simulations are provided in the supplementary materials. Figure 2A shows that the equilibrium cohesive energy densities (CED_{eq}) (curve a), as well as the end-to-end chain distances (r_{eq}) (curve a') both reach maxima for self-healable 45/55 to 50/50 MMA/nBA compositions (range II). Copolymer interchain packing (Fig. 2B) is greater within self-healing compositional range II, whereas non–self-healable ranges I and III exhibit less interwinding chains. Further, representative copolymer chains extracted from each range shown in Fig. 2C indicate that, within self-healing range II, the chains exhibit extended helix-like conformations with average r_{eq} values of ~34 Å, whereas within ranges I and III, globular shapes with r_{eq} values in the ~25 to 29 Å range are observed. The results of MD simulations are summarized in Table 1 and show that the CED values upon reaching equilibrium (CED_{eq}) increase

[1]Department of Materials Science and Engineering, Clemson University, Clemson, SC 29634, USA. [2]Department of Chemistry, Clemson University, Clemson, SC 29634, USA. [3]Center for Optical Materials Science and Engineering Technologies (COMSET), Clemson University, Clemson, SC 29634, USA.
*Corresponding author. Email: mareku@clemson.edu

for self-healing compositions (range II). Within range II, the vdW_{eq} density values reach 1.96×10^5 kJ/m^3, thus indicating that the extended-chain helix-like conformations are energetically preferable. It is also useful to examine chain conformations equilibrated in the absence of interchain vdW interactions for all compositions. The results of MD simulations for single isolated 30/70, 45/55, and 70/30 p(MMA/nBA) chains shown in Fig. 2D illustrate that, regardless of the copolymer composition, globular conformations, similar to non–self-healable compositions (ranges I and III) are preferable, and the single-chain end-to-end distances (r_{eq}) are within the 21.7 to 27.8 Å range. Thus, without interchain vdW interactions, globular chain conformations prevail regardless of the copolymer topology.

In all MD simulations, an experimental average copolymer density of 1.125 g/cm^3 was used. In separate simulations conducted under the same conditions, copolymer chains were allowed to have excess free volume by assuming an initial density of 0.5 g/cm^3, thus enabling chain motion in and out of the physical cell boundaries upon reaching an equilibrium. The premise behind these simulations was to examine the role, if any, of vdW interactions as a function of copolymer composition in their ability to assume higher- or lower-density states. With an initial density of 0.50 g/cm^3, respective copolymer chains were isothermally equilibrated. Only for self-healing compositions (range II) did the density increase to the 0.529- to 0.562-g/cm^3 range, whereas for non–self-healing compositions, the density decreased (fig. S9), supporting the hypothesis that enhanced vdW forces facilitate favorable interchain interactions and return to denser packing upon physical separation. The question then arises from these experimental and modeling exercises: What are the molecular entities within this narrow compositional range that lead to stronger interchain interactions and subsequent self-healing?

To determine the role of the monomer sequences and the vdW contributions to self-healing, we examined vdW forces and cohesive energies (CE_p) for model pentads containing selected sequences of M and B monomer units (where M and B represent MMA and nBA monomers, respectively). Under NVT MD conditions, selected pentads were placed into one cell and equilibrated. Figure 3A illustrates BMBMB/BMBMB, BMBMB/BMBBM, BMBMB/BMMBB, and BMBMB/MMBBB pentad pairs and the CE_p values due to their interactions. The highest CE_p value (313.6 kJ/mol) exhibits an alternating BMBMB/BMBMB pair (1:1). By contrast, more "blocky-type" MMBBB/MMBBB pentads (Fig. 3B, pair 4-4) have the lowest CE_p value (258.2 kJ/mol). Similarly, other "blocky-type" combinations (Fig. 3B) also exhibit lower CE_p values, thus indicating that the alternating BMBMB-type monomer sequences of the neighboring chains favor overall higher CE_p values. Notably, for alternating BMBMB-type segments composed of MMA (M) and nBA (B) units, there is an average ~120 Å^3 space (~7.1 Å by 4.2 Å by 4.0 Å) between two neighboring nBA monomers separated by one MMA unit along one chain, thus being spatially capable and energetically favorable for hosting an nBA unit of an adjacent chain and thereby enabling the key-and-lock interactions stabilized by vdW forces.

Helix-like chain conformations may also contribute to the high CED_{eq} values within self-healable compositions (range II; Fig. 2). To examine this hypothesis, we analyzed cohesive energy densities for fixed helix-like conformations (CED_{hl}) as a function of copolymer composition. All copolymers across the compositional range were forced to retain a 34.0 ± 0.2 Å end-to-end distance (Fig. 2C") and extended helix-like chain conformations of the 45/55 self-healable copolymer. The results are summarized in Table 1, and CED_{hl} and end-to-end distance values are

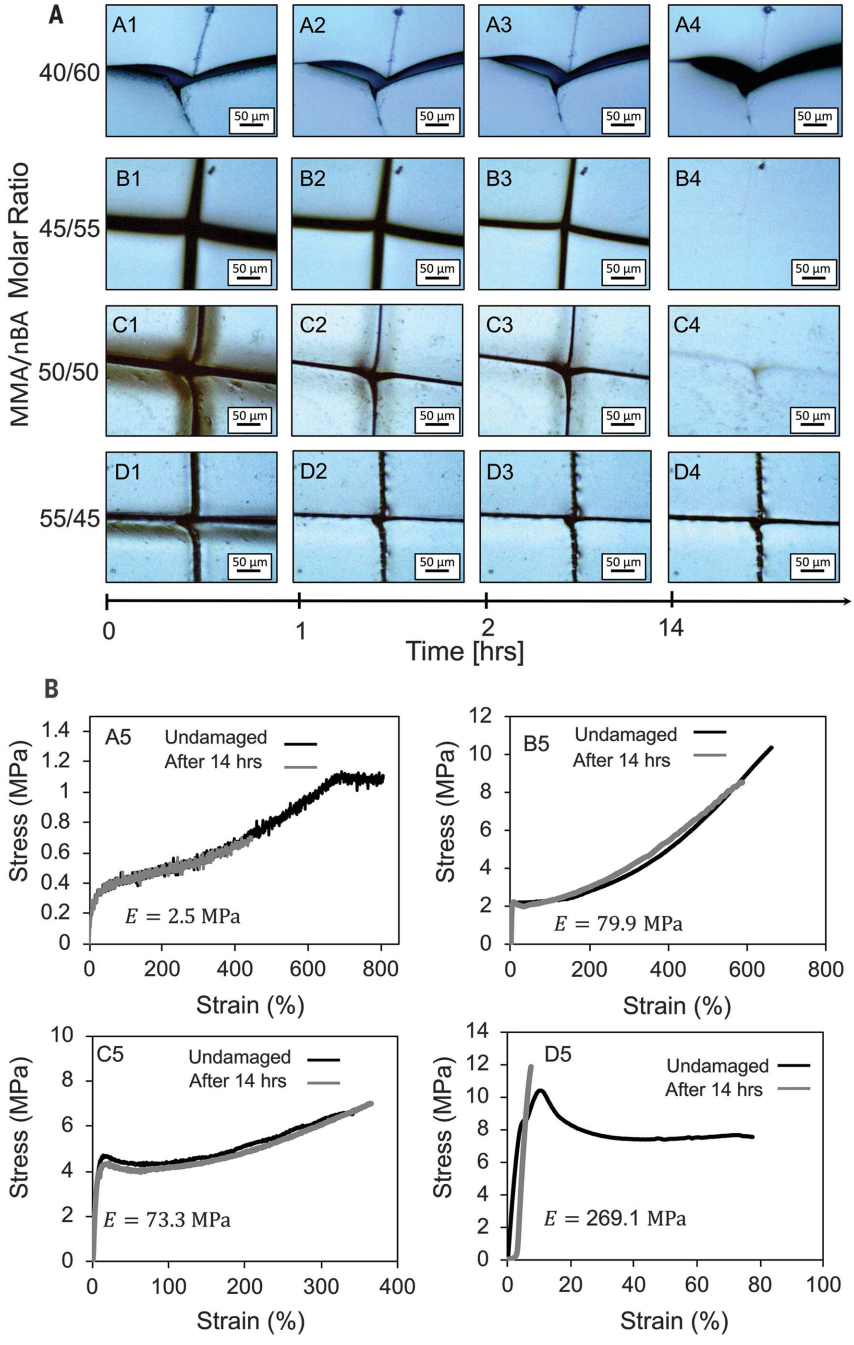


Fig. 1. Self-healing of copolymer films and their mechanical analysis. (**A**) Optical images of damaged p(MMA/nBA) copolymers with the following MMA/nBA molar ratios: 40/60 (A1 to A4), 45/55 (B1 to B4), 50/50 (C1 to C4), and 55/45 (D1 to D4). The copolymers were allowed to repair under ambient conditions. A video of the self-healing process is shown in movie S1. (**B**) A5 to D5: The corresponding stress-strain curves before damage and 14 hours after repair for each copolymer composition in (A).

plotted in Fig. 2A (curves b and b', respectively). As shown, regardless of p(MMA/nBA) composition, the CED_{hl} values are higher compared to their corresponding CED_{eq} counterparts, suggesting that the helix-like conformations resulting from alternating monomer sequences are the main contributing factors to higher CED_{eq} and strong vdW interchain forces. The monomer sequence contributions to self-healing (range II, Fig. 2) are also supported by the most negative ΔH_{eq} values (Table 1). Because pMMA-b-nBA block copolymers do not self-heal and exhibit lower CED_b values (table S7B), these results further substantiate that the presence of alternating/random BMBMB-like sequences favors strong interchain vdW interactions reflected in higher vdW densities that facilitate self-healing (Table 1, italicized rows). Average MMA and nBA reactivity ratios (r_1 = 1.75 to 3.15 and r_2 = 0.2 to 0.39) indicate that it is unlikely that copolymers in range II will form purely alternating copolymers. However, ^{1}H NMR analysis shows the presence of minute homopolymer blocks manifested by the presence of CH_3 protons due to MMA triads (fig. S6 and table S5) for self-healing compositions (range II), but their content is small compared to non–self-healing compositions. MD simulations conducted for average r_1 and r_2 values (2.61 and 0.36, respectively) showed that the maximum CED values are still reached for self-healing compositions (table S7), the probability of finding alternating topologies are also greater, and chain conformations follow the same trend.

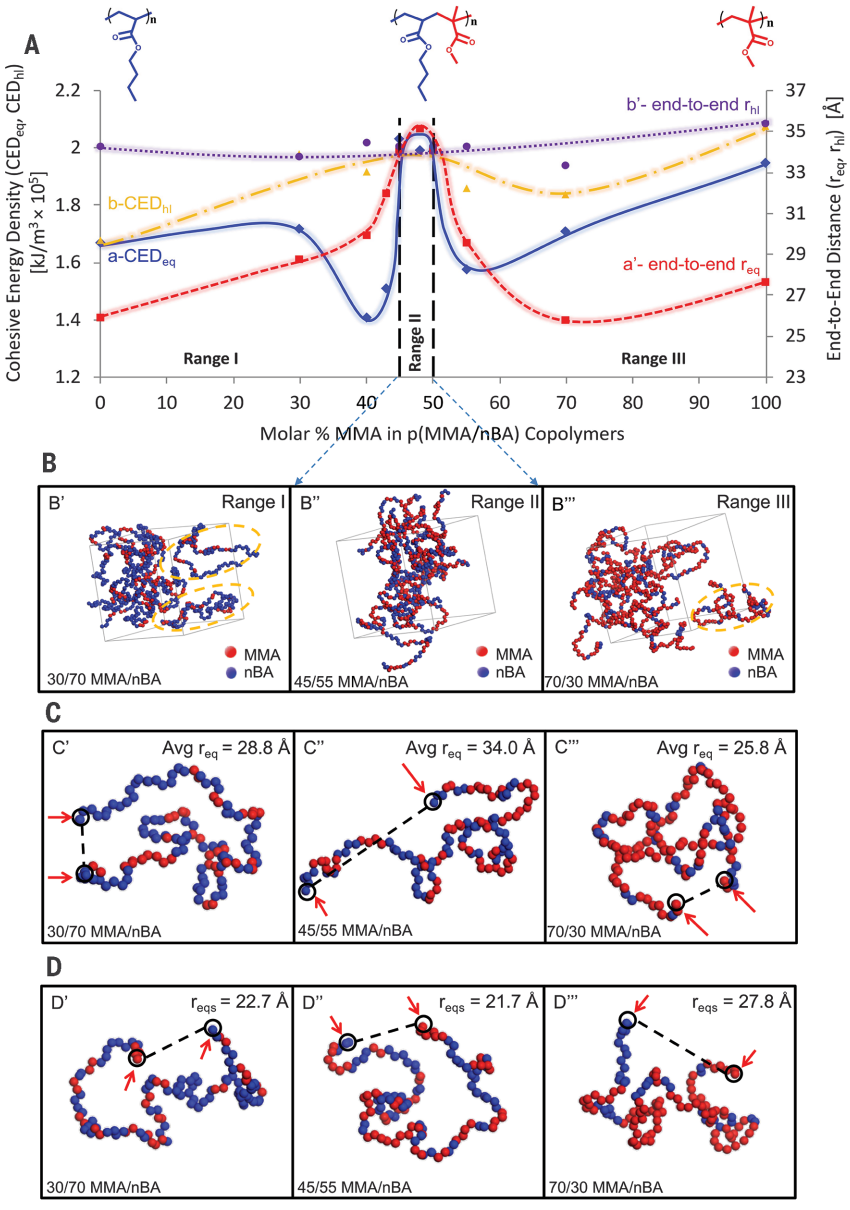


Fig. 2. The results of MD simulations as a function of copolymer composition. (**A**) Cohesive energy densities at equilibrium (CED_{eq}) (curve a), end-to-end equilibrium distances (r_{eq}) (curve a'), cohesive energy densities (CED_{hl}) of forced helical conformations (curve b), and end-to-end chain distances for forced helical conformations (r_{hl}) (curve b') as a function of molar % of MMA in p(MMA/nBA) copolymers. (**B**) Representative examples of copolymer morphologies in range I (MMA/nBA molar ratio: 30/70), range II (MMA/nBA molar ratio: 45/55), and range III (MMA/nBA molar ratio: 70/30); circles denote examples of non-interwinding chains. (**C**) Average end-to-end distances for macromolecular chains extracted from MD simulations in (B). (**D**) Average end-to-end distances for single isolated chains (r_{eqs}) in range I (MMA/nBA molar ratio: 30/70), range II (MMA/nBA molar ratio: 45/55), and range III (MMA/nBA molar ratio: 70/30). The r_{eq} and r_{eqs} values were measured from 3D chain images and may appear not to scale.

A lack of interfacial fluidity attributed to the elevated T_g at damage on the MMA-rich compositional end (range III), and limited quantities of vdW interactions on the MMA-poor end (range I), inhibit self-healing outside the 50/50 to 45/55 region (range II). Because the increase in the CED_{eq} values parallels the increasing number of neighboring MMA/nBA units (table S7C), the formation of key-and-lock configurations between adjacent chains will be favorable for alternating/random copolymer topologies within region II, as reflected by higher junction densities. Assuming that chain entanglements (E) and side-by-side (S) chains are the primary contributors to enhanced junction densities (ν_j) experimentally obtained in DMA measurements (table S6), we extracted both types of interactions from MD simulations and examined the distribution of the induced dipoles due to vdW interactions that contribute to the enhanced ν_j values. Figure 4A-1 illustrates that within the self-healing range II, δ^+- δ^-–induced dipole interactions dominate the entanglement (MD-E') and side-by-side (MD-S') chain interactions. By contrast, Fig. 4A-2 shows extracted copolymer chains just outside the self-healable range (range III) in which randomized orientation of induced dipoles for entangled (MD-E") and side-by-side (MD-S") chains dominate. ν_j values significantly increased for self-healing compositions, clearly supporting MD predictions. Enhanced segmental chain mobility within interfacial regions generated during damage may also aid the self-healing process, owing to lower T_g values near surfaces (*24*, *25*) which can be boosted by collective structural rearrangements at the interfacial regions (*26*).

Further evidence for interchain interactions can be found in determining the flexibility parameter (f_{eq}) (*27*), defined as the fraction of bonds capable of bending out of the collinear direction of previous segments expressed as $f_{eq} = r_{max}/[r_{eq}^2(l\,(2-f))]$, where: r_{max} is fully extended chain length, r_{eq} is the end-to-end distance obtained from MD simulations, and l is length of the repeat unit. The f_{eq} values as a function of copolymer composition are summarized in Table 1. When chains are in the equilibrium state (f_{eq}), the chain flexibility is the smallest for self-healing compositions, indicating that if chains are

deformed as a result of external forces, they will store energy and act like mechanical springs capable of returning to the original state. As was shown for pentad model MD simulations (Fig. 3), these interactions are stabilized by BMBMB/BMBMB key-and-lock junctions between neighboring chains, resulting in recovery upon displacement. Similar behavior is observed for methylmetacrylate/n-pentyl acrylate (PA)– and methylmetacrylate/n-hexyl acrylate) (HA)–based pentads (table S8), in which also alternating copolymer compositions favor enhanced CE_p. The optical images (fig. S10) of selected copolymer compositions show similar self-healing behavior, and stress-strain curves recorded before and after damage are strong indicators of mechanical property recovery after ~14 hours (fig. S10). Comparison of mechanical properties before damage and after self-healing for selected p(MMA/nBA), p(MMA/nPA), and p(MMA/nHA) copolymers is summarized in tables S10 and S11. To illustrate that vdW interactions can be highly effective in self-healing of thermoplastic materials, we severed and physically reattached ~200-μm-thick 46/54 p(MMA/nBA) film. After reattachment, self-healing occurred within a few minutes, but to regain ~70 to 85% mechanical properties took ~80 hours under ambient conditions. The tensile strength of these materials before damage and after self-healing is in the range of 6 to 9 MPa (fig. S11). Repetitive damage and self-healing by making parallel cuts over the same area does not affect self-healing efficiency (fig. S12).

The presence of strong vdW interchain forces for predominantly alternating/random copolymer compositions forming helix-like conformations

Table 1. Cohesive energy density of equilibrated (CED_{eq}) and forced helix-like (CED_{hl}) p(MMA/nBA) copolymer conformations, van der Waals (vdW) density, end-to-end distance (r_{eq}), flexibility parameter (f_{eq}), and enthalpy changes (ΔH_{eq}) as a function of MMA/nBA molar ratios [italics indicate self-healing (range II of Fig. 2) copolymer compositions].

MMA/nBA molar ratio	CED_{eq} ± 0.05 (10^5 kJ/m^3)	CED_{hl} ± 0.05 (10^5 kJ/m^3)	vdW_{eq} density ± 0.05 (10^5 kJ/m^3)	r_{eq} ± 0.2 (Å)	f_{eq} ± 0.1	ΔH_{eq} ± 0.08 (10^3 kJ/mol)
100/0	1.95	2.08	1.87	27.7	0.696	−6.64
70/30	1.71	1.84	1.53	25.8	0.762	−6.58
55/45	1.58	1.86	1.35	29.6	0.64	−6.45
50/50	*1.99*	*1.99*	*1.91*	*34.1*	*0.521*	*−8.11*
45/55	*2.03*	*2.01*	*1.96*	*34.0*	*0.523*	*−8.32*
40/60	1.41	1.92	1.30	30.0	0.625	−5.88
30/70	1.72	1.98	1.44	28.8	0.66	−7.26
0/100	1.67	1.68	1.49	25.9	0.758	−7.93

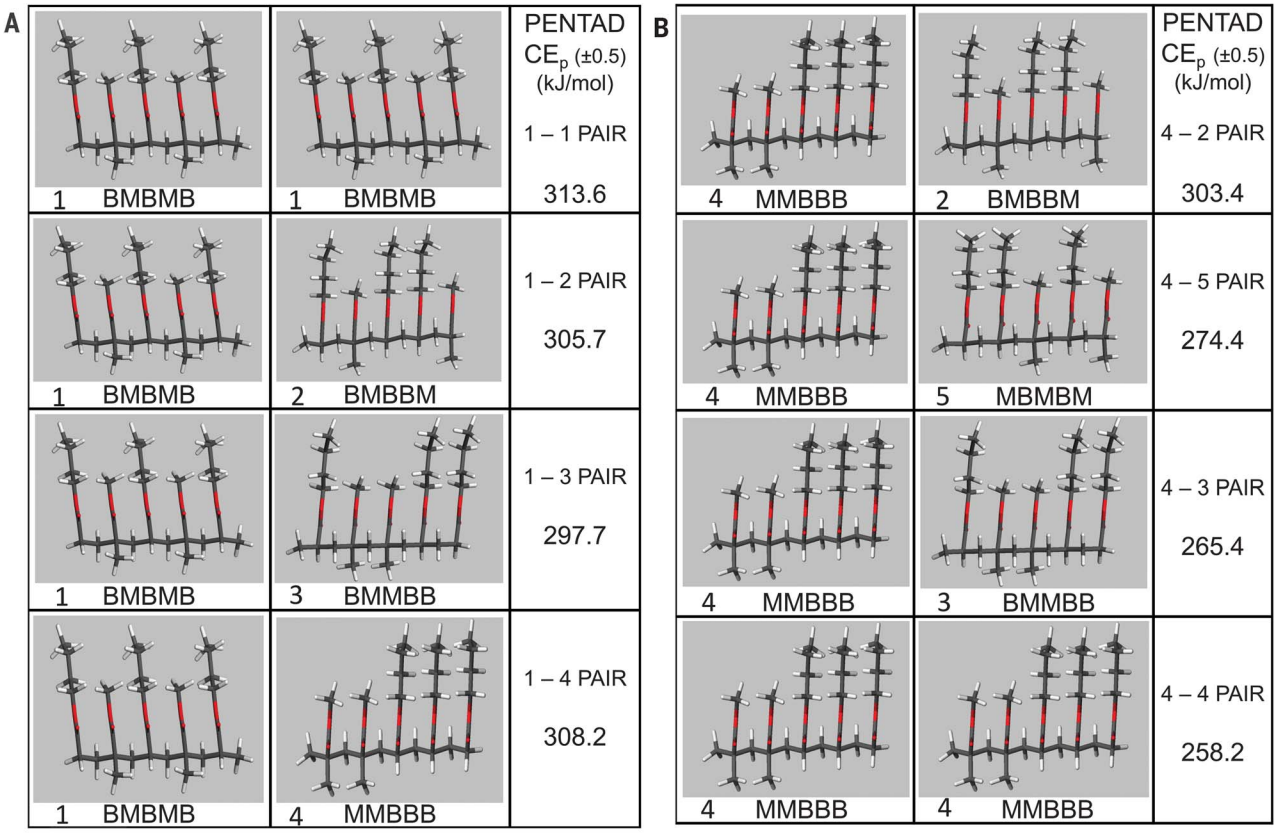


Fig. 3. Cohesive energies (CE_p) for selected pentad pair combinations. (**A**) CE_p values for (1–1), (1–2), (1–3), (1–4). (**B**) (4–2), (4–5), (4–2), (4–4) pentad pair interactions (1, BMBMB; 2, BMBBM; 3, BMMBB; 4, MMBBB 5, MBMBM).

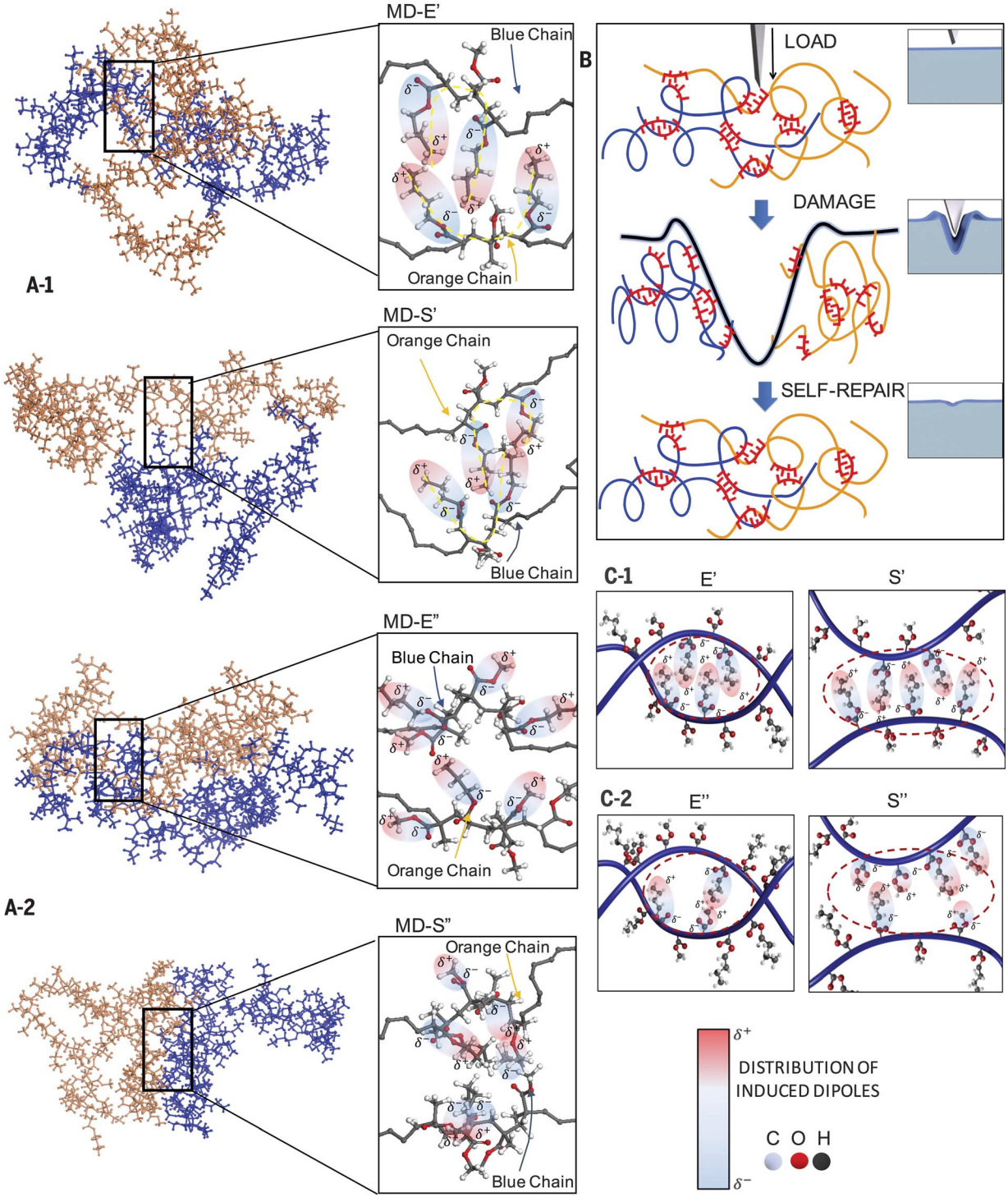


Fig. 4. Visual representation of MD simulations and proposed self-healing mechanism. (**A1**) Extracted interchain interactions from MD simulations for self-healable compositions (range II) of entangled (MD-E') and side-by side (MD-S') chains. (**A2**) Extracted interchain interactions from MD simulations for non–self-healable compositions (range III; 55/45 MMA/nBA ratio) of entangled (MD-E") and side-by side (MD-S") chains. To visually differentiate copolymers, the neighboring chains were colored in orange and blue. The color scale represents relative distributions of induced dipoles (red, high; blue, low). (**B**) Proposed self-healing mechanism responsible for the restoration of vdW interactions; the presence of key-and-lock associations (red) facilitates chain recovery upon mechanical damage. (**C1**) Pictorial representation of the distribution of induced dipole moments in self-healable entangled (E') and side-by side (S') chains. (**C2**) Pictorial representation of distribution of induced dipole moments for none self-healable entangled (E'') and side-by side (S'') chains.

creates a viscoelastic response that energetically favors self-recovery upon chain separation because of key-and-lock associations of neighboring chains (Fig. 4B). In the presence of these interactions, vdW forces stabilize key-and-lock neighboring junctions reflected in the enhanced CED_{eq} values. When chains are separated as a result of mechanical damage and an external force is removed, copolymer chains return to their initial conformations by restoring helix-like chain conformations in a spring-like manner and reforming key-and-lock junctions manifested by increased CED_{eq} and r_{eq} distances for self-healing compositions (range II). Outside self-healing compositions (ranges I and III), irreversible chain dislocations and insufficient interchain vdW forces inhibit complete chain recovery. Thus, the presence of directional vdW forces due to induced dipole interactions enhances CED_{eq} of entangled or side-by-side chains (Fig. 4C).

For comparison with vdW forces, when supramolecular interactions, such as H-bonding, were employed in self-healing of rubber, the tensile strength ~3.5 MPa at similar elongation levels was reached (*7*). Although the underlying mechanisms responsible for self-healing using supramolecular and vdW forces are substantially different, they may result in somewhat similar responses. Considering directionality and polarity as commonly accepted differences between H-bonding and vdW interactions, the former facilitates localized bonding directionality because of the orientation of interacting molecular orbitals and high polarity (hydrophilicity). The main feature of vdW interactions is high polarizability (hydrophobicity) with a tendency to form ubiquitous nondirectional contacts between neighboring macromolecular segments. However, in layered systems with large individual atomic planes, individual weak vdW attractive forces in two-dimensional materials (e.g., graphene, others) are directional and become collectively strong. In amorphous polymers, at first approximation, vdW interactions are nondirectional, but the magnitude of vdW forces will strongly depend on the proximity of the neighboring units (*28*). As extended semihelix macromolecules are in closer proximity to their alternating/random copolymer neighbors, vdW forces will increase because of the preferable bearings of the side groups, resulting in interdigitated key-and-lock interchain morphologies that facilitate self-healing.

REFERENCES AND NOTES

1. S. R. White *et al*., *Nature* **409**, 794–797 (2001).
2. B. Ghosh, M. W. Urban, *Science* **323**, 1458–1460 (2009).
3. K. Imato *et al*., *Angew. Chem. Int. Ed.* **51**, 1138–1142 (2012).
4. Y. Yang, M. W. Urban, *Angew. Chem. Int. Ed.* **53**, 12142–12147 (2014).
5. H. Ying, Y. Zhang, J. Cheng, *Nat. Commun.* **5**, 3218 (2014).
6. X. Chen *et al*., *Science* **295**, 1698–1702 (2002).
7. P. Cordier, F. Tournilhac, C. Soulié-Ziakovic, L. Leibler, *Nature* **451**, 977–980 (2008).
8. M. Burnworth *et al*., *Nature* **472**, 334–337 (2011).
9. N. Holten-Andersen *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **108**, 2651–2655 (2011).
10. M. Nakahata, Y. Takashima, H. Yamaguchi, A. Harada, *Nat. Commun.* **2**, 511 (2011).
11. Y. Chen, A. M. Kushner, G. A. Williams, Z. Guan, *Nat. Chem.* **4**, 467–472 (2012).
12. C.-H. Li *et al*., *Nat. Chem.* **8**, 618–624 (2016).
13. C. C. Corten, M. W. Urban, *Adv. Mater.* **21**, 5011–5015 (2009).
14. L. Huang *et al*., *Adv. Mater.* **25**, 2224–2228 (2013).
15. Y. Yang, D. Davydovich, C. C. Hornat, X. Liu, M. W. Urban, *Chem* **4**, 1928–1936 (2018).
16. H. M. Jonkers, A. Thijssen, G. Muyzer, O. Copuroglu, E. Schlangen, *Ecol. Eng.* **36**, 230–235 (2010).
17. F. S. Bates *et al*., *Science* **336**, 434–440 (2012).
18. Y. Yang, M. W. Urban, *Chem. Soc. Rev.* **42**, 7446–7467 (2013).
19. M. W. Urban, *Stimuli-Responsive Materials: From Molecules to Nature Mimicking Materials Design* (Royal Society of Chemistry, 2016), pp. 390–391.
20. M. W. Urban, *Vibrational Spectroscopy of Molecules and Macromolecules on Surfaces* (Wiley, New York, 1993), pp. 261–262.
21. F. Bovey, *High Resolution NMR of Macromolecules* (Elsevier, 2012), pp. 145–160.
22. H. N. Cheng, T. Asakura, A. D. English, *NMR Spectroscopy of Polymers: Innovative Strategies for Complex Macromolecules* (ACS, 2011), pp. 247–249.
23. C. C. Hornat, Y. Yang, M. W. Urban, *Adv. Mater.* **29**, 1603334 (2017).
24. K. Dalnoki-Veress, J. A. Forrest, C. Murray, C. Gigault, J. R. Dutcher, *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* **63**, 031801 (2001).
25. R. D. Priestley, C. J. Ellison, L. J. Broadbelt, J. M. Torkelson, *Science* **309**, 456–459 (2005).
26. G. Adam, J. H. Gibbs, *J. Chem. Phys.* **43**, 139–146 (1965).
27. P.-J. Flory, *Proc. R. Soc. London Ser. A* **234**, 60–73 (1956).
28. J. de Boer, *Trans. Faraday Soc.* **32**, 10–37 (1936).

ACKNOWLEDGMENTS

We thank K. Ivey for technical assistance in GPC, DSC, and DMA measurements. **Funding**: This work was supported by the National Science Foundation under Award DMR 1744306 and partially by the J.E. Sirrine Foundation Endowment at Clemson University. **Author contributions:** The experiment was designed by M.W.U., D.D., Y.Y., and L.C. (EPR). Experimental work was conducted by D.D., T.D., and Y.Z. Data analysis was performed by M.W.U., Y.Y, D.D., and L.C. M.W.U. wrote the manuscript. **Competing interests:** None declared. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper or the supplementary materials. Patent application no. 62/702,410 was filed 24 July 2018; contact: C. Gesswein, Clemson University Research Foundation (CURF); email: agesswe@clemson.edu.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/362/6411/220/suppl/DC1
Materials and Methods
Figs. S1 to S12
Tables S1 to S11
References (*29–40*)
Movies S1 and S2

27 February 2018; accepted 30 August 2018
10.1126/science.aat2975

ORGANIC CHEMISTRY

Ketyl radical reactivity via atom transfer catalysis

Lu Wang, Jeremy M. Lear, Sean M. Rafferty, Stacy C. Fosu, David A. Nagib*

Single-electron reduction of a carbonyl to a ketyl enables access to a polarity-reversed platform of reactivity for this cornerstone functional group. However, the synthetic utility of the ketyl radical is hindered by the strong reductants necessary for its generation, which also limit its reactivity to net reductive mechanisms. We report a strategy for net redox-neutral generation and reaction of ketyl radicals. The in situ conversion of aldehydes to α-acetoxy iodides lowers their reduction potential by more than 1 volt, allowing for milder access to the corresponding ketyl radicals and an oxidative termination event. Upon subjecting these iodides to a dimanganese decacarbonyl precatalyst and visible light irradiation, an atom transfer radical addition (ATRA) mechanism affords a broad scope of vinyl iodide products with high *Z*-selectivity.

The ketyl coupling of carbonyls offers a mechanistically inverted approach to constructing C-C bonds versus classic, polar mechanisms (*1*). However, a major limitation of this valuable method is its reliance on strong, stoichiometric reductants (e.g., Na, K, Ti) (*2*–*4*) to overcome the large reduction potential of carbonyls (peak cathodic potential $E_{p,c}$ > −2 V versus SCE) (*5*, *6*). As shown in Fig. 1A, a powerful tool used to overcome the thermodynamic barrier for ketyl generation is Kagan's reagent (SmI_2) (*7*, *8*). This single-electron reductant can even be used catalytically when coupled with strong, stoichiometric reductants (e.g., Zn/Hg) (*9*). Recently, photochemical approaches have been applied to address this fundamental thermodynamic challenge in a complementary manner. For example, Knowles and co-workers developed a proton-coupled electron transfer strategy to reduce ketones to ketyl radicals using Ir or Ru photocatalysts (< −1.5 V), Brønsted acids, and a milder stoichiometric reductant (e.g., Hantzsch ester) (*10*, *11*). The groups of Yoon, Ngai, and Huang (*12*–*15*) have also shown that concerted use of Lewis acids enables photocatalytic reduction of carbonyls to access ketyl radicals and their vinylogous analogs. Additional metal-catalyzed strategies for carbonyl-alkyne coupling promote complementary reactivity (by Ni, Ru, or Ir) with alternate reductants (e.g., Et_3B, Mg, H_2) (*16*–*18*). However, in each of the ketyl-based strategies, a redox-neutral approach remains mechanistically unfeasible because catalysts capable of overcoming the high reduction potential of a carbonyl are necessarily also prone to reducing the resulting ketyl-coupling adduct.

Our complementary strategy, outlined in Fig. 1, addresses this challenge by replacing the carbonyl reduction step with a halogen abstraction event via an atom transfer radical addition (ATRA) mechanism. In this case, a Mn catalyst promotes mild ketyl radical formation from an in situ–generated intermediate containing a weak C-I bond. The resulting radical is capable of coupling with a range of alkynes. Catalyst turnover occurs (via formal oxidation of the ensuing vinyl radical intermediate) by atom transfer (*19*, *20*) with Mn-I, in a net redox-neutral mechanism that enables access to synthetically versatile (*21*) *Z*-vinyl halides.

In our synthetic design, we focused on addressing the challenge of ketyl radical generation from aliphatic aldehydes. We envisioned that their conversion to α-oxy iodides would benefit from hyperconjugative donation of the nonbonding oxygen electrons into the C-I antibonding orbital (n → σ*) to further weaken this bond. Among several α-iodo radical precursors we investigated, the α-acetoxy derivative proved simplest to generate under mild conditions. This in situ activation is performed by combining AcI (AcCl, NaI) with carbonyls [5% $Zn(OTf)_2$, 0°C, 15 min] in a modified version of Adams' nearly century-old procedure (*22*). The resulting aldehyde derivatives are conveniently handled in an aerobic atmosphere at room temperature and are stable to basic, aqueous washes without elimination. This practical synthetic accessibility allowed us to electrochemically validate our hypothesis that α-acetoxy iodides ($E_{p,c}$ = −1.1 V; fig. S2) are more easily reduced than their aldehyde precursors (> −2.2 V) (*5*) as well as other alkyl iodides (*23*–*26*). Next, we explored the generation of ketyl radicals—and their redox-neutral coupling with alkynes—by using a variety of atom transfer catalysts.

In accordance with our design plan, we observed that several photocatalysts promote this redox-neutral alkyne coupling with ketyl radicals derived from aliphatic aldehydes (**1**), as illustrated in Fig. 1C. Expecting the nucleophilicity of α-OAc ketyl radicals to be attenuated, we explored their combination with several alkynes of varying electronic character. In particular, coupling of silyl acetylene **2** affords silyl vinyl iodide **3** (containing orthogonal handles for further synthetic manipulation) in a streamlined approach that obviates an alkynylation–Al reduction–iodination sequence. Among the atom transfer catalysts investigated, several photocatalysts afforded new, redox-neutral coupling adduct **3** preferentially over the classic, reduced coupling product **4**. These catalysts included complexes of earth-abundant, first-row metals (e.g., Mn, Fe) as well as more reducing photocatalysts (e.g., Ru, Ir; see table S1 for further details). Ultimately, a $Mn_2(CO)_{10}$ catalyst was found to provide excellent selectivity for redox-neutral coupling (**3**:**4**, >20:1) along with high *Z*:*E* diastereoselectivity (>20:1). The latter feature is notable, as there are few methods to access vinyl iodides with high *Z*-selectivity (*27*). Although Mn-mediated reactions are typically associated with oxidative mechanisms, atom transfer pathways are accessible via $Mn_2(CO)_{10}$ (*28*–*30*). Unlike typical photocatalysts, a photon is not necessary for turnover of the Mn catalyst. However, we observed that continual irradiation is necessary to access high efficiency and selectivity, likely due to an equilibrium between the precatalyst dimer and the active catalyst. Given the high chemo- and stereoselectivity afforded by the Mn catalyst, we decided to further explore its synthetic potential in this redox-neutral mechanism.

In our mechanistic proposal, shown in Fig. 1D, $Mn_2(CO)_{10}$ precatalyst is homolyzed to $Mn(CO)_5$ (**5**, Mn•) by irradiation with a blue LED. This 17-electron species is a competent ATRA catalyst that can abstract I• from the weak C-I bond [bond dissociation energy (BDE), 58 kcal/mol] (*31*) of in situ–generated α-acetoxy iodide **6**. Combination of the resultant ketyl radical **7** with alkyne **8** affords vinyl radical **9**. This open-shell intermediate is formally oxidized by $Mn(CO)_5I$ (**10**, [Mn]-I) via atom transfer to regenerate the Mn• catalyst (**5**), while also forming vinyl iodide **3′** in a redox-neutral mechanism (*32*). The net conversion of **6** to **3′** is exothermic because of the formation of a strong vinyl C-I bond (BDE, 61 to 68 kcal/mol). An observed post-reaction isomerization of the vinyl iodide products—from 1:1 to >20:1 *Z*:*E* selectivity—is also thermodynamically favored by up to 3 kcal/mol (*33*). This Mn-catalyzed isomerization likely occurs via an intermediate vinyl radical, which is consistent with reports of photoinduced, single-electron reduction of aryl iodides for sp^2 radical generation (*34*, *35*). This Mn-catalyzed ATRA mechanism precludes an alternate pathway, in which vinyl radical **9** is further reduced to vinyl anion **11** [$E_{p,c}$ = −0.1 V (*36*), which is at least 2 V more favorable than carbonyl to ketyl reduction] to afford allyl ester **4**. Instead, by coupling catalyst turnover with product formation, the classic reductive mechanism can be overridden by this redox-neutral pathway.

In probing the synthetic utility of this strategy, we were pleased to find that aliphatic aldehydes, which are challenging to reduce to ketyls ($E_{p,c}$ > −2 V) (*5*), efficiently combine with a range of alkynes, as shown in Fig. 2. Alkynes with broad electronic character are coupled to the stabilized ketyl radicals, providing *Z*-vinyl

Department of Chemistry and Biochemistry, The Ohio State University, Columbus, OH, USA.
*Corresponding author. Email: nagib.1@osu.edu

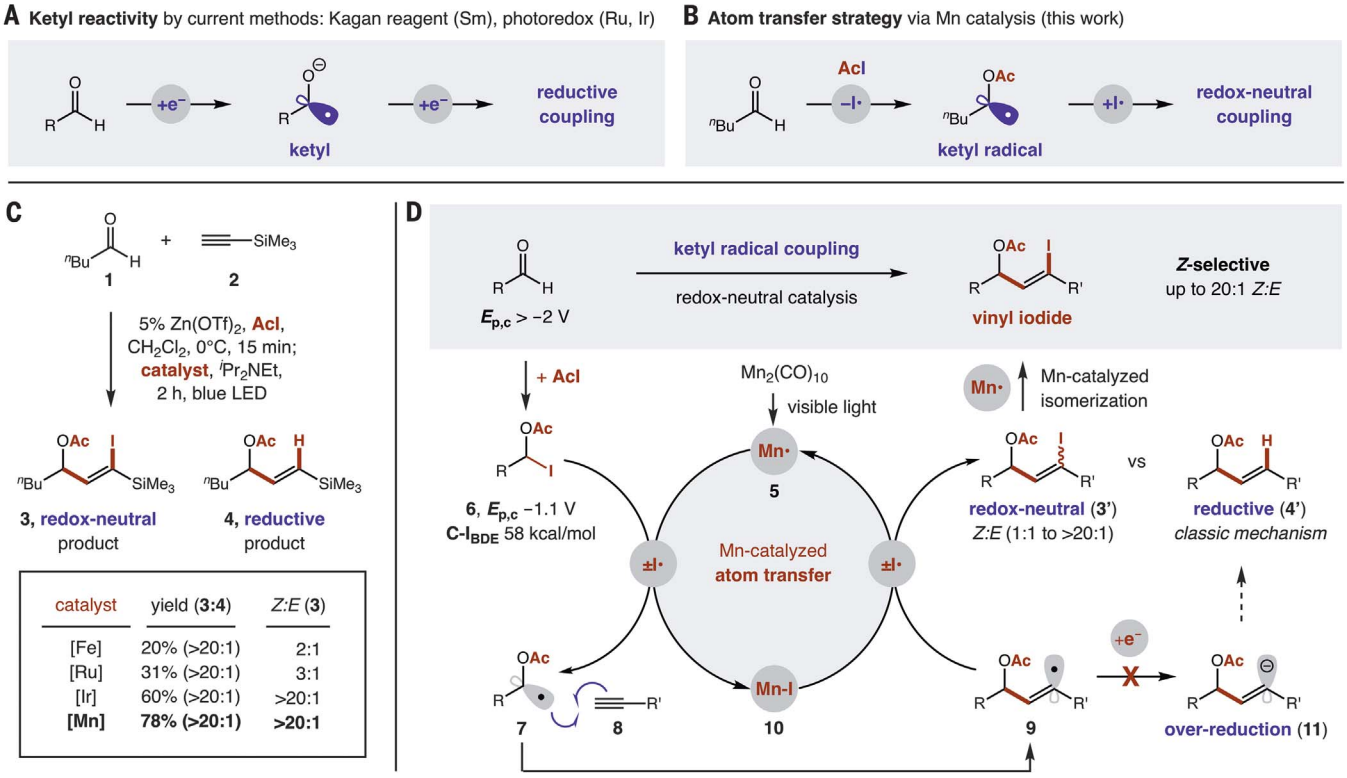


Fig. 1. Discovery of a redox-neutral ketyl radical coupling. (**A**) Current synthetic approaches to ketyl radicals rely on strong reductants that reduce both the carbonyl and subsequent radical intermediates. (**B**) An alternate strategy based on atom transfer of an α-acetoxy iodide enables milder reductive initiation, coupled with oxidative termination, to provide a redox-neutral ketyl coupling mechanism. (**C**) Several catalysts enable this atom transfer strategy, including $[CpFe(CO)_2]_2$, $Ru(bpy)_3Cl_2$, $Ir(ppy)_2(dtbbpy)PF_6$, and $Mn_2(CO)_{10}$, which provide a range of chemo- and stereoselectivities. (**D**) Proposed mechanism: Photoinitiated homolysis of $Mn_2(CO)_{10}$ occurs with visible light (blue LED). The $Mn(CO)_5$ catalyst then formally reduces the in situ formed α-acetoxy iodide via inner-sphere atom transfer. The resulting ketyl radical combines with an alkyne to form a vinyl radical, which is oxidized by the Mn catalyst to form the vinyl iodide product and regenerate $Mn(CO)_5$. This catalyst turnover step precludes an alternate reductive pathway, wherein a second reduction of the vinyl radical affords the classic, non-iodo product. ^{n}Bu, *n*-butyl; Me, methyl; ^{i}Pr, isopropyl; Et, ethyl; Ac, acetyl; $Zn(OTf)_2$, zinc trifluoromethylsulfonate; Cp, cyclopentadienyl; bpy, 2,2′-bipyridine; ppy, 2-phenylpyridine; dtbbpy, 4,4′-di-*tert*-butyl-2,2′-bipyridyl.

iodides in up to 20:1 diastereomeric ratio. For example, alkynes with highly electron-releasing substituents, such as $SiEt_3$ or BPin, are competent partners (**12**, **13**), affording versatile vinyl silanes or boronates (*37*, *38*). Aryl alkynes and 1,3-enynes afford valuable styrene and diene adducts (**14**, **15**), and electron-deficient propiolates also undergo ketyl coupling to merge these electrophiles (**16**, **17**).

In addition to conjugated alkynes, we found propargyl esters to be suitable ketyl radical acceptors that afford *Z*-vinyl iodide **18**, which contains orthogonal allyl esters. An internal competition between alkynes (alkyl- versus ester-substituted) selectively converts the propiolate alkyne to vinyl iodide **19**, leaving the unsubstituted alkyne intact. The mild, radical addition conditions are also tolerant of alcohols (**20**) and ketones (**21**, **22**), illustrating the orthogonality and synthetic utility of this ATRA-based strategy.

To further demonstrate the functional group compatibility of this ketyl-alkyne coupling, we also explored a wide range of aliphatic aldehydes. Aldehyde-derived ketyl radical precursors with both small (**23**) and large (**24**) steric footprints provide *Z*-vinyl iodides in >20:1 selectivity. The tolerated functionality on the aldehyde component includes arenes of varying electronic character (**25**–**27**), ethers (**28**–**29**), and amides (**30**–**32**). Alkyl halides (e.g., Cl, **33**; I, **34**) are also tolerated in the reaction, illustrating selective reaction of the Mn catalyst with the hyperconjugatively activated α-acetoxy iodides.

Finally, although some ketone-derived ketyl radicals are too hindered for efficient cross-coupling, we found that trifluoroacetone provides efficient access to *Z*-vinyl iodide **35** bearing an adjacent tertiary ester. Electron-rich alkenes are also capable of ketyl radical coupling, including those substituted with Si (**36**) or B (**37**). An antioxidant, vitamin E, was also incorporated in the alkyne acceptor (**38**) without inhibiting this mild ketyl radical coupling.

Our mechanistic hypothesis for the redox-neutral catalytic reaction described above is supported by reaction intermediate isolation and resubjection, competition experiments, and kinetic measurements (Fig. 3). To probe the origin of the high *Z*:*E* diastereoselectivity, we monitored product formation of vinyl iodides **12** and **16** over the reaction course (Fig. 3A). In each case, we observed rapid product formation (>70% yield) within 15 min, albeit with modest *Z*:*E* selectivity (7:1, 1:1). For the first 2 hours, product yields remained within 5 to 10% of their 15-min values, but *Z*:*E* selectivity increased markedly (>20:1, 14:1), suggesting a product isomerization mechanism. Additional evidence of post-reaction epimerization was found when a 1:1 mixture of vinyl iodide **16** was resubjected to the Mn photocatalyst for 2 hours, affording the *Z*-isomer selectively: 4:1 (10% catalyst) or >20:1 (20% catalyst) (Fig. 3B).

Further insights were obtained for each elementary step of the atom transfer radical addition via the mechanistic experiments shown in Fig. 3. First, a competition experiment between an electron-poor and an electron-rich alkyne resulted in exclusive ketyl radical coupling to the electron-deficient propiolate alkyne (**16** versus **12**), indicating that the α-OAc radical has nucleophilic character (Fig. 3C). Next, we validated the viability of the key oxidation step required for catalyst turnover by stoichiometric trapping of an aryl radical with

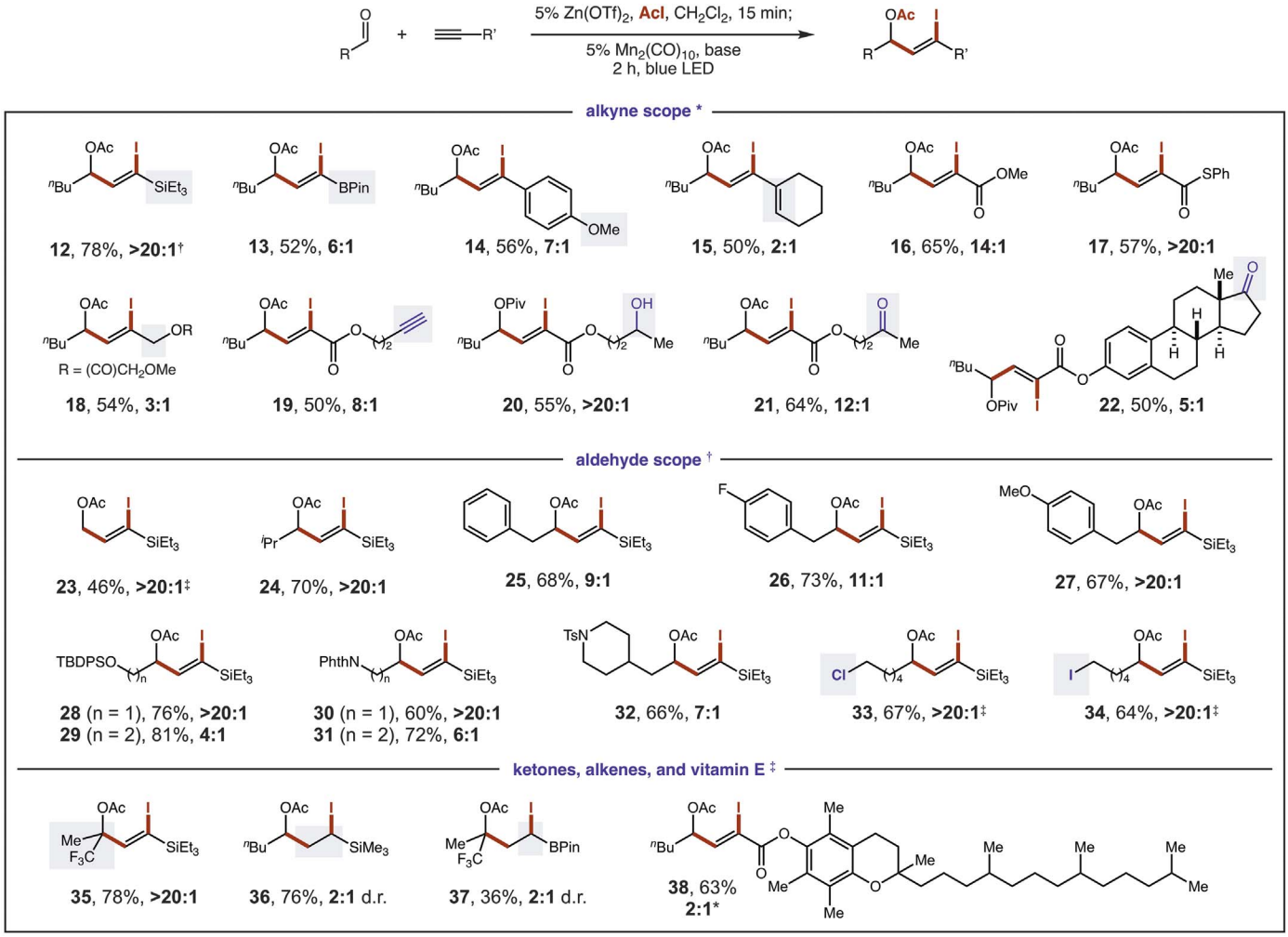


Fig. 2. Synthetic scope of redox-neutral ketyl radical coupling. See supplementary materials for experimental details. Isolated yield and *Z*:*E* ratio are indicated below each entry. *20% $Mn_2(CO)_{10}$, 1 equiv KOAc. †5% $Mn_2(CO)_{10}$, 0.5 equiv $^iPr_2NEt$. ‡10% $Mn_2(CO)_{10}$. Abbreviations: Ac, acyl; Pin, pinacol; Piv, pivaloyl; TBDPS, *tert*-butyldiphenylsilyl; Phth, phthalyl; Ts, toluenesulfonyl; BPin, bis(pinacolato)diboron.

[Mn]-I to form Ar-I (Fig. 3D). The sp^2 C• radical precursor, diazonium **39**, affords aryl iodide **40** in 70% yield when combined with $Mn(CO)_5I$ and eosin Y photocatalyst (3% without light, 68% with only thermal initiation). Next, the rate of atom transfer was compared to known radical clocks (*39*) appended to propiolate (Fig. 3E). For each of the three intramolecular traps investigated, only atom transfer (i.e., vinyl C-I termination, **41**–**43**) was observed, in lieu of radical cyclization, which occurs rapidly with these acceptors (>1×10^8 s^{-1}). Finally, cascade reactions were facilitated by incorporation of intramolecular traps within the ketyl radical precursors (Fig. 3F). Unlike the radical clock experiments that did not manifest cyclization, these ring closures occurred selectively, albeit with divergent outcomes dependent on the hybridization of the coupled radical intermediate. For example, alkyne **44**, which forms an sp^2 C•, selectively traps iodine, affording two-component coupling product **45**, even in the presence of alkyne traps (e.g., $-CO_2Et$, $-SiEt_3$). Alternatively, an sp^3 C• formed by ketyl radical cyclization of alkene **46** selectively combines with Et_3Si-alkyne to afford three-component coupling product **47**. In this case, intramolecular sp^3 to sp^3 translocation of the carbon radicals (α-OAc to α-CO_2Et) precedes intermolecular trapping of the alkyne. The resultant sp^3 C• is long-lived enough (or exists as a living radical) (*40*) to then combine with an alkyne, forming an sp^2 C• that is rapidly trapped as the vinyl iodide.

Finally, to demonstrate the synthetic utility of these functionally rich ketyl-alkyne adducts (bearing a geminal vinyl iodide/silane), we manipulated the atom transfer product **3**, as shown in Fig. 4. First, Pd-catalyzed cross-coupling of the vinyl iodide with a range of boronic acids afforded alkylation (**48**), vinylation (**49**), and arylation (**50**–**52**) products with complete stereoretention. Next, the resulting *E*-vinyl silane **48** was converted to a *Z*-vinyl iodide with *N*-iodosuccinimide (NIS), and subsequent arylation with $PhB(OH)_2$ afforded **53** via a modular, five-component coupling. We expect that this atom transfer method for accessing ketyl radicals—and selectively combining them with an electronically diverse range of π-acceptors—through the use of earth-abundant metal catalysts will have broad utility in various synthetic arenas.

REFERENCES AND NOTES

1. D. J. Hart, *Science* **223**, 883–887 (1984).
2. B. E. Kahn, R. D. Rieke, *Chem. Rev.* **88**, 733–745 (1988).
3. J. E. McMurry, *Chem. Rev.* **89**, 1513–1524 (1989).
4. M. Szostak, N. J. Fazakerley, D. Parmar, D. J. Procter, *Chem. Rev.* **114**, 5959–6039 (2014).
5. H. G. Roth, N. A. Romero, D. A. Nicewicz, *Synlett* **27**, 714–723 (2016).
6. All potentials are reported versus SCE (saturated calomel electrode).
7. P. Girard, J. L. Namy, H. B. Kagan, *J. Am. Chem. Soc.* **102**, 2693–2698 (1980).
8. G. A. Molander, C. R. Harris, *Chem. Rev.* **96**, 307–338 (1996).
9. E. J. Corey, G. Z. Zheng, *Tetrahedron Lett.* **38**, 2045–2048 (1997).
10. K. T. Tarantino, P. Liu, R. R. Knowles, *J. Am. Chem. Soc.* **135**, 10022–10025 (2013).
11. L. J. Rono, H. G. Yayla, D. Y. Wang, M. F. Armstrong, R. R. Knowles, *J. Am. Chem. Soc.* **135**, 17735–17738 (2013).
12. M. A. Ischay, M. E. Anzovino, J. Du, T. P. Yoon, *J. Am. Chem. Soc.* **130**, 12886–12887 (2008).
13. J. Du, K. L. Skubi, D. M. Schultz, T. P. Yoon, *Science* **344**, 392–396 (2014).

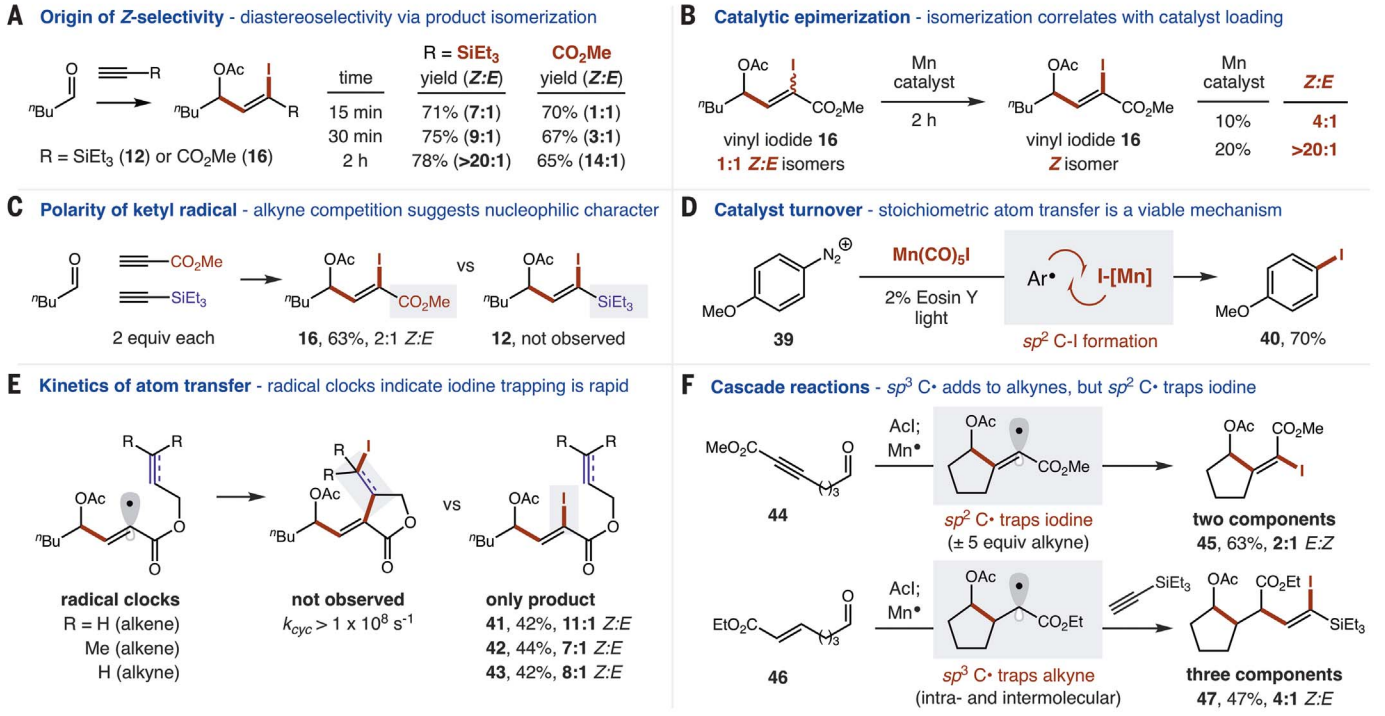
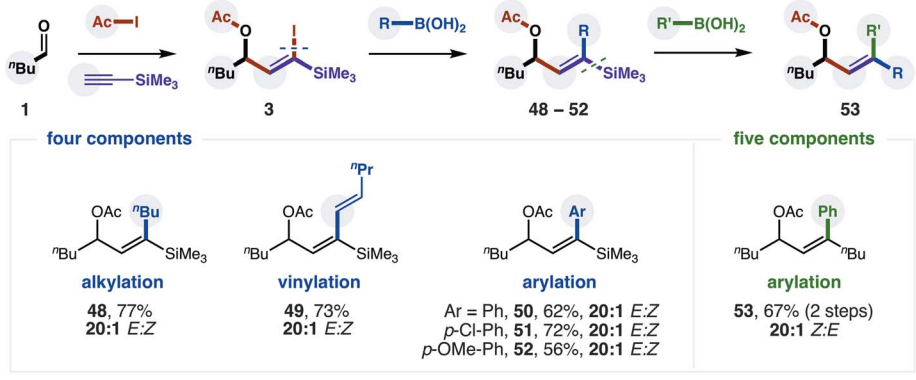


Fig. 3. Mechanistic experiments. (**A**) *Z*:*E* selectivity increases over the reaction course for both alkyne classes, suggesting a product isomerization pathway. (**B**) A 1:1 mixture of vinyl iodide isomerizes to increased ratios of the *Z*-isomer correlating with Mn catalyst loading, illustrating the catalyst's role in isomerization. (**C**) Selective combination with an electron-deficient alkyne indicates that the ketyl radical is nucleophilic. (**D**) The catalyst turnover step involving sp^2 C-I formation is recapitulated by trapping an aryl radical with [Mn]-I to form Ar-I. (**E**) Selective recombination of the vinyl radical with I• (versus intramolecular traps) suggests that this step is rapid. (**F**) Intramolecular traps with sp^2 or sp^3 C• intermediates afford either two- or three-component cascade couplings.

Fig. 4. Synthetic utility of the atom transfer product. Silyl vinyl iodide **3** was coupled to alkyl, vinyl, and aryl boronic acids with 10% $Pd(PPh_3)_2Cl_2$ in a net four-component coupling. The resulting *E*-vinyl silane was converted to a *Z*-vinyl iodide with NIS, then also coupled with phenyl boronic acid in a net five-component coupling.



14. K. N. Lee, Z. Lei, M.-Y. Ngai, *J. Am. Chem. Soc.* **139**, 5003–5006 (2017).
15. C.-X. Ye *et al.*, *Nat. Commun.* **9**, 410 (2018).
16. K. D. Nguyen *et al.*, *Science* **354**, aah5133 (2016).
17. J. Montgomery, *Angew. Chem. Int. Ed.* **43**, 3890–3908 (2004).
18. S. Z. Tasker, E. A. Standley, T. F. Jamison, *Nature* **509**, 299–309 (2014).
19. T. M. Williams, C. R. J. Stephenson, in *Visible Light Photocatalysis in Organic Chemistry* (Wiley-Blackwell, 2018), pp. 73–92.
20. J. C. Theriot *et al.*, *Science* **352**, 1082–1086 (2016).
21. H. Renata, Q. Zhou, P. S. Baran, *Science* **339**, 59–63 (2013).
22. H. French, R. Adams, *J. Am. Chem. Soc.* **43**, 651–659 (1921).
23. J. D. Nguyen, E. M. D'Amato, J. M. R. Narayanam, C. R. J. Stephenson, *Nat. Chem.* **4**, 854–859 (2012).
24. C. J. Wallentin, J. D. Nguyen, P. Finkbeiner, C. R. J. Stephenson, *J. Am. Chem. Soc.* **134**, 8875–8884 (2012).
25. B. M. Monks, S. P. Cook, *Angew. Chem. Int. Ed.* **52**, 14214–14218 (2013).
26. Y. Shen, J. Cornella, F. Juliá-Hernández, R. Martin, *ACS Catal.* **7**, 409–412 (2017).
27. M. J. Koh, T. T. Nguyen, H. Zhang, R. R. Schrock, A. H. Hoveyda, *Nature* **531**, 459–465 (2016).
28. B. B. Snider, *Chem. Rev.* **96**, 339–364 (1996).
29. T. J. Meyer, J. V. Caspar, *Chem. Rev.* **85**, 187–218 (1985).
30. M. C. Baird, *Chem. Rev.* **88**, 1217–1227 (1988).
31. See table S4 for calculated BDEs.
32. Radical chain propagation could also provide **3'**; however, this pathway is unlikely, as initiation by 50% Et_3B/O_2 only affords 29% product. Moreover, competition experiments between α-OAc bromides and α-OAc iodides do not provide crossover products. See supplementary materials for full experimental details.
33. Despite the larger size of iodine (and higher priority in *E*/*Z* notation), its longer C-I bond (and smaller A-value) likely contributes to the thermodynamic favorability of the *Z*-vinyl iodide.
34. S. E. Creutz, K. J. Lotito, G. C. Fu, J. C. Peters, *Science* **338**, 647–651 (2012).
35. I. Ghosh, T. Ghosh, J. I. Bardagi, B. König, *Science* **346**, 725–728 (2014).
36. C. P. Andrieux, J. Pinson, *J. Am. Chem. Soc.* **125**, 14801–14806 (2003).
37. L. Zhang *et al.*, *Science* **351**, 70–74 (2016).
38. M. Kischkewitz, K. Okamoto, C. Mück-Lichtenfeld, A. Studer, *Science* **355**, 936–938 (2017).
39. A. L. J. Beckwith, D. M. O'Shea, *Tetrahedron Lett.* **27**, 4525–4528 (1986).
40. K. Koumura, K. Satoh, M. Kamigaito, *Macromolecules* **41**, 7359–7367 (2008).

ACKNOWLEDGMENTS

We thank M. He and A. Chen for assistance with cyclic voltammetry experiments and density functional theory studies. **Funding:** Supported by NSF CAREER award 1654656, NIH grant R35 GM119812, and the American Chemical Society Petroleum Research Fund. S.C.F. is grateful for an HHMI Gilliam Fellowship. **Author contributions:** L.W. and D.A.N. designed this strategy. All authors contributed to designing, performing, and analyzing experiments, as well as to the writing of this manuscript. **Competing interests:** Authors declare no competing interests. **Data and materials availability:** All data are available in the main text or the supplementary materials.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/362/6411/225/suppl/DC1

Figs. S1 to S9
Tables S1 to S4
NMR Spectra
References (*41*–*60*)

15 May 2018; accepted 13 August 2018
10.1126/science.aau1777

MAGNETISM

Electrical generation and detection of spin waves in a quantum Hall ferromagnet

Di S. Wei[1], Toeno van der Sar[2]*, Seung Hwan Lee[2], Kenji Watanabe[3], Takashi Taniguchi[3], Bertrand I. Halperin[2], Amir Yacoby[1,2]†

Spin waves are collective excitations of magnetic systems. An attractive setting for studying long-lived spin-wave physics is the quantum Hall (QH) ferromagnet, which forms spontaneously in clean two-dimensional electron systems at low temperature and in a perpendicular magnetic field. We used out-of-equilibrium occupation of QH edge channels in graphene to excite and detect spin waves in magnetically ordered QH states. Our experiments provide direct evidence for long-distance spin-wave propagation through different ferromagnetic phases in the *N* = 0 Landau level, as well as across the insulating canted antiferromagnetic phase. Our results will enable experimental investigation of the fundamental magnetic properties of these exotic two-dimensional electron systems.

Quantum Hall (QH) ferromagnetism arises from the interaction of electrons in massively degenerate, quantized energy levels known as Landau levels (LLs) (*1*). When disorder is low enough for Coulomb interactions to manifest, the electrons in partially filled LLs spin-polarize spontaneously to minimize their exchange energy, with the single-particle Zeeman effect dictating their polarization axis (*2*, *3*). In graphene, these phenomena give rise to ferromagnetic phases when the $N = 0$ LL is at one-quarter and three-quarters filling (*4*–*8*). Such QH ferromagnets have an insulating topological bulk and spin-polarized edge states. Furthermore, a canted antiferromagnetic (CAF) state is believed to emerge at one-half filling, with a canting angle determined by the competing valley anisotropy and Zeeman energy (*9*, *10*). Spin waves, also known as magnons, are the lowest-energy excitation in both the QH ferromagnet and the CAF state (*1*, *11*, *12*) and could provide crucial information about these topologically nontrivial magnetic states.

In our experimental setup, we generate magnons by creating an imbalance of chemical potential between two edge states of opposite spin that run along the boundary of a QH magnet. If this imbalance is smaller than the energy required for generating magnons in the QH magnet (and there are no thermal magnons already present in the system), scattering between these two edge states is forbidden because the change in angular momentum of a scattered electron cannot be absorbed by the system. Indeed, previous measurements have shown that oppositely spin-polarized edge channels do not equilibrate as long as the imbalance is small (*13*, *14*). However, we find that edge-channel equilibration commences when the imbalance exceeds the minimum energy required for exciting magnons in the QH ferromagnet. Because the magnetization of the QH ferromagnet is extremely dilute, there are negligible demagnetizing fields and the minimum energy to excite magnons is given by the Zeeman energy $E_Z = g\mu_B B$ (*1*, *15*), where g is the electron g-factor, μ_B is the Bohr magneton, and B is the external magnetic field. Although magnon generation does not directly affect the conductance of the system, the reverse process of magnon absorption by faraway edge states does, allowing us to detect the propagation of magnons electrically, in close analogy to the conventional detection of magnons in insulators via the inverse spin Hall effect (*16*–*19*).

To demonstrate spin-wave propagation, we begin with a dual-gated monolayer graphene device (device 1) where the central region can be tuned to a filling factor different from that of the adjacent regions (Fig. 1A). Connecting the two leads is a chiral edge state that carries spin-polarized electrons aligned with the magnetic

[1]John A. Paulson School of Engineering and Applied Sciences, Harvard University, Cambridge, MA 02138, USA. [2]Department of Physics, Harvard University, Cambridge, MA 02138, USA. [3]National Institute for Materials Science, Tsukuba, Ibaraki 305-0044, Japan.
*Present address: Kavli Institute of Nanoscience, Delft University of Technology, 2628CJ Delft, Netherlands.
†Corresponding author. Email: yacoby@physics.harvard.edu

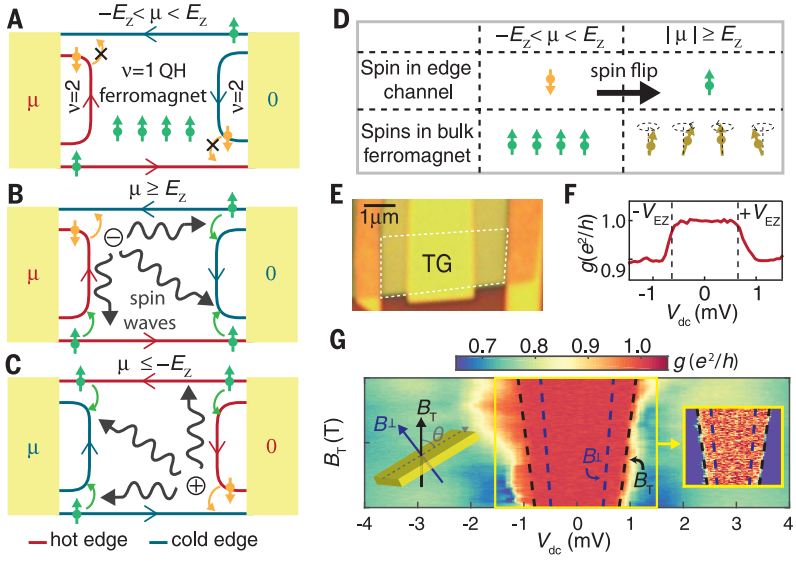


Fig. 1. Magnons in a quantum Hall ferromagnet. (**A** to **C**) A chemical potential difference (μ) is applied between the left and right leads. Edge channels with high and low chemical potential are labeled "hot" and "cold," respectively. Spin-up and spin-down polarization are denoted by the green and orange arrows, respectively. The central region is tuned to $\nu = 1$ and adjacent regions are tuned to $\nu = 2$. (A) The chemical potential difference between the spin-up and spin-down edge channel is less than the Zeeman energy (E_Z), and scattering is suppressed. (B) $\mu \geq E_Z$: Electrons have enough energy to flip their spins and transfer spin angular momentum (magnons) into the bulk (at the encircled minus sign). These magnons are absorbed at distant corners, causing electrons to flip from spin-up into spin-down channels. (C) $\mu \leq -E_Z$: Magnons are generated at the location denoted by the encircled plus sign. (**D**) Bulk spin polarization before and after magnon creation, conserving total spin angular momentum. (**E**) Optical micrograph of device 1; graphene is outlined in white. TG, top gate. (**F**) A dc voltage (V_{dc}) and a 50-μV ac excitation voltage (V_{ac}) are applied to the left contact, and the differential conductance (dI/dV, where $V = V_{ac} + V_{dc}$) is measured through the right contact ($B_\perp = 4$ T, $V_{TG} = -0.18$ V, $V_{BG} = 3$ V). Conductance is quantized to e^2/h until $|\mu| \geq E_Z$. (**G**) dI/dV as a function of bias and magnetic field. The blue dashed line is the Zeeman energy $E_{Z\perp} = g\mu_B B_\perp$ calculated using the perpendicular magnetic field $B_\perp$; the black dashed line is the Zeeman energy $E_{ZT} = g\mu_B B_T$ calculated using the total field B_T. Both the top gate (V_{TG}) and the back gate (V_{BG}) are swept to stay at $\nu = 1$ throughout the device from 7 T ($V_{TG} = 0.16$ V, $V_{BG} = 0.73$ V) to 5 T ($V_{TG} = 0.12$ V, $V_{BG} = 0.44$ V). The decrease in conductance from e^2/h evolves linearly with the magnetic field, coinciding with E_{ZT} rather than $E_{Z\perp}$. Left inset: The sample is fixed at a 45° angle to $B_\perp$. Right inset: A saturated color plot (from 0.98 to 1.02 e^2/h) of the region enclosed by the yellow box. All measurements are conducted in a cryostat with a base temperature of 20 mK.

field, which we call spin-up. We tune the central region to a three-quarters filled LL (ν = 1), whereas the outer regions are tuned to a nonmagnetic fully filled LL (ν = 2). We apply a source-drain voltage V_{dc} to induce a difference in chemical potential $\mu = -eV_{dc}$ between the edge channels emerging from the two contacts, where e is the electron charge. Once $|\mu| \geq E_Z$, an electron traveling in a high-energy ("hot") spin-down edge state can relax into a low-energy ("cold") spin-up edge state by emitting a magnon into the ferromagnetic bulk (Fig. 1, B and C). Because equilibration must occur close to the ferromagnetic bulk in order to launch magnons, the edge states must equilibrate over short length scales at localized "hot spots" where the hot and cold edges meet. This makes graphene an ideal platform to observe this phenomenon, because edge state equilibration in graphene can occur over length scales of <1 μm (*13*, *20*, *21*) [see (*22*) for further discussion]. Because only spin-down angular momentum can be propagated into the spin-up bulk, magnon generation occurs at the location denoted by an encircled minus sign when $\mu \geq E_Z$ (Fig. 1B) and at the location denoted by an encircled plus sign when $\mu \leq -E_Z$ (Fig. 1C). These magnons propagate through the insulating QH ferromagnet and can be absorbed by the reverse process between other edge channels (Fig. 1, B and C), which causes a deviation in the conductance from a well-quantized ν = 1 QH state.

When we measure the conductance of the graphene device (Fig. 1E; atomic force microscopy image in fig. S3) as a function of V_{dc}, we find that the ν = 1 QH ferromagnet remains precisely quantized at the expected value of e^2/h, and then changes once the applied bias reaches the Zeeman threshold ($V_{dc} = \pm V_{EZ} = \mp E_Z/e$), as expected from our model (Fig. 1F). Interestingly, we find that thanks to contact doping (*22*, *23*) we can tune the entire device to ν = 1 and find the same phenomenon of conductance deviation at the Zeeman threshold (*22*) (fig. S4).

By tilting the external magnetic field with respect to the sample-plane normal axis, we verify that the change in conductance occurs when the applied chemical potential exceeds the bare Zeeman energy $E_Z = g\mu_B B_T$ ($g = 2$), which is given by the total field B_T (Fig. 1G; sample is tuned entirely to ν = 1). In contrast, previous transport studies of spin and valley excitations in graphene and GaAs have only found excitations related to the exchange energy gap (*2*, *3*, *24*), which depends on the component of the field perpendicular to the sample plane ($B_\perp$). Our tilted-field measurements therefore corroborate our magnon-based interpretation of the observed change in sample conductance. All further experiments described in this work are done at perpendicular field.

The conductance change at E_Z can either be positive or negative, depending on the number of magnons absorbed at each contact. To examine this, we use different sets of leads in the same device (Fig. 2A, device 2) to perform two-terminal conductance measurements. We start with leads L_2 and L_1 in Fig. 2B. We label the amount of redistributed chemical potential at each of the absorption sites ε_i, with i indexing the absorption site (note that $\varepsilon_i = 0$ for $-E_Z < \mu < +E_Z$), where ε_i is proportional to the number of magnons absorbed at site i. Absorption at ε_1 and absorption at ε_2 have opposite effects on the conductance, as magnon absorption transfers chemical potential from the outer edge to the inner edge. Therefore, for $\mu \geq E_Z$, magnon absorption at ε_1 decreases the particle current ($I_P = -I/e$, where I is the charge current), whereas magnon absorption at ε_2 increases I_P (Fig. 2B). For $\mu \leq -E_Z$, the hot and cold reservoirs are reversed, and we now consider the change to the negative particle current $-I_P$. Although ε_1 still decreases the particle current, I_P is now negative, and so ε_1 actually increases the magnitude of the particle current ($|-I_P|$); similarly, for $\mu \leq -E_Z$, ε_2 decreases $|-I_P|$ (Fig. 2C). We can quantify this using current conservation to formulate the differential conductance as a function of ε_i and μ:

$$\frac{dI}{dV} = \frac{dI_P}{d\mu} = \frac{1}{R_Q}\left(1 + \frac{d\varepsilon_2}{d\mu} - \frac{d\varepsilon_1}{d\mu}\right) \quad (1)$$

where $R_Q = h/e^2$ is the resistance quantum, $V = V_{ac} + V_{dc}$, and we neglect contact resistance [see (*22*) for a derivation that takes contact resistance into account]. We find that the conductance decreases at negative bias and increases at positive bias (Fig. 2D), indicating that $\varepsilon_1 > \varepsilon_2$ for both positive and negative bias. This implies that more magnons are absorbed at ε_1 than at ε_2. Because our contacts have all been fabricated identically, we conclude that this is because ε_1 is closer to magnon generation than ε_2 (for both positive and negative bias; Fig. 2, A and B). Using different sets

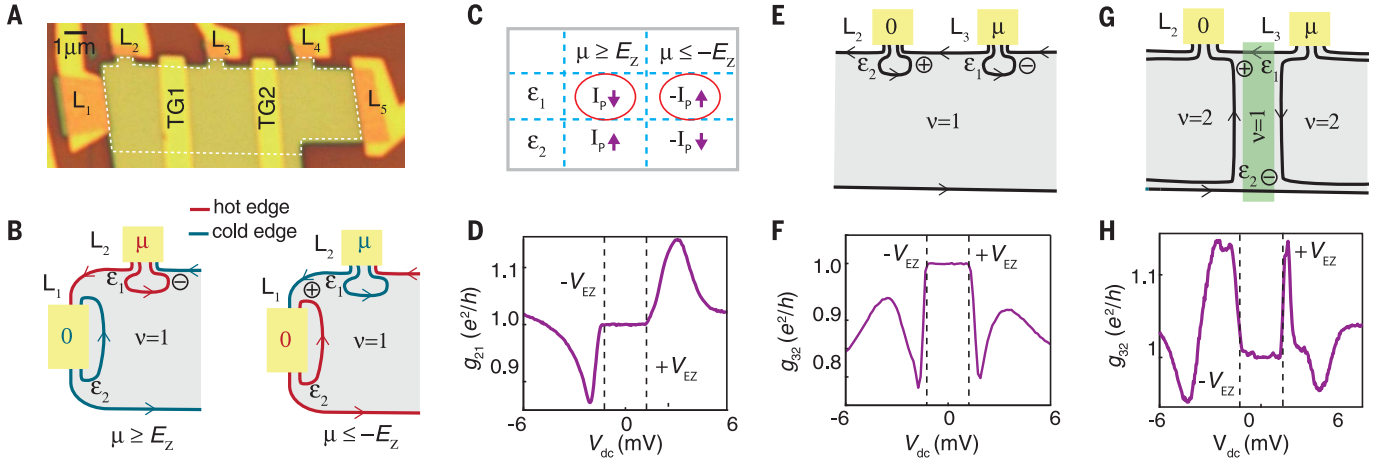


Fig. 2. Effects of relative magnon absorption on conductance. (**A**) Optical micrograph of device 2. Graphene is outlined in white. (**B**) Schematic of a two-terminal conductance measurement using leads L_2 and L_1 where hot and cold edges are colored red and blue, respectively, for both $\mu \geq E_Z$ (left) and $\mu \leq -E_Z$ (right), and the magnon generation site is labeled by the encircled plus or minus sign indicating positive or negative bias. $\mu \geq E_Z$: Magnon absorption at ε_1 transfers chemical potential from a forward-moving edge to a backward-moving edge, causing the particle current ($I_P = -I/e$) to decrease. Conversely, magnon absorption at ε_2 transfers chemical potential from a backward-moving edge to a forward-moving edge, increasing I_P. $\mu \leq -E_Z$: Magnon absorption at ε_1 causes an increase in $|-I_P|$; absorption at ε_2 causes a decrease in $|-I_P|$. (**C**) The effects of ε_1 and ε_2 at $\mu \geq E_Z$ and $\mu \leq -E_Z$. The current changes caused by ε_1 are dominant and are circled in red. The purple arrows indicate an increase (up) or decrease (down) in the magnitude of the signed particle current. (**D**) Conductance from L_2 to L_1 ($g_{21} = dI/dV = dI_P/d\mu$) decreases at $V_{dc} = -V_{EZ}$ and increases at $V_{dc} = V_{EZ}$, indicating that ε_1 has a larger effect than ε_2 (B = 8 T, V_{BG} = 4 V). See (*22*) for full circuit analysis. (**E** and **F**) Conductance from L_3 to L_2 (g_{32}) where the entire device is tuned to ν = 1 (V_{BG} = 4 V; TG1 = 0 V is not shown). At positive bias, $\varepsilon_2 > \varepsilon_1$, and at negative bias, $\varepsilon_1 > \varepsilon_2$, resulting in a conductance drop for both biases. (**G** and **H**) Conductance from L_3 to L_2 (g_{32}) where TG1 is tuned to $\nu_{TG1} = 1$ (TG1 = −0.36 V), whereas the regions outside are set to $\nu_{BG} = 2$ (V_{BG} = 6.5 V). At positive bias, $\varepsilon_1 > \varepsilon_2$, and at negative bias, $\varepsilon_2 > \varepsilon_1$, resulting in a conductance rise for both biases. See fig. S5 for a detailed analysis.

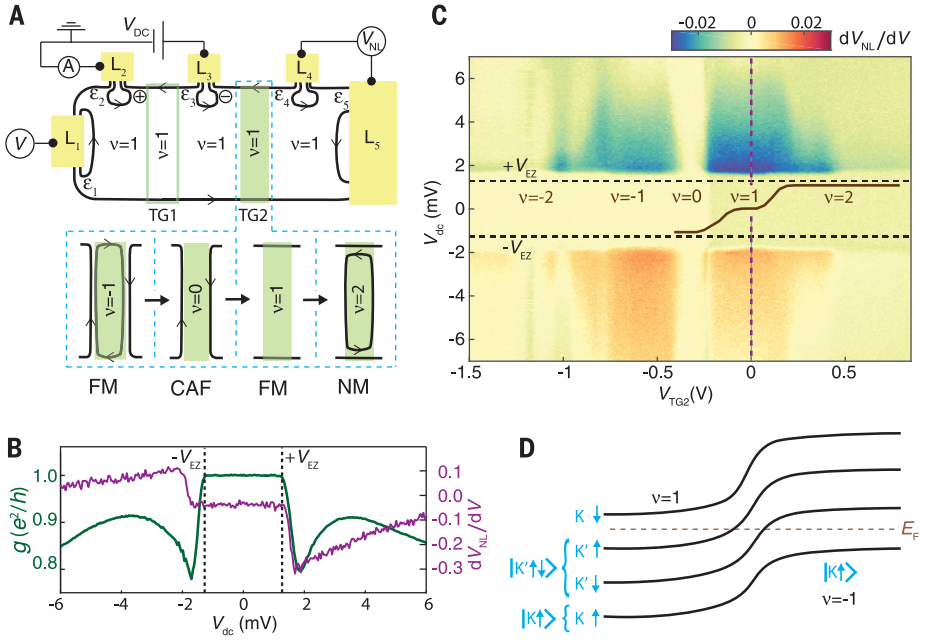


Fig. 3. Nonlocal voltage signal due to magnon absorption. Shown are the data from device 2. (**A**) Schematic circuit configuration for measuring a nonlocal voltage in device 2. The filling factor under TG1 (ν_{TG1}) is 1 for all measurements, whereas the filling factor under TG2 (ν_{TG2}) is swept from −2 to 2, and the rest of the device is kept at ν_{BG} = 1 (V_{BG} = 4 V). The bottom panel highlights the magnetic properties of different cases of ν_{TG2}: nonmagnetic (NM), ferromagnetic (FM), or canted antiferromagnetic (CAF). (**B**) S_{NL} (purple) superimposed onto d*I*/d*V* (green) as a function of V_{dc} when ν_{TG2} = 1 (B = 8 T). The onset of S_{NL} is slightly offset from the decrease in conductance, indicating that magnon generation needs to reach a threshold before being absorbed in distant contacts. (**C**) A pronounced S_{NL} signal when ν_{TG2} = 1 and ν_{TG2} = −1 (see fig. S8 for similar measurements using TG1). Tuning TG2 to the nonmagnetic QH phases (ν_{TG2} = 2 and ν_{TG2} = −2), as well as the ν_{TG2} = 0 CAF state, strongly suppresses S_{NL}. There is a small finite background S_{NL} when edge states pass through TG2, discussed in fig. S7B. Solid brown curve indicates where ν_{TG2} = 0, 1, and 2 (fig. S7, C and D). (**D**) The spatial variation of the LLs at a ν = 1 / ν = −1 junction, with the expected valley and spin polarizations of each level labeled.

of contacts and top gates (Fig. 2, E to H), we can change the relative distances of ε_i to the locations of magnon generation. We confirm that for each configuration, the conductance values after E_Z correspond to a greater number of magnons absorbed at the site closer to magnon generation.

This change to the conductance is not a consequence of QH breakdown. Conductance deviations after the Zeeman threshold that depend on the sign of V_{dc} are not explained by any current breakdown theories (*25*). Additionally, we find that the threshold voltage bias does not depend on the lead configuration (Fig. 2), the size of the ν = 1 region (fig. S4), or the density of the ν = 1 region (fig. S6). All of these observations are inconsistent with trivial QH breakdown but are consistent with our magnon model. In total, we have measured this ν = 1 conductance deviation occurring at the Zeeman energy for eight devices of widely varying geometries (figs. S3, S4, and S11).

Thus far, we have established that we are able to generate and absorb magnons at current carrying contacts. If these chargeless excitations propagate through the insulating bulk, we also expect to see signatures of magnon propagation and absorption via nonlocal voltage measurements (dV_{NL}/dV, referred to as the nonlocal signal S_{NL}), away from the source-drain current. To measure S_{NL}, we use L_3 and L_2 in device 2 as source-drain contacts, and use contacts L_4 and L_5 as voltage probes (Fig. 3A). These contacts are separated from the source-drain contacts by a top gate (TG2), which we tune between ν_{TG2} = −2 and ν_{TG2} = 2, whereas all other regions are tuned to ν = 1. The conductance g between L_3 and L_2 drops at $\pm V_{EZ}$ in accordance with our model (Fig. 3B), whereas magnon generation is largely unaffected by TG2 (fig. S7A). At ν_{TG2} = 1, we measure a change in S_{NL} at $\pm V_{EZ}$ owing to the relative absorption at each magnon absorption site (ε_i).

The sign of S_{NL} indicates that there is more magnon absorption at sites closer to where magnon generation occurs. Through current conservation (*22*), we find that the measured differential voltage (unitless) is

$$\frac{dV_{NL}}{dV} = \left(\frac{d\varepsilon_4}{d\mu} - \frac{d\varepsilon_5}{d\mu}\right) \quad (2)$$

The site labeled by ε_4 is closer to magnon generation than ε_5 for both negative and positive bias, so $d\varepsilon_4 > d\varepsilon_5$. However, the differential change in voltage ($d\varepsilon_i/d\mu$) is negative for $V_{dc} \geq V_{EZ}$ and positive for $V_{dc} \leq -V_{EZ}$, corresponding to an overall negative value for S_{NL} at $V_{dc} \geq V_{EZ}$ and a positive value at $V_{dc} \leq -V_{EZ}$ (Fig. 3C).

The device geometry used for our nonlocal measurements allows us to tune TG2 away from ν_{TG2} = 1, and thereby examine magnon transmission through different filling factors. We observe that when ν_{TG2} = −1, the signal S_{NL} is almost identical to when ν_{TG2} = 1 (Fig. 3C and fig. S8). This signal arises in the absence of any charge leakage across the ν_{TG2} = −1 region (fig. S9), so that changes in S_{NL} can be attributed to magnon transport through the ν_{TG2} = −1 ferromagnet. This suggests that there is neither spin nor valley mismatch between the ferromagnetic states on either side of the boundary. We therefore propose an ordering of the LLs that does not require a spin or valley flip for magnons to travel across the interface between ν_{BG} = 1 and ν_{TG2} = −1 [Fig. 3D; see (*22*) for a theoretical discussion].

In addition, we unexpectedly find that S_{NL} is suppressed at $\pm V_{EZ}$ when ν_{TG2} = 0. For nonmagnetic regions such as ν_{TG2} = 2, it is expected that magnons will be blocked from passing through, as experimentally confirmed in Fig. 3C (the nonlocal signal occurring at the transition between ν = 1 and ν = 2 is explained in fig. S7E). However, ν = 0 is purportedly a CAF, which is theoretically capable of hosting even zero-energy magnons (*12*). It appears that the probability for an incident magnon to be transmitted across the junction between the ν = 0 and ν = 1 regions is very small for energies close to E_Z. This may be caused by, in part, the mismatch in propagation velocities in the two phases, or a barrier due to the complex nature of the interface region. Close to the boundary with a ν = 1 phase, the ground state of the ν = 0 phase may not have canted spins but may instead be in an aligned antiferromagnet state, where spins are parallel to the magnetic field on one sublattice and antiparallel on the other. Eventually, far from the boundary, we may expect the local spin arrangement to rotate into the CAF orientation (Fig. 4B). In the transition region, the minimum magnon energy will be larger than E_Z, thanks to effects of the valley-dependent interaction terms (*10*), which were initially responsible for the antiferromagnet arrangement to be favored over the ferromagnetic arrangement. To cross from the ν = 1 region to the CAF region, a magnon with energy close to E_Z would have to tunnel through the barrier region, and we would expect the transmission rate to be low. If the magnons have enough energy to overcome this barrier, they should be able to more easily enter the CAF region. Figure 4C shows that we can experimentally exceed this barrier, where we see nonlocal signals at higher $|V_{dc}|$ with signs in agreement with our magnon model. The onset of this magnon

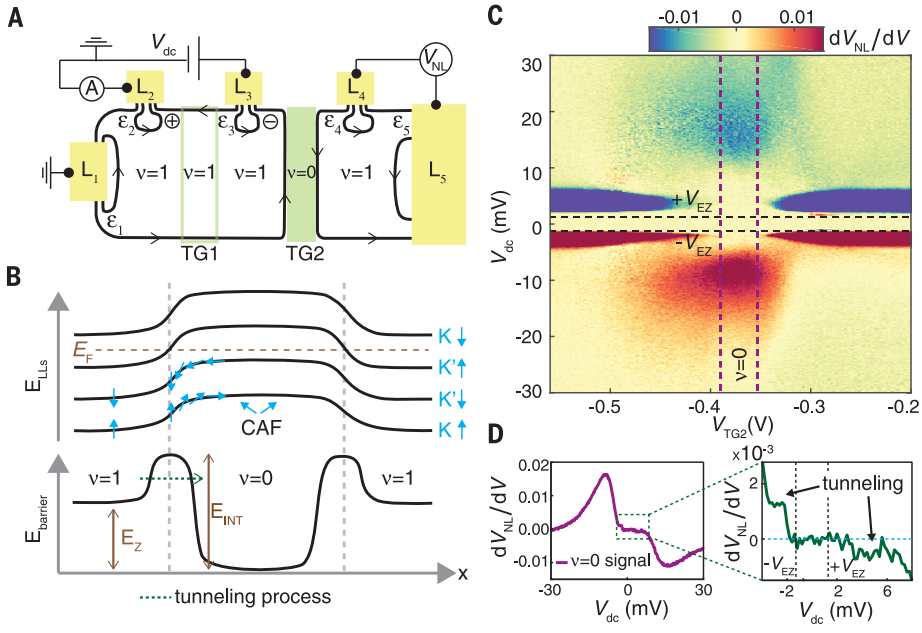


Fig. 4. Nonlocal voltage signal caused by magnon propagation through the ν = 0 CAF. (**A**) Schematic of the circuit used to measure S_{NL} in device 2 across a ν = 0 region; ν_{TG1} = 1 for all measurements, whereas ν_{TG2} is swept from −1 to 1 (ν_{BG} = 1, V_{BG} = 4 V). (**B**) Top: Postulated spatial variation of the LLs and spin arrangement in a ν = 1, ν = 0, ν = 1 geometry. Close to the interface between ν = 1 and ν = 0, spins in the two filled LLs prefer to be in an aligned antiferromagnetic (AF) arrangement. Deeper in the ν = 0 region, the spins rotate into the CAF phase. Because the minimum magnon energy in the aligned AF region is higher than E_Z, it should present a barrier for incident magnons close to the energy threshold. Bottom: Energy barrier seen by the magnons as a function of position, where E_{INT} is the energy barrier of the interface. (**C**) When magnons are generated, we see another onset of S_{NL} at energies exceeding $\pm V_{EZ}$ (B = 8 T), indicating that higher-energy magnons have overcome E_{INT} and have propagated through the ν = 0 region underneath TG2. Purple dashed lines indicate a region where vertical line cuts were taken and averaged to obtain the line trace in (D). (**D**) A clear onset of S_{NL} is shown at biases exceeding $\pm V_{EZ}$ when ν_{TG2} = 0. It is not currently understood why the signal is asymmetric in both energy of onset and strength of signal. The zoomed-in region shows a clear increase in S_{NL} at $-V_{EZ}$ and a signal consistent with a decrease, slightly offset from $+V_{EZ}$, indicating that magnons can tunnel through the interface barrier at lower energies (S_{NL} is offset by 0.01 μV at V_{dc} = 0 and is manually corrected for).

signal is unaffected by any charge transport across the ν_{TG2} = 0 region (fig. S10). Closely examining the signal at ν_{TG2} = 0, we see signals commencing at $\pm V_{EZ}$, which we attribute to tunneling events across this barrier between ν_{BG} = 1 and ν_{TG2} = 0 (Fig. 4D).

Note that all nonlocal signals (occurring at ν_{TG2} = −1, 0, and 1) appear only in a finite band of V_{dc}. This suppression of the differential voltage signal indicates that magnon generation is suppressed, or alternatively, that the differently spaced contacts begin to see identical amounts of magnon absorption once the system has reached a certain magnon density threshold. We further speculate that this cutoff could be related to the magnon bandwidth, but we leave this to a future investigation.

The experiments presented here introduce a method of using magnons to probe the SU(4) spin and valley anisotropies of graphene QH systems, which can be used to probe highly correlated states such as the fractional QH regime (*26*) or the quantum spin Hall phase of monolayer graphene (*9*). Owing to the theoretical prediction for spin superfluidity in the CAF state (*12*), this study paves the way for exploring and realizing dissipationless spin waves in a Bose-Einstein condensate (BEC) of magnons. Such condensates should result in a coherent precession of the spins in the QH magnet, which may be probed through emitted microwave radiation. Furthermore, coherent spin waves associated with a BEC may be able to propagate long distances with negligible dissipation, which could be tested by careful length dependence measurements.

REFERENCES AND NOTES

1. S. M. Girvin, in *Topological Aspects of Low Dimensional Systems*, A. Comtet, T. Jolicœur, S. Ouvry, F. David, Eds. (Springer, 2000), pp. 53–175.
2. A. F. Young *et al.*, *Nat. Phys.* **8**, 550–556 (2012).
3. S. L. Sondhi, A. Karlhede, S. A. Kivelson, E. H. Rezayi, *Phys. Rev. B* **47**, 16419–16426 (1993).
4. J. Alicea, M. P. A. Fisher, *Phys. Rev. B* **74**, 075422 (2006).
5. K. Yang, S. Das Sarma, A. H. MacDonald, *Phys. Rev. B* **74**, 075423 (2006).
6. Y. Zhang *et al.*, *Phys. Rev. Lett.* **96**, 136806 (2006).
7. K. Nomura, A. H. MacDonald, *Phys. Rev. Lett.* **96**, 256602 (2006).
8. M. O. Goerbig, *Rev. Mod. Phys.* **83**, 1193–1243 (2011).
9. A. F. Young *et al.*, *Nature* **505**, 528–532 (2014).
10. M. Kharitonov, *Phys. Rev. B* **85**, 155439 (2012).
11. A. G. Green, N. R. Cooper, *Phys. Rev. B* **65**, 125329 (2002).
12. S. Takei, A. Yacoby, B. I. Halperin, Y. Tserkovnyak, *Phys. Rev. Lett.* **116**, 216801 (2016).
13. F. Amet, J. R. Williams, K. Watanabe, T. Taniguchi, D. Goldhaber-Gordon, *Phys. Rev. Lett.* **112**, 196601 (2014).
14. D. S. Wei *et al.*, *Sci. Adv.* **3**, e1700600 (2017).
15. C. Kittel, *Phys. Rev.* **73**, 155–161 (1948).
16. Y. Kajiwara *et al.*, *Nature* **464**, 262–266 (2010).
17. L. J. Cornelissen, J. Liu, R. A. Duine, J. Ben Youssef, B. J. Van Wees, *Nat. Phys.* **11**, 1022–1026 (2015).
18. A. V. Chumak, V. I. Vasyuchka, A. A. Serga, B. Hillebrands, *Nat. Phys.* **11**, 453–461 (2015).
19. D. Wesenberg, T. Liu, D. Balzar, M. Wu, B. L. Zink, *Nat. Phys.* **13**, 987–993 (2017).
20. J. R. Williams, L. Dicarlo, C. M. Marcus, *Science* **317**, 638–641 (2007).
21. B. Özyilmaz *et al.*, *Phys. Rev. Lett.* **99**, 166804 (2007).
22. See supplementary materials.
23. G. Giovannetti *et al.*, *Phys. Rev. Lett.* **101**, 026803 (2008).
24. A. Schmeller, J. P. Eisenstein, L. N. Pfeiffer, K. W. West, *Phys. Rev. Lett.* **75**, 4290–4293 (1995).
25. G. Nachtwei, *Physica E* **4**, 79–101 (1999).
26. A. H. MacDonald, J. J. Palacios, *Phys. Rev. B* **58**, R10171 (1998).

ACKNOWLEDGMENTS

We thank A. H. Macdonald, J. D. Sanchez-Yamagishi, S. L. Tomarken, and S. P. Harvey for helpful discussions and feedback, X. Liu for fabrication help, and P. Kim for providing the transfer setup. **Funding:** Supported by the Gordon and Betty Moore Foundation's EPiQS Initiative through grant GBMF4531; the U.S. Department of Energy, Basic Energy Sciences Office, Division of Materials Sciences and Engineering under award DE-SC0001819 (D.S.W. and T.v.d.S.); NSF Graduate Research Fellowship grant DGE1144152 (D.S.W.); the STC Center for Integrated Quantum Materials, NSF grant DMR-1231319 (B.I.H.); and the Elemental Strategy Initiative conducted by MEXT, Japan, and JSPS KAKENHI grant JP15K21722 (K.W. and T.T.). Nanofabrication was performed at the Center for Nanoscale Systems at Harvard, supported in part by NSF NNIN award ECS-00335765. **Author contributions:** D.S.W., T.v.d.S., B.I.H., and A.Y. conceived and designed the experiments; D.S.W. fabricated the devices; D.S.W. and A.Y. performed the experiments; D.S.W., T.v.d.S., S.H.L., B.I.H., and A.Y. analyzed the data and wrote the paper; and K.W. and T.T. synthesized the hexagonal boron nitride crystals. **Competing interests:** The authors declare no competing financial interests. **Data and materials availability:** All measured data are available in the supplementary materials.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/362/6411/229/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S11
References (*27*–*39*)
Data Files

5 November 2017; accepted 19 August 2018
10.1126/science.aar4061

PLANT SCIENCE

Systemic control of legume susceptibility to rhizobial infection by a mobile microRNA

Daniela Tsikou[1]*†, Zhe Yan[1]*‡, Dennis B. Holt[1]§, Nikolaj B. Abel[1]¶, Dugald E. Reid[1], Lene H. Madsen[1], Hemal Bhasin[2], Moritz Sexauer[2], Jens Stougaard[1], Katharina Markmann[1,2]#

Nitrogen-fixing root nodules on legumes result from two developmental processes, bacterial infection and nodule organogenesis. To balance symbiosis and plant growth, legume hosts restrict nodule numbers through an inducible autoregulatory process. Here, we present a mechanism where repression of a negative regulator ensures symbiotic susceptibility of uninfected roots of the host *Lotus japonicus*. We show that microRNA miR2111 undergoes shoot-to-root translocation to control rhizobial infection through posttranscriptional regulation of the symbiosis suppressor TOO MUCH LOVE in roots. miR2111 maintains a susceptible default status in uninfected hosts and functions as an activator of symbiosis downstream of LOTUS HISTIDINE KINASE1–mediated cytokinin perception in roots and HYPERNODULATION ABERRANT ROOT FORMATION1, a shoot factor in autoregulation. The miR2111-*TML* node ensures activation of feedback regulation to balance infection and nodulation events.

Development of nitrogen-fixing nodules in roots of legumes such as soybean, pea, and *Lotus japonicus* is induced by rhizobial signal molecules. Nodule numbers are systemically controlled by host plants through a process called autoregulation of nodulation (AON). In the current AON model, the onset of nodulation leads to the transport of root-synthesized *CLE* peptides to *Lotus* shoots, where they are perceived by the CLAVATA1-like leucine-rich repeat receptor–like kinase HYPERNODULATION ABERRANT ROOT FORMATION1 (HAR1) (*1*–*3*). The resulting shoot-to-root signaling was proposed to involve cytokinins as mobile molecules (*4*). Restriction of nodule emergence is effected by TOO MUCH LOVE (TML), a root-active kelch-repeat F-box protein (*5*), by an unknown mechanism. This model assumes that AON is induced by rhizobia. In this study, we have shown that shoot-derived miR2111 acts in the absence of infection to unlock root responsiveness to rhizobial nodulation (Nod) factors, ensuring susceptibility to infection.

Lotus mature miR2111 levels negatively respond to infection with *Mesorhizobium loti* in both leaves and roots within 2 days (Fig. 1A and fig. S1). Screening of the *Lotus* genome (*6*, *7*) revealed three genomic loci representing potential *MIR2111* precursor genes (fig. S2A) encoding three miR2111 isoforms, miR2111a to miR2111c (fig. S2A). The *MIR2111-1* transcript contains a single copy of miR2111a (fig. S2, A and B), whereas *MIR2111-2* and *-3* are polycistronic, containing two isoforms each in consecutive stem-loops (fig. S2, A, C, and D).

To understand the biological relevance of the decline in miR2111 abundance during rhizobial infection, we expressed the *MIR2111-2* and *MIR2111-3* precursor genes under the control of the *LjUBIQUITIN1* (*UBQ1*) promoter. *pUBQ1*::*MIR2111-3*–expressing roots (fig. S3A) were hyperinfected (Fig. 1B). Overexpression of three other miRNAs by the same system had no phenotypic effect (*8*, *9*), making hypernodulation of *pUBQ1*::*MIR2111-3* roots unlikely to reflect indirect effects of miRNA overexpression. No qualitative defects in root hair responses or infection were apparent (fig. S3, B and C). At 21 days postinfection (dpi), *pUBQ1*::*MIR2111-3* roots had more nodules than controls (Fig. 1C). The mature nodules were smaller than those on wild-type roots but were infected and morphologically normal (fig. S3, D to G), and the proportion of immature nodules was increased to 72.8% from 20.7% in control roots. Infection events (fig. S4, A and B) and nodules (Fig. 1, D and E) were present along a wider fraction of *pUBQ1*::*MIR2111-3* roots, suggesting an expanded susceptibility zone where infections can take place. Overexpression of the 5′ stem-loop of *MIR2111-2*, encoding a single copy of miR2111b, suggests a dosage-dependent effect of miR2111 (Fig. 1, B and C, and fig. S3A).

To gain insight into the molecular context of miR2111 activity, we queried a degradome dataset prepared from nodule and root tissues (*10*) for targets of miR2111-directed endonucleolytic cleavage. This revealed *TML* as an in vivo target of miR2111 (Fig. 1F and fig. S5A). miR2111, as well as target sites in putative *TML* orthologs, is conserved in dicotyledonous plants (fig. S5, B and C). *TML* mRNA was rare in *pUBQ1*::*MIR2111-3* roots (fig. S3A), and its abundance increased as miR2111 in roots decreased at 2 to 4 dpi (Fig. 1G and fig. S5D). *MIR2111-3* overexpression roots (Fig. 1, B, C, and E) thus phenocopy *tml* loss-of-function mutants (*5*, *11*) (fig. S5E). *TML* promoter activity is constitutive in *Lotus* roots (*5*), and our data demonstrate miR2111-dependent posttranscriptional regulation of *TML* mRNA abundance. To investigate miR2111 roles in symbiosis control, we overexpressed a short tandem target mimic (STTM) construct to sequester mature miR2111 (*pUBQ1*::*miR2111STTM*). This reduced miR2111 abundance, and *TML* levels were increased (Fig. 1H). In line with a positive effect of miR2111, *pUBQ1*::*miR2111STTM* roots formed fewer nodules than controls (Fig. 1, I and J).

TML mRNA was not detected in leaf tissue (fig. S5F) (*12*). Yet, the *M. loti*–induced miR2111 response was systemic (Fig. 1A). We thus hypothesized that miR2111 might undergo shoot-to-root translocation. To locate miR2111 source tissues, we generated stable *pMIR2111-3*::*GUS*–expressing plants. Four independent lines showed GUS activity in leaf phloem but none in roots (Fig. 2, A to D, and fig. S6, A to F). GUS signal in *pMIR2111-2*::*GUS* lines was below the visual detection limit, but quantitative reverse transcription–polymerase chain reaction (qRT-PCR) detected *GUS* transcripts (fig. S6F). *Agrobacterium rhizogenes*–induced *pMIR2111-1*::*GUS*, *-2*::*GUS*, or *-3*::*GUS* roots showed no detectable GUS signal (fig. S6, G to L) and low *GUS* transcript levels (fig. S6F). These data suggest that whereas mature miR2111 is present in roots, it is synthesized primarily in shoots.

If root miR2111 was of shoot origin, shootless roots should have low miR2111 levels. In line with this hypothesis, 3 days after mechanical separation from shoots, roots showed less miR2111 (Fig. 2E and fig. S7A), whereas induction of the infection-responsive miR172 (*9*) demonstrated that de novo miRNA synthesis was unimpaired (fig. S7B). Consistent with reduced miR2111 content, shootless roots had more *TML* mRNA (Fig. 2E and fig. S7C). Infection led to further increase in *TML* mRNA (fig. S7C), pointing to additional transcriptional regulation of *TML* mRNA abundance. miR2111 observed in shootless roots (fig. S7A) may reflect miRNA stability or low-level local expression. We further identified miR2111 in flow-out liquid collected from cut-off shoots (table S1), suggesting its presence in phloem sap. These shoots contained miR2111 levels similar to those in shoots of intact plants (fig. S7D), suggesting that miR2111 biosynthesis remained active in cut-off shoots. Thus, miR2111 was translocated from shoots to roots in vivo, and shoot-root separation interrupted translocation (fig. S7E).

In uninfected plants, active *TML* transcription (*5*) may be counterbalanced posttranscriptionally by shoot-derived miR2111, ensuring susceptibility. To test whether *TML* loss or miR2111-mediated transcript reduction can alleviate the TML-imposed blockage, we removed shoots from *tml-1* or wild-type roots expressing *pUBQ1*::*MIR2111-3*, respectively. ôSmall but infected nodules were

[1]Department of Molecular Biology and Genetics, Aarhus University, Aarhus, Denmark. [2]Zentrum für Molekularbiologie der Pflanzen, Tübingen University, Tübingen, Germany. *These authors contributed equally to this work. †Present address: Department of Biochemistry and Biotechnology, University of Thessaly, Larissa, Greece. ‡Present address: School of Life Science and Engineering, Southwest University of Science and Technology, Mianyang, China. §Present address: DNA Diagnostic, Risskov, Denmark. ¶Present address: Department of Plant Cell Biology, University of Freiburg, Freiburg, Germany. #Corresponding author. Email: katharina.markmann@zmbp.uni-tuebingen.de

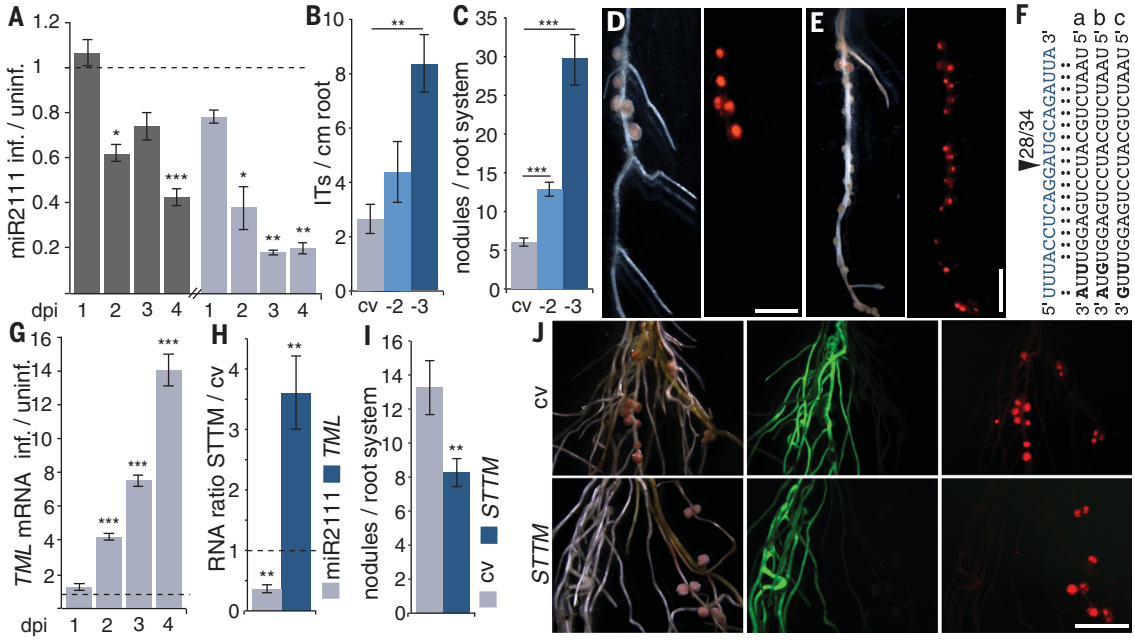


Fig. 1. miR2111 regulates *TML* posttranscriptionally. (**A**) miR2111 abundance in *Lotus* leaves (dark bars) and roots (light bars) at 1 to 4 dpi with *M. loti*. inf., infected; uninf., uninfected. (**B**) Infection thread (IT) (10 dpi) and (**C**) nodule (21 dpi) numbers in *pUBQ1*::*MIR2111-2* (-2) and *pUBQ1*::*MIR2111-3* (-3) roots compared to control (cv) roots. (**D** and **E**) Nodulation in control (D) and *pUBQ1*::*MIR2111-3* (E) roots (21 dpi). Right-hand panels visualize *M. loti DsRED* in nodules. Scale bars, 2 mm. (**F**) miR2111 directs *TML* cleavage. Bold font marks polymorphisms between miR2111 isoforms a to c (black). Numbers: degradome 5′ ends at arrowhead/total within *TML* target region (blue). (**G**) *TML* mRNA in *M. loti*–infected roots (1 to 4 dpi). (**H** to **J**) *miR2111STTM* (*STTM*) expression reduced miR2111, increased *TML* (H), and reduced nodulation [(I) and (J)] compared to those in control roots (cv). (I) $n = 23/26$ (*miR2111STTM* roots/control roots). Green fluorescence [(J), center] shows cotransformation; red [(J), right] indicates nodules with *M. loti DsRED*. Scale bar, 5 mm. Transgenic roots [(B) to (E) and (H) to (J)] were *A. rhizogenes* induced. [(A), (G), and (H)] qRT-PCR analyses. RNA levels are relative to those for two reference genes. Dashed lines mark 1 as the reference value in ratio graphs. Error bars [(A) to (C) and (G) to (I)]: SEM of at least three biological replicates. Student's *t* test *P* values: $*P \le 0.05$; $**P \le 0.01$; $***P \le 0.001$. (I) Results represent one biological replicate ($P = 0.006$) and were similar in a second ($P = 0.001$).

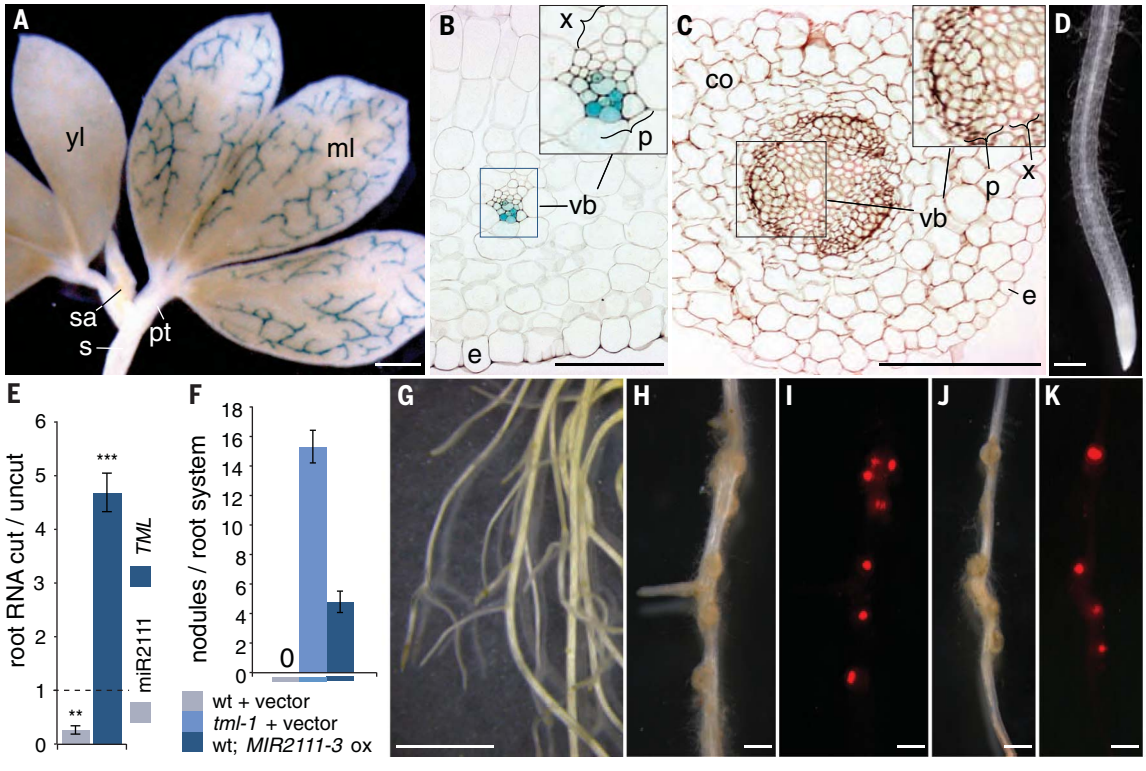


Fig. 2. *MIR2111-3* expression and translocation to roots. (**A** to **D**) GUS activity in *pMIR2111-3*::*GUS*–expressing plants (2 weeks). [(A) and (B)] GUS activity in phloem (p) cells of higher-degree veins of mature leaves (A) and cotyledons (B). No GUS was detected in leaf xylem (x), shoot apices (sa), petioles (pt), and stems (s) (A) or roots [(C) and (D)]. co, cortex; e, epidermis; ml, mature leaflets; vb, vascular bundle; yl, young leaflets. Scale bars, 1 mm [(A) and (D)]; 50 μm [(B) and (C)]. (**E**) qRT-PCR analyses of miR2111 and *TML* levels in uninfected roots following root-shoot separation. Error bars: SEM of three biological replicates. $**P \le 0.01$; $***P \le 0.001$. (**F** to **K**) *TML* loss [(H) and (I)] or *pUBQ1*::*MIR2111-3* expression [(J) and (K)] rescues the asymbiotic phenotype of shootless wild-type roots (G). Nodule counts (F) were at 3 weeks postinoculation with *M. loti* expressing *DsRED* on *A. rhizogenes*–induced wild-type (wt) (G) or *tml-1* [(H) and (I)] roots (expressing control vector) and wild-type roots expressing *pUBQ*::*MIR2111-3* [(J) and (K)]. (F) Error bars show SEM of two biological replicates ($n = 12$, 13, and 11 total root systems, respectively). ox, *pUBQ1*-mediated overexpression. Scale bars, 1 cm (G); 1 mm [(H) to (K)].

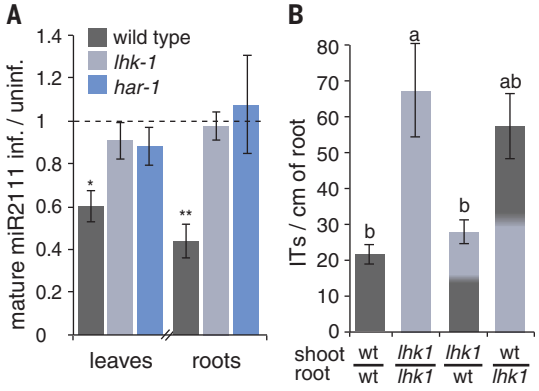


Fig. 3. Systemic regulation of miR2111. (**A**) Infection-dependent reduction of mature miR2111 levels in wild-type leaves and roots depends on *HAR1* and *LHK1*. Plants were harvested uninfected or 3 dpi with *M. loti* and were analyzed by qRT-PCR. (**B**) Infection thread (IT) counts on roots of grafted wild-type, *lhk1-1*, and chimeric plants. n = 3, 5, 6, 8 root systems (left to right). [(A) and (B)] Error bars show SEM of at least three biological replicates. Comparisons used Student's t test (infected versus uninfected; $^{*}P \leq 0.05$, $^{**}P \leq 0.01$) (A) or analysis of variance (ANOVA) and post hoc Tukey testing (B) (P = 0.011), with distinct letters indicating significant differences.

formed in the presence of *M. loti* (Fig. 2, F and H to K), but we never observed nodules on wild-type roots expressing a control vector (Fig. 2, F and G). Together, these results demonstrate that interactions between *TML* and miR2111 control root susceptibility and that shoot-produced miR2111 contributes to regulating root *TML* levels.

Like *TML* loss (*11*) (fig. S8A) or *MIR2111* overexpression (Fig. 1B), knockout of the cytokinin receptor *LHK1* results in hyperinfection (*13*) (fig. S8A). Infected *lhk1-1* mutants maintained miR2111 levels upon infection in both roots and shoots (Fig. 3A and fig. S8, B and C), suggesting a systemic negative effect of *LHK1*-dependent cytokinin signaling on miR2111 accumulation. Grafting revealed root genotype dependence of *lhk1-1*–associated hyperinfection (Fig. 3B). As cytokinin biosynthesis is first detected in the root cortex rather than in the epidermal cells where infection is initiated (*14*), these results suggest a connection between the miR2111-*TML* module and the inducible AON system. The latter involves induction of *CLE-RS1*, *-2*, and *-3* upon infection (*3*, *15*). In 10-day-old plants that retained seed-derived nutrient supplies and were moderately nitrogen sufficient, we found that *lhk1-1* roots showed no infection-dependent *CLE-RS3* and impaired *CLE-RS1* accumulation (fig. S8D). This suggests a role of *LHK1* in inducing autoregulation in response to infection. Supporting this, root expression of selected *CLE-RS* genes is inducible by cytokinin application in an *LHK1/CRE1*-dependent manner (*16*, *17*). Dependence of *CLE* gene expression on *LHK1* places root cytokinin perception upstream of *HAR1* in AON, in line with nodulation inhibition in *lhk1* roots grafted to *CLE-RS1* overexpression shoots (*4*), and does not exclude involvement of shoot-derived cytokinin (*4*) at a later stage. Consistent with *lhk1-1* hyperinfection (*13*) (Fig. 3B and fig. S8A), *TML* mRNA levels are reduced in *lhk1-1* mutants compared to those in the wild type at 3 dpi (fig. S8E).

TML functions downstream of *HAR1* in symbiotic autoregulation. Infection-triggered miR2111 regulation was *HAR1* dependent (Fig. 3A and fig. S8, B and C), and *TML* mRNA levels were reduced in *har1-3* mutant roots compared with the wild type (fig. S8E). Although *har1* mutants are hypernodulating, uninfected *har1-3* plants contained less miR2111 than the wild type (fig. S8, B and C). This suggests that *HAR1* positively affects miR2111 abundance in uninfected plants and that additional factors contribute to *TML* regulation.

Thus, the miR2111-*TML* module operating in uninfected roots connects to the symbiotically triggered AON (fig. S9). This is dependent on LHK1-mediated cytokinin signaling in roots, and CLE-activated shoot HAR1 modifies miR2111 production and/or shoot-root translocation, thereby contributing to the release of root *TML* from posttranscriptional miR2111 control. These interconnected regulatory systems ensure that transcription of *TML* in uninfected roots (*5*) does not prevent symbiosis but restricts its progression immediately upon contact with compatible rhizobial bacteria (fig. S9).

Apart from infection, nitrogen availability also controls symbiosis in a *HAR1*-dependent manner (*18*). In line with this, miR2111 abundance was reduced in plants supplied with nitrate (fig. S10A). Whereas nitrogen homeostasis may induce AON once fixing nodules are present, both infection and organogenesis control set in before nodule-derived nitrogen is available. We found that Nod factor receptor *nfr1-1* (*19*) and *nfr5-1* (*19*) mutants failed to downregulate miR2111 (fig. S10B), suggesting that the release of miR2111-mediated *TML* regulation is triggered by bacterial signaling molecules at the initiation of symbiotic cross talk, resulting in symbiosis control being promptly effective upon infection. In summary, miR2111 is a systemic symbiosis activator constitutively repressing AON in uninfected roots, thereby maintaining susceptibility to rhizobial infection. This mechanism highlights the importance of immediate host control during symbiosis establishment.

REFERENCES AND NOTES

1. L. Krusell *et al.*, *Nature* **420**, 422–426 (2002).
2. I. R. Searle *et al.*, *Science* **299**, 109–112 (2003).
3. S. Okamoto, H. Shinohara, T. Mori, Y. Matsubayashi, M. Kawaguchi, *Nat. Commun.* **4**, 2191 (2013).
4. T. Sasaki *et al.*, *Nat. Commun.* **5**, 4983 (2014).
5. M. Takahara *et al.*, *Plant Cell Physiol.* **54**, 433–447 (2013).
6. S. Sato *et al.*, *DNA Res.* **15**, 227–239 (2008).
7. A. Małolepszy *et al.*, *Plant J.* **88**, 306–317 (2016).
8. A. De Luis *et al.*, *Plant Physiol.* **160**, 2137–2154 (2012).
9. D. B. Holt *et al.*, *New Phytol.* **208**, 241–256 (2015).
10. D. B. Holt, "Small RNAs involved in bacterial root symbiosis and its regulation in *Lotus japonicus*," thesis, Aarhus University, Aarhus, Denmark (2014).
11. S. Magori *et al.*, *Mol. Plant-Microbe Interact.* **22**, 259–268 (2009).
12. J. Verdier *et al.*, *Plant J.* **74**, 351–362 (2013).
13. J. D. Murray *et al.*, *Science* **315**, 101–104 (2007).
14. D. Reid *et al.*, *Plant Physiol.* **175**, 361–375 (2017).
15. H. Nishida, Y. Handa, S. Tanaka, T. Suzaki, M. Kawaguchi, *J. Plant Res.* **129**, 909–919 (2016).
16. V. Mortier, E. De Wever, M. Vuylsteke, M. Holsters, S. Goormachtig, *Plant J.* **70**, 367–376 (2012).
17. T. Soyano, Y. Shimoda, M. Hayashi, *Plant Cell Physiol.* **56**, 368–376 (2015).
18. S. Okamoto, M. Kawaguchi, *Plant Signal. Behav.* **10**, e1000138 (2015).
19. S. Radutoiu *et al.*, *Nature* **425**, 585–592 (2003).

ACKNOWLEDGMENTS

We thank M. Kawaguchi for *tml-1* seeds; F. Pedersen and K. A. Kristensen for plant care; M. Nadzieja for help with imaging; and Y. Kawaharada, S. Radutoiu, and C. Gutjahr for discussions. We apologize to authors whose work could not be cited due to space limitations. **Funding:** This work was funded by the Danish National Research Foundation (grant DNRF79); ERC (advanced grant 268523); German Research Foundation (grant CRC1101, project C07); and Ministry of Science, Research and Art of Baden-Wuerttemberg (Az:7533-30-20/1). **Author contributions:** D.T., Z.Y., K.M., D.B.H., N.B.A., D.E.R., L.H.M., H.B., and M.S. performed experiments and analyzed data; K.M., D.T., Z.Y., and J.S. conceived of and designed research; K.M. wrote the paper with input from co-authors. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** All data are available in the main text or supplementary materials. For material requests please contact the corresponding author.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/362/6411/233/suppl/DC1
Materials and Methods
Figs. S1 to S10
Tables S1 and S2
References (*20*–*38*)

26 March 2018; accepted 21 August 2018
Published online 30 August 2018
10.1126/science.aat6907

MOLECULAR BIOLOGY

Systematic discovery of natural CRISPR-Cas12a inhibitors

Kyle E. Watters[1], Christof Fellmann[1,2], Hua B. Bai[1], Shawn M. Ren[1], Jennifer A. Doudna[1,2,3,4,5,6,7]*

Cas12a (Cpf1) is a CRISPR-associated nuclease with broad utility for synthetic genome engineering, agricultural genomics, and biomedical applications. Although bacteria harboring CRISPR-Cas9 or CRISPR-Cas3 adaptive immune systems sometimes acquire mobile genetic elements encoding anti-CRISPR proteins that inhibit Cas9, Cas3, or the DNA-binding Cascade complex, no such inhibitors have been found for CRISPR-Cas12a. Here we use a comprehensive bioinformatic and experimental screening approach to identify three different inhibitors that block or diminish CRISPR-Cas12a–mediated genome editing in human cells. We also find a widespread connection between CRISPR self-targeting and inhibitor prevalence in prokaryotic genomes, suggesting a straightforward path to the discovery of many more anti-CRISPRs from the microbial world.

CRISPR-Cas systems represent the only known adaptive mechanism by which prokaryotes protect themselves from biological attackers (*1*). Although diverse in composition, all CRISPR-Cas pathways use RNA-guided enzymes that recognize and destroy foreign nucleic acids, commonly double-stranded DNA (dsDNA) (*2*). The ease of changing the RNA guide molecule, and hence the DNA targeting specificity, has enabled use of CRISPR-Cas9 and CRISPR-Cas12 for programmable genome editing in a wide range of cells and organisms (*3*, *4*). To control Cas9, bacterial protein inhibitors referred to as anti-CRISPRs (Acrs) have been found to limit or block Cas9 functions (*5*–*9*). However, these inhibitors have been found only sporadically, and no such inhibitors have been reported for Cas12a (Cpf1).

The known CRISPR-Cas inhibitors have been identified either through isolation of CRISPR-resistant phages (*9*–*13*) or by proximity to anti-CRISPR associated (*aca*) genes (*14*, *15*). As an alternative inhibitor discovery strategy, the presence of stable self-targeting CRISPR sequences has been proposed as a potential indicator of genomes or mobile genetic elements (MGEs) harboring CRISPR inhibitors (Fig. 1A) (*16*, *17*). Self-targeting CRISPR sequences in CRISPR arrays are expected to be lethal to the host cell by directing cleavage and subsequent degradation of the microbial genome (*18*). However, such self-targeting CRISPR sequences could exist in cells harboring CRISPR inhibitors (*16*). To test whether CRISPR inhibitors can be discovered systematically by flagging CRISPR self-targeting genomes, we built a bioinformatic pipeline to search across the National Center for Biotechnology Information (NCBI) prokaryotic sequence database to locate self-targeting examples within predicted CRISPR arrays (Fig. 1B). The self-targeting spacer searcher (STSS) first predicts all possible CRISPR arrays using the CRISPR recognition tool (CRT) (*19*) and BLASTs each spacer against the host genome and any associated plasmids. Additionally, STSS collects information to gauge the likelihood that the self-targeting sequence would be lethal to the organism and whether the target sequence occurs in a MGE (fig. S1) (*20*, *21*).

Using STSS, we collected self-targeting data for 150,291 genomes and observed 22,125 cases of predicted self-targets, representing 8917 unique sequences across 9155 genomes (fig. S2 and data S1). Focusing initially on three species in which multiple Acrs have been previously identified (*Pseudomonas aeruginosa*, *Listeria monocytogenes*, and *Neisseria meningitidis*), we determined the number of genomes that contained at least one lethal self-targeting CRISPR spacer and the number of those genomes that also contained an Acr using a blastp search (Fig. 1C). In *N. meningitidis* only 6% of the genomes were observed to contain a potential anti-CRISPR, whereas in *P. aeruginosa* and *L. monocytogenes*, the number exceeded 80 or 90%. The self-targeting genomes devoid of known Acrs may also contain inhibitors that have yet to be discovered, especially in *N. meningitidis*, where the number of self-targeting genomes is high but the number containing known Acrs is low.

On the basis of our observation that self-targeting genomes frequently contain CRISPR inhibitors, we sought to determine whether screening genes in genomes containing self-targeting spacers could uncover new inhibitors. We focused our efforts on the CRISPR-Cas12a system (*22*–*24*), which has so far eluded discovery of inhibitory proteins despite the increasing use of Cas12a in gene editing and diagnostic applications (*22*, *25*, *26*). From the STSS results, we identified four strains of *Moraxella bovoculi* (strains 22581, 33362, 58069, and 283689) that contain self-targeting CRISPR-Cas12a systems as top candidates for containing anti-CRISPRs (see methods). These strains each contain at least one perfectly matched self-targeting sequence in or near a predicted MGE with a correct TTV (where V is A, C, or G) protospacer-adjacent motif (PAM) sequence and intact Cas12 open reading frame, which should render the self-targeting spacers lethal in the absence of anti-CRISPRs (Fig. 1D and fig. S3) (*27*).

To test whether the *Moraxella* genomes encode type V-A anti-CRISPRs (AcrVA), we used a cell-free transcription-translation (TXTL) system (*28*, *29*) to express gene products from *Moraxella* genomic fragments. As an initial test of *M. bovoculi* Cas12a (MbCas12a) protein activity, we used polymerase chain reaction (PCR) to amplify a genomic fragment containing the promoter region and all of the Cas proteins (Cas12a, Cas1, Cas2, and Cas4) from *M. bovoculi* strain 22581. This amplicon was added to TXTL reactions with two reporter plasmids encoding green or red fluorescent protein (GFP or RFP) (Fig. 2A). When supplied with CRISPR RNAs (crRNAs) with base pairing complementary to the GFP and RFP genes, the presence of the MbCas12a-containing genomic fragment greatly reduced expression of both reporters (Fig. 2B and fig. S4). This result demonstrates that MbCas12 is active in *M. bovoculi*, suggesting the existence of a CRISPR inhibitor or inhibitors to prevent the self-targeting spacers from killing the cell.

To identify potential AcrVA-encoding genes, we used a directed screening approach to search the predicted MGEs within three of the *M. bovoculi* strains (strain 283689 was unavailable) that contain self-targeting sequences from a type V-A CRISPR array. Interestingly, we also observed 13 self-targeting CRISPR type I-C spacers in strain 58069 that strongly suggest the presence of I-C anti-CRISPRs in that strain (fig. S5). For each of the *M. bovoculi* genomes, pairs of PCR primers were used to make overlapping amplicons of 2 to 10 kb in length spanning all of the predicted MGEs in the three strains (generally excluding highly similar sequences) (table S1). These genomic fragments (GFs) were then added to the TXTL cleavage reactions described above.

From a total of 67 GFs that we tested for type V-A CRISPR inhibition activity, four correlated with increased levels of gene expression for both reporters (Fig. 2, C and D, and figs. S6 to S8). We then cloned the individual open reading frames within these fragments (Fig. 2D and table S3) downstream of the tetracycline-inducible promoter to separately induce transcription and translation of each gene and assessed them for CRISPR inhibition activity using the TXTL Cas12a cleavage assay. From the pool of candidates, three proteins supported high levels of expression for both reporters (Fig. 2, E and F, and fig. S9): GF36

[1]Department of Molecular and Cell Biology, University of California, Berkeley, Berkeley, CA 94720, USA. [2]Gladstone Institutes, San Francisco, CA 94158, USA. [3]Department of Chemistry, University of California, Berkeley, Berkeley, CA 94720, USA. [4]Innovative Genomics Institute, University of California, Berkeley, Berkeley, CA 94720, USA. [5]Howard Hughes Medical Institute, University of California, Berkeley, Berkeley, CA 94720, USA. [6]Li Ka Shing Biomedical and Health Sciences Center, University of California, Berkeley, Berkeley, CA 94720, USA. [7]Molecular Biophysics and Integrated Bioimaging Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA.
*Corresponding author. Email: doudna@berkeley.edu

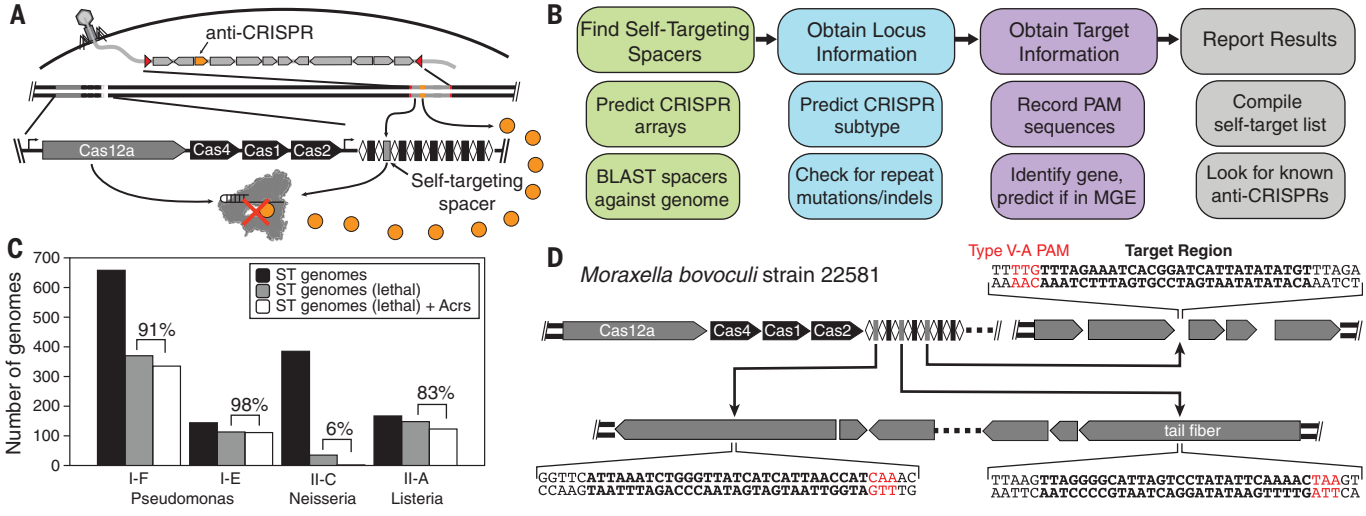
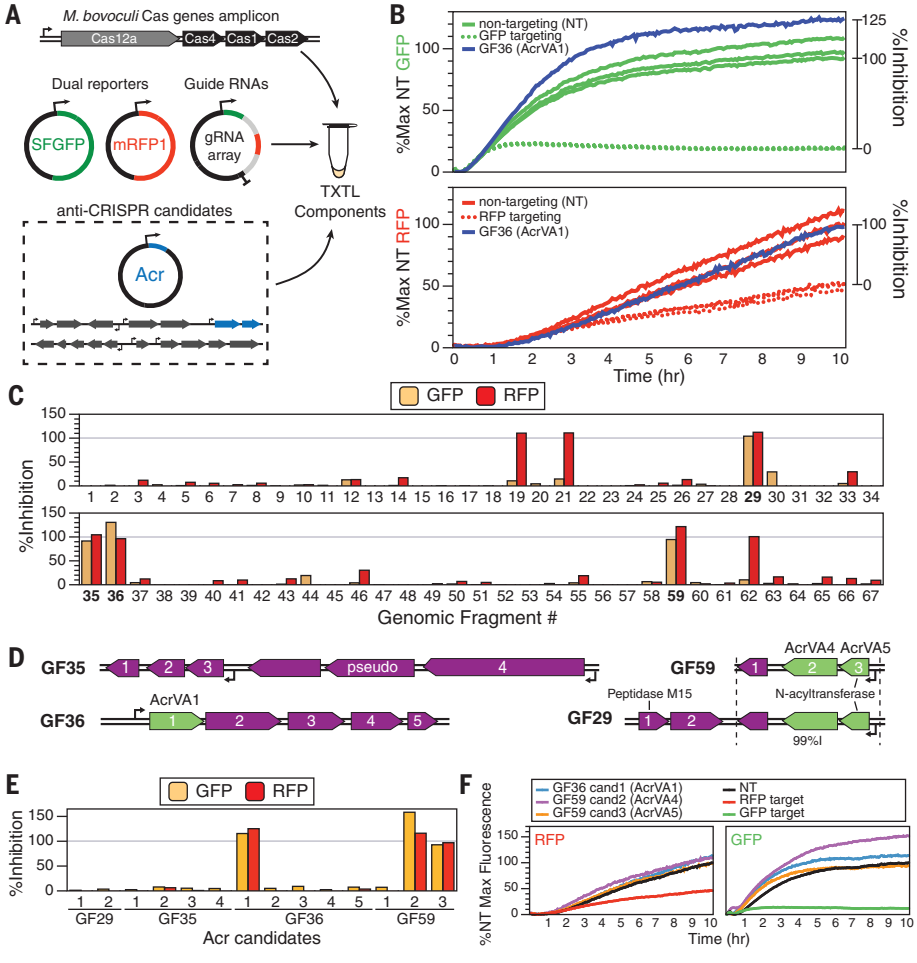


Fig. 1. Bioinformatic approach for discovering Acr genes. (**A**) Anti-CRISPRs allow cells containing self-targeting CRISPR arrays to survive. Expression of anti-CRISPRs inhibits endogenous CRISPR systems and prevents targeting and cleavage of the host genome. (**B**) STSS finds self-targeting CRISPR spacers in genomic DNA, predicts the type of CRISPR system involved, and obtains information about the targeted sequence. (**C**) A large percentage of genomes containing self-targeting (ST) spacers that are predicted to be lethal also contain previously identified Acr genes that inhibit the type I-F, I-E, II-C, or II-A CRISPR systems from the indicated organisms. (**D**) *M. bovoculi* strain 22581 contains three self-targeting spacers in two different prophages in the genome. All of the protospacers contain a TTV PAM (*22*).

Fig. 2. TXTL screening for Acr gene candidates. (**A**) Overview of the TXTL reaction. DNA expressing Cas12a, two fluorescent reporters (SFGFP and mRFP1), and two guide RNAs (gRNAs) are mixed with or without DNA that potentially contains Acr genes. (**B**) Cleavage of the reporter plasmid results in a reduced fluorescent output that is rescued by Acr genes [Acr-absent (NT) data in triplicate]. NT, nontargeting. (**C**) Percent inhibition observed for 67 GFs across three self-targeting *M. bovoculi* strains, where percent inhibition is the level of reporter expression observed relative to the NT control. Four GFs (bold) exhibited inhibition in both fluorescence channels. (**D**) GF29, GF35, GF36, and GF59 [99% nucleotide identity (99%I) to GF29] exhibited high levels of expression for both reporters. The genes in each fragment that were tested for Acr activity are numbered as in (E). One predicted pseudogene is marked as pseudo. (**E**) Testing the individual genes from the fragments in (D) (table S2) resulted in the identification of AcrVA1 (GF36 candidate 1), AcrVA4 (GF59 candidate 2), and AcrVA5 (GF59 candidate 3). (**F**) Kinetic TXTL data for the AcrVA genes measured over the course of 10 hours of gene expression.



candidate 1, GF59 candidate 2, and GF59 candidate 3—hereafter referred to as AcrVA1, AcrVA4, and AcrVA5, respectively—to complement the other AcrVA genes discovered concurrently with this work (*30*).

To confirm the CRISPR inhibition activity of AcrVA1, AcrVA4, and AcrVA5 in vitro, we first overexpressed and purified each putative Cas12a inhibitor, MbCas12a, and two additional Cas12a enzymes (AsCas12a and LbCas12a) commonly used in genome editing or DNA detection applications (fig. S10) (*25*, *26*). We then generated crRNA-protein (RNP) complexes for each of the purified Cas12a enzymes and added the RNPs to a linearized target plasmid that was preincubated with increasing concentrations of each candidate AcrVA protein. We observed that AcrVA1 inhibited DNA cleavage by all three Cas12a enzymes; the strongest inhibition was observed for MbCas12a, and the weakest was observed for AsCas12a (Fig. 3). AcrVA4 and AcrVA5 inhibited dsDNA cleavage for both MbCas12a and LbCas12a but did not inhibit AsCas12a. Interestingly, we also observed that AcrVA4 more strongly inhibited the MbCas12a from strain 58069 (fig. S11) than the MbCas12a from strain 22581 (Fig. 3A) and that AcrVA5 was unable to inhibit the MbCas12a from strain 58069 (fig. S11). None of the AcrVA proteins inhibited *Streptococcus pyogenes* Cas9 (SpyCas9) cleavage (fig. S12).

Having confirmed robust DNA cleavage inhibition by AcrVA1, AcrVA4, and AcrVA5 using purified protein samples, we next tested whether these Cas12a inhibitors could block or reduce Cas12a-mediated genome editing in human cells. We cloned each AcrVA candidate, AcrIIA4 (a SpyCas9 inhibitor), or negative controls into a lentiviral expression vector and stably transduced human embryonic kidney (HEK) 293T–derived genome editing reporter cells containing a doxycycline-inducible GFP marker. Purified AsCas12a, LbCas12a, MbCas12a, or SpyCas9 protein was assembled with a GFP-targeting guide RNA and transfected into the AcrVA-expressing reporter cell lines (Fig. 4A). At 24 hours after RNP delivery, cells were induced by doxycycline for another 24 hours before quantifying editing efficiency by flow cytometry (Fig. 4B and fig. S13A) and a T7 endonuclease 1 assay (Fig. 4C and fig. S13B). We observed high levels of genome editing induced by SpyCas9 with no inhibition by any of the AcrVA proteins or negative controls but virtually complete inhibition by AcrIIA4. We also observed strong Cas12a RNP-editing inhibition that generally matched the in vitro cleavage results. Mirroring their biochemical behavior, AcrVA1 provided the broadest inhibition of Cas12a and fully blocked AsCas12a with efficiencies

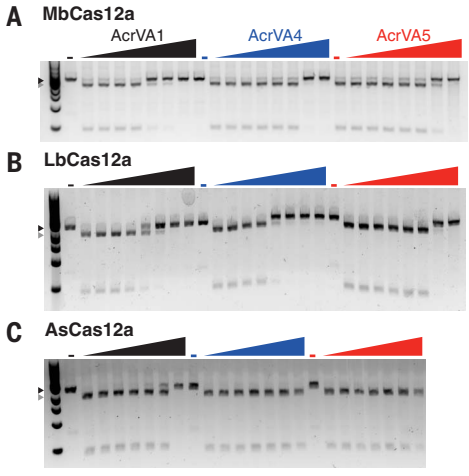


Fig. 3. Biochemical validation of AcrVA inhibitors. (**A**) MbCas12a in vitro dsDNA cleavage is inhibited by increasing concentrations of AcrVA1, AcrVA4, and AcrVA5 (0 to 1.25 μM; see methods). The "–" indicates a control lane without Cas12. (**B**) LbCas12a (*4*, *22*, *26*) is also inhibited by all three AcrVA proteins. (**C**) High concentrations of AcrVA1 inhibit AsCas12a-mediated dsDNA cleavage, but AcrVA4 and AcrVA5 have no effect. The triangles to the left of each image indicate uncleaved (black) or cleaved (gray) DNA.

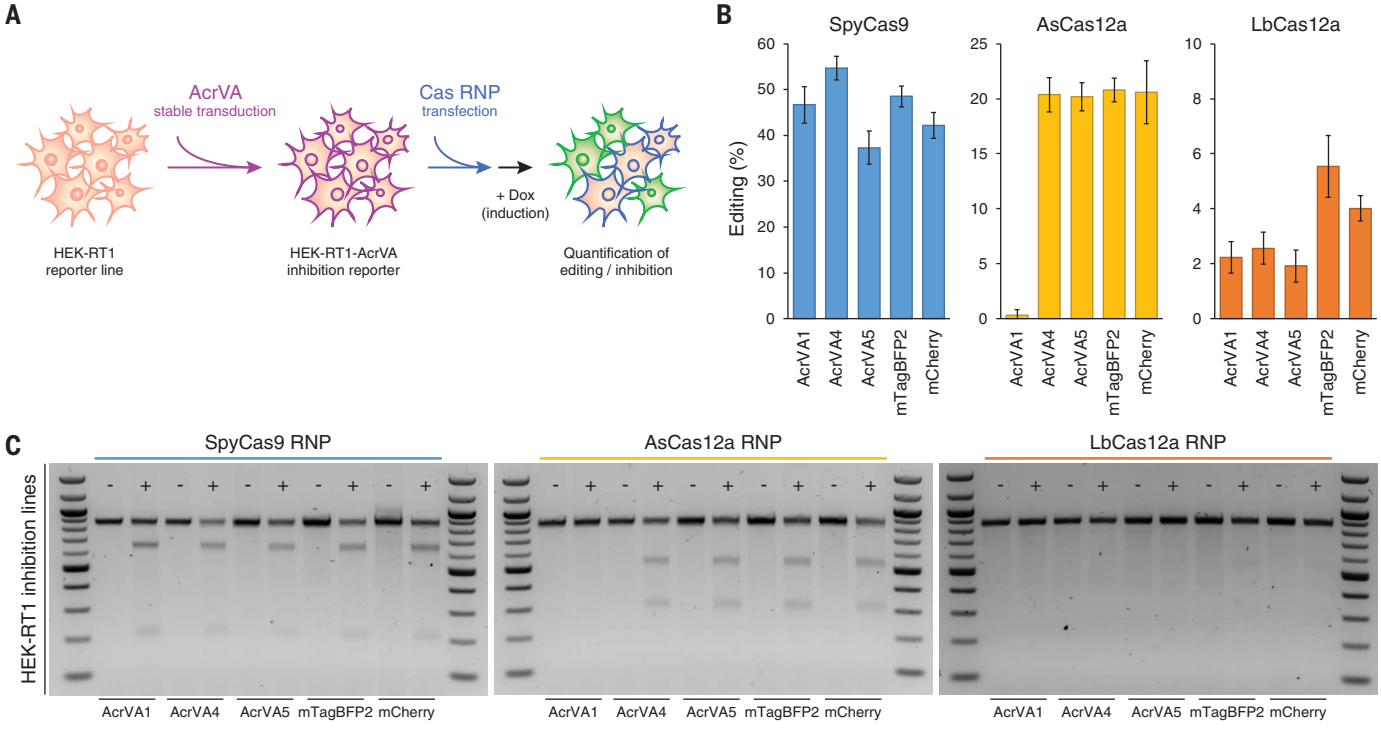


Fig. 4. AcrVAs robustly inhibit genome editing by specific CRISPR-Cas12a nucleases in mammalian cells. (**A**) Overview of the editing reporter assay in human cells. Dox, doxycycline. (**B**) Quantification of genome editing in reporter cell lines stably expressing the indicated CRISPR-Cas12a inhibitors (AcrVAs) or a negative control (mTagBFP2 and mCherry). The scale of each plot is adjusted to compensate for differences in editing efficiency. Error bars indicate standard deviations of triplicates. (**C**) Biochemical analysis of AcrVA-mediated inhibition in representative samples shown in (B) in the presence (+) or absence (–) of each RNP. Editing was assessed by the T7 endonuclease 1 assay.

comparable to that of AcrIIA4's inhibition of SpyCas9. AcrVA4 and AcrVA5 only inhibited LbCas12a. RNP-based delivery of MbCas12a did not edit efficiently enough to determine the effectiveness of the AcrVA genes on its activity (fig. S13), consistent with previous findings (*22*).

Together, these results establish a new approach for systematic discovery and validation of CRISPR-Cas inhibitors hidden within self-targeting genomes. Importantly, the Cas12a inhibitors revealed by this approach are only found within a few genomes within the NCBI database, with AcrVA4 and AcrVA5 being particularly rare genes, only co-occurring with each other and thus intractable to an *aca*-based search approach (fig. S14). Although we expect the extensive set of yet-to-be-analyzed CRISPR self-targeting genomes (data S1) will lead to the discovery of many more Acrs across all CRISPR subtypes, the AcrVAs discovered and validated here provide a toolkit for selective Cas12a regulation both in vitro and in mammalian systems, with the potential to advance synthetic biology, CRISPR diagnostics, and therapeutic genome editing.

REFERENCES AND NOTES

1. S. A. Jackson *et al*., *Science* **356**, eaal5056 (2017).
2. A. Plagens, H. Richter, E. Charpentier, L. Randau, *FEMS Microbiol. Rev.* **39**, 442–463 (2015).
3. C. Fellmann, B. G. Gowen, P.-C. Lin, J. A. Doudna, J. E. Corn, *Nat. Rev. Drug Discov.* **16**, 89–100 (2017).
4. A. V. Wright, J. K. Nuñez, J. A. Doudna, *Cell* **164**, 29–44 (2016).
5. J. Shin *et al*., *Sci. Adv.* **3**, e1701620 (2017).
6. H. Yang, D. J. Patel, *Mol. Cell* **67**, 117–127.e5 (2017).
7. D. Dong *et al*., *Nature* **546**, 436–439 (2017).
8. E. M. Basgall *et al*., *Microbiology* **164**, 464–474 (2018).
9. A. P. Hynes *et al*., *Nat. Commun.* **9**, 2919 (2018).
10. J. Bondy-Denomy, A. Pawluk, K. L. Maxwell, A. R. Davidson, *Nature* **493**, 429–432 (2013).
11. A. Pawluk, J. Bondy-Denomy, V. H. W. Cheung, K. L. Maxwell, A. R. Davidson, *mBio* **5**, e00896 (2014).
12. A. P. Hynes *et al*., *Nat. Microbiol.* **2**, 1374–1380 (2017).
13. F. He *et al*., *Nat. Microbiol.* **3**, 461–469 (2018).
14. A. Pawluk *et al*., *Nat. Microbiol.* **1**, 16085 (2016).
15. A. Pawluk *et al*., *Cell* **167**, 1829–1838.e9 (2016).
16. B. J. Rauch *et al*., *Cell* **168**, 150–158.e10 (2017).
17. J. Bondy-Denomy, *ACS Chem. Biol.* **13**, 417–423 (2018).
18. G. E. Heussler, G. A. O'Toole, *J. Bacteriol.* **198**, 1481–1486 (2016).
19. C. Bland *et al*., *BMC Bioinformatics* **8**, 209 (2007).
20. D. Arndt *et al*., *Nucleic Acids Res.* **44**, W16–W21 (2016).
21. C. M. Hudson, B. Y. Lau, K. P. Williams, *Nucleic Acids Res.* **43**, D48–D53 (2015).
22. B. Zetsche *et al*., *Cell* **163**, 759–771 (2015).
23. S. Shmakov *et al*., *Mol. Cell* **60**, 385–397 (2015).
24. D. Burstein *et al*., *Nature* **542**, 237–241 (2017).
25. J. S. Gootenberg *et al*., *Science* **360**, 439–444 (2018).
26. J. S. Chen *et al*., *Science* **360**, 436–439 (2018).
27. A. M. Dickey *et al*., *Vet. Res.* **47**, 31 (2016).
28. R. Marshall *et al*., *Mol. Cell* **69**, 146–157.e3 (2018).
29. Z. Z. Sun *et al*., *J. Vis. Exp.* **2013**, e50762 (2013).
30. N. D. Marino *et al*., *Science*, **362**, 240–242 (2018).

ACKNOWLEDGMENTS

We thank R. Lew for technical support; D. Burstein for Cas protein hidden Markov models; M. West and the CIRM/QB3 Shared Stem Cell Facility/High-Throughput Screening Facility for flow cytometry equipment support; A. Borges for discussions on self-targeting in *Pseudomonas*; G. Knott for purified 33362 MbCas12; and Doudna lab members for thoughtful discussion. **Funding:** This research was developed with funding from the Defense Advanced Research Projects Agency (DARPA) award HR0011-17-2-0043. The views, opinions, and/or findings expressed are those of the author and should not be interpreted as representing the official views or policies of the Department of Defense or the U.S. government. Approved for public release; distribution is unlimited. This research was also supported by the Allen Distinguished Investigator Program, through the Paul G. Allen Frontiers Group, and the National Science Foundation (MCB-1244557 to J.A.D.). C.F. is supported by a U.S. National Institutes of Health K99/R00 Pathway to Independence Award (K99GM118909 and R00GM118909) from the National Institute of General Medical Sciences. J.A.D. is an investigator of the Howard Hughes Medical Institute (HHMI), and this study was supported in part by HHMI; J.A.D is also a Paul Allen Distinguished Investigator. **Author contributions:** Conceptualization, K.E.W.; methodology, K.E.W. and C.F.; software, K.E.W.; investigation, K.E.W., C.F., H.B.B., and S.M.R.; data curation, K.E.W.; writing–original draft, K.E.W., C.F., and J.A.D.; writing–review and editing, K.E.W., C.F., and J.A.D.; funding acquisition, K.E.W., C.F., and J.A.D.; and supervision, K.E.W. and J.A.D. **Competing interests:** J.A.D. is a cofounder of Caribou Biosciences, Editas Medicine, Intellia Therapeutics, Scribe Therapeutics, and Mammoth Biosciences; a scientific adviser to Caribou, Intellia, Scribe, Synthego, Metagenomi, Inari, and eFFECTOR Therapeutics; and a director of Driver and Johnson & Johnson. The Regents of the University of California have patents pending for CRISPR-related technologies on which the authors are inventors. **Data and materials availability:** All data are available in the main text or the supplementary materials. Plasmids for expression of AcrVA1, AcrVA4, AcrVA5, and MbCas12 are available on Addgene (IDs 115653 to 115664, 115669, 115670, and 115795 to 115797) under a Uniform Biological Material Transfer Agreement. Code for STSS is available on GitHub. *M. bovoculi* strains are available from the University of Nebraska (strain 58069) or the U.S. Meat Animal Research Center (strains 22581 and 33362) under a Material Transfer Agreement.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/362/6411/236/suppl/DC1
Materials and Methods
Figs. S1 to S14
Tables S1 to S4
References (*31*–*44*)
Data S1

18 June 2018; accepted 23 August 2018
Published online 6 September 2018
10.1126/science.aau5138

MOLECULAR BIOLOGY

Discovery of widespread type I and type V CRISPR-Cas inhibitors

Nicole D. Marino[1]*, Jenny Y. Zhang[1]*, Adair L. Borges[1]*, Alexander A. Sousa[2,3,4], Lina M. Leon[1], Benjamin J. Rauch[1]†, Russell T. Walton[2,3,4], Joel D. Berry[1]‡, J. Keith Joung[2,3,4,5], Benjamin P. Kleinstiver[2,3,4,5], Joseph Bondy-Denomy[1,6]§

Bacterial CRISPR-Cas systems protect their host from bacteriophages and other mobile genetic elements. Mobile elements, in turn, encode various anti-CRISPR (Acr) proteins to inhibit the immune function of CRISPR-Cas. To date, Acr proteins have been discovered for type I (subtypes I-D, I-E, and I-F) and type II (II-A and II-C) but not other CRISPR systems. Here, we report the discovery of 12 *acr* genes, including inhibitors of type V-A and I-C CRISPR systems. AcrVA1 inhibits a broad spectrum of Cas12a (Cpf1) orthologs—including MbCas12a, Mb3Cas12a, AsCas12a, and LbCas12a—when assayed in human cells. The *acr* genes reported here provide useful biotechnological tools and mark the discovery of *acr* loci in many bacteria and phages.

The discovery of bacterial CRISPR-Cas systems that prevent infection by bacterial viruses (phages) has opened a paradigm for bacterial immunity while yielding exciting tools for targeted genome editing. CRISPR systems destroy phage genomes, and in turn, phages express anti-CRISPR (Acr) proteins that directly inhibit Cas effectors (*1*, *2*). Six distinct types (I to VI) of CRISPR systems are spread widely across the bacterial world (*3*), but Acr proteins have only been discovered for type I and II CRISPR systems (*1*, *3*–*6*). Given the prevalence and diversity of CRISPR systems, we predict that Acr proteins against other types await discovery.

Acr proteins do not have conserved sequences or structures and only share their relatively small size, making de novo prediction of *acr* function challenging (*6*). However, *acr* genes often cluster together with other *acr* genes or are adjacent to highly conserved Acr-associated genes (*aca* genes) in "*acr* loci" (*7*, *8*). In this work, we sought to identify *acr* genes in bacteria and phages that are not homologous to previously identified *acr* or *aca* genes.

Acr proteins were first discovered in *Pseudomonas aeruginosa*, inhibiting type I-F and I-E CRISPR systems (*1*, *9*). *P. aeruginosa* strains also encode a third CRISPR subtype (type I-C), which lacks known inhibitors (*10*). We engineered *P. aeruginosa* to target phage JBD30 with type I-C CRISPR-Cas (fig. S1A) and used it in parallel with existing type I-E (strain SMC4386) and I-F (strain PA14) CRISPR strains to screen for additional *acr* candidates.

Homologs of *aca1* were searched for in *Pseudomonas* genomes, and seven gene families not previously tested for Acr function were identified upstream of *aca1* (Fig. 1A). Three genes inhibited the type I-E CRISPR-Cas system (*acrIE5-7*), one inhibited type I-F (*acrIF11*), restoring the plaquing of a targeted phage, and two genes had no inhibitory activity (*orf1* and *orf2*) (Fig. 1B, fig. S1B, and tables S1 and S2). Another gene exhibited dual I-E and I-F inhibition, and domain analysis revealed a chimera of previously identified *acrIE4* and *acrIF7* (*acrIE4-F7*). No type I-C inhibitors were identified. The type I-F inhibitor *acrIF11* was commonly represented in both the *P. aeruginosa* mobilome and in more than 50 species of diverse Proteobacteria (fig. S2 and table S2). *acrIF11* is often associated with genes that encode DNA-binding motifs, which we have designated *aca4* to *aca7* (fig. S2 and tables S1, S3, and S4). To confirm that these *aca* genes can be used to facilitate *acr* discovery, we used *aca4* to discover an additional *Pseudomonas* Acr, *acrIF12* (Fig. 1, A and B).

Given the widespread nature of *acrIF11*, we next used it to discover Acr proteins against CRISPR systems in which they have not yet been found: type I-C, a minimal class 1 system, and type V-A CRISPR-Cas12a (Cpf1), a class 2 single effector system that has high efficiency in genome editing (*11*–*13*). To find AcrIC and AcrVA proteins, we first searched for genomes that encode CRISPR spacers that match a target protospacer elsewhere in the same genome (Fig. 2A). The tolerance of this "self-targeting" in viable bacteria indicates potential inhibition of the CRISPR system (*4*) because genome cleavage would result in bacterial death.

The Gram-negative bovine pathogen *Moraxella bovoculi* (*14*, *15*) is a Cas12a-containing organism (*11*) in which four of the seven genomes feature type V-A self-targeting (table S5), and one strain (58069) also features self-targeting by type I-C (table S6). Although no previously described *acr* or *aca* genes were present in this strain, an *acrIF11* homolog was found in phages that infect the human pathogen *Moraxella catarrhalis* (*16*), a close relative of *M. bovoculi*. Genes adjacent to *acrIF11* in *M. catarrhalis* had homologs in the self-targeting *M. bovoculi* strains (Fig. 2B), and together, these genes were selected as candidate *acr* genes. Each gene was first tested against the type I-C and I-F systems introduced above because both subtypes are found in *Moraxella*. Gene *AAX09_07415* (now *acrICI*) inhibited the type I-C system, explaining the tolerance of self-targeting

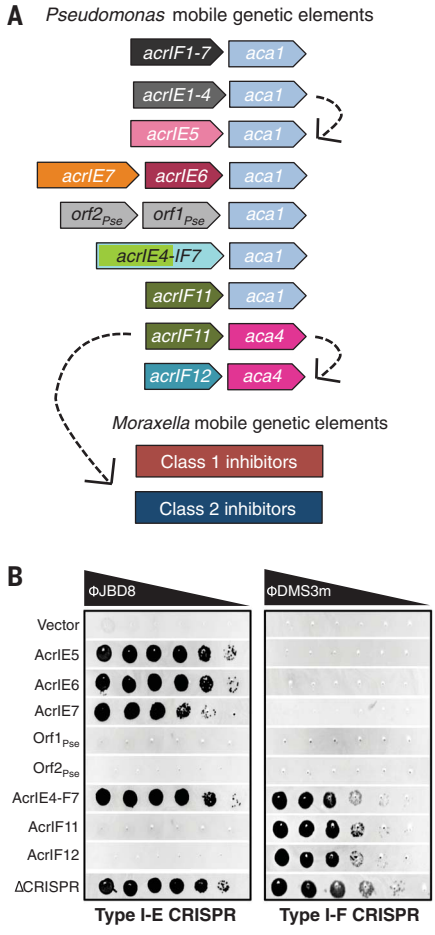


Fig. 1. The discovery of a widespread type I inhibitor. (**A**) Schematic of type I-E and type I-F Acrs with Acr-associated (*aca1* and *aca4*) genes in *Pseudomonas* sp. mobile genetic elements, with dotted lines indicating the "guilt-by-association" relationships used to discover new *acr* genes in *Pseudomonas* sp. and *Moraxella* sp. from known *acr* genes (top two rows). (**B**) Phage plaque assays to assess CRISPR-Cas inhibition. Tenfold serial dilutions of a type I-E or type I-F CRISPR-targeted phage (JBD8 or DMS3m, respectively) titered on lawns of *P. aeruginosa* with naturally active type I-E or type I-F CRISPR-Cas systems expressing candidate inhibitors. Δ*CRISPR* strains measure phage replication in the absence of CRISPR immunity (top row).

[1]Department of Microbiology and Immunology, University of California, San Francisco, San Francisco, CA 94143, USA. [2]Molecular Pathology Unit, Massachusetts General Hospital, Charlestown, MA 02129, USA. [3]Center for Cancer Research, Massachusetts General Hospital, Charlestown, MA 02129, USA. [4]Center for Computational and Integrative Biology, Massachusetts General Hospital, Charlestown, MA 02129, USA. [5]Department of Pathology, Harvard Medical School, Boston, MA 02115, USA. [6]Quantitative Biosciences Institute, University of California, San Francisco, San Francisco, CA 94143, USA.
*These authors contributed equally to this work. †Present address: Zymergen, Emeryville, CA 94608, USA. ‡Present address: Caribou Biosciences, Berkeley, CA 94710, USA.
§Corresponding author. Email: joseph.bondy-denomy@ucsf.edu

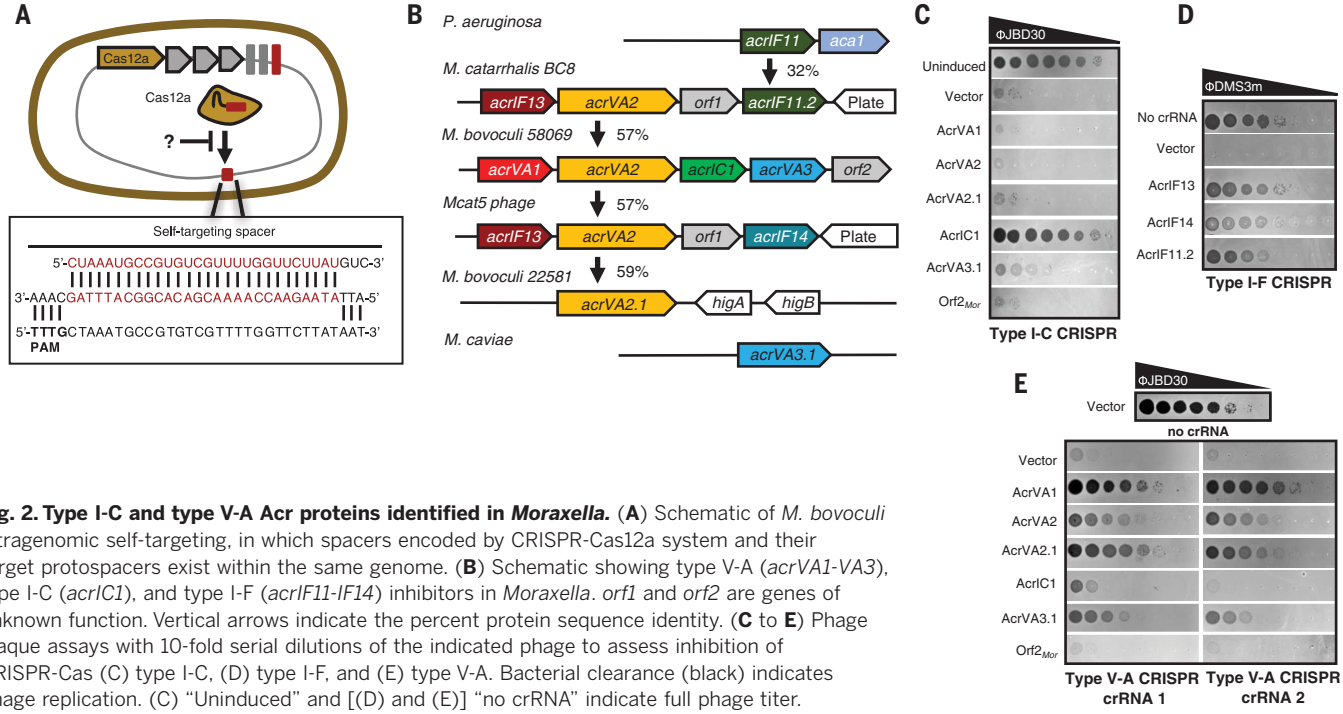


Fig. 2. Type I-C and type V-A Acr proteins identified in *Moraxella*. (**A**) Schematic of *M. bovoculi* intragenomic self-targeting, in which spacers encoded by CRISPR-Cas12a system and their target protospacers exist within the same genome. (**B**) Schematic showing type V-A (*acrVA1-VA3*), type I-C (*acrIC1*), and type I-F (*acrIF11-IF14*) inhibitors in *Moraxella*. *orf1* and *orf2* are genes of unknown function. Vertical arrows indicate the percent protein sequence identity. (**C** to **E**) Phage plaque assays with 10-fold serial dilutions of the indicated phage to assess inhibition of CRISPR-Cas (C) type I-C, (D) type I-F, and (E) type V-A. Bacterial clearance (black) indicates phage replication. (C) "Uninduced" and [(D) and (E)] "no crRNA" indicate full phage titer.

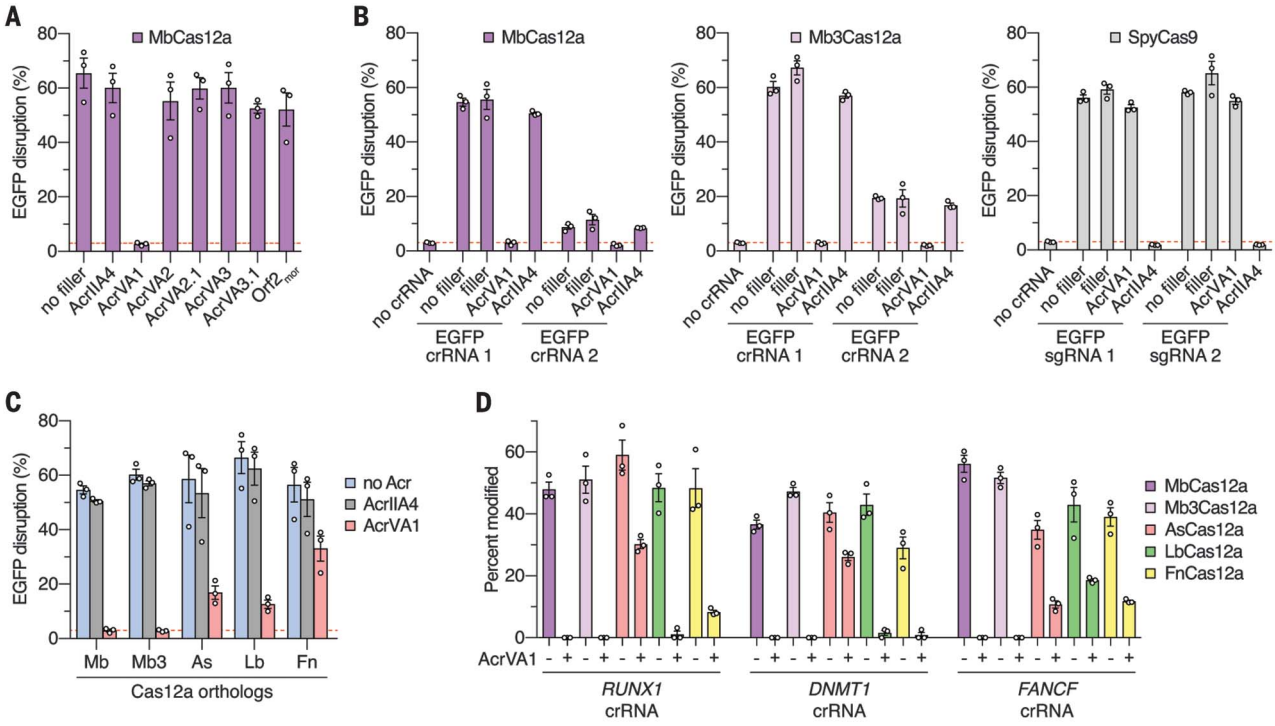


Fig. 3. AcrVA1 blocks Cas12a-mediated gene editing in human cells. (**A** to **C**) Human cell U2-OS-EGFP disruption experiments to assess AcrVA-mediated inhibition of Cas12a activities. (A) Inhibition of MbCas12a activity with various AcrVA constructs; the "no filler" condition contained only plasmids for Cas12a and crRNA expression. (B) Comparisons between the inhibitory activities of AcrVA1 and AcrIIA4 against MbCas12a, Mb3Cas12a, and SpyCas9. Controls that use "filler" plasmid in lieu of Acr plasmids were included in order to equalize amounts of DNA. (C) Assessment of AcrVA1 activity against Cas12a orthologs, with AcrIIA4 used as control. Background EGFP disruption is indicated by the red dashed line; error bars indicate SEM for $n = 3$ independent biological replicates. (**D**) Inhibition of activity of Cas12a orthologs against endogenous sites in human cells (*RUNX1*, *DNMT1*, or *FANCF* genes). Gene modification assessed by means of T7E1 assay 72 hours after transfection; error bars indicate SEM for $n = 3$ independent biological replicates.

in strain 58069 (Fig. 2C). Additionally, gene *E9U_08473* (*acrIF13*) from the *M. catarrhalis* BC8 prophage completely inhibited I-F function, as did AKI27193.1 (*acrIF14*), which is found in phage Mcat5 at the same genomic position as that of *acrIF11* in BC8 (Fig. 2, B and D). These Acr proteins possess broad spectrum activity; the type I-C and I-F systems in *Moraxella* and *Pseudomonas* only share an average pairwise identity of 30 and 36%, respectively (fig. S3)

Because of the limited tools available for the genetic manipulation of *Moraxella* sp., the remaining genes were tested for type V-A Acr function in *P. aeruginosa* PAO1 engineered to express MbCas12a and a CRISPR-RNA (crRNA) that targets *P. aeruginosa* phage JBD30. Two distinct crRNAs were used, showing strong reduction of titer by more than four orders of magnitude (Fig. 2E). The first gene in the *M. bovoculi* 58069 *acr* locus, *AAX09_07405* (*acrVA1*), restored phage titers nearly to levels seen with the crRNA-minus control, indicating that it robustly inhibits Cas12a. This is in good agreement with the independent discovery of AcrVA1 reported in a companion paper (*17*). The adjacent gene, *acrVA2*, also inhibited targeting, as did its ortholog (*acrVA2.1*) (Fig. 2E). An additional gene from this locus, *acrVA3*, possessed subtle anti-Cas12a activity but was toxic to cells and adversely affected JBD30 phage growth independently of Cas12a (fig. S5). We therefore tested an ortholog with 43% sequence identity, *B0181_04965* (*acrVA3.1*), which showed stronger Cas12a inhibition with no toxicity or adverse effects on phage growth (Fig. 2E and fig. S5). Surprisingly, *acrVA3.1* also showed partial restoration of phage titer during type I-C targeting, suggesting that it may inhibit the type I-C as well as type V-A system (Fig. 2, C and E). Although these two CRISPR subtypes do not share any protein components, a dual-specificity inhibitor may use distinct protein-interaction interfaces or modulate an undiscovered host process required for CRISPR immunity. We used the Acr "key" *acrIF11* to unlock *acr* loci that encode seven distinct *acr* genes inhibiting type I-C, I-F, and V-A CRISPR. Below, we focus on the evolutionary analysis of type V-A inhibitors and their function in mammalian cells.

The gene *acrVA1* encodes a 170–amino acid protein found only in *Moraxella* sp. and *Eubacterium eligens* (fig. S6), both type V-A CRISPR-containing organisms. By contrast, *acrVA2* (322 amino acids) and *acrVA3* (168 amino acids) orthologs are found broadly distributed throughout multiple classes of bacteria. For example, *acrVA2* orthologs are present in *Lachnospiraceae* and *Leptospira* (fig. S7), which contain type V-A CRISPR, as well as in *Moraxella, Leptospira,* and *Lactobacillus* phages. Distant orthologs of *acrVA2* were also identified on plasmids and conjugative elements in *Escherichia coli* (fig. S7), although the importance of a bacterium that lacks type V-A CRISPR encoding a putative *acrVA* gene is unknown. Orthologs of *acrVA3* were identified in many Proteobacteria and in *Eubacterium* and *Clostridium* species, which encode type V-A CRISPR (fig. S8).

Given the inhibitory effect of *acrVA1-3.1* on MbCas12a in bacteria, we sought to determine whether AcrVA proteins could block MbCas12a activity in human cells. Human U2-OS–enhanced green fluorescent protein (EGFP) cells (*18*) transiently expressing MbCas12a, *EGFP*-targeting crRNA, and human codon-optimized *acrVA1-3.1* were assessed for EGFP disruption by using flow cytometry. Coexpression of MbCas12a and crRNA resulted in ~60 to 70% disruption of *EGFP* expression relative to background (Fig. 3A). AcrVA1 expression reduced *EGFP* disruption to background levels, indicating inhibition of MbCas12a, whereas the other *acrVA* genes showed little evidence of activity here (Fig. 3A). We additionally found that *acrVA1* inhibited another Cas12a ortholog (Mb3Cas12a) while having no impact on SpyCas9 editing in the same assay (Fig. 3B). Titration of the Acr plasmid relative to the Cas12a expression plasmid revealed comparable dose-dependent responses to inhibition between MbCas12a or Mb3Cas12a with AcrVA1 and SpyCas9 with AcrIIA4 (fig. S9). Furthermore, AcrVA1 was found to be a broad-spectrum inhibitor of other commonly used Cas12a orthologs (*11*), providing strong inhibition of AsCas12a and LbCas12a and modest inhibition of FnCas12a (Fig. 3C).

Last, to determine whether AcrVA1 could inhibit Cas12a-mediated modification of endogenous loci in human cells, U2-OS cells were cotransfected with plasmids expressing *acrVA1*, Cas12a, and crRNAs that target endogenous genes (*RUNX1*, *DNMT1*, or *FANCF*) and assessed for gene disruption by means of T7 endonuclease I (T7E1) assay. We found that AcrVA1 completely inhibited gene disruption by MbCas12a and Mb3Cas12a, with modest to strong inhibition of As-, Lb-, and FnCas12a orthologs (Fig. 3D and fig. S10).

Here, we report the discovery of a broadly distributed type I-F Acr protein (AcrIF11) that served as a marker for *acr* loci and led to the identification of type I-C and V-A CRISPR inhibitors. One of these *acrVA* genes (*acrVA1*) potently inhibits Cas12a in bacteria and in human cells, providing a new tool for Cas12a regulation. Our findings show that mobile genetic elements can tolerate bacteria with more than one CRISPR-Cas type by possessing multiple Acr proteins in the same locus. The strategy described here enabled the identification of many widespread Acr proteins, which may prove useful in future Acr discovery.

REFERENCES AND NOTES

1. J. Bondy-Denomy, A. Pawluk, K. L. Maxwell, A. R. Davidson, *Nature* **493**, 429–432 (2013).
2. J. Bondy-Denomy *et al.*, *Nature* **526**, 136–139 (2015).
3. E. V. Koonin, K. S. Makarova, F. Zhang, *Curr. Opin. Microbiol.* **37**, 67–78 (2017).
4. B. J. Rauch *et al.*, *Cell* **168**, 150–158.e10 (2017).
5. A. Pawluk *et al.*, *Cell* **167**, 1829–1838.e9 (2016).
6. A. L. Borges, A. R. Davidson, J. Bondy-Denomy, *Annu. Rev. Virol.* **4**, 37–59 (2017).
7. A. Pawluk, A. R. Davidson, K. L. Maxwell, *Nat. Rev. Microbiol.* **16**, 12–17 (2018).
8. A. Pawluk *et al.*, *Nat. Microbiol.* **1**, 16085 (2016).
9. A. Pawluk, J. Bondy-Denomy, V. H. W. Cheung, K. L. Maxwell, A. R. Davidson, *MBio* **5**, e00896, e00896–e14 (2014).
10. A. van Belkum *et al.*, *MBio* **6**, e01796–e15 (2015).
11. B. Zetsche *et al.*, *Cell* **163**, 759–771 (2015).
12. I. Fonfara, H. Richter, M. Bratovič, A. Le Rhun, E. Charpentier, *Nature* **532**, 517–521 (2016).
13. B. P. Kleinstiver *et al.*, *Nat. Biotechnol.* **34**, 869–874 (2016).
14. J. A. Angelos, P. Q. Spinks, L. M. Ball, L. W. George, *Int. J. Syst. Evol. Microbiol.* **57**, 789–795 (2007).
15. A. M. Dickey *et al.*, *Vet. Res.* **47**, 31 (2016).
16. A. Ariff *et al.*, *BMC Genomics* **16**, 860 (2015).
17. K. E. Watters, C. Fellmann, H. B. Bai, S. M. Ren, J. A. Doudna, *Science* **362**, 236–239 (2018).
18. D. Reyon *et al.*, *Nat. Biotechnol.* **30**, 460–465 (2012).

ACKNOWLEDGMENTS

We thank J. M. Peters and C. A. Gross (University of California, San Francisco) for providing the pTN7C130 entry vector for cloning MbCas12a. We thank D. Loy (University of Nebraska) and M. Clawson (U.S. Department of Agriculture) for providing *M. bovoculi* strain 58069 and 22581 for genomic DNA extraction. **Funding:** Work in the Bondy-Denomy laboratory was supported by the University of California, San Francisco Program for Breakthrough in Biomedical Research, funded in part by the Sandler Foundation, and a National Institutes of Health (NIH) Office of the Director Early Independence Award (DP5-OD021344). Acr discovery efforts were specifically supported by DARPA Safe Genes grant HR0011-17-2-0043. Work in the Joung laboratory was supported by the Desmond and Ann Heathwood MGH Research Scholar Award (to J.K.J.) and the NIH awards K99 CA218870 (B.P.K.) and R35 GM118158 (J.K.J.). **Author contributions:** N.D.M. conducted *Moraxella* AcrVA and AcrI discovery, characterization, and bioinformatics. J.Y.Z. and A.L.B. conducted *Pseudomonas* AcrI and *aca* discovery, characterization, and bioinformatics. L.M.L. built the *Pseudomonas* type I-C strain, and J.D.B. and N.D.M. built the *Pseudomonas* type V-A strain. B.J.R. conducted self-targeting analysis. J.B.-D conceptualized the project and supervised all bioinformatics and bacterial experiments. B.P.K., A.A.S., and R.T.W. constructed AcrVA expression plasmids and performed human cell experiments. Funding for this work was procured by J.B.-D., B.P.K., and J.K.J. The manuscript was written by N.D.M., A.L.B., and J.B.-D., with editing and contributions from all authors. **Competing interests:** A patent has been filed pertaining to AcrVA genes and their applications. J.K.J. has financial interests in Beam Therapeutics, Blink Therapeutics, Editas Medicine, Endcadia, Monitor Biotechnologies (formerly known as Beacon Genomics), Pairwise Plants, Poseida Therapeutics, and Transposagen Biopharmaceuticals. J.K.J.'s interests were reviewed and are managed by Massachusetts General Hospital and Partners HealthCare in accordance with their conflict-of-interest policies. **Data and materials availability:** All data are available in the main text or the supplementary materials. Plasmids described in this work are available through Addgene. Phages and bacterial strains will be made available upon request to joseph.bondy-denomy@ucsf.edu.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/362/6411/240/suppl/DC1
Materials and Methods
Figs. S1 to S10
Tables S1 to S11
References (*19*–*26*)

18 June 2018; accepted 25 August 2018
Published online 6 September 2018
10.1126/science.aau5174

Enzo

Visit Us at SfN Booth #901

Detect Your Critical Biomarkers

Multiplex and Singleplex IHC Detection Solutions

Improve Your Workflow with a Complete Range of Innovative Tools for IHC Detection

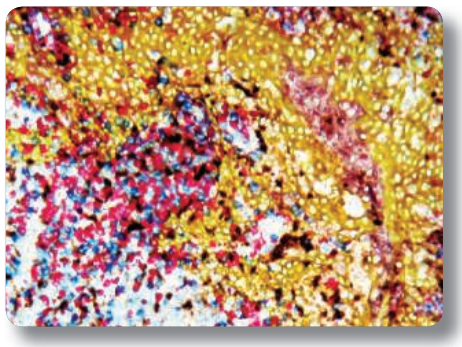
POLYVIEW® PLUS non-biotinylated nanopolymer detection reagents for minimal background without sacrificing signal

MULTIVIEW® PLUS kits for detection of multiple antigens in a single tissue sample

HIGHDEF® Chromogens for high-quality staining and high intensity

Detection Reagents and Chromogens validated to produce impeccable results both manually and on automated stainers

Over 1,000 IHC Validated Antibodies



Four-color multiplex IHC of human tonsil tissue using HIGHDEF chromogens.

scientists ***enabling*** scientists™

www.enzolifesciences.com/IHC

For Research Use Only
Not for Use in Diagnostic Procedures

RISING STAR AWARD
IN NEURODEGENERATIVE RESEARCH

CALL FOR SUBMISSIONS

The Mahoney Institute for Neurosciences (MINS) at the University of Pennsylvania is pleased to announce an international call for submissions for the annual Rising Star Award in neuroscience research. To highlight the "Year of Neurodegenerative Research" on Penn's campus, the award honors a young researcher for outstanding contributions to neurodegenerative research with a USD 10,000 personal honorarium at the *MINS 35th Annual Retreat and Symposium* on April 3, 2019.

"Neurodegenerative disorders are major health problems for the elderly, and there are currently no treatments for any of these diseases," said Virginia Lee, PhD, director of Penn's Center for Neurodegenerative Disease and a MINS faculty member. *"I am proud that there is a large research community at Penn, including MINS, at the forefront of research elucidating the etiology and pathogenesis of these devastating disorders. We look forward to honoring and facilitating a young researcher to enhance our understanding of age-related neurodegeneration and to advance future therapies."*

ELIGIBILITY

- Researchers who received their first advanced degree, such as the PhD, in 2005 or more recently are invited to submit an entry.

AWARD

- In addition to the USD 10,000 personal honorarium, the award recipient will present a research seminar at the symposium and, before the seminar, a separate introductory lecture on neurodegenerative disease.
- Also at the symposium, Prof. Pietro De Camilli, Yale School of Medicine, will give the Sprague Lecture, and Prof. Beth Stevens, Harvard Medical School, will give the Adler Lecture.

SUBMISSION

- Applicants should submit a single PDF file with a one-page description of their contribution to neurodegenerative research, full curricula vitae, and names of three references.
- The deadline for submissions is **December 1, 2018** to **MINSRisingStarAward@lists.upenn.edu**. More information can be found at **www.med.upenn.edu/ins**.

Penn MINS
Mahoney Institute *for* Neurosciences

L O O K
D E E P E R

Decoding disease.
Just another day at the beach.

Standing on the shores of a country and a continent, UC San Diego scientists also find themselves standing on the precipice of changing the world. They're studying how to "update" mosquitos with a gene that prevents the spread of malaria. Imagine that. At the Tata Institute for Genetics and Society, we already have.

UC San Diego

lookdeeper.ucsd.edu

Frontiers in Research Excellence and Discovery

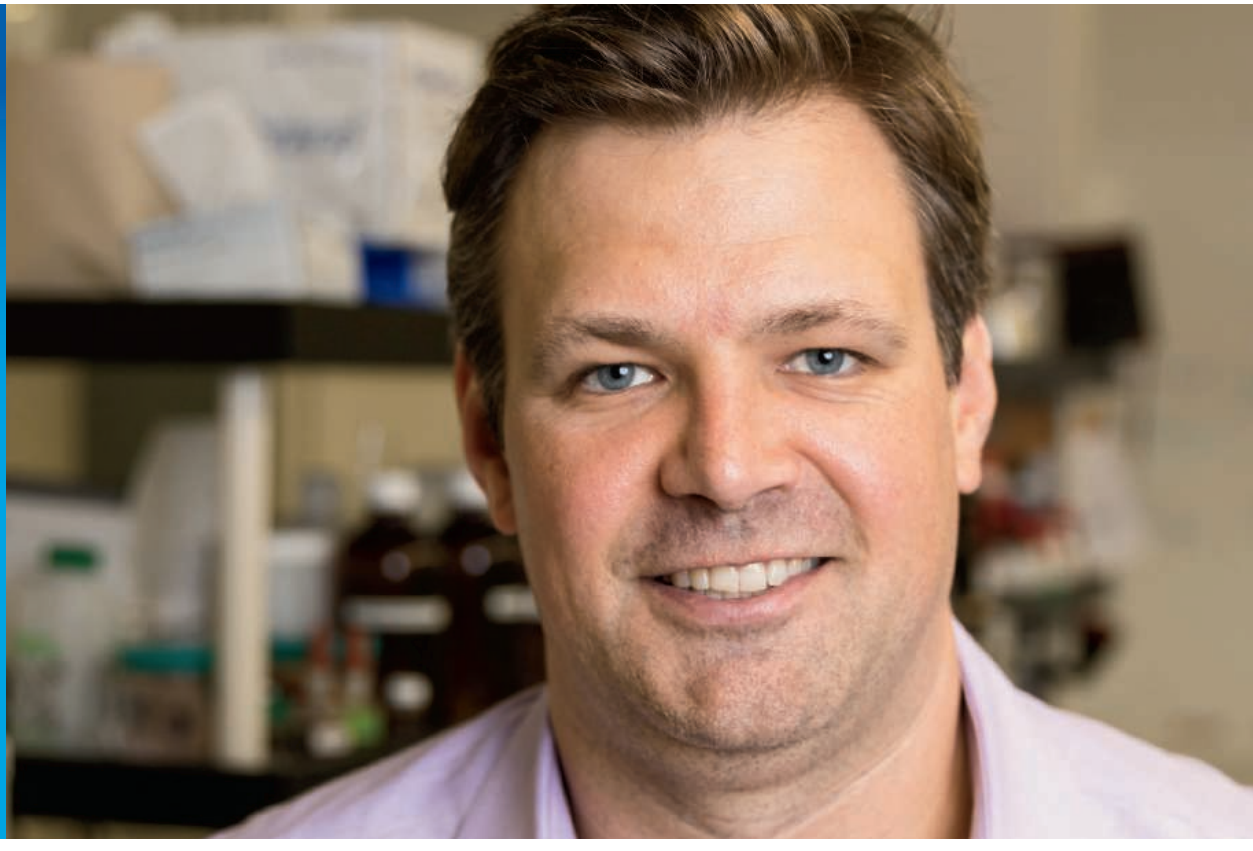
Chemist Will Dichtel, Northwestern University, is the 2018 recipient of the Cottrell Frontiers in Research Excellence and Discovery (FRED) Award from Research Corporation for Science Advancement. His work addresses the most important contemporary challenges at the interface of organic chemistry and materials science. Specifically, the $250,000 FRED Award goes to support Dichtel's pioneering efforts to develop new methods to polymerize diverse monomers into ordered, two-dimensional polymers with unprecedented control of their covalent structure, crystallite size and shape, and higher-order assembly.

The Cottrell FRED Award supports early stages of exceptional, high-risk/high-reward research that may potentially transform a field of scientific inquiry. It is presented to highly creative Cottrell Scholars whose ideas and potential solutions, though not readily funded by conventional grant programs, address major current challenges in the recipients' areas of research expertise.

By developing unique perspectives for solving key challenges, Cottrell FRED awardees create new approaches that accelerate basic science research for the benefit of society.

Since 1994 the Cottrell Scholar program has developed outstanding teacher-scholars recognized by their scientific communities for the quality and innovation of their research programs and their academic leadership skills

2018 Cottrell Frontiers in Research Excellence and Discovery Awardee



WILL DICHTEL
Professor of Chemistry
Northwestern University

RESEARCH CORPORATION
for SCIENCE ADVANCEMENT
A foundation dedicated to science since 1912.

gairdner
LES PRIX CANADA GAIRDNER AWARDS

2018 GAIRDNER SYMPOSIA

Gairdner.org/events-upcoming

2018 CANADA GAIRDNER LAUREATE LECTURES

OCTOBER 25, 2018
9:00 AM – 3:00 PM

MacLeod Auditorium, Medical Sciences Building, University of Toronto

LIVE WEBCAST:
bit.ly/Gairdner2018

OPTOGENETICS

Peter Hegemann
2018 Canada Gairdner International Award Laureate

Karl Deisseroth
2018 Canada Gairdner International Award Laureate

Edward S. Boyden
2018 Canada Gairdner International Award Laureate

LEADERSHIP IN ONCOLOGY

Frances A. Shepherd
2018 Canada Gairdner Wightman Award Laureate

GENOMIC IMPRINTING

Davor Solter
2018 Canada Gairdner International Award Laureate

Prof. Azim Surani
2018 Canada Gairdner International Award Laureate

GLOBAL HEALTH "GLOBAL BURDEN OF DISEASE"

Alan D. Lopez
2018 John Dirks Canada Gairdner Global Health Award Laureate

Christopher J.L. Murray
2018 John Dirks Canada Gairdner Global Health Award Laureate

LET THERE BE LIGHT: Optogenetics in Neuroscience and Beyond

OCTOBER 24, 2018
9:00 AM – 1:00 PM

MacLeod Auditorium, Medical Sciences Building, University of Toronto

LIVE WEBCAST:
bit.ly/Gairdner2018

FEATURING

Leslie Vosshall
Rockefeller University
(New York, NY, USA)

Hailan Hu
Zhejang University
(Hangzhou, China)

Yang Dan
University of California Berkeley
(Berkeley, CA, USA)

Susuma Tonegawa
MIT
(Cambridge, MA, USA)

Jared Toettcher
Princeton University
(Princeton, NJ, USA)

Sheena Josselyn
Hospital for Sick Children
(Toronto, ON, CAN)

HUMAN VACCINES PROJECT

MICHELSON MEDICAL RESEARCH FOUNDATION

MICHELSON PRIZES

EXPANDED CALL FOR ENTRIES

The Michelson Prizes are scientific awards that support young investigators from across the globe applying disruptive research concepts to significantly advance the development of vaccines and immunotherapies for major global diseases.

2019 Prizes will be awarded in three focus areas, all related to vaccine discovery:

- *Human Immunology*
- *Computational Biology and Protein Engineering*
- *Neglected Parasitic Diseases*

Proposals will be reviewed by a distinguished committee of internationally-recognized scientists.

Researchers under the age of 35 are invited to apply.

APPLY TODAY

www.humanvaccinesproject.org/MichelsonPrizes

DEADLINE OCTOBER 29, 2018

Prize Money of $150,000

Opportunity to present at the annual conference on *'The Future of Vaccine Discovery'*

CHALLENGE INEQUALITY

NTNU is Norway's leading science and technology university. Here, equality, stability and long-term government research funding create freedom to innovate, collaborate and explore new ideas. Together, we challenge — challenge established truths, naysayers and ourselves. We create knowledge for a better world.

www.ntnu.edu

NTNU

Knowledge for a better world

eppendorf & Science

PRIZE FOR NEURO BIOLOGY

And the Winner Is...

Eppendorf & Science Prize for Neurobiology
Congratulations to Johannes Kohl on winning the 2018 Eppendorf & Science Prize for his work on neural mechanisms underlying parental care. Dr. Kohl's research has revealed how a small population of genetically defined neurons controls the motor, motivational, hormonal and social aspects of parental behavior in males and females. These findings provide a new model for how specific components of a social behavior are generated at the neural circuit level. Unravelling the functional architecture of such circuits will advance our understanding of how the brain coordinates complex behaviors.

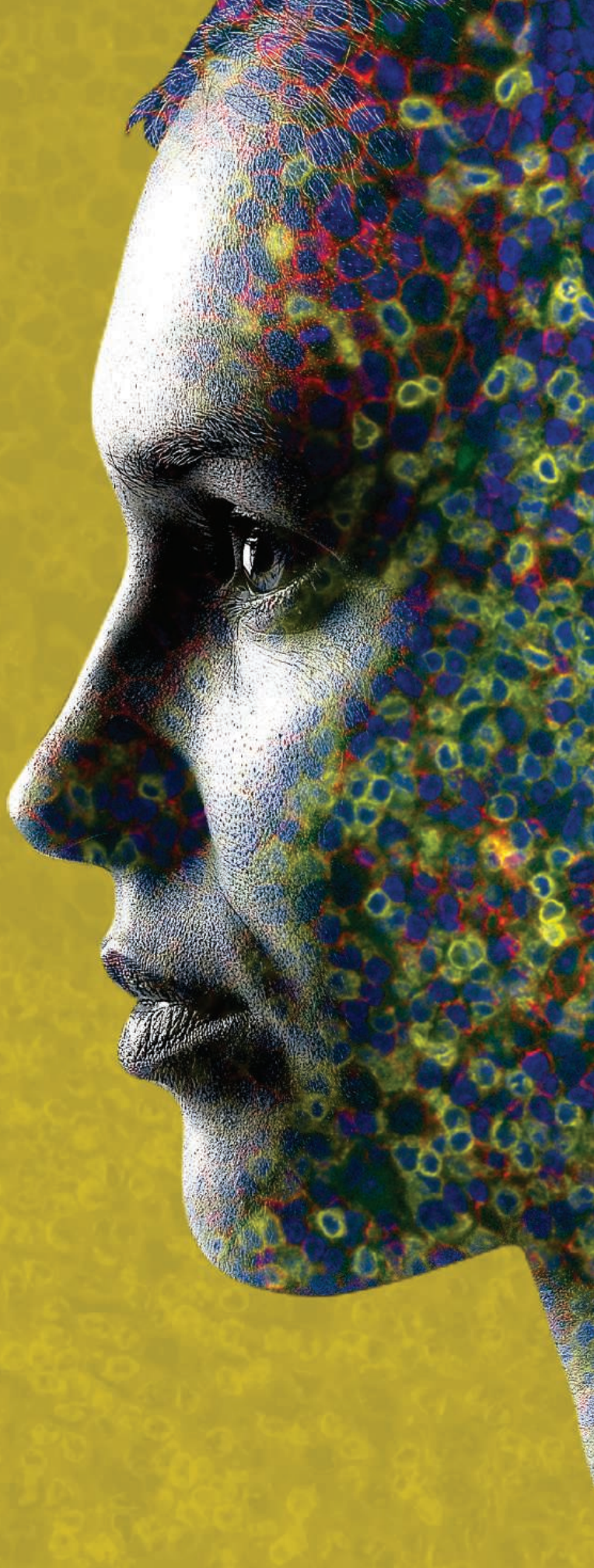
The annual US$25,000 Eppendorf & Science Prize for Neurobiology honors scientists, like Dr. Kohl, for their outstanding contributions to neurobiological research. Johannes Kohl is the 17th recipient of this international award. He will be presented with the Prize at a ceremony held during the week of the 2018 Annual Meeting of the Society for Neuroscience in San Diego.

You could be next to win this prize.
If you are 35 years of age or younger and currently performing neurobiological research, you could be next to win the 2019 Prize. Deadline for entries is June 15, 2019.

Learn more at: www.eppendorf.com/prize

eppendorf Science AAAS

Ignite your breakthrough.



Ensure reproducible results with antibodies that are 100% guaranteed to work.

Every great discovery requires a bold approach and reliable tools. With each cell-labeling experiment, you can trust in our more than 45 years of experience in antibody manufacturing and on-site validation. You need antibodies with performance that exceeds the 50% failure rate common in the industry.* That's why, from polyclonals to recombinant rabbit monoclonal abs, we guarantee 100% target specificity and sensitivity to help you continue moving toward your next breakthrough. Our antibodies have been trusted in over 10,000 independent citations in the last 15 years. Because at Bethyl, we put a lot in every drop.

See our data at bethyl.com/breakthrough

Weller, MG, Analytical Chemistry Insights:11, 21-27 (2016).

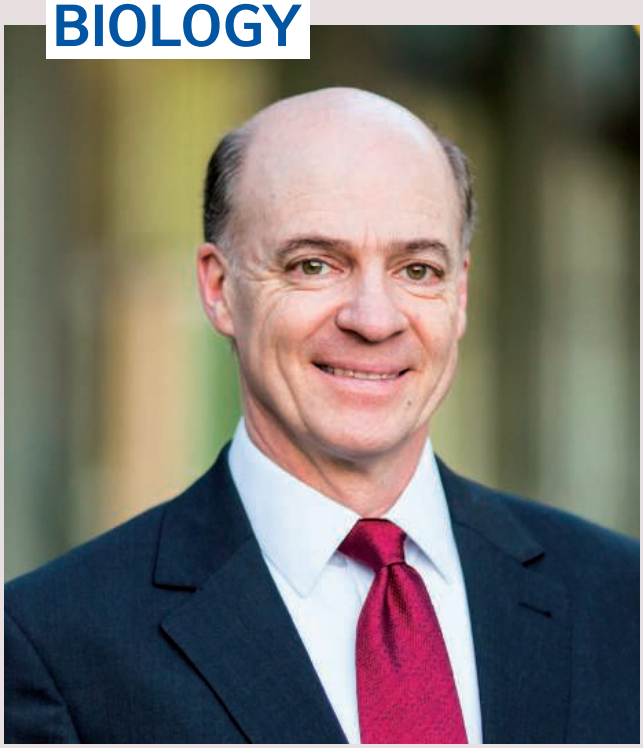
Antibodies shown: Rabbit anti-CD3e recombinant monoclonal (yellow, A700-016) & mouse anti-CD20 monoclonal (red, A500-017A) in FFPE lymph node.

BETHYL
LABORATORIES, INC

Really Good Antibodies

eppendorf & Science

PRIZE FOR NEURO BIOLOGY



Dr. Peter Stern, *Science* Editor & Prize Jury Chair

Dr. Flavio Donato, 2017 Prize Winner

Neuroscience 2018, San Diego
Sunday, November 4, 2018
2:00 – 3:00 pm, Eppendorf Booth #1201

Meet & Greet

Neuroscience 2018
The Eppendorf & Science Prize for Neurobiology is an international research prize of US$ 25,000. It is awarded annually to one young scientist of 35 years of age or younger for the most outstanding neurobiological research based on methods of molecular and cell biology.

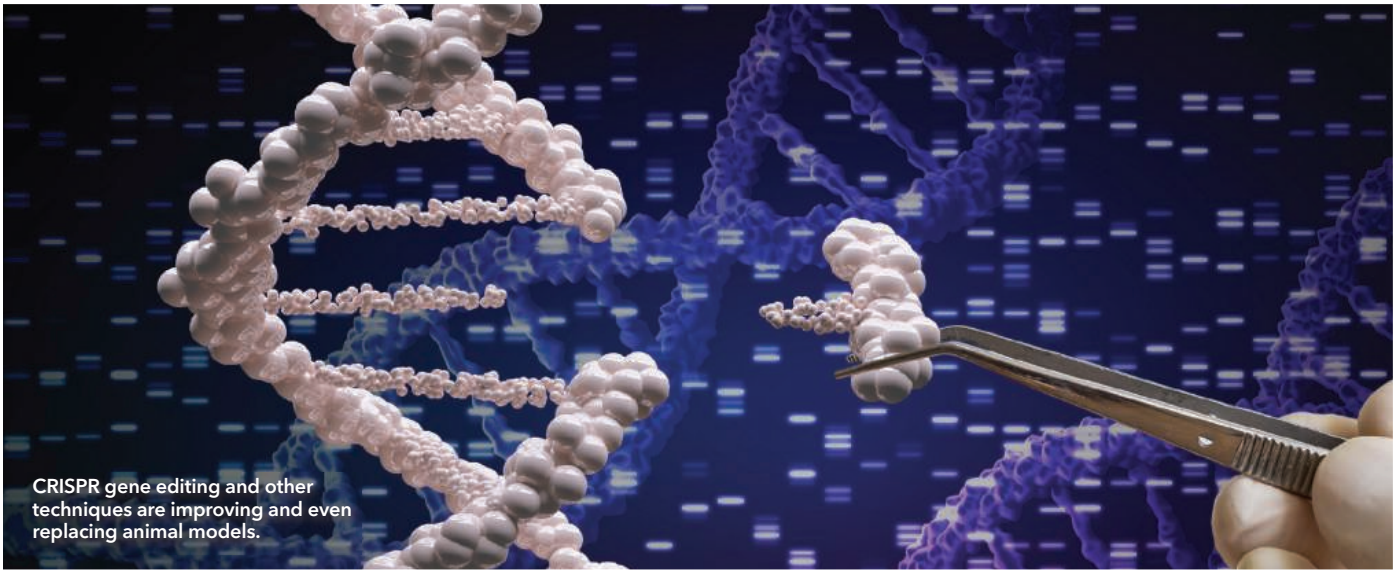
Learn more about the prize at:
www.eppendorf.com/prize

You could be next to win this prize.
Stop by the Eppendorf booth #1201 and meet the 2017 Winner, Dr. Flavio Donato from the Norwegian University of Science and Technology in Trondheim. Learn more about his research on how neural networks mature during development to represent space in the brain. Ask the Prize Jury Chair, Dr. Peter Stern, about the prize application process and what makes a winning paper.

Register for prize updates to receive an Eppendorf & Science Prize for Neurobiology t-shirt*!

*Subject to prize eligibility while supplies last.

eppendorf Science AAAS



CRISPR gene editing and other techniques are improving and even replacing animal models.

Replacing the replacements: Animal model alternatives

When it comes to mimicking human disease or predicting the human body's response to candidate drugs, traditional laboratory animal models are woefully inadequate. New technologies—3D cell culturing, human induced pluripotent stem cells, and gene editing—are leading to new solutions for replacing, refining, and reducing animal models. **By Kendall Powell**

Six years ago, Ping Yeh's oncologist told him his Hodgkin's lymphoma was resistant to the standard chemotherapy regimen and he would need a more potent, seven-drug cocktail. The treatment knocked his cancer into remission, but afterward Yeh needed an ultrasound to check whether the treatment, which can be cardiotoxic in some patients, had damaged his heart.

"The treatment could have cured me. Or cured me and killed me," says Yeh, a nanotechnologist. "It was a pretty scary experience and it planted the seed for me to figure out if there was a better way to test for drug safety." That seed sprouted into Minneapolis-based **StemoniX,** which Yeh cofounded in 2014 for the purpose of combining advances in engineering, manufacturing, and human stem cells to develop drug screening and testing platforms with more relevance to human physiology.

Fortunately, Yeh's heart was spared. Unfortunately, patients like Yeh encounter toxic or ineffective drugs all too often, because the animal models used to test drugs before they go into patients are imperfect in many ways. Of all the drugs that enter clinical trials, only about 10% go on to be approved (*1*). The other 90% fail during trials—for reasons ranging from off-target, undesirable effects and problematic dosing, to low or no efficacy, and the worst-case scenario, toxicity.

"The closer you can get to mimicking the human situation, the better the research is going to be in [terms of] understanding the fundamental pathology of disease and also in predicting patient efficacy and toxicity of drug therapies," says Richard Eglen, vice president and general manager of **Corning Life Sciences** in Boston, Massachusetts.

However, the challenge will be to design models that hold significant advantages over current approaches. That means producing models that give robust, reproducible data, that are predictive of human biology, and that will not greatly increase costs. StemoniX, Corning, and several other companies are coming up with innovative ways to meet that challenge.

Imposing structure on cultures

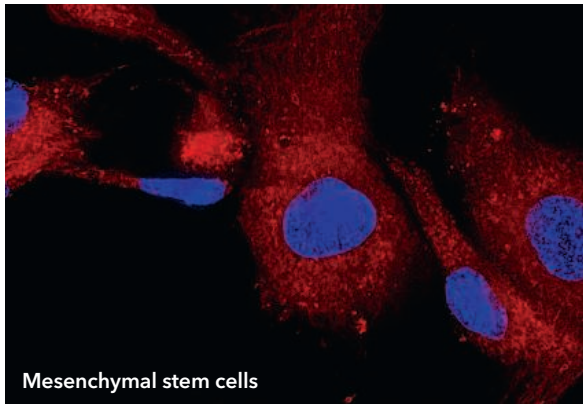
Culturing methods have continued to improve, and are now yielding 3D cell cultures, spheroids, and even more complex organoids that more accurately reflect human tissues. Corning has led the way in providing researchers with the best surfaces on which to grow these cultures. Now, the company has added a 1,536-well spheroid plate that can work at the highest levels of automated screening. The rounded-bottom wells, which are coated with an ultralow attachment surface, encourage plated cells to aggregate with each other and form a sphere.

"That is going to enable researchers to do ultra-high-throughput screening of up to 100,000 compounds per day," says Eglen. The plates have small well volumes and therefore produce small spheroids that range from 500 to 2,000 cells, but they give reproducible responses that are linear relative to cell number.

A group at the Scripps Research Institute in La Jolla, California, has used the plates with pancreatic tumor spheroids to identify an inhibitor of those cells' RAS oncogene mutation. A parallel 2D-cell-culture–based screen did not turn up the same inhibitor (*2*). Corning's own research using known liver toxins on human liver spheroids shows they can be used to screen cont.>

compounds for liver toxicity. "What everyone would like is the ability to better predict the metabolism of drugs before putting them into the patient," says Eglen.

Corning has developed another technology to support the 3D growth of mesenchymal stem cells (MSCs). Researchers would like to develop MSCs—also known as adult stem cells—into individualized, cell-based patient therapies, such as replacements for faulty insulin-producing islet cells in diabetes, for instance. But it's been a challenge growing enough MSCs under pristine conditions to then administer the cells back to the patient. Corning's digestible microcarrier technology uses an inert, nonanimal-derived polymer to grow the cells in a 3D orientation in solution on the microcarrier's surface. Then, by adding an enzyme, the microcarrier can be dissolved away to leave just the MSCs for isolation.


Mesenchymal stem cells

Researchers would like to develop MSCs—also known as adult stem cells—into individualized, cell-based patient therapies, such as replacements for faulty insulin-producing islet cells in diabetes.

Human cell mass production

StemoniX offers "off-the-shelf" structured 2D microHeart and microBrain plates, as well as microBrain 3D spheroid plates, for high-throughput screening that offers precise measurement of human tissue responses. The company chose to focus on brain and heart cells in part because neurotoxicity and cardiotoxicity are the top two reasons why drugs fail for safety reasons during clinical trials.

The microHeart cultures are grown along microchannels to form sarcomeric unit structures like those found in heart muscle—complete with physical markers, correct ion-channel formation, and a unidirectional contraction.

The microBrain 2D and 3D cells are a mix of astrocytes and neurons that form synapses as well as being rudimentary neural networks. The 2D culture is key for measuring visual changes in the cells, such as neural projections. The 3D spheroid cultures are harder to visualize, but they exhibit spontaneous, synchronous neural firing, which can be quantitatively screened for drug responses that influence that firing. Similarly, the cardiomyocyte contractions can be quantitatively assayed for drug responses, such as arrhythmias.

"No model's perfect," acknowledges Yeh, but he finds StemoniX's heart and brain models extremely useful, "because they are human cells that are structured correctly, easily measurable, and give more predictive and reproducible data at a fraction of the cost of transgenic animals."

Microenvironment mimics

Micropatterning allows researchers to add sophistication to cultures, for example, by manipulating signaling molecules or growth factors in specific patterns or gradients, but with more experimental control to influence signaling than in whole-animal models.

While micropatterning techniques have been available for use with 2D cultures, the current techniques are crude when it comes to designing gradients or layering two or more proteins. One such technique, microcontact printing, uses a stamp dipped into protein solution to imprint it onto a culture surface, but lacks precision for printing specific quantities or aligned proteins. To address those limitations, a trio of academic researchers at France's National Center for Scientific Research (CNRS) dreamt up PRIMO, a contactless and maskless technique for micropatterning, and started up **Alvéole.**

PRIMO is based on light-induced molecular adsorption of proteins (LIMAP) and combines a UV illumination system with a photoactive reagent. The PRIMO box can be attached to an inverted microscope. It contains a UV laser and a micromirror device projector that can project any image file from a computer onto the surface of culture plates through the microscope's objective lens, with a resolution of 1.2 µm.

Glass slides, coverslips, plastic plates, and also hydrogels can first be treated with polyethylene glycol (PEG) as an antifouling agent, then the photoactive phospholipid phosphatase (PLPP) photoinitiator reagent is added in solution. When combined with the projected UV light in the pattern of choice, the PLPP degrades the PEG, leaving a "hole" to be filled by the protein(s) of choice. PRIMO's Leonardo software controls the movements of the motorized microscope stage for aligning or shifting patterns. Also, using the software's 256 grayscale levels to vary the UV intensity allows researchers to control how much PEG is degraded and therefore how much protein adheres to the surface.

For 3D applications, protein micropatterns laid down on a coverslip can be transferred to the surface of a hydrogel, since illuminating hydrogels directly with UV light can change their rigidity. Researchers have also used PRIMO to coat the bottom and sides of a polydimethylsiloxane polymer microwell to grow hepatocytes in a particular 3D configuration. Another group used PRIMO to grow human induced pluripotent stem cells (hiPSCs) into astrocytes oriented along thin, straight lines to measure microtubule growth dynamics in a patient with the neurodegenerative disease Rett syndrome (*3*).

"Using PRIMO, we helped that group try different shapes and widths of lines in just a few weeks—with other micropatterning techniques it would have taken months," says Marie-Charlotte Manus, operational marketing manager at Alvéole in Paris.

The system is not well suited to high-throughput imaging or screening yet, says Manus, although the company is working on a version to use with multiwell plates. For now, individual researchers can use PRIMO to get closer to the in vivo microenvironment by optimizing the signal patterns that their cell cultures experience.

More precise CRISPR models

Gene editing technology is bringing a whole new level of genetic manipulation to human cell cultures, whether they are transformed cell lines, patient-derived primary cultures, or iPSCs. While the research world is not morally ready to consider making knockout humans yet, technologies like CRISPR/Cas9 allow

Featured participants

Alvéole
www.alveolelab.com

Corning Life Sciences
www.corning.com/worldwide/en/products/life-sciences.html

Horizon Discovery
www.horizondiscovery.com

StemoniX
stemonix.com

researchers to knock out specific genes quickly from human cell cultures and then screen for desired properties.

"In the postgenome era, we can interrogate biological systems in ways we have never thought about before," says Benedict Cross, head of functional genomic screening for **Horizon Discovery** in Cambridge, United Kingdom. "Taking a whole-genome, systems approach is not easy to do in a reasonably controlled way in whole animals."

Horizon's CRISPR Knockout Screening platform can knock out every single gene, or suites of genes related to a specific process, such as apoptosis, in a population of human cells, and then measure how those altered cells respond to drug compounds. It can be used to identify novel drug targets, to find genes related to drug sensitivity or resistance, or to select patients who are the best candidates for clinical trials.

Each CRISPR knockout guide RNA encodes a barcode unique to the gene it is knocking out, and all the cells are kept together as a mixed population of knockouts, whether they are grown in 2D monolayers, 3D spheroids, or in suspension. One set of cells is subjected to an experimental treatment, such as a drug, and another set is kept as a control. The treated cells that are enriched for the desired activity above the control cells are isolated and their barcodes sequenced to reveal their missing gene.

"It's quite precise and quantitative at that point," says Cross. The higher the frequency of a gene's barcode, the more likely the inhibition of that gene leads to the desired biological activity. "It's a phenotypic screen with a whole-genome level analysis."

Horizon has also tweaked the CRISPR/Cas9 system to create two other screening platforms, CRISPR Interference (CRISPRi) and CRISPR Activation (CRISPRa). CRISPRi knocks gene message levels down, but not entirely out, and CRISPRa amps up the targeted gene's expression.

All three CRISPR screening platforms can be used in parallel to study knockouts, loss-of-function, and gain-of-function of the same genes. "Interrogating a biological process is massively improved when you are able to look in both directions, to find things that activate and inhibit the pathway endogenously," says Cross.

Whole-animal editing

Through its acquisition of Sigma Advanced Genetic Engineering (SAGE) in 2014, Horizon Discovery has also deployed CRISPR technology to make improved and "humanized" rat models for drug discovery and safety testing.

As Kevin Forbes, manager of Horizon Discovery's in vivo R&D group in St. Louis, Missouri, explains, the rise of mouse models in drug discovery was tied to the ease of making genetic knockouts through embryonic stem cell (ESC) technology. However, prior to that, rats were the preferred models because their larger size allows for multiple blood draws and longer-term compound studies, and because their prosocial behavior and slower metabolisms are more analogous to human biology.

Horizon can now make custom gene-edited rat, mouse, and rabbit models, with modifications ranging from small point mutations to megabase deletions, including knockout–knockin models in which the animal gene is replaced with the human version of a gene. Horizon also has "off-the-shelf" rat knockout models for use in toxicology, oncology, and cardiovascular disease studies as well as models for research into neurological disorders, including autism, Parkinson's disease, and Alzheimer's disease.

One rat model allows the study of drug absorption, distribution, metabolism, excretion, and toxicology (ADME/Tox), because it has been gene-edited to remove the rat versions of three nuclear receptors, PXR, CAR, AHR, and the major drug-metabolizing enzyme cytochrome P450 CA4. The receptors act as xenobiotic sensors and turn on the cytochrome P450 genes to metabolize "foreign" drugs. In these ADME/Tox rats, the human versions of those four genes were edited back in. Those rats, in theory, should "see" novel drugs in the same way that human small-intestine and liver cells would detect drugs to set their metabolism in motion.

Forbes describes CRISPR/Cas9 as a more efficient and faster process for making complex transgenic models such as the ADME/Tox rat when compared to ESC-based transgenic methods, estimating that CRISPR shaves off as much as a year from the process.

These "humanized" rats will soon provide a whole-animal solution for preclinical drug safety testing that is much more advantageous than testing on human liver cells in vitro. For example, cell cultures won't catch tumors formed elsewhere in the body, nor pick up cardiac or cognitive problems caused by drugs.

Of course, drug development may never eliminate the use of animals, because they represent the biological complexity of integrated organ systems, Yeh notes. But these new technologies should greatly reduce both numbers and types of animals needed for drug testing in the future.

As models improve in their "humanness," drug developers should be able to better predict how compounds will act in the human body—avoiding the problems that doom so many candidate drugs and harm patients like Yeh. Inspired by his recent close call, he keeps working to build better models: "If a model has enough of the right parts, connected correctly, then we'll continue to improve the functionality."

References

1. M. Hay *et al., Nat. Biotech.* **32**, 40–51 (2014).
2. S. Kota *et al., Oncogene* **37**, 4372–4384 (2018); doi: 10.1038/s41388-018-0257-5.
3. C. Delépine *et al., Hum. Mol. Genetics.* **25**, 146–157 (2016); doi: 10.1093/hmg/ddv464.

Kendall Powell is a freelance science writer based in Lafayette, Colorado.

In Vivo Imaging Systems

Leveraging hardware and software advances and smart accessory design, the IVIS Lumina S5 and X5 high-throughput benchtop 2D optical in vivo imaging systems provide researchers a streamlined workflow to expedite the understanding of disease progression and to help develop treatments for a wide range of cancers, infectious diseases, and other disorders. Both systems provide high-sensitivity, high-throughput bioluminescence as well as and fluorescence imaging with spectral unmixing. IVIS Lumina X5 integrates high-resolution X-ray imaging, allowing scientists to explore multiple facets of a disease by providing both molecular-level and highly detailed anatomical information in a single image.

PerkinElmer
For info: 800-762-4000
www.perkinelmer.com/invivo

Natural Killer Cell Expansion Medium

PRIME-XV NK Cell CDM from Irvine Scientific is a chemically defined (CDM), animal component-free medium for the ex vivo expansion of natural killer (NK) cells. Designed for use in NK cell–based immunotherapy research and translational applications, this advanced medium delivers strong cell growth and maintains NK cell potency and functionality without the need for serum supplementation in either the presence or absence of feeder cells. Our innovative formula helps scientists avoid many of the disadvantages associated with the use of feeder cells while maintaining the characteristic morphology and cytotoxicity of NK cells.

Irvine Scientific
For info: 800-577-6097
www.irvinesci.com

Pulse Oximeter

The MouseOx Plus Pulse Oximeter for rodents provides real-time, continuous monitoring of oxygen saturation. This system can be used with either anesthetized or conscious animals. An array of sensors is available, making this system appropriate for rodents sized from neonatal mice to adult rats. MouseOx Plus has a patented design enabling it to detect heart rates of 90 bpm–900 bpm. Many options are available, allowing for a highly customizable system. MouseOx provides the following measurements: arterial oxygen saturation, heart rate, temperature, pulse distention, and breath distention.

Harvard Apparatus
For info: 800-272-2775
www.harvardapparatus.com

Micropatterned Hepatocyte Co-Cultures

HepatoPac Kits are in vitro tools used for predicting likely in vivo outcomes and mechanisms of toxicity with respect to transport and metabolism of drugs and chemicals. HepatoPac plates contain micropatterned co-cultures of primary hepatocytes and stromal cells. This technology replicates the physiological microenvironment of the liver and allows hepatocytes to exhibit normal metabolic activity for over four weeks, for both short- and long-term toxicology and efficacy studies during preclinical drug discovery. Hepatocyte health, functionality, and liver enzyme activity can be extensively characterized on the HepatoPac platform. HepatoPac products are available with human, rat, dog, or monkey hepatocytes, or in multispecies formats.

BioIVT
For info: 516-483-1196
www.bioivt.com

CAR T-Cell Immunotherapy Reagents

AMS Biotechnology produces a growing range of experimental cell lines, recombinant proteins, and screening services to advance the search for new chimeric antigen receptor (CAR T) cell therapies to treat cancers. A therapeutic CAR is a transmembrane protein designed with an extracellular domain based on an antibody single-chain variable fragment (scFv) and intracellular signaling domains derived from the T-cell receptor gamma (TCRg) chain, along with other costimulatory receptors. The scFv provides a specific binding domain that recognizes target proteins on cancer cells. A patient's own T cells are isolated and activated, then transfected with a gene expressing the CAR. This reprograms the T cells to identify and attack tumor cells expressing the target protein, creating personalized immune cells designed to specifically target the patient's cancer.

AMS Biotechnology
For info: 617-945-5033
www.amsbio.com/car-t-cell-research.aspx

Animal-Free Recombinant Proteins

Cell Sciences offers Animal-Free Recombinant Proteins for researchers concerned about trace amounts of contaminating growth factors and infectious agents that can sometimes be found in commercially available purified proteins. These proteins are produced with specific prokaryotic or eukaryotic expression systems in a dedicated animal-free laboratory, to ensure these reagents are not exposed to potential contamination by animal components or byproducts during production. All procedures use animal-free media, with dedicated labware and equipment. These products are ready for preclinical research use in life science fields, such as developmental biology, oncology, rare diseases, immunology, and hematology. They are delivered priority overnight in lyophilized or liquid form and come with quality control documentation that includes details on purity, endotoxin levels, and biological activity.

Cell Sciences
For info: 888-769-1246
www.cellsciences.com

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.

gsk do more feel better live longer

Shape the future

GSK has been helping millions of people do more, feel better and live longer for over 300 years. Our goal is to be one of the world's most innovative companies, delivered by people who are proud to say they work in R&D at GSK.

Our new approach to R&D has at its core science related to the immune system, and will leverage human genetics, functional genomics and advanced analytics to improve the probability of successfully developing new medicines. Used on a genome-wide scale, these technologies have the power to uncover insights either previously hidden within obtainable genetic data or out of the reach of evolution. We are poised to transform drug discovery by taking a deeper look into the genetic clues behind disease, which has never before been possible.

Our early experience with UKBioBank, Open Targets and The Altius Institute, and our world-leading internal capabilities, position GSK to respond quickly. We recently signed an exclusive partnership with 23andMe giving us access to one of the largest sets of genotypic and phenotypic information.

Join our team

We are continuing to build our capability and are looking for two outstanding senior leaders in Functional Genomics and Human Genetics to put us at the forefront of this technology revolution. We are also looking for scientists at all levels as we grow and invest both teams. Are you passionate about helping patients, have the courage to take smart-risks, and have ambitions to develop and grow in a company where you will be supported?

To apply please go to gsk.com/careers and search for 'genetics'

Illustrating use of digital technologies to understand biology

linkedin.com/company/glaxosmithkline

facebook.com/GSKcareers

youtube.com/GSKvision

@GSK_careers

R&D SYSTEMS
a bio-techne brand

Consistency Is

N21-MAX and N-2 Neural Media Supplements from R&D Systems

Each lot of serum-free media is checked for performance consistency by our in-house quality team.

Learn more | rndsystems.com/neuralmedia

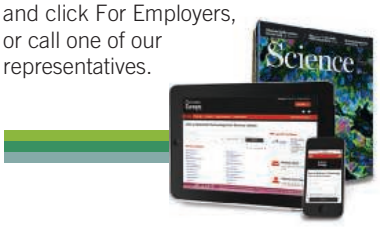
bio-techne®

Global bio-techne.com info@bio-techne.com TEL +1 612 379 2956 **North America** TEL 800 343 7475
Europe | Middle East | Africa TEL +44 (0)1235 529449 **China** info.cn@bio-techne.com TEL +86 (21) 52380373

Science Careers

SCIENCE CAREERS ADVERTISING

For full advertising details, go to ScienceCareers.org and click For Employers, or call one of our representatives.



AMERICAS

+1 202 326-6577
+1 202 326-6578
advertise@sciencecareers.org

EUROPE, INDIA, AUSTRALIA, NEW ZEALAND, REST OF WORLD

+44 (0) 1223 326527
advertise@sciencecareers.org

CHINA, KOREA, SINGAPORE, TAIWAN, THAILAND

+86 131 4114 0012
advertise@sciencecareers.org

JAPAN

+81 3-6459-4174
advertise@sciencecareers.org

CUSTOMER SERVICE

AMERICAS

+1 202 326-6577

REST OF WORLD

+44 (0) 1223 326528

advertise@sciencecareers.org

ScienceCareers
FROM THE JOURNAL SCIENCE AAAS

ScienceCareers.org

McGOVERN INSTITUTE
FOR BRAIN RESEARCH AT MIT

Call for Nominations: Scolnick Prize in Neuroscience

The McGovern Institute for Brain Research is accepting nominations for the 16th annual Edward M. Scolnick Prize in Neuroscience. The Prize recognizes an outstanding discovery or significant advance in the field of neuroscience. The prize is $150,000. The recipient presents a public lecture at MIT, hosted by the McGovern Institute and followed by a dinner in Spring 2019.

Nomination Deadline: December 15, 2018

Nomination procedures: Candidates for the award must be nominated by individuals affiliated with universities, hospitals, medicals schools, or research institutes, with a background in neuroscience. Self-nomination is not permitted. Each nomination should include:

- A biosketch or CV of the nominee;
- A letter of nomination with a summary and analysis of the major contributions of the nominee to the field of neuroscience.
- Up to two representative reprints will be accepted.

Selection Procedure:

- Members of the selection committee and faculty affiliated with MIT are not eligible.
- Announcement of the award recipient will be made in January 2018
- Recipient must attend all events to be awarded the prize.

Past Scolnick Prize Recipients:
Dr. Masakazu Konishi, Dr. Judith L. Rapoport, Dr. Michael E. Greenberg, Dr. David Julius, Dr. Michael Davis, Dr. Jeremy Nathans, Drs. Lily and Yuh-Nung Jan, Dr. Bruce McEwen, Dr. Roger Nicoll, Dr. Thomas Jessell, Dr. Huda Zoghbi, Dr. Charles Gilbert, Dr. Cornelia Bargmann, Dr. Catherine Dulac, Dr. David J. Anderson,

Send nomination packet to: gwolf@mit.edu or Attn: Scolnick Prize Nomination, McGovern Institute for Brain Research, Massachusetts Institute of Technology, 77 Massachusetts Avenue 46-3160, Cambridge, MA 02139.

For more information: http://mcgovern.mit.edu

POSITIONS OPEN

SEATTLE PACIFIC UNIVERSITY
ASSISTANT PROFESSOR OF BIOLOGY

Tenure-track position in Biology to start September 1, 2019. The optimal candidate will have expertise in cell biology, physiology or systematics and be able to teach at least two of the following courses: Cell Biology, Neurobiology, Physiology or Human Anatomy & Physiology. Candidates should provide evidence of excellence in teaching and mentoring undergraduates, a focused research program appropriate for undergraduate involvement, and the ability to advise and mentor underrepresented, and first-generation students. A Ph.D. is required by the start date. Members of historically underrepresented groups, as defined by the National Science Foundation, are encouraged to apply. SPU has been committed to building diversity since 1991 resulting in 50% of incoming freshman belonging to underrepresented groups in 2018. SPU is a Christian university, and applicants will submit a faith statement with their application. The online application requests a cover letter specifically addressing the mentoring of underrepresented students, a Curriculum Vitae, an application form, and three, one-page statements--of faith, teaching philosophy, and research goals. Applications should be submitted online at website: **https://tinyurl.com/SPUJOB**. Informal inquiries can be made to Dr. Derek Wood, Chair, Department of Biology, woodd1@spu.edu.

UT Health
San Antonio

Postdoctoral Fellowships in Texas

Multiple postdoctoral research positions are immediately available in the laboratory of James Stockand, PhD at the University of Texas Health Science Center at San Antonio. Positions will be funded with competitive salaries supported by National Institutes of Health Grants R01 DK987460, R01 DK117909, R01 DK103758 and T32 HL007446. Highly motivated early-career scientists interested in neuroscience, ion transport, renal physiology, cardiovascular biology and/or mechanisms of hypertension are encouraged to apply. Priority will be given to those with experience with electrophysiology. At the time of appointment, candidates must hold an advanced professional degree (e.g. MD or PhD). To apply, please send a curriculum vitae and cover letter, which includes a brief statement of interest and qualifications and the names of three professional referees, to Dr. Stockand at **Stockand@uthscsa.edu**.

All postdoctoral appointments are designated as security sensitive positions. The University of Texas Health Science Center at San Antonio is an Equal Employment Opportunity/Affirmative Action Employer, including protected veterans and persons with disabilities.

Shenzhen: The New Center for China's Top Schools

In just 40 years, Shenzhen has grown from a small, unknown fishing village to a thriving megalopolis, and is now one of the world's most commercially competitive cities (referred to collectively along with Beijing, Shanghai, and Guangzhou as "BeiShangGuangShen"). But if you asked people on the streets of China a few years ago about the Shenzhen Special Economic Zone—the country's first—from which came the well-known phrase "Shenzhen Speed," they would tell you that Shenzhen was simply an "upstart" with no cultural foundation. Because of their rigid opinion, they would probably say, "There is no culture in Shenzhen, not even a few universities. What it has most is money."

Those who hold this view, which Shenzhen's citizens reject as "sour grapes," may be embarrassed now, because the city's financial gains are gradually helping to fill its educational vacuum. By February 2018, 17 of China's highest-level universities, including Tsinghua University, Peking University, Shandong University, Sun Yat-sen University, Tianjin University, and Xiamen University—which account for almost half of the best universities in China—had established a presence in Shenzhen in the form of research institutes, graduate schools, branch campuses, etc. That number does not include famous universities, such as the University of the Chinese Academy of Sciences, Shanghai Jiaotong University, and Renmin University of China, which have signed agreements with Shenzhen. At the pace of current development, most of China's Project 985 colleges and universities will have outposts in Shenzhen in the future, among them one former Project 211 university and six Hong Kong universities, including the Chinese University of Hong Kong.

Attracting talent with money and vision

As part of China's overall talent recruitment strategy, other cities are rushing to acquire talent—Shenzhen, however, is not only fighting to procure talent, but colleges and universities as well. In China's traditional higher education system, colleges and universities are typically clustered in municipalities and capital cities, but Shenzhen is slowly trying to change this situation. The logic behind this is very clear: Colleges and universities are one of the most important reservoirs for high-level talent in cities. In the first half of 2018, Shenzhen gained 38 full-time academicians, most of whom are in colleges and universities. During March to September of this year, three of the four Nobel Prize winners who have moved to Shenzhen also signed with universities, including the Chinese

Shixin Wang
Deputy Chief Editor of China Education Online, Chief Executive Editor of AcaBridge

University of Hong Kong (in Shenzhen).

Why is Shenzhen favored by China's top universities? The figures tell the story. The city's fiscal budget for education in 2018 is RMB 63.4 billion (USD 9.3 billion), accounting for about 16% of the city's fiscal expenditure, an increase of 25% over 2017. On the official website of the Shenzhen Science and Technology Innovation Commission, a public list of funding for 2018 scientific research indicates that projects supporting local universities in Shenzhen, plus famous universities and scientific research institutions at home and abroad cooperating with Shenzhen (including 17 foreign Project 985 universities, one Project 211 university, and six Hong Kong universities), account for about 70% of funded projects. Tsinghua University and the Harbin Institute of Technology in Shenzhen have received nearly RBM 30 million (USD 4.4 million) in grants. The graduate school and research institute of Peking University in Shenzhen have received more than RMB 25 million (USD 3.65 million). The Shenzhen Institute of Hong Kong University of Science and Technology has received RMB 13.3 million (USD 1.9 million).

These figures certainly explain why local universities in Shenzhen are developing so rapidly. The research funding for Shenzhen University is nearly RMB 90 million (USD 13.1 million), and Southern University of Science and Technology is receiving more than RMB 30 million (USD 4.4 million). Most remarkably, Shenzhen Institute of Information Technology, a vocational college, receives a research grant of up to RMB 10.8 million (USD 1.6 million), which is higher than the funding for many Project 985 colleges and universities in Shenzhen.

These figures confirm an aphorism that applies to the development of colleges and universities: "Money may not be omnipotent, but no money makes things impossible." Of course, Shenzhen's success in this area cannot be simply attributed to money, but rather to Guangdong Province and Shenzhen's forward-looking strategic vision and various supporting policies, including funds, talent acquisition, and scientific research.

Thinking globally, acting locally

China officially proposed to build world-class universities and first-class disciplines in 2015. In April of that year, Guangdong had taken the lead in launching the construction of high-level universities in science and engineering in China (referred to as "Double High" for short), uniting the formation of national Double First-Class universities with that of Double-High universities, and dealing quickly with all aspects of organization, funding, and project management. The province invested more than RMB 30 billion (USD 4.4 billion) in Double-High universities in three years. The speed of this implementation is undoubtedly amazing. By contrast, three years later, there are still many provinces in China that are more than willing to construct Double First-Class universities, but lack the necessary resources. In these provinces, the greatest asset is undoubtedly their talent recruitment policy. However, with no high-level personnel handling hardware and material accumulation, the policy amounts to little more than a "castle in the air."

Last year, Guangdong issued the document "Opinions on Deepening the Reform of Mechanisms for Talent Development System Implementation," and proposed a series of measures to promote the establishment of a globally competitive talent system and accelerate the construction of special areas where talent can live and work. Shenzhen has been formulating more active and open talent policies, ranging from the "1 + 6" policy for high-level professionals, to the "Peacock Plan" for introducing highly skilled talent from overseas, to the Ten Thousand Talents Project and the Talent Work Regulations. The "four beams and eight pillars" of the talent policy have been continuously improved and perfected, and more policy refinements are underway. Under these policies, the number of high-level experts in Shenzhen is now over 10,000, and the total number of professionals in the city exceeds 5.1 million.

Thanks to well-organized and firm implementation, the talent development policies have achieved excellent results. The most direct proof of this is that Zhongshan University, South China University of Technology, Jinan University, Guangzhou University of Traditional Chinese Medicine, and South China Normal University are now listed as Double First-Class universities, making Shenzhen seventh in the country in terms of the number of these institutions.

By May 2017, 545 state-level experts and 371 provincial experts had been acquired by Shenzhen's Double-High universities; 50 disciplines had entered the top 1% of the Essential Science Indicators (ESI) database, with five subjects entering the top 1%; and 2,024 projects had been funded by the National Natural Science Foundation, with a growth of 10%. The number of approved projects in eight of Shenzhen's universities ranked among the top 100 in the nation's colleges and universities, marking the greatest achievement in the city's history.

Recognizing local universities

Local universities in Shenzhen have also benefited enormously. In March 2017, the Guangdong Provincial Department of Education announced funding of RMB 1.5 billion (USD 220 million) for high-level university construction during that same year. According to the announcement, Shenzhen University received nearly RMB 30 million (USD 4.4 million) in financial subsidies, with nearly RMB 20 million (USD 2.9 million) in award funds. It should be said that this is in line with the rapid speed of the city's growth. Shenzhen University has six disciplines that have entered the top 1% of ESI's global rankings. This swift increase in academic rankings at home and abroad reflects the fact that this young university, founded only 35 years ago, is now in full vigor.

According to media reports, Shenzhen will strengthen cooperation with prestigious universities at home and abroad on the basis of its effectively managed local universities, including Shenzhen University and Southern University of Science and Technology, and strive to develop several top-notch universities as well as more high-level, distinctive disciplines and professions. By 2025, the number of colleges and universities in Shenzhen is projected to reach 20, with full-time students numbering more than 200,000, three to five colleges and universities ranking among China's top 50, and more than 30 disciplines entering the top 1% of the ESI global rankings, making it one of the strongest cities for higher education in China.

AcaBridge invites outstanding scholars from home and abroad to reach out to us. We'll help you contact colleges and universities in Guangdong, provide one-on-one, personal consultation, and help you learn about and apply for talent-recruitment programs in Guangdong. If you need any help, please contact our recruitment consultant at consultant@acabridge.edu.cn. For more details, visit our website at www.edu.cn/jjgd.

Postdoctoral Fellow in Atmospheric Sciences

暨南大學 JINAN UNIVERSITY

Job description

3-5 full-time Postdoctoral Research Fellow positions are available at Center for Air Pollution and Climate Change Research (APCC), Jinan University, Guangzhou, China. The successful applicants will work on projects related to atmospheric multiphase chemistry and air pollution (more details see http://apcc.jnu.edu.cn/ and https://sites.google.com/view/jnu-apcc/). The positions are for a period of 2 years and can be extended upon mutual agreement.

Qualified Postdoc candidates are expected to

- have a Ph.D. degree in meteorology, atmospheric physics/chemistry, or a related field
- be under 35 years of age and no more than 3 years beyond receiving Ph.D. degree
- with experiences in aerosol measurements (physical or chemical), instrument development, or model simulation (WRF-Chem, GEOS-Chem or Molecular Dynamics)

Remuneration and Benefits

- Basic salary of 250,000 to 350,000 RMB (approx. 36,000 to 50,000 USD) per year depending on performance
- Housing allowance of 2000 RMB per month

About application materials:

- CV
- Copies of educational certificates
- A complete list of publications
- Names and contact details of 2-3 references (name, relationship to candidate, e-mail and telephone number)

Contact information

lan.zou.eci@jnu.edu.cn

The recruitment information above is long-term valid.

Faculty Positions Available at Harbin Institute of Technology, Shenzhen (HITSZ)

Founded in 1920, ***Harbin Institute of Technology (HIT)***, which is under the Ministry of Industry and Information Technology, is a national key university offers specializations in science, engineering, management and many other fields. It is a member of the C9 League and also the first universities to be part of Project 985 and one in list of "Double First-Class" University project.

In the engineering field, HIT ranked No.2 in China and No.6 in the world according to US NEWS 2018. What is more, according to ARWU 2017, HIT ranked No.7 in China, top 200 in the world. In the same year, HIT ranked ESI top 1/10000 in the world, in China only three universities have ESI top 1/10000 disciplines in the world, and HIT is one of them.

Together with the Shenzhen Municipal Government, HIT created Harbin Institute of Technology, Shenzhen (HITSZ), which was established in 2002 as HIT Shenzhen Graduate School and currently serves as a key campus of HIT. It is the first university in Shenzhen which belong to C9 League member, also in the list of Project 985 and "Double First-Class" University project which starts to enroll undergraduate students.

Shenzhen is a city born to innovate, Shenzhen has become a frontrunner in promoting innovation-driven development when China's economy steps into the new normal. Shenzhen is now widely known as a 'City of Makers' and a 'City of Innovation.' With a brand-new look, HITSZ will follow the HIT tradition, keep on the philosophy of high starting point, high standard to attract more international excellent talents and carries forward the Shenzhen spirit to serve as a contributor to national and regional economic and social development.

For more details, please refer to www.hitsz.edu.cn.

Fields Open for Recruitment:

(1) Computer Science and Technology
(2) Electronic Science and Technology
(3) Materials Science and Engineering
(4) Control Science and Engineering
(5) Power Engineering and Engineering Thermophysics
(6) Mechanics
(7) Mathematics
(8) Practical Economics
(9) Business Administration
(10) Civil Engineering
(11) Mechanical Engineering
(12) Architecture
(13) Space Science and Technology
(14) Information and Communication Engineering
(15) Environmental Science and Engineering
(16) Management Science and Engineering
(17) Biomedical Engineering
(18) Chemistry
(19) Physics
(20) Biology
(21) Design
(22) Sociology
(23) Marxism
(24) Marine Science
(25)Aeronautical and Astronautical Science and Technology
(26) Urban and Rural Planning
(27) Linguistics
(28) English

Qualifications and Requirements:

(1) Ph.D. completed in a related field before the start of employment.
(2) Overseas working experience or postdoctoral training is preferred.
(3) All strong candidates are encouraged to apply.

Salary and Benefits:

(1) Applicants will be appointed as 'Professor','Associate Professor' or'Assistant Professor' according to their qualifications and backgrounds.
(2) Salary: approximately from 300 thousands to 1.5 million.
(3) Research funds will be provided according to different positions and subjects.
(4) Applicants could apply for high-level talents allowance in Shenzhen (approximately from 1.6 million to 3 million), or rent teachers' welfare house, the rent is lower than the market price.

To Apply:

Each application must include the following documents:
(1) Application Form for Faculty Position (the form could be downloaded from website: http://www.hitsz.edu.cn/job/view/2.html; applicants should indicate their main research area to facilitate the application process).
(2) A cover letter including three parts: (1)self-introduction (explaining why the applicant should be considered for the job), (2)the expected contribution to the School based in terms of research, and (3)future work plan, if hired.
(3) Three Letters of Recommendation
(4) Electronic copy of supporting documents (diploma, achievements, list of publications, etc.)

Application materials should be sent to YANG Zhixi at: hrsz@hit.edu.cn.

Contact:

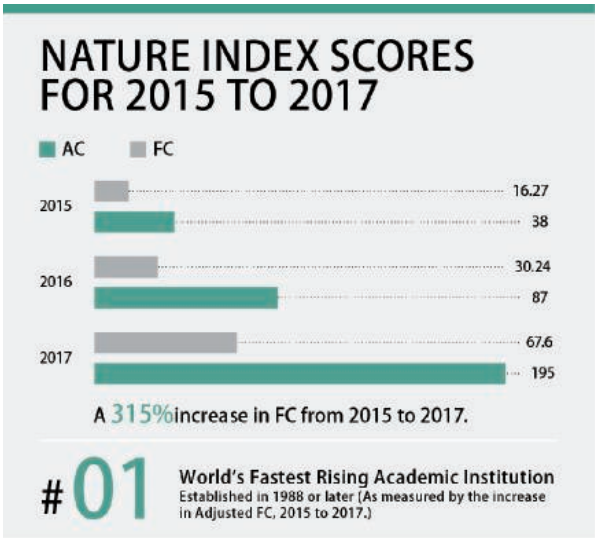
Ms. YANG Zhixi
Human Resources Department
Harbin Institute of Technology, Shenzhen
Tel: +86-755-26033365
E-mail: hrsz@hit.edu.cn

Faculty Positions at all ranks at Southern University of Science and Technology (SUSTech)

The SOUTHERN UNIVERSITY OF SCIENCE AND TECHNOLOGY (SUSTech) is offering exciting opportunities for scientists and engineers in a variety of disciplines.

The **Southern University of Science and Technology (SUSTech)** in Shenzhen, is actively seeking outstanding candidates. Applications are invited for all major science, engineering and medical disciplines.

Successful applicants will be appointed to the faculty of SUSTech at a level commensurate with their background and experience, from tenure-track assistant professor to tenured chair professor.



SUSTech offers a competitive salary and start-up package (compared with US and HK universities) for recipients of relevant national talents programmes. Benefits include: a competitive starting salary; a living subsidy of 2.75 million RMB (young talents programmes) or 4.5 million RMB (senior talents programmes) over three to five years; a start-up fund of up to 12 million RMB; principal investigator and tenure-track systems; a housing allowance up to 8,000 RMB per month; and social insurance and welfare.

Applicants should have a PhD in a relevant science, engineering or medical field. Applicants must have a proven track record of high-quality peer-reviewed academic publications, as well as excellent communication skills and the capacity to teach in English.

If interested, please submit the following material electronically to hiring@sustc.edu.cn: 1) Curriculum Vitae (with a complete list of publications); 2) Statement of research interests; 3) Statement of teaching philosophy; 4) Selected reprints of three recent papers; and 5) Names and contact information of five references. Review of applications will begin immediately and continue until the positions are filled.

Phone: +86-755-88010968
+86-755-88010945
Email: talents@sustc.edu.cn
www.sustc.edu.cn/en

Marine Biological Laboratory
THE UNIVERSITY OF CHICAGO

Marine Biological Laboratory

2019 Whitman Center Fellowships

THE MARINE BIOLOGICAL LABORATORY, a hub for research and education and an affiliate of the University of Chicago, convenes biologists from around the world each year to advance the mission of biological discovery. We are now accepting applications for **Whitman Center Fellowships** for 2019. Support is available for scientists to come to the Marine Biological Laboratory for 4 to 10 weeks to conduct research, year-round.

We particularly encourage applications from individuals or collaborative groups focused on the following:

- Evolutionary, genetic, and genomic approaches in regenerative and developmental biology, microbiomes, and neuroscience with an emphasis on marine organisms
- Integrated imaging and computational approaches to illuminate cellular function and biology emerging from the study of marine and other organisms
- Integrated approaches to the study of microbial communities in coastal communities.

The 2019 Whitman Center Fellowships include **Whitman Center Early Career Fellowships**, specifically designated for individuals less than 10 years from their doctoral degree who wish to focus on these areas of biological discovery.

Whitman Center Fellowships cover laboratory rental and housing costs. The MBL offers access to state-of-the-art instrumentation, innovative imaging technology, DNA sequencing and informatics, year-round availability of model freshwater and marine organisms, and modern laboratory facilities.

The Marine Biological Laboratory hosts more than 1,000 researchers, postdocs, and graduate students from around the world to participate in scientific discovery courses, research, lectures, and field studies. As a convener of biology, the Marine Biological Laboratory is well known for fostering a highly collaborative environment, with scientists and students engaged in intensive research in a collegial and informal atmosphere.

Eligible applicants must hold appointments at accredited universities, colleges or research institutions anywhere in the world. While applications will be evaluated on the basis of scientific merit, we are especially interested in individuals from diverse backgrounds, experiences and perspectives, especially those under represented in science.

mbl.edu/research/whitman-fellowships

research@mbl.edu

Application Deadline: **December 15, 2018**

Independent Research Fellowships Leading to Tenured Faculty Positions at the John Innes Centre

John Innes Centre
Unlocking Nature's Diversity

The John Innes Centre (JIC), Norwich, UK is a world leading centre of excellence in plant and microbial sciences based on the Norwich Research Park. We are inviting applications from outstanding researchers who either hold, or wish to apply for Independent Research Fellowships [such as a UKRI Future Leaders Fellowship (**https://www.ukri.org/funding/funding-opportunities/future-leaders-fellowships/**), a Royal Society University Research Fellowship (**http://royalsociety.org/grants/schemes/university-research/**), or a BBSRC David Phillips Fellowship (**https://bbsrc.ukri.org/funding/filter /david-phillips/**)]. Shortlisted candidates will be invited to give a seminar at a Fellows Conference at the JIC on **04 February 2019**. At the conference, you will also be able to discuss your proposals, the development of your group and your future career plans in depth with JIC faculty.

After the conference, we will select and mentor outstanding candidates in writing Fellowship applications and/or offer the opportunity to move existing Fellowships to the JIC. **Candidates who win Fellowships will be considered for transfer onto tenure-track at 3 years, and if transferred, for tenure at 5 years. Considerable additional resources will be provided to Fellows by the Centre. For further information, please contact mark.buttner@jic.ac.uk**

Further details and particulars can be found at https://www.jic.ac.uk/training-careers/vacancies/

To apply please e-mail a 2-page summary of your research plan, a copy of your CV and arrange for three letters of recommendation to be e-mailed to fellows@jic.ac.uk by Friday 30 November 2018. Before applying please read our Privacy Notice.

Athena SWAN Gold Award
GENDER CHARTER

The John Innes Centre is a registered charity (No223852) grant-aided by the Biotechnology and Biological Sciences Research Council and is an Equal Opportunities Employer and supports flexible working.

Northwestern University

Assistant Professor of Molecular Biosciences, Tenure-Track Faculty Position

The Department of Molecular Biosciences seeks to recruit a tenure-track faculty member at the level of Assistant Professor, although exceptional applicants at other ranks will be considered. We particularly seek individuals who utilize quantitative or systems approaches to address fundamental questions in biology. We are interested in individuals holding a Ph.D. and/or M.D. degree who show significant potential for innovative scholarship, and demonstrated commitment to excellence in research and teaching.

Northwestern University offers a rich and collaborative intellectual environment and state-of-the-art support facilities. Appointees will join a broad, highly interactive research community that is committed to collaboration across disciplinary boundaries. The Department of Molecular Biosciences has close ties with other Northwestern departments and centers including Physics, Statistics, Applied Math, the NSF-Simons Center for Quantitative Biology, the Northwestern Institute on Complex Systems, and the Chemistry of Life Processes Institute.

Applicants should submit (in PDF format) a cover letter, CV, description of research plans, and a statement of teaching interests. Applications must be submitted electronically. Please request at least three letters of recommendation; details on preparing and submitting the application can be found at the Molecular Biosciences homepage (http://www.molbiosci.northwestern.edu). The anticipated start date is Fall of 2019. To ensure full consideration, all materials should be submitted by **December 1, 2018**. Inquiries can be sent to **molbioscisearch@northwestern.edu**.

Northwestern University is an Equal Opportunity, Affirmative Action Employer of all protected classes, including veterans and individuals with disabilities. Women, racial and ethnic minorities, individuals with disabilities, and veterans are encouraged to apply. Hiring is contingent upon eligibility to work in the United States.

香港中文大學
The Chinese University of Hong Kong

Applications are invited for:-

Centre for Cardiovascular Genomics and Medicine
Associate Professors / Assistant Professors
(Ref. 180001UN)

The Chinese University of Hong Kong (CUHK), an internationally ranked top-tier university in Asia, has recently established a Centre for Cardiovascular Genomics and Medicine (CCGM) with research emphases in the following programmatic areas:

- Genetics and genomics of pediatric and adult cardiovascular diseases
- Stem cells in cardiac development, regeneration, and repair
- Metabolomics in heart failure and cardiovascular development and disease
- Translational medicine and therapeutics

Hong Kong is a cosmopolitan city rich in culture, has long played a leading role in international commerce, finance and education, and is now fast becoming a hub of high tech innovation in the Greater Bay Initiative. In this regard, CUHK is strategically situated adjacent to the Hong Kong Science Park, which houses biotech start-ups and major scientific industries, with strong history of academic-industry collaborations.

CCGM is jointly sponsored by the Department of Medicine and Therapeutics and the Department of Paediatrics in the CUHK Faculty of Medicine. Applicants seeking positions at the Associate or Assistant Professor level are welcome to apply. More senior appointment may be considered for exceptional candidates. Applicants should have (i) a PhD, MD, DVM or equivalent biomedical degree and (ii) a track record of research excellence with evidence of ability to lead a vigorous, independent research program and success in securing independent grant funding.

Appointments will initially be made on contract basis for 2-3 years commencing November 2018, which, subject to mutual agreement, may lead to longer-term appointment or substantiation later. Generous start-up package, competitive salary, housing allowance, and access to core facilities such as genomics and computational resource, imaging, and animal resources will be offered.

Application Procedure
Applicants please complete online application form, upload the CV and contact information of three referees. Reference letters can also be sent to Professor Cecilia W. Lo by e-mail to daisy.chan@cuhk.edu.hk.

The University only accepts and considers applications submitted online for the post above. For more information and to apply online, please visit http://career.cuhk.edu.hk.

HARVARD UNIVERSITY

Harvard University
Faculty of Arts and Sciences
Department of Organismic and Evolutionary Biology

FACULTY POSITION IN PLANT BIOLOGY

The Department of Organismic and Evolutionary Biology at Harvard University, in partnership with the Arnold Arboretum of Harvard University, invites applications for a tenure-track faculty position in **plant biology**. We seek to appoint an individual whose scholarship will complement the research and teaching missions of the Department of Organismic and Evolutionary Biology. The area of scholarship within plant biology is open, and we especially encourage applications from those studying genomics, molecular genetics, developmental biology, whole plant physiology, ecophysiology, organismic biology, biomechanics, phylogenetics, evolution, ecosystems biology, ecology, community biology, and urban ecology. It is hoped that the successful candidate will have a research program that can incorporate and leverage the extraordinary living collections of the Arnold Arboretum. The research laboratory will be situated in the recently opened 43,000 square foot state-of-the-art research facility in the Arnold Arboretum of Harvard University. Potential course offerings will be in the Department of Organismic and Evolutionary Biology, but may also serve students in Earth and Planetary Sciences, Environmental Science and Public Policy, and Molecular and Cellular Biology. The successful candidate will be expected to have a strong commitment to undergraduate and graduate teaching. The Department and Arboretum have strong linkages to a number of allied institutions, including the Harvard Forest, Harvard University Herbaria, Harvard University Center for the Environment, and the Museum of Comparative Zoology.

Special instructions: Applications should include a curriculum vitae, statements of research and teaching interests, 3-5 representative publications, and the names, institutional affiliations, and e-mail addresses of 3-5 references. Please submit these materials to http://academicpositions.harvard.edu/postings/8469. Review of applications and nominations will begin November 15, 2018 and conclude when the position is filled. Letters of nomination from third parties are also welcome and should be sent by e-mail to William (Ned) Friedman at ned@oeb.harvard.edu.

Contact information: Further information about the Department of Organismic and Evolutionary Biology and the Arnold Arboretum are available at https://oeb.harvard.edu and https://www.arboretum.harvard.edu.

Contact email: Address questions about the position to Professor William (Ned) Friedman and about the application/nomination process to Christian Flynn in OEB at faculty_search@oeb.harvard.edu.

Harvard University is an Equal Opportunity Employer and 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.

POSITIONS OPEN

U.S. POSTAL SERVICE

Statement required by the Act of 12 August 1970, Section 3685, Title 39, United States Code, showing the ownership, management, and circulation of:

1–9. Science, Publication No. 0036-8075, is published weekly on Friday, except the last week in December, at 1200 New York Avenue, N.W., Washington, DC 20005. Date of filing: 25 September 2018. This is also the address of the publisher, the editor, and the managing editor, who are, respectively, Bill Moran, Jeremy Berg, and Monica M. Bradford. 10. The owner is the American Association for the Advancement of Science, 1200 New York Avenue, N.W., Washington, DC 20005. Stockholders: None. 11. Known bondholders, mortgages, and other security holders owning or holding 1 percent or more of total amount of bonds, mortgages, or other securities: None. 12. The purpose, function, and nonprofit status of this organization and the exempt status for federal income tax purposes have not changed during the preceding 12 months. 13–15. The average number of copies of each issue during the preceding 12 months is (A) Total number of copies printed: 64,909; (B) Paid circulation: 60,543; (1) Paid/Requested outside-county mail subscriptions stated on form 3541: 52,853; (2) Paid/Requested in-county subscriptions stated on form 3541: 0; (3) Sales through dealers and carriers, street vendors, counter sales: 7,679. (4) Other classes mailed through USPS: 11; (C) Total paid circulation: 60,543; (D) Free distribution: samples, complimentary, and other free copies: 3,236; (1) Outside-county as stated on form 3541: 2,449; (2) In-county as stated on form 3541: 0; (3) Other classes mailed through the USPS: 2; (4) Free distribution outside of mail carrier or other means: 786; (E) Total free distribution: 3,236; (F) Total distribution: 63,779; (G) Copies not distributed: 1,130; (H) Total: 64,909; (I) Percent paid and/or Requested Circulation: 94.9%.

Actual number of copies of single issue (9/21/2018) published nearest to filing date are (A) Total number of copies printed: 64,448; (B) Paid circulation: 59,365; (1) Paid/Requested outside-county mail subscriptions stated on form 3541: 51,889; (2) Paid/Requested in-county subscriptions stated on form 3541: 0; (3) Sales through dealers and carriers, street vendors, counter sales: 7,465; (4) Other classes mailed through USPS: 11; (C) Total paid circulation: 59,365; (D) Free distribution: Samples, complimentary, and other free copies: 2,973; (1) Outside-county as stated on form 3541: 2,516; (2) In-county as stated on form 3541: 0; (3) Other classes mailed through the USPS: 2; (4) Free distribution outside of mail: Carrier or other means: 455; (E) Total free distribution: 2,973; (F) Total distribution: 62,338; (G) Copies not distributed: 2,110; (H) Total: 64,448; (I) Percent paid and/or Requested Circulation: 95.2%.

I certify that the statements made above are correct and complete. (signed) Bill Moran, Publisher.

ILLINOIS
UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN

FACULTY POSITION
THEORETICAL GRAVITATIONAL PHYSICS
Department of Physics

The Department of Physics invites applicants for a full-time tenured or tenure-track position in theoretical gravitational physics beginning in August 2019. We seek individuals with a strong background in general relativity, numerical relativity, gravitational wave physics, physics of compact objects, relativistic astrophysics, or cosmology. A successful applicant will be expected to conduct a vigorous and significant research program. Ideal candidates include those who demonstrate evidence of a commitment to diversity, equity, and inclusion through research, teaching, and/or service endeavors. For full consideration, application materials must be received by **November 14, 2018**. Please visit **http://jobs.illinois.edu** to view the complete position announcement and application instructions.

The University of Illinois conducts background checks on all job candidates upon acceptance of a contingent offer.

The U of I is an EEO Employer/Vet/Disabled
http://www.inclusiveillinois.illinois.edu

POSITIONS OPEN

TENURE-TRACK BIOCHEMIST FACULTY POSITION

The Department of Biochemistry, Microbiology and Immunology at Wayne State University School of Medicine is searching for an outstanding biochemist for a tenure-track Assistant or Associate Professor position. Candidates should have a Ph.D. or M.D./Ph.D. degree with postdoctoral research experience. Although all qualified candidates will be considered, the search is targeting researchers using biochemical approaches to study problems associated with human disease. Candidates will be expected to establish a high-impact, extramurally funded research program, and to participate in graduate and medical teaching. A competitive start-up package commensurate with the candidate's experience will be provided.

For further information and to submit your application, please visit our application site at website: **www.jobs.wayne.edu** (Posting 043896). Applications should include a curriculum vitae, a 2-3 page summary discussing current and future research plans, and the names and addresses of three referees.

In you have questions, please contact Dr. Bharati Mitra at bmitra@wayne.edu

Wayne State University is a premier, public, urban research university located in the heart of Detroit where students from all backgrounds are offered a rich, high quality education. Our deep rooted commitment to excellence, collaboration, integrity, diversity and inclusion creates exceptional educational opportunities preparing students for success in a diverse, global society. WSU encourages applications from women and minorities. WSU is an Affirmative Action/Equal Opportunity Employer.

EPFL
ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Two Faculty Positions in Quantum Science and Technology

at the Ecole polytechnique fédérale de Lausanne (EPFL)

The School of Basic Sciences (Physics, Chemistry, and Mathematics) at EPFL seeks to appoint two tenure track Assistant Professors in experimental and theoretical Quantum Science and Technology (QST). The appointments are offered at the Tenure Track level, but in exceptional cases, appointments at other ranks might be considered.

Areas of interest in experimental QST include superconducting quantum circuits, solid-state qubits in semiconductors, defects centers, or other solid-state quantum systems for sensing, communication, or computation. Photonic quantum technologies are also considered.

Areas of interest in theoretical QST include tensor networks, quantum field theory, and other advanced theoretical and computational approaches to quantum matter, artificial quantum structures, open quantum systems, quantum information processing or quantum simulation.

We expect candidates to establish a leadership in one of these areas and strengthen the EPFL endeavor in Quantum Science and Technology. Priority will be given to the overall originality and promise of the candidate's work over any particular specialization area.

Candidates should hold a PhD and have an excellent record of scientific accomplishments in quantum science. In addition, commitment to teaching general courses at the undergraduate level and specialized courses at the Master and doctoral levels is expected. Proficiency in French teaching is not required, but willingness to learn the language is expected.

EPFL, with its main campus located in Lausanne, Switzerland, on the shores of lake Geneva, is a dynamically growing and well-funded institution fostering excellence and diversity. It has a highly international campus with first-class infrastructure, including facilities for high performance computing. As a technical university covering essentially the entire palette of engineering and science, EPFL offers a fertile environment for research cooperation between different disciplines. The EPFL environment is multi-lingual and multi-cultural, with English often serving as a common interface.

Applications should include a cover letter, a CV with a list of publications, a concise statement of research (maximum 3 pages) and teaching interests (one page), and the names and addresses (including e-mail) of at least three references.

Applications should be uploaded (as PDFs) by **November 30, 2018** to:

https://facultyrecruiting.epfl.ch/position/10977291

Enquiries may be addressed to:

Prof. Harald Brune, Chair of the Search Committee
E-mail: IPHYSDirector@epfl.ch

For additional information please consult www.epfl.ch, sb.epfl.ch

EPFL is an equal opportunity employer and family friendly university. It is committed to increasing the diversity of its faculty. It strongly encourages women to apply.

By **Sasha Nikolaeva**

A normal student parent

"I don't know if you'll be able to reach her," the student said. I was in the first year of my master's program, preparing to be a teaching assistant by contacting students who had taken the course in previous years. This student had recommended that I connect with the graduate student who had taught the course when he took it, but now he was backtracking. "She was pregnant when she taught the class. She must be on leave now." There was some curiosity in his voice, mixed with what I perceived as a hint of judgment. I should probably keep the fact that I am a single parent to myself, I thought.

I had my daughter several years before I had any thoughts of pursuing a scientific career. My marriage didn't work out, and I moved with my daughter from Russia to the United States. I started working at a tech company, where having a family did not feel out of the ordinary.

Once I entered the academic world, on the other hand, I found that speaking about my daughter with my colleagues and professors usually elicited surprise—or worse. When I finished my master's degree and decided to apply for a Ph.D., for example, a professor told me that pursuing a doctoral degree would be challenging in my "situation."

"I need to make myself visible as a single parent and to advocate for change."

In some ways, that professor was right. It is impossible to support myself and my daughter on the standard Ph.D. stipend. My university offers a student parent supplement, which helps, but I still struggle to pay the bills. I miss many social and professional development activities because I want to spend time with my daughter, and academic culture does not leave much room for personal time. Even if the events are scheduled at times when I can attend, going means that I have to work at night or on weekends to catch up.

Even so, I want to do science. Academic research shouldn't only be for people who have the privilege of time and money. So I'm pushing forward with my academic training.

I have also decided that, despite my discomfort and the potential for backlash, I need to make myself visible as a single parent and to advocate for change. I am not shy about telling people that sometimes I don't have money to pay rent. I started bringing my daughter to some social activities. I wrote an open letter to my department about the challenges that student parents face. I created a survey to gather more information about other student parents in the department.

I think my efforts are making some headway. (And, to be fair, my department is much more family-friendly than some others.) Emails about social events are now more likely to include explicit statements that family members, including kids, are welcome. My department chair is always willing to discuss any problems I might be having. Members of my department's graduate student association emailed me recently to ask for input about how they could use university wellness funds to support student parents.

Yet I still worry that my efforts make me more visible as a parent than as a scientist. In August, a fellow Ph.D. student wished me a happy Mother's Day. He is from Costa Rica, where they celebrate mothers at the end of the summer, and he thought that I would appreciate the sentiment—which I did. But it didn't alleviate my concerns about the reputation I'm creating for myself. Studies show that academic mothers tend to be viewed as less serious, whereas fathers are seen as more grounded and more organized.

I figure that, for a while, I will have to accept the fact that my role as a mother may be more interesting to some of my colleagues than my research. But I hold on to the hope that one day, thanks to my efforts and those of many others, grad student parenthood won't seem remarkable at all. ■

Sasha Nikolaeva is a Ph.D. student in the Department of Environmental Science, Policy, and Management at the University of California, Berkeley. Do you have an interesting career story that you would like to share? Send it to SciCareerEditor@aaas.org.

ILLUSTRATION: ROBERT NEUBECKER