Exquisite molecular portraits—at an affordable price p. 354

Suppression of movement during sleep pp. 366 & 440

Probing surfaces with ultrafast microscopy pp. 368 & 411

Science

$15
24 JANUARY 2020
SPECIAL ISSUE
sciencemag.org
AAAS

CHEMISTRY FOR TOMORROW'S EARTH

It's a matter of expression.

For over 40 years, New England Biolabs has been developing and using recombinant protein technologies in *E.coli* for our own manufacturing processes.

Protein expression can be a very complex, multi-factorial process. Each protein requires a specific environment to correctly and efficiently achieve its secondary and tertiary structures. Proteins may also require post-translational modifications or insertion into a cellular membrane for proper function. Other proteins, once expressed, may be toxic to the host. Thus, no single solution exists for the successful production of all recombinant proteins, and a broad range of expression tools is often required to ensure the successful expression of your target protein.

Our NEBExpress™ portfolio of products include solutions for expression and purification of a wide range of proteins, and is supported by access to scientists with over 40 years of experience in developing and using recombinant protein technologies in *E. coli.* We use these solutions in our own research and manufacturing processes, and know that quality and performance are critical – all of our products are stringently tested so that you can be sure they will work optimally for your solution, just as we rely on them to work in ours.

Featured products include:

- **Cell-free expression systems** – express analytical amounts of protein in approximately two hours
- ***E. coli* expression and purification kits** – generate and purify high yields of recombinant proteins
- **Competent cells** – express a variety of proteins in *E. coli*, including difficult targets, proteins with multiple disulfide bonds and His-tagged proteins
- **Purification beads, columns and resins** – available for CBD-, MBP- and His-tagged proteins

Access our entire portfolio and request your sample at **www.neb.com/ProteinExpression**.

NEW ENGLAND BioLabs® Inc.

be INSPIRED
drive DISCOVERY
stay GENUINE

CONTENTS

24 JANUARY 2020 • VOLUME 367 • ISSUE 6476

SPECIAL SECTION

Chemistry for Tomorrow's Earth

Can linings are a challenge for chemists.

INTRODUCTION

NEWS

REVIEWS

ON THE COVER

Synthetic chemicals are important components of modern life, but the current system of global production, distribution, and disposal has, in many instances, caused damage to the environment and human health. Innovations in chemical synthesis, environmental monitoring, and analysis of exposure and risk will ensure that tomorrow's chemicals are less persistent, have greater potency, and move us toward a sustainable future. See page 378. *Illustration: Adam Simpson*

SEE ALSO POLICY FORUM P. 360
PERSPECTIVE P. 369

NEWS

IN BRIEF

IN DEPTH

FEATURES

INSIGHTS

POLICY FORUM

PERSPECTIVES

BOOKS *ET AL.*

PHOTO: ROB LARSON

Published by AAAS

AAAS ANNUAL MEETING

Seattle, WA | February 13–16, 2020

aaas.org/meetings **#AAASmtg**

AAAS, publisher of *Science*, thanks the sponsors and supporters of the 2020 Annual Meeting.

European Union

Johnson & Johnson INNOVATION

UK Research and Innovation

University of Washington HOST UNIVERSITY

Bristol-Myers Squibb

FRED HUTCH CURES START HERE®

Arizona State University

UF University of Florida

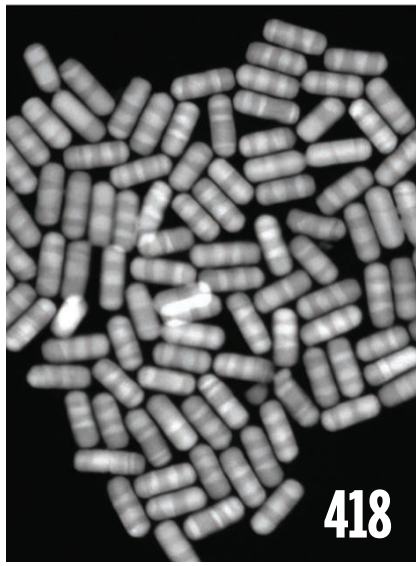
NIH National Human Genome Research Institute

Seattle Children's HOSPITAL · RESEARCH · FOUNDATION

SCIENCE ON TAP AAAS

As of December 2019

Download the meeting app: aaas.org/app



CREDITS: (LEFT TO RIGHT) XINHUA/ALAMY STOCK PHOTO; B. C. STEIMLE *ET AL.*

SCIENCE (ISSN 0036-8075) is published weekly on Friday, except last week in December, by the American Association for the Advancement of Science, 1200 New York Avenue, NW, Washington, DC 20005. Periodicals mail postage (publication No. 484460) paid at Washington, DC, and additional mailing offices. The title SCIENCE is a registered trademark of the AAAS. Domestic individual membership, including subscription (12 months): $165 ($74 allocated to subscription). Domestic institutional subscription (51 issues): $2148; Foreign postage extra: Air assist delivery: $98. First class, airmail, student, and emeritus rates on request. Canadian rates with GST available upon request, GST #125488122. Publications Mail Agreement Number 1069624. **Printed in the U.S.A.**
Change of address: Allow 4 weeks, giving old and new addresses and 8-digit account number. **Postmaster:** Send change of address to AAAS, P.O. Box 96178, Washington, DC 20090–6178. **Single-copy sales:** $15 each plus shipping and handling available from backissues.sciencemag.org; bulk rate on request. **Authorization to reproduce** material for internal or personal use under circumstances not falling within the fair use provisions of the Copyright Act can be obtained through the Copyright Clearance Center (CCC), www.copyright.com. The identification code for *Science* is 0036-8075. *Science* is indexed in the *Reader's Guide to Periodical Literature* and in several specialized indexes.

Published by AAAS

INNOVATIVE EDUCATION DRIVES ADVANCES AT WASEDA UNIVERSITY

Japanese higher education institutions are not well known outside the country. Waseda University in Tokyo is resolved to change that and is implementing a plan to transform itself into a stellar global university within the next three decades.

"Yes, it's a lofty goal, but with determination, we can achieve it," says Aiji Tanaka, who assumed the Waseda presidency in 2018.

Asked how Waseda hopes to achieve this, Tanaka says the answer is to employ an innovative strategy that also makes use of unique aspects of Japanese culture. The strategy he's advocating has two pillars: systematization of courses and syllabuses, and recruitment of top faculty members both domestically and overseas who can conduct outstanding research and teach well.

Before becoming president, Tanaka was dean of academic affairs and vice-provost, and introduced such changes as systematic course numbering, student evaluation of lecturers and courses, a restructured tenure-track system, and a program of international academic recruitment, building an educational framework new to Japanese universities. Regarding systematic course numbering, Tanaka says, "By streamlining the number of courses, the university is able to combine those that overlap. This helps not only students but also faculty, because they will have more time to enhance the level of research and quality of education they provide."

When it comes to recruiting outstanding scholars, Waseda is seeking younger academics—both Japanese and foreign—who display great potential. And as for attracting better-known scholars from abroad, the university has hatched novel recruitment plans, such as the Joint Appointment System, which bring them to Waseda for short periods to teach and conduct joint research with Waseda's faculty. "In this way, we can create a course taught by several outstanding scholars working in rotation," explains Tanaka. "The students, the scholars, and our own faculty all gain greatly."

Another innovation is the *zemi* (short for seminar) system of teaching. It brings together a group of undergraduate students in their junior or senior year who study and research as a group under one professor for two years. The *zemi* style can also be applied at the graduate level, allowing a mix of Master's and doctoral students to learn together. This approach teaches the value of teamwork and collaboration, while working under a single mentor helps create a sense of bonding.

Waseda's determination to become an internationally respected university coincides with the Japanese government's efforts to upgrade and internationalize selected universities by funding a multimillion-dollar program called the Top Global University Project. In its bid to participate, Waseda drew up a plan dubbed "Waseda Goes Global," an initiative to accelerate the international mobility of its students and researchers, and to build a worldwide academic network that is open, dynamic, and diverse.

Waseda President Aiji Tanaka: "With determination, we can transform ourselves into a stellar global university."

"Waseda Goes Global prioritizes funding for seven model research units, each established around fields of study we are particularly strong in," says Tanaka. "These fields, which consistently rank highly in the QS [Quacquarelli Symonds] World University Rankings annual report, are global Japanese studies, positive/empirical analysis of political economy, health promotion, information and communications technology [ICT] and robotics, energy and nanomaterials, mathematics and physics, and global Asia studies. We hope that our initiative will have a spillover effect in other fields as well."

The initiative was accepted by the government in 2014, and Waseda was placed in the highest category of universities to receive funding from the government project. Subsequently, Waseda has made great strides in internationalizing itself. The number of foreign students enrolled in the 2018 academic year reached almost 8,000, representing 125 countries, exemplifying Waseda's success at boosting admissions of both undergraduate and graduate students and making it one of the leading universities in Japan for attracting students from abroad.

"In these ways, we've begun the journey to become a top global university," says Tanaka. "It may take us a couple of decades or more, but get there we will."

PHOTO: COURTESY OF WASEDA UNIVERSITY

Kan Hatakeyama (center) and Niccolo Giannetti (right) share with Hironori Kasahara how their research could help address the issue of pollution.

THEORETICAL AND APPLIED RESEARCH HELP CUT POLLUTION

Pollution is a global problem that requires global solutions. At Waseda University, researchers are addressing this important issue on multiple fronts.

Next-generation batteries

Kan Hatakeyama, an assistant professor in Waseda's Department of Applied Chemistry, is researching more eco-friendly, plastic batteries. He is using electrically conducting polymers to develop batteries that are flexible, transparent, and rechargeable. The materials in these new devices—which can be designed to be biodegradable—might one day replace the toxic and environmentally harmful chemicals in today's batteries.

Hatakeyama and his colleagues have already developed the world's thinnest battery: a 1-μm storage device that powers a pocket calculator. They have also developed an ultrathin, rechargeable, stretchable device using organic polymer nanosheets, which can be attached to the skin for biomonitoring purposes.

"Previous biomonitors were too thick and toxic to use on the skin," says Hatakeyama. "We've developed a power source from redox-active organic polymers just 100 nm thick, which can work even when attached to a finger joint."

Hironori Kasahara, senior executive vice president for research at Waseda, says that if Hatakeyama and his colleagues are successful in developing such batteries, it would dramatically impact the electronics industry, enabling new applications like roll-up displays and wearable electronics as well as storage devices attached to solar cells.

The secret to efficient heat control

Research conducted by Niccolo Giannetti, an assistant professor at the Waseda Institute for Advanced Study, could also benefit the environment. He is working on mathematically and physically modeling heat-pump systems to gain a deeper understanding of their theoretical underpinnings.

"Heat pumps control the flow of heat in industrial processes and also in our homes for thermal comfort," explains Giannetti, who is from Italy. "Working as air conditioners, they extract heat from our homes when it's too hot, or warm them up when it's cold."

Giannetti notes, however, that there is a gap between the technology of heat pumps and the underlying theoretical knowledge, resulting in inefficient design and operation. Designing factory equipment based on a strong theoretical grounding would better address waste-heat recovery and reduce CO_2 and pollutant emissions, he explains. He analyzes and models heat-pump systems to generate design formulas that engineers can apply to improve efficiency and reduce pollution.

The research of Hatakeyama and Giannetti encompasses just some of the projects taking place at Waseda's Energy and Nanomaterials Model Unit and at the Mathematics and Physics Model Unit, respectively.

Competition vs. collaboration and mentoring

When Hatakeyama studied abroad at Texas A&M University as a Waseda graduate student, he was struck by how intensely U.S. students competed against each other.

"Competition is good, but for me, collaboration and mentoring work even better," says Hatakeyama. "And it helps not only academically but also in our daily lives." He prefers the Japanese way of *senpai–kouhai* (senior–junior mentoring).

Giannetti also finds Waseda's emphasis on collaboration stimulating. His lab is partnering with government research institutes and other universities in Japan as well as universities in Europe, the United States, the Philippines, Indonesia, and Malaysia. He also enjoys the weekly visits from retired researchers in different Japanese industries who lecture on their experiences.

"In terms of research, I could hardly ask for more, given all these opportunities," says Giannetti. "I feel like I'm exposed to research from everywhere."

"Collaborative research is a feature of Japanese academia, and of Waseda in particular," says Kasahara. Moreover, Waseda takes a flexible approach to research in that graduates and researchers can collaborate with industry to address societal needs, creating products that people around the world will find useful. "This helps produce a good spirit in the labs and allows researchers to enjoy their work," he adds.

PHOTO: COURTESY OF WASEDA UNIVERSITY

Science AAAS

Editor-in-Chief Holden Thorp, hthorp@aaas.org

Executive Editor Monica M. Bradford

Editors, Research Valda Vinson, Jake S. Yeston **Editor, Insights** Lisa D. Chong

DEPUTY EDITORS Julia Fahrenkamp-Uppenbrink (UK), Stella M. Hurtley (UK), Phillip D. Szuromi, Sacha Vignieri **SR. EDITORIAL FELLOW** Andrew M. Sugden (UK) **SR. EDITORS** Gemma Alderton (UK), Caroline Ash (UK), Brent Grocholski, Pamela J. Hines, Paula A. Kiberstis, Marc S. Lavine (Canada), Steve Mao, Ian S. Osborne (UK), Beverly A. Purnell, L. Bryan Ray, H. Jesse Smith, Keith T. Smith (UK), Jelena Stajic, Peter Stern (UK), Valerie B. Thompson, Brad Wible, Laura M. Zahn **ASSOCIATE EDITORS** Michael A. Funk, Priscilla N. Kelly, Tage S. Rai, Seth Thomas Scanlon (UK), Yury V. Suleymanov **LETTERS EDITOR** Jennifer Sills **LEAD CONTENT PRODUCTION EDITORS** Harry Jach, Lauren Kmec **CONTENT PRODUCTION EDITORS** Amelia Beyna, Jeffrey E. Cook, Chris Filiatreau, Julia Katris, Nida Masiulis, Suzanne M. White **SR. EDITORIAL COORDINATORS** Carolyn Kyle, Beverly Shields **EDITORIAL COORDINATORS** Aneera Dobbins, Joi S. Granger, Jeffrey Hearn, Lisa Johnson, Maryrose Madrid, Ope Martins, Shannon McMahon, Jerry Richardson, Alana Warnke, Alice Whaley (UK), Anita Wynn **PUBLICATIONS ASSISTANTS** Jeremy Dow, Alexander Kief, Ronmel Navas, Hilary Stewart (UK), Brian White **EXECUTIVE ASSISTANT** Jessica Slater **ASI DIRECTOR, OPERATIONS** Janet Clements (UK) **ASI SR. OFFICE ADMINISTRATOR** Jessica Waldock (UK)

News Editor Tim Appenzeller

NEWS MANAGING EDITOR John Travis **INTERNATIONAL EDITOR** Martin Enserink **DEPUTY NEWS EDITORS** Elizabeth Culotta, Lila Guterman, David Grimm, Eric Hand (Europe), David Malakoff **SR. CORRESPONDENTS** Daniel Clery (UK), Jon Cohen, Jeffrey Mervis, Elizabeth Pennisi **ASSOCIATE EDITORS** Jeffrey Brainard, Catherine Matacic **NEWS REPORTERS** Adrian Cho, Jennifer Couzin-Frankel, Jocelyn Kaiser, Kelly Servick, Robert F. Service, Erik Stokstad, Paul Voosen, Meredith Wadman **INTERNS** Eva Frederick, Rodrigo Perez Ortega **CONTRIBUTING CORRESPONDENTS** Warren Cornwall, Ann Gibbons, Mara Hvistendahl, Sam Kean, Eli Kintisch, Kai Kupferschmidt (Berlin), Andrew Lawler, Mitch Leslie, Eliot Marshall, Virginia Morell, Dennis Normile (Shanghai), Elisabeth Pain (Careers), Charles Piller, Michael Price, Tania Rabesandratana (Barcelona), Emily Underwood, Gretchen Vogel (Berlin), Lizzie Wade (Mexico City) **CAREERS** Donisha Adams, Rachel Bernstein (Editor), Katie Langin(acting editor) **COPY EDITORS** Julia Cole (Senior Copy Editor), Cyra Master (Copy Chief) **ADMINISTRATIVE SUPPORT** Meagan Weiland

Creative Director Beth Rakouskas

DESIGN MANAGING EDITOR Marcy Atarod **GRAPHICS MANAGING EDITOR** Alberto Cuadra **PHOTOGRAPHY MANAGING EDITOR** William Douthitt **WEB CONTENT STRATEGY MANAGER** Kara Estelle-Powers **SENIOR DESIGNER** Chrystal Smith **DESIGNER** Christina Aycock **GRAPHICS EDITOR** Nirja Desai **INTERACTIVE GRAPHICS EDITOR** Xing Liu **SENIOR SCIENTIFIC ILLUSTRATORS** Valerie Altounian, Chris Bickel **SCIENTIFIC ILLUSTRATOR** Alice Kitterman **SENIOR GRAPHICS SPECIALISTS** Holly Bishop, Nathalie Cary **SENIOR PHOTO EDITOR** Emily Petersen **PHOTO EDITOR** Kaitlyn Dolan

Chief Executive Officer and Executive Publisher Sudip Parikh

Publisher, Science Family of Journals Bill Moran

DIRECTOR, BUSINESS SYSTEMS AND FINANCIAL ANALYSIS Randy Yi **DIRECTOR, BUSINESS OPERATIONS & ANALYSIS** Eric Knott **DIRECTOR OF ANALYTICS** Enrique Gonzales **MANAGER, BUSINESS OPERATIONS** Jessica Tierney **SENIOR BUSINESS ANALYST** Cory Lipman, Meron Kebede **FINANCIAL ANALYST** Alexander Lee **ADVERTISING SYSTEM ADMINISTRATOR** Tina Burks **SENIOR SALES COORDINATOR** Shirley Young **DIGITAL/PRINT STRATEGY MANAGER** Jason Hillman **QUALITY TECHNICAL MANAGER** Marcus Spiegler **ASSISTANT MANAGER DIGITAL/PRINT** Rebecca Doshi **SENIOR CONTENT SPECIALISTS** Steve Forrester, Jacob Hedrick, Antoinette Hodal, Lori Murphy **DIGITAL PRODUCTION MANAGER** Lisa Stanford **CONTENT SPECIALIST** Kimberley Oster **ADVERTISING PRODUCTION OPERATIONS MANAGER** Deborah Tompkins **DESIGNER, CUSTOM PUBLISHING** Jeremy Huntsinger **SR. TRAFFIC ASSOCIATE** Christine Hall **SPECIAL PROJECTS ASSOCIATE** Sarah Dhere

ASSOCIATE DIRECTOR, BUSINESS DEVELOPMENT Justin Sawyers **GLOBAL MARKETING MANAGER** Allison Pritchard **DIGITAL MARKETING MANAGER** Aimee Aponte **JOURNALS ARKETING MANAGER** Shawana Arnold **MARKETING ASSOCIATES** Tori Velasquez, Mike Romano, Ashley Hylton **DIGITAL MARKETING SPECIALIST** Asleigh Rojanavongse **SENIOR DESIGNER** Kim Huynh

DIRECTOR AND SENIOR EDITOR, CUSTOM PUBLISHING Sean Sanders **ASSISTANT EDITOR, CUSTOM PUBLISHING** Jackie Oberst

DIRECTOR, PRODUCT & PUBLISHING DEVELOPMENT Chris Reid **DIRECTOR, BUSINESS STRATEGY AND PORTFOLIO MANAGEMENT** Sarah Whalen **ASSOCIATE DIRECTOR, PRODUCT MANAGMENT** Kris Bishop **SR. PRODUCT ASSOCIATE** Robert Koepke **DIGITAL PRODUCT STRATEGIST** Michael Hardesty **SPJ ASSOCIATE** Samantha Bruno Fuller

DIRECTOR, INSTITUTIONAL LICENSING Iquo Edim **ASSOCIATE DIRECTOR, RESEARCH & DEVELOPMENT** Elisabeth Leonard **MARKETING MANAGER** Kess Knight **SENIOR INSTITUTIONAL LICENSING MANAGER** Ryan Rexroth **INSTITUTIONAL LICENSING MANAGER** Marco Castellan **MANAGER, AGENT RELATIONS & CUSTOMER SUCCESS** Judy Lillibridge **SENIOR OPERATIONS ANALYST** Lana Guz **FULFILLMENT COORDINATOR** Melody Stringer **SALES COORDINATOR** Josh Haverlock

DIRECTOR, GLOBAL SALES Tracy Holmes **US EAST COAST AND MID WEST SALES** Stephanie O'Connor **US WEST COAST SALES** Lynne Stickrod **US SALES MANAGER, SCIENCE CAREERS** Claudia Paulsen-Young **US SALES REP, SCIENCE CAREERS** Tracy Anderson **ASSOCIATE DIRECTOR, ROW** Roger Goncalves **SALES REP, ROW** Sarah Lelarge **SALES ADMIN ASSISTANT, ROW** Bryony Cousins **DIRECTOR OF GLOBAL COLLABORATION AND ACADEMIC PUBLISHING RELATIONS, ASIA** Xiaoying Chu **ASSOCIATE DIRECTOR, INTERNATIONAL COLLABORATION** Grace Yao **SALES MANAGER** Danny Zhao **MARKETING MANAGER** Kilo Lan **ASCA CORPORATION, JAPAN** Kaoru Sasaki (Tokyo), Miyuki Tani (Osaka) **COLLABORATION/CUSTOM PUBLICATIONS/JAPAN** Adarsh Sandhu

DIRECTOR, COPYRIGHT, LICENSING AND SPECIAL PROJECTS Emilie David **RIGHTS AND LICENSING COORDINATOR** Jessica Adams **RIGHTS AND PERMISSIONS ASSOCIATE** Elizabeth Sandler **CONTRACTS AND LICENSING ASSOCIATE** Lili Catlett

---

MAIN HEADQUARTERS
Science/AAAS
1200 New York Ave. NW
Washington, DC 20005

SCIENCE INTERNATIONAL
Clarendon House
Clarendon Road
Cambridge, CB2 8FH, UK

SCIENCE CHINA
Room 1004, Culture Square
No. 59 Zhongguancun St.
Haidian District, Beijing, 100872

SCIENCE JAPAN
ASCA Corporation
Sibaura TY Bldg. 4F, 1-14-5
Shibaura Minato-ku
Tokyo, 108-0073 Japan

EDITORIAL
science_editors@aaas.org

NEWS
science_news@aaas.org

INFORMATION FOR AUTHORS
sciencemag.org/authors/science-information-authors

REPRINTS AND PERMISSIONS
sciencemag.org/help/reprints-and-permissions

MEDIA CONTACTS
scipak@aaas.org

MULTIMEDIA CONTACTS
SciencePodcast@aaas.org
ScienceVideo@aaas.org

INSTITUTIONAL SALES AND SITE LICENSES
sciencemag.org/librarian

PRODUCT ADVERTISING & CUSTOM PUBLISHING
advertising.sciencemag.org/products-services
science_advertising@aaas.org

CLASSIFIED ADVERTISING
advertising.sciencemag.org/science-careers
advertise@sciencecareers.org

JOB POSTING CUSTOMER SERVICE
employers.sciencecareers.org
support@sciencecareers.org

MEMBERSHIP AND INDIVIDUAL SUBSCRIPTIONS
sciencemag.org/subscriptions

MEMBER BENEFITS
aaas.org/membercentral

Science serves as a forum for discussion of important issues related to the advancement of science by publishing material on which a consensus has been reached as well as including the presentation of minority or conflicting points of view. Accordingly, all articles published in *Science*—including editorials, news and comment, and book reviews—are signed and reflect the individual views of the authors and not official points of view adopted by AAAS or the institutions with which the authors are affiliated.

EDITORIAL

Drop the chalk

A diverse scientific workforce; policy-makers who recognize the importance of science; a voting public that understands the scientific process (and even some facts about the ancient universe and climate change, too)—this has been the mantra of the scientific community, but these words don't match actions. Look no further than the way science is still taught at universities—it's no wonder that there is a struggle to recruit students to science. That is costing society a generation of researchers, educators, a population that better grasps science, and maybe more.

Despite evidence to the contrary, most science departments believe that the only way to produce scientists is to bludgeon young people with too much material, in the wrong setting, at the wrong time, and with the wrong kinds of assessments. The rationale is that only by mastering an abundance of facts and quantitative skills can someone become a scientist—and that's where one must start, not finish.

Lecture halls have been fitted with audience response system technology (like clickers) to improve student engagement.

I understand the people who make this case because I used to be one of them, a professor filled with righteousness about all of the things that students had to learn. Then I went into administration, saw the data on how poorly universities were doing, and changed my mind. A smart social science colleague told me that instead of weeding out, we should be weeding in. That rocked my world. Even so, I mostly failed at getting the PowerPoint and the big chalk out of the hands of my colleagues who were convinced that the old way was the right way.

The irony of a Ph.D. education is that the research part actually does the right thing by getting students to learn material when they need it for their research. Lots of research on learning supports the idea that we learn new things best when we need them. Legendary biologist E. O. Wilson sounded off on this matter to *The Chronicle of Higher Education* last year. He was unhappy about science teaching and lamented the "intellectual triathlon" that is used to turn off inquisitive young minds, referring to the dogmatic belief that all scientists have to start by running the gauntlet of biology, chemistry, and physics as "nonsense." "The right way to create a young scientist who's going to be on fire by the time they're in college," he said, "is to let them pick something, some subject that has really excited them."

We know from Nobel laureate Carl Wieman and other scholars that active learning that is heavy on group work and discussion creates a better opportunity for students of all identities to succeed in science. Yet 55% of all science instruction in the United States is still traditional lecturing. Only 27% has a modest intervention such as multiple-choice questions with clickers to engage students during class. And only 18% is designed for students to work through problems together in class after learning didactic material in advance. This is the face of modern science education, even when a Nobel laureate has shown that lectures don't work effectively.

Whom do lectures work for? They work for those who love to memorize facts and equations and can "plug and chug" on exams. They work for those who are not subjected to social cues that make sitting in a large lecture hall and taking high-stakes tests intimidating. They work for those who look like the people who were in the classrooms when this method of teaching was invented. That's not fair and it's not going to cut it for the future of science and the planet.

There are plenty of green shoots, especially Kelly Hogan's reinvention of large lectures by including active learning for groups of up to 400 students, wrap-around programs like Freeman Hrabowski's Meyerhoff Scholars that surround students with support and place them in a cohort that increases their chances of success, and institutions like Xavier University of Louisiana that are lapping well-known research universities in educating African Americans who are admitted to prestigious medical schools. But despite these impressive examples, the future mostly languishes in lecture halls while the PowerPoints and big chalk drone on.

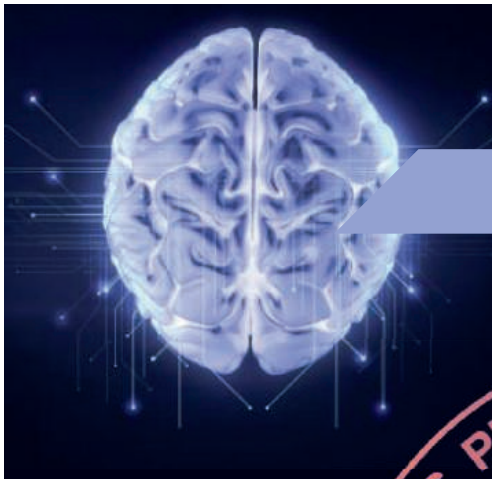
How do we get more folks into science? Let's start by not running them off.

–H. Holden Thorp

H. Holden Thorp
Editor-in-Chief, *Science* journals.
hthorp@aaas.org;
@hholdenthorp

CREDITS (TOP TO BOTTOM): CAMERON DAVIDSON; DON HEUPEL/AP IMAGES

10.1126/science.aba9156



Apply for our exciting research Prize!

SCIENCE-PINS PRIZE FOR NEUROMODULATION

SCIENCE-PINS

$25, 000 Grand Prize!
Get published in *Science*!

The *Science*-PINS Prize is a highly competitive international prize that honors scientists for their excellent contributions to neuromodulation research. For purposes of the Prize, neuromodulation is any form of alteration of nerve activity through the delivery of physical (electrical, magnetic, or optical) stimulation to targeted sites of the nervous system with impications for translational medicine.

For full details, judging criteria and eligibility requirements, visit:

www.sciencemag.org/prizes/pins

Submission Deadline: March 15, 2020

Science AAAS

PINS 品驰

Science Translational Medicine AAAS

A wildfire burned perilously close to endangered Wollemi pines before firefighters came to the rescue.

PHOTO: NEW SOUTH WALES DEPARTMENT OF PLANNING, INDUSTRY & ENVIRONMENT

CONSERVATION

Firefighters save 'dinosaur' pines from Australian wildfire

Australia got a rare bit of good news last week about the wildfires raging there: Firefighters have rescued the only known wild stand of Wollemi pine, a critically endangered type of conifer that dates to the age of dinosaurs. Workers used aircraft to drop water and flame retardant into a single canyon where the pines live, in the Blue Mountains west of Sydney. This unusual kind of *Araucaria* pine was once common across the ancient southern supercontinent of Gondwana but was known only from fossils until 1994, when a park ranger discovered living specimens. Experts fear that Australia's wildfires may have pushed more than 70 other species toward extinction. The blazes have begun to recede in the drought-stricken country as rains have returned to some areas of its east coast. But authorities also worry that in coming months, ash will be washed into rivers, lakes, and watersheds, causing fish kills and potentially tainting drinking water.

greenhouse gas emissions, saying the plaintiffs lacked standing. In a two-to-one decision in *Juliana v. United States*, a panel of the Ninth Circuit Court of Appeals ruled that the youths' challenge was not a matter for the courts. "The plaintiffs' case must be made to the political branches or to the electorate at large," wrote Judge Andrew Hurwitz, an appointee of former President Barack Obama. The lawsuit, filed in 2015 with support from environmental groups and scientists, is based on a legal argument that the government must reduce the emissions to protect natural resources from the effects of emissions-induced climate change, for the sake of future generations. An appeal is possible.

A modern diphtheria treatment

PUBLIC HEALTH | The standard treatment for diphtheria, developed in the 19th century, is produced by injecting a toxin from the diphtheria bacterium into horses and then harvesting their antibodies. But some say the antiquated method is cruel to the horses, and their antibodies induce a life-threatening immune response in some human patients. Now, scientists funded by an international animal welfare consortium funded by People for the Ethical Treatment of Animals have used lab-grown cells to make the antibodies, avoiding the need for horses. A combination of the new antibodies protected guinea pigs from the toxin's effects, the scientists reported last week in *Scientific Reports*; they hope trials in humans will be next. Funding those may be a challenge, though, because diphtheria's rarity makes it of low interest to pharmaceutical companies. Routine childhood immunizations with an inactivated form of the toxin have reduced the number of cases to several thousand annually worldwide, mostly in children who missed their vaccinations.

Fly brain mapping speeds up

NEUROSCIENCE | Researchers have used computing help from Google to assemble the largest and most detailed wiring diagram yet of an animal brain. The partially complete data set, released online this week, maps roughly 20 million connections between neurons in about one-third of the poppy seed–size brain of the fruit fly *Drosophila melanogaster*. A complete neural map, or connectome, has been achieved for only one animal so far—the 300-neuron nematode *Caenorhabditis elegans*; tracing the paths of *D. melanogaster*'s 100,000 neurons by hand would have taken decades. A team at Janelia Research Campus worked with Google scientists, who developed algorithms that more quickly analyze terabytes of data from microscope images; the programs recognize neurons in the brain slices and document their junctions with other cells. The researchers now expect to complete the fly connectome within 2 years. The algorithms were not released, but researchers outside Janelia can use the public data to more easily trace neural circuits that underlie fly behaviors.

AGRICULTURE

United Kingdom breaks from EU farm subsidies

"Revolutionary" change will focus payments on environmental benefits, not food production

By **Erik Stokstad**

When the United Kingdom leaves the European Union at the end of the month, it will sever ties with Europe's farm subsidy policies—and to many researchers, that is a good thing. Last week, the U.K. government proposed radical changes to £3 billion a year in agricultural spending that will focus the money on benefits to climate, ecosystems, and the public. "It's dramatic and utterly critical," says Dieter Helm, an economist at the University of Oxford. "This is an agricultural revolution."

Under the bill, introduced to Parliament on 16 January and expected to become law within a few months, farmers will be given subsidies not simply for cultivating land—the current EU system—but only for delivering "public goods." These include sequestering carbon in trees or soil, enhancing habitat with pollinator-friendly flowers, and improving public access to the countryside. To ease the transition, direct subsidies will be phased out over 7 years beginning in 2021, and new payments for environmental services will be tested in pilot projects. "It certainly could have really positive benefits for the environment," says Lynn Dicks, an animal ecologist at the University of Cambridge who studies wild pollinator conservation.

After the destruction and starvation of World War II, European tariffs helped protect farmers from foreign competition while subsidies boosted their yields. "It was just about production, it didn't matter what you did to the environment," says Ian Bateman, an environmental economist at the University of Exeter. New lands were brought under the plow and hedgerows were ripped up, leading to erosion. Excessive fertilizer and pesticides polluted air and water. And the loss of habitat harmed pollinators and other wildlife. The cost of the EU common agricultural policy (CAP) wasn't just environmental: Up through the 1990s, the subsidies consumed 80% of the EU budget. Even today, the €59 billion CAP represents about 40% of EU public spending.

> **"If it's been successful, that will be a very powerful argument for the Europeans to follow."**
>
> **Alan Matthews**, Trinity College Dublin

Brexit will now let the United Kingdom go its own way. The new bill addresses only England, because the United Kingdom allows Wales, Scotland, and Northern Ireland to determine their own agriculture policies, but Helm expects they will move in the same direction.

Under the new scheme, to be overseen by a body created by the Department for Environment, Food & Rural Affairs (DEFRA), total funding will not change, but some farmers will be impacted more than others. To be profitable, beef and sheep farms rely on subsidies more than dairy and wheat farms and, without them, they might be abandoned in hard-scrabble places such as in Scotland. But Helm sees lifelines for some of these farms, such as payments for sequestering carbon with tree plantations or restored peatlands. Grants for restoring heritage buildings or enhancing landscape beauty could also help sustain farms while boosting tourism. Other payments will help farmers adapt

PHOTO: PETER MULLIGAN/GETTY IMAGES

U.K. farm subsidies will require efforts to support public goods, such as recreation.

to climate change or reduce their environmental impact. Subsidies for equipment to inject manure into the soil, for example, could reduce both air pollution and the need for chemical fertilizers.

About one-eighth of existing U.K. farm subsidies pay for environmentally friendly activities such as maintaining hedgerows and other habitat, and those efforts will expand. To get more value for money, DEFRA plans to use auctions, in which farmers or other land managers would bid for government contracts for environmental services. Water companies have already used auctions to select farmers who are paid to use less fertilizer and different pesticides, lowering water treatment costs. "The impact has been amazing," Bateman says.

DEFRA wants to tailor payment schemes for different regions but figuring out how to maximize the benefits will require research. Carbon sequestration payments could backfire if used in the wrong places. For example, planting trees in peatlands can dry them out, releasing more greenhouse gases than would ever be sequestered by the trees, Bateman says.

Socio-economic models will be needed to study the impact of the policy changes on farms and rural communities, adds David Harvey, an agricultural economist at Newcastle University. Much remains to be determined. Will farmers only get payments if air and water quality are shown to improve? And over what time scale? Who will measure it? "You're left with more questions than answers," says Mark Sutton, a nitrogen expert at the Centre for Ecology & Hydrology.

Farmers—especially owners of vulnerable small operations—have eyed all these changes warily. The National Farmers Union, the United Kingdom's biggest agricultural trade group, lobbied for more emphasis on supporting food production. The bill stipulates that the government will "take regard to the need to encourage the production of food by producers in England," which the union calls a "robust starting point" for designing the new support programs. But Bateman and others worry about backsliding toward payments that support private profits, rather than environmental progress.

Other countries will be watching closely, too, says Alan Matthews, an agricultural economist at Trinity College Dublin, who studies European agricultural policy. "If it's been successful, that will be a very powerful argument for the Europeans to follow." ■

PHOTO: REUTERS/ANUSHREE FADNAVIS

SCIENCE AND POLITICS

Campus attack and police violence alarm Indian academics

Many scientists are at odds with Hindu nationalists

By **Sanjay Kumar**, *in New Delhi*

Early in the evening on 5 January, more than 70 masked intruders armed with iron rods, stones, and sticks entered the campus of Jawaharlal Nehru University (JNU) here. They set upon teachers and students who were holding a peaceful political gathering, and marched into student hostels, terrorizing and injuring dozens. Panicked students posted videos on social media and called the police for help, which didn't arrive.

The attack, apparently by Hindu nationalists, on one of India's most prestigious universities is the latest sign that the political forces tearing apart Indian society are also affecting the country's academic community. Students at JNU, a liberal bastion, had been on strike for months against both a major hike in student fees and the government's controversial Citizenship Amendment Act (CAA), widely decried as discriminatory against Muslims. From the inaction of both campus security and the New Delhi police, many concluded that the mob acted with the consent of India's Hindu nationalist government.

For many academics, the rampage—which came on the heels of a brutal police response to several other university protests last month—felt like an assault on freedom of speech and democracy itself. "It looks like we are living in an era [of] textbook fascist methodology," says Dinesh Abrol, a spokesperson for the Delhi Science Forum, a nonprofit organization that promotes science. "The space for dissent, free thinking, and contrarian views has already shrunk," says geographer Sucharita Sen, a professor of regional development at JNU who was hit on the head with a brick during the attack.

Many say they're worried about being branded as "antinationals" or communists and targeted on TV and social media. "I downplay my identity now and don't express any opinions that may sound political," says a Muslim scientist at JNU who asked not to be identified. Although Krishnaswamy VijayRaghavan, the Indian government's principal scientific adviser, "unhesitatingly and unequivocally" condemned the violence at JNU, many say politicians have fanned the flames. Days before the attack, Home Minister Amit Shah—to whom the New Delhi police report—said "antinational gang members" at JNU "should be taught a lesson."

The violence came on the heels of other clashes sparked by the CAA, which became law on 12 January and is designed to provide citizenship to persecuted minorities from Pakistan, Bangladesh, and Afghanistan. Muslims are excluded, which critics say violates India's secular constitution. (India's prime minister, Narendra Modi, has argued that

Aishe Ghosh (center), president of JNU's student union, suffered multiple injuries during a 5 January attack.

"people are being misled" over the act.) A National Register of Citizens and a National Population Register, which would force every Indian to produce documents proving their citizenship, are expected to follow. Many Muslims fear they will be declared stateless.

On 15 December 2019, police beat and teargassed students protesting the CAA at Jamia Millia Islamia (JMI), a government-funded university here; one student was blinded. The same day, Aligarh Muslim University in Uttar Pradesh state was closed after police forcefully quelled a protest. A fact-finding team concluded that tear gas shells, stun grenades, and bullets left more than 100 students "with shattered bones, grave injuries, deep bruises, and severe psychological trauma." The right hand of a doctoral student in chemistry, Mohammad Tariq, had to be amputated after he was hit by a shell.

India's scientific establishment, traditionally apolitical, has also spoken out against the CAA. The past months saw protests at leading research centers, including the Indian Institutes of Technology, the Indian Institutes of Science Education and Research, and the Jawaharlal Nehru Centre for Advanced Scientific Research, many spearheaded by young scientists. Nearly 2000 scientists and science students have signed an open letter denouncing the bill, including leading researchers such as Sandeep Trivedi, director of the Tata Institute of Fundamental Research, and Rajesh Gopakumar, director of the International Center for Theoretical Sciences.

Researchers of Indian descent abroad have spoken out as well, including Venki Ramakrishnan, president of the United Kingdom's Royal Society. Abhijit Banerjee, a JNU alumnus at the Massachusetts Institute of Technology who won an economics Nobel in 2019, warned that the attack on JNU "has too many echoes of the years when Germany was moving towards Nazi rule."

The CAA isn't the only problem scientists have with the Modi government. Many also abhor the rise of pseudoscience rooted in Hindu nationalism (*Science*, 15 February 2019, p. 679) and the budgetary neglect of science and higher education. The central government's budget for universities has plummeted from 0.6% of gross domestic product in 2013–14, the year Modi came to power, to 0.2% in 2018–19.

Some fear the growing divisions will slow India's recent progress in science and technology. "Look at where our next-door rival China is going," says one physicist at JMI. "This government is taking us hundreds of years backwards, to medieval ages. It's pitting the entire population against each other and setting the country on fire." ■

Sanjay Kumar is a science journalist in New Delhi.

IMAGE: DANA BERRY/NAT GEO CREATIVE

SPACE SCIENCE

Planetary turmoil unleashed during Solar System infancy

Models and observations suggest that an earlier date for the upheaval can explain puzzles including a puny Mars

By **Paul Voosen**

An early maelstrom shaped our Solar System. Sometime after the planets took shape from primordial gas and dust, resonant tugs between the giant planets threw their orbits out of kilter. The gravity of the errant giants blasted Pluto and its many icy neighbors into the far-out Kuiper belt. The instability also scattered oddball moons and asteroids and triggered smaller bodies to pummel the inner planets.

Now, that scenario is experiencing some upheaval of its own.

Scars on the Moon had convinced many planetary scientists that the storm hit about 3.95 billion years ago, 650 million years after the Solar System formed. But this model has long had a flaw: Mercury, Venus, Earth, and Mars would likely not have survived such a late assault. And over the past few years, a new timeline has begun to emerge, one that shifts the chaos earlier, to less than 100 million years after the system's creation—and perhaps as few as 10 million years. "The tides are moving and people are now more and more convinced that the instability happened early," says David Nesvorný, a planetary scientist at the Southwest Research Institute (SwRI) in Boulder, Colorado. Several new papers explore what triggered this early instability and how it can explain a host of Solar System quirks.

Two decades ago, scientists recognized that planets must have migrated to create the modern Solar System. A group including Alessandro Morbidelli, a planetary scientist at the University of Côte d'Azur, gathered in Nice, France, for 1 year to hash out the idea, creating what's known as the Nice model (*Science*, 17 July 2009, p. 262). As the model now goes, after the giant planets formed out of the gas disk, Jupiter drew its fellow giants into a resonant chain of orbits where, for example, Saturn orbited the Sun three times for two turns of Jupiter. The surrounding gas acted as a damping agent, calming any instability like an air conditioner in a room of irritable siblings. But once the gas dissipated, the collective push and pull of giant planets' masses, agitated by nearby planetary building blocks, unleashed chaos.

The turmoil came relatively late, suggested lunar rocks collected from impact craters by

Soon after the planets formed, the gravity of the gas giants flung debris around the Solar System.

the Apollo astronauts. The ages of the rocks seemed to indicate that the Moon suffered a cataclysmic assault, dubbed the Late Heavy Bombardment (LHB), 3.95 billion years ago, sandwiched by hundreds of millions of years of quiet. But over the past few years this story has evaporated, says Nicolle Zellner, a lunar geochemist at Albion College. New work suggests rocks collected by astronauts at multiple craters, once believed to represent simultaneous strikes, are instead debris from a single impact, 3.95 billion years ago, that created Imbrium Basin. More precise dating of meteorites ejected from the Moon shows that the impacts responsible took place as many as 4.3 billion years ago—or well after the supposed LHB. "The idea of a very strong cataclysm has gone away," Zellner says.

Planetary dynamicists have welcomed the LHB's vanishing act. Their models had highlighted a puzzle: A late catastrophe would have either destroyed the rocky planets of the inner Solar System or disrupted their stately, nearly circular orbits, flat with the Solar System's plane. "This was my first red flashing light," says Kevin Walsh, a planetary scientist at SwRI.

Now, in a paper accepted for publication in *Icarus*, Morbidelli and co-authors show that such a late instability wouldn't work in any case. Their computer modeling indicates that, for a late instability to have created the current Solar System, a large gap would have had to exist between Neptune and the encircling disk of planetary building blocks outside its orbit. But the gap rarely appears in the models. And without the gap, it's impossible to delay the catastrophe, Morbidelli says.

Freed from the late constraint, planetary scientists are now exploring how an earlier cataclysm could explain odd features of the Solar System. Over the past few years, Matthew Clement of the Carnegie Institution for Science, Walsh, and others have shown in computer simulations that an instability less than 10 million years after Solar System formation would allow the inner planets to coalesce in peace. An early instability would also scour away planet-forming material near Mars and the asteroid belt, explaining their weirdly low masses. And in a paper published last month in the *Monthly Notices of the Royal Astronomical Society: Letters*, they show that as Saturn moved away from Jupiter near the end of the instability, a final tug between them might have flung away asteroids in orbits far removed from the orbital plane, giving the asteroid belt its current compact structure. "We kind of simplify the whole story," Clement says. "We can have one event explain all these problems."

Still, "The details are strongly debated," says Thomas Kruijer, a geochemist at Lawrence Livermore National Laboratory. There's little direct evidence for such an early instability, and at least two other scenarios that could explain how the rocky planets survived. Clement also has yet to reconcile a similarity between noble gases measured by the Rosetta spacecraft around the comet 67P and features of Earth's atmosphere, which suggests the instability likely caused Earth to be bombarded with a hail of comets after it was solid—not before.

But Kruijer says a bombardment within the first 100 million years of the Solar System is plausible. Perhaps the best evidence for it is now found near Jupiter, Nesvorný adds. There, following Jupiter in its orbit, spins a binary asteroid named Patroclus-Menoetius. The icy composition of its two bodies indicates they formed in the far reaches of the Solar System and were implanted into Jupiter's wake during the instability. In a 2018 paper, Nesvorný and co-authors showed there's no way the binary would have survived 600 million years in the outer Solar System—collisions would have ground it up after only 100 million years. "That's a very solid constraint" supporting an early instability, Morbidelli says.

The hunt is on for more observations that can parse what happened during those first 100 million years, whether from asteroid samples, clusters of primordial asteroid families, or craters on the Moon and Mars. "Now, the question is, was it a few million years after or 80 million years?" Morbidelli says. "Honestly we don't know." ■

FOREIGN INFLUENCE

Florida center details fired scientists' links to China

Report finds six cancer researchers didn't disclose Thousand Talents deals

By **Jeffrey Mervis**

Six Florida cancer researchers who were dismissed last month for hiding their ties to a Chinese medical university appear to have been motivated by simple greed and a disregard for both institutional and federal rules.

A report last week from the Moffitt Cancer Center describes how one longtime Moffitt researcher, a Chinese American, recruited four colleagues, including Moffitt CEO Alan List, to participate in Chinese foreign talent recruitment programs. But none of the researchers disclosed those deals—or the personal payments they received—to Moffitt, a nonprofit private entity created by Florida in 1981, or to the National Institutes of Health (NIH), which last year awarded Moffitt $36 million in grants. The report notes that some researchers saw List's involvement as an indication that Moffitt condoned such behavior.

The six Moffitt researchers, none of whom would comment to *Science*, add to a growing list of scientists found to have violated institutional and federal policies on disclosing who is funding their research. In August 2018, NIH began to send letters to more than 60 grantee institutions about nearly 200 individuals NIH believed had skirted its rules. In spring of 2019, the University of Texas MD Anderson Cancer Center and Emory University went public with their investigations, but most institutions have kept mum about what they have found. Last week, the University of Florida acknowledged it had dismissed four faculty members but did not release their names nor details of their infractions.

Moffitt says it did not receive any letters from NIH. Its own compliance office initiated the investigation in January 2019, after some employees provided information that awakened its suspicions. This month, the

Florida legislature launched its own probe of foreign research collaborations at state-funded universities.

China's Thousand Talents Program, launched in 2008, is at the center of both investigations. "All Moffitt faculty participants in the Talents programs acknowledged receiving personal payments that they did not promptly disclose to Moffitt," Moffitt officials wrote on 17 January to state Representative Chris Sprowls (R), who leads a special legislative committee created this month to probe the collaborations. "They also acknowledged having opened or maintained personal bank accounts in China to receive Talents program compensation."

In announcing the departures of the six researchers on 18 December 2019, Moffitt identified only List and the director of its research program, Thomas Sellers. The letter to Sprowls names the others, and offers a detailed description of how the researchers became involved with the talent recruitment program and the payments they received from Tianjin Medical University (TMU) and its affiliated cancer institute and hospital. Moffitt also submitted nearly 1400 pages of material, some of it confidential, to the special panel, which will hold its first hearing this week.

The former CEO of Florida's Moffitt Cancer Center did not report personal payments from a Chinese institution.

A summary of Moffitt's investigation, released publicly last week, found that the central figure was immunologist Sheng Wei, a TMU graduate who came to Moffitt in 1992. Wei began to receive support from the Thousand Talents Program in 2011 and over the next several years tried to recruit others, at Moffitt and elsewhere, to join the program, investigators say. At Moffitt, four researchers signed on: List; Sellers; Daniel Sullivan, head of the center's clinical science program; and cancer biologist Pearlie Epling-Burnette. All have resigned.

The sixth scientist who left, pharmacogenomicist Howard McLeod, was already participating in the Thousand Talents Program at another Chinese institution when he came to Moffitt in 2013 to lead an institute on personalized medicine. According to the report, McLeod enlisted a Chinese scientist, Yijing He, as his "agent in China." McLeod managed to put He on Moffitt's payroll for 5 years, although He never worked at the Florida cancer center. Instead, the report says, "Dr. He facilitated a wide variety of opportunities and activities in China, both commercial and academic, for himself and Dr. McLeod."

The report notes that Moffitt already receives $500,000 a year from the Tianjin cancer institute and hospital for joint research, training, and consultation. But the fired scientists erred in not reporting their personal remuneration from that institution. Nor did they tell NIH in their grant applications. The report notes that the researchers also failed to disclose the existence of personal bank accounts in China in which the unreported funds were deposited.

The report leaves unanswered the question of what the scientists did in return for the payments. It describes how McLeod and Wei appear to have used existing federal grants from U.S. agencies to help carry out research in China, a practice that NIH has called running a "shadow lab." At the same time, the report also suggests several Moffitt scientists were scamming their Chinese counterparts by not putting in the time they had promised under their Thousand Talents deals.

"Dr. Wei stated that the substantial [2 or 3 months per year] time commitments he elicited on recorded videos from all Moffitt faculty participants, including Dr. List, were an open 'lie' necessary to apply for the Talents programs and to enable the participating Moffitt faculty to receive their Talents programs payments," the report says. ■

DATA

NSF rolls out huge makeover of science statistics

Agency wants its biennial *Indicators* to be more timely in tracking global trends

By **Jeffrey Mervis**

In a world awash with data on demand, a report on the global scientific enterprise that appears every 2 years is an anachronism. So the National Science Foundation (NSF) has made its beloved *Science and Engineering Indicators* more timely and nimble.

The 2020 version of *Indicators*, officially released last week, still covers international trends in science education, the demographics of the technical workforce, funding, the nature of industrial and academic research, and public attitudes toward science. But NSF is releasing this report's eight chapters, totaling almost 2000 pages, on a rolling basis to speed up its analysis of the dynamic scientific landscape. The first chapter, on precollege math and science education, appeared online in the fall of 2019, for example, and the final chapter, on public attitudes, will include the latest polling results when it comes out this spring. Last week's release of a brief summary was intended to satisfy a federal mandate for a biennial report to Congress.

Going forward, the flexible production schedule will allow NSF to issue more specialized reports as the opportunity arises, notes Beethika Khan, who heads NSF's in-house *Indicators* crew. NSF has also revamped the online interface to make it easier for researchers and the public to explore and download the data, and now provides quarterly updates on research activity at the state level that are gobbled up by state and federal legislators.

The new *Indicators* still suffers from a significant time lag in some areas. Last year, for example, China likely topped the United States in overall research spending for the first time in history. But the 2020 *Indicators* doesn't record that milestone. Instead, the report shows the United States retaining the top spot, at $548 billion, because it is based on 2017 data.

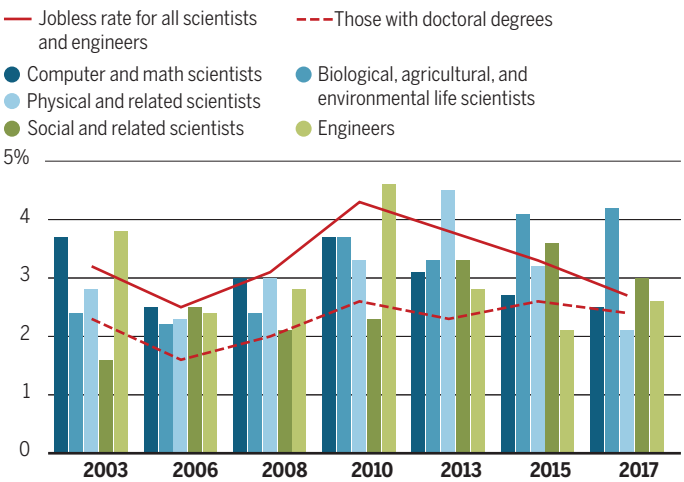
PHOTO: MOFFITT CANCER CENTER

Vital statistics on global science

This year's version of the National Science Foundation's *Science and Engineering Indicators* has a new look (see main story, p. 352), but it still delivers a cornucopia of statistics on the global scientific enterprise. The graphs below highlight important trends in the U.S. scientific workforce. But *Indicators* also puts those U.S. trends in a global context, as seen in the graph on publications.

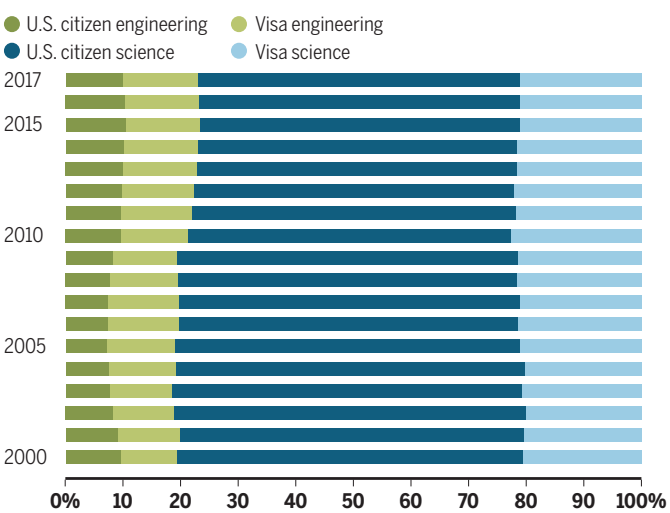
A uneven recovery from the great recession

Although the overall jobless rate for U.S. scientific workers has dropped sharply since the global 2008 downturn, those in the physical sciences have benefited the most from the economic recovery, whereas U.S.-trained Ph.D.s and biologists feel the lingering effects.



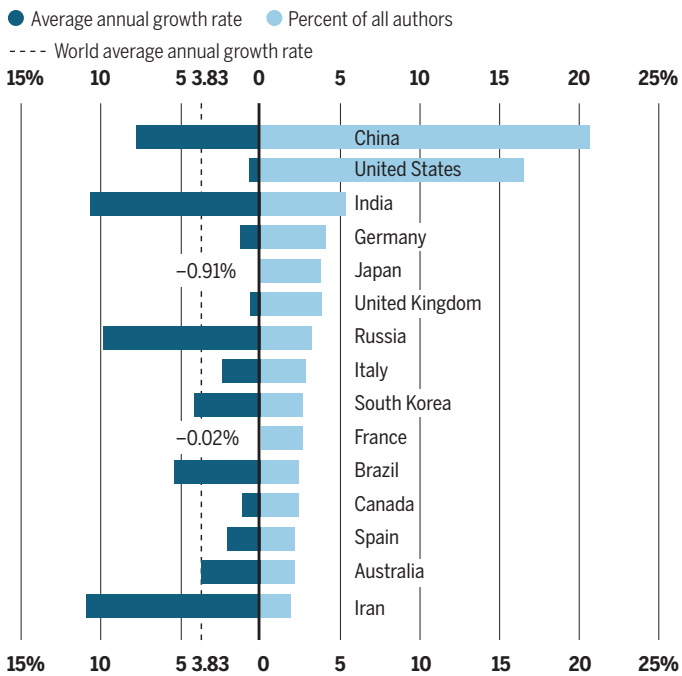
Domestic U.S. students are not endangered

Contrary to popular belief, the share of domestic students earning Ph.D.s from U.S. institutions has changed very little since 2000 in both engineering—although they remain a minority—and across the natural and social sciences, where they have roughly a three-to-one edge over foreign students holding visas.



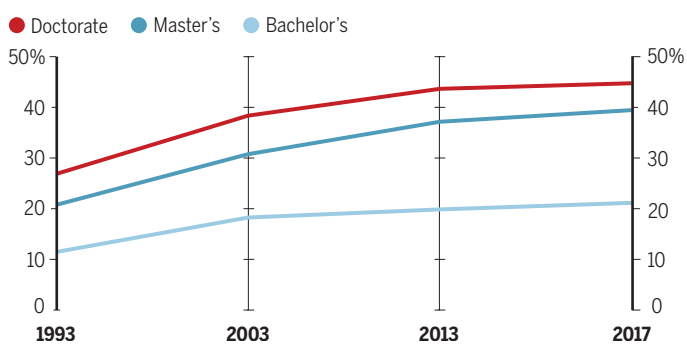
The paper race

Iranian scientists represent fewer than 2% of all authors globally, but they have enjoyed the fastest growth in market share over the past decade. China's 21% share of all papers not only puts it on top of the heap, but its annual growth in publications is 10 times that of the United States, which holds second place.



Immigrants bolster the U.S. technical workforce

Foreign-born scientists represent nearly 30% of the U.S. science and engineering workforce, nearly double their numbers in 1993. Their presence is largest among workers holding doctoral degrees, at almost 45%.

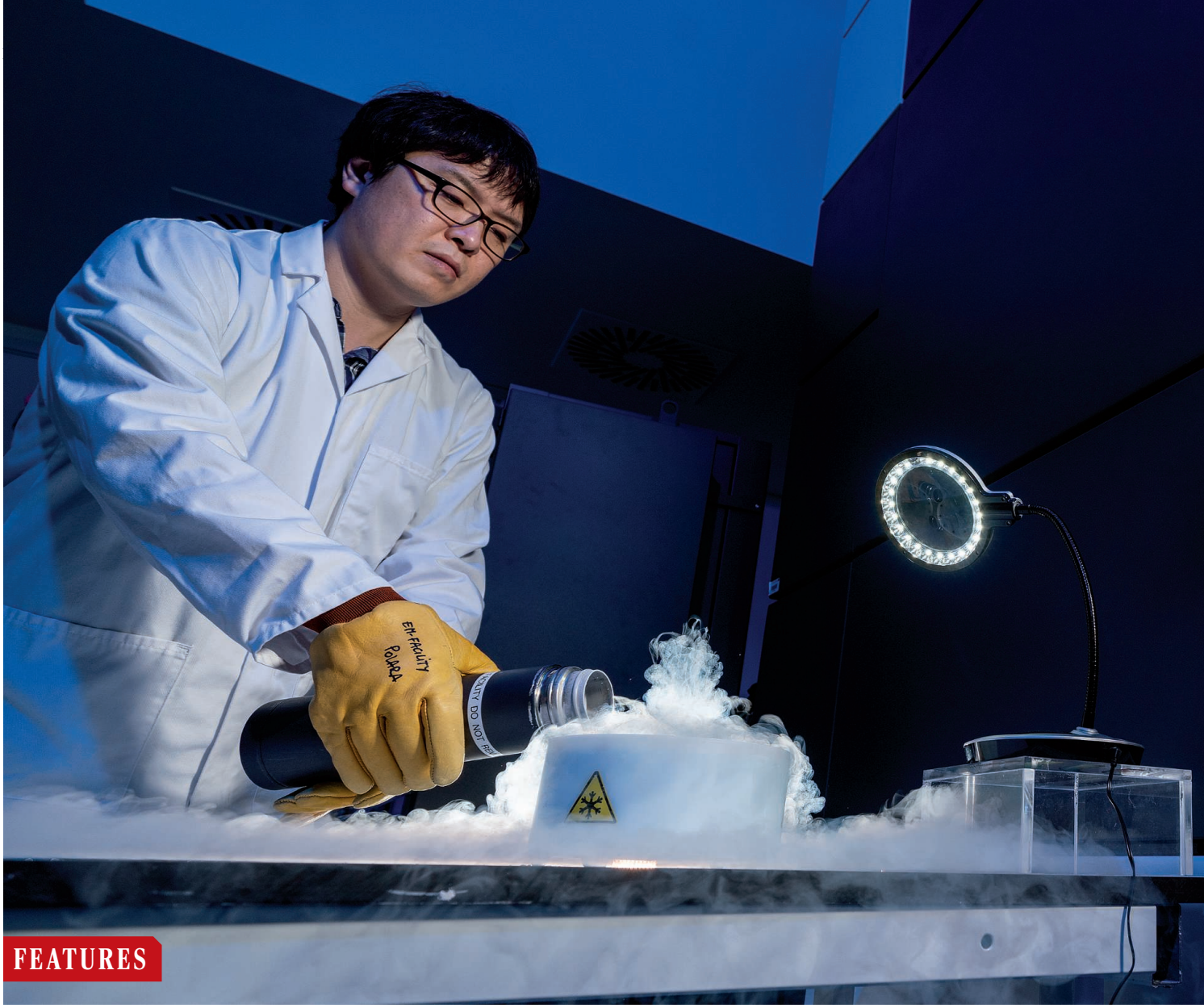


CREDITS: (GRAPHICS) X. LIU/*SCIENCE*; (DATA) 2020 SCIENCE AND ENGINEERING INDICATORS

"We think that China may have overtaken the U.S. at some point in 2019," says Julia Phillips of the National Science Board, NSF's presidentially appointed oversight body, which issues *Indicators*. The board had previously predicted that the changing of the guard might occur in 2018, she notes, but 2018 data, which NSF published earlier this month in a separate report, show that a stronger-than-expected investment by U.S. industry allowed the country to retain its lead over China. NSF analysts expect upcoming 2018–19 data from the Organisation for Economic Co-operation and Development will reveal a new global spending king.

The report does not say what China's ascent might mean for global innovation, or for U.S. economic health and national security. That's by design, Khan says. "We do not look at the why questions."

Those questions fall in the realm of the science board, Phillips says. And she hopes U.S. leaders see the heightened competition from China and other nations as a call to action. "The United States is not likely to regain its past position of dominance," Phillips says. "But we can use these data to lay out a vision of where we want to lead, and then adopt policies and make investments to achieve those goals." ■



CHEAP SHOTS

Cryo-EM reveals exquisite molecular structures—at high cost. A cheaper microscope could bring the resolution revolution to the masses

By **Eric Hand**

The Laboratory for Molecular Biology (LMB), clad in glass the color of sea ice, rises like a futuristic factory above the rapeseed fields of Cambridge, U.K. It is the crown jewel of the U.K. Medical Research Council, a storied government lab that has garnered more than a dozen Nobel Prizes. One of the first came in 1962 after LMB researchers, having pioneered x-ray crystallography, used the technique to decipher the first atomic structures for proteins—those of myoglobin and hemoglobin, which carry oxygen in muscle tissue and blood. X-ray crystallography has dominated structural biology ever since, but it has an Achilles' heel: Some proteins just can't be coaxed to form crystals, which scatter x-rays to reveal structure.

One of LMB's most recent Nobel laureates is Richard Henderson, an unassuming Scotsman who slouches in his office chair in socks and sandals, surrounded by mountains of paperwork. In 2017, Henderson won a share of the chemistry prize for

PHOTO: JON MOLD

Corrected 24 January 2020. See full text.

Published by AAAS

Byung-Gil Lee, wearing thick gloves to protect him from frigid temperatures, prepares samples for cryo–electron microscopy. He hopes to get a close-up view of cohesin, a protein involved in cell division.

PHOTO: JON MOLD

work in developing detectors for cryo–electron microscopy (cryo-EM). The technique, also pioneered at LMB, is the brash upstart in structural biology, challenging crystallography in resolution and surpassing it in purview: It opens up far more proteins to inspection and captures many more of their natural configurations.

Cryo-EM dodges the problem of crystallization with a flash-freezing process that fixes proteins in thin films of glassy ice. Then, an electron microscope takes thousands of 2D snapshots of the proteins caught in random orientations. A computer stitches them together to reveal the 3D structure, so important in understanding how a protein works—and how a drugmaker might target it.

But cryo-EM has a big problem of its own: long waits to use extraordinarily expensive microscopes. Moreover, tricky sample preparation means that even when researchers get access, much of their time ends up wasted. "There's this dirty little secret of the field," says LMB physicist Chris Russo. "It's almost as hard as making a crystal. There's a lot of trial and error in it."

Some researchers are aiming to fix those problems with automation: robots that can make icy protein samples more reliably and with less waste. Other scientists are developing materials that protect protein molecules during freezing. Still others are working on software that could gather data more efficiently.

But Henderson sees those efforts as mere Band-Aids. For him, what holds back the technique is the forbidding cost of a microscope. Henderson, Russo, and a small group of confederates are trying to make cryo-EM affordable.

A top machine costs about $7 million. Preparing a room and installing a microscope can cost just as much. Then come the operational costs—a torrent of electricity, dedicated troubleshooting staff—that can rise to $10,000 per day.

Roughly 130 Krios machines—the microscopes widely considered the best—have been sold by Thermo Fisher Scientific and installed around the world. LMB has the luxury of three for a relatively small staff, and yet even its researchers must wait a month or more to get time.

Most structural biologists have no access at all. "The wait can be from 3 months to infinity," says Bridget Carragher, co-director of the Simons Electron Microscopy Center in New York City, a mecca for cryo-EM. "It's becoming the haves and the have-nots."

No one complains about the quality of the Krios machines, which take months to assemble by hand from thousands of parts in a Dutch factory. They are Cadillacs. But science would benefit from something less

Corrected 24 January 2020. See full text.

Published by AAAS

posh, Henderson says. "We need a people's cryo-EM for maybe 10 times less: a Volkswagen Beetle."

For decades, microscopists have opted for machines that run at high energies and require expensive parts and precautions. After revisiting the basic physics, Henderson, Russo, and colleagues showed that a cheaper, lower-energy microscope can take pictures that are just as good, if not better.

Now, Henderson hopes to persuade manufacturers to make his cut-rate machine. At less than $1 million, it should put cryo-EM within reach of thousands of labs. Doing so would democratize the field, he says, and accelerate the discovery of protein structures. At the very least, Carragher says, researchers could use a cheap machine to screen out bad samples, preventing wasted time on a Krios. "People need to learn the trade and try things out," she says. "They shouldn't be doing it on the glamour machine."

CRYO-EM'S RISE SEEMS unstoppable. At first, researchers worried that the technique could not go small enough: Before 2010, it could attain resolutions below 4 angstroms—four times the diameter of a hydrogen atom—only for a few symmetric, easy-to-solve structures. But machines like the Krios, along with new detectors that record the path of pertinent electrons before the electron beam fries the sample, have changed the game. In the best cases, researchers can now make maps with resolutions below 2 angstroms, putting cryo-EM on par with crystallography.

Cryo-EM also has distinct advantages over the older technique. Consider the study of the cell's gatekeepers, membrane proteins, which are drugmakers' most popular targets. The proteins are tough to crystallize because they flex to let things in and out of the cell. Crystallographers must put them in chemical straitjackets to stabilize them and get them to crystallize. But doing that means losing valuable information about how a floppy protein functions, says Melanie Ohi, a structural biologist at the University of Michigan, Ann Arbor.

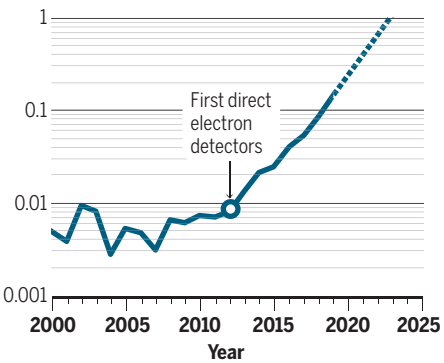
By contrast, cryo-EM can freeze proteins in any of their shapes, showing how they act. This year, for example, Sriram Subramaniam, a structural biologist at the University of British Columbia, Vancouver, and colleagues used cryo-EM to obtain snapshots of Cas9, the enzymatic scissors of CRISPR gene editing, as it snipped and cut DNA. A protein "is not a single, static thing," he says. "Cryo-EM tells you the story of how things work."

X-ray crystallography isn't going away anytime soon: It is buttressed by the sunk costs of billion-dollar synchrotrons, government-owned facilities that supply bright x-ray beams for crystallographers and other users. But trends at the Protein Data Bank, a repository of protein structures, reflect cryo-EM's growing popularity (see chart, below). The number of proteins solved by x-rays peaked in 2017, whereas the number of cryo-EM solutions is nearly doubling every year. Cryo-EM will soon surpass x-rays, predicts Jim Naismith, director of the Rosalind Franklin Institute, a biological imaging center that will open in 2021 in Didcot, U.K., and will task many of its 200 researchers with improving cryo-EM methods. "In 5 years or less, EM will be the dominant method," he says.

Icy tool gets hot

X-ray crystallography still leads cryo–electron microscopy in solving protein structures. But a turning point came in 2012, when the first high-resolution cryo-EM cameras appeared. Cryo-EM could surpass crystallography in several years.

Ratio of proteins solved by cryo-EM to proteins solved by crystallography



At LMB, it already is: The lab crossed that threshold in 2016.

To alleviate some of the pent-up demand for access, the U.S. National Institutes of Health (NIH) in 2018 announced $130 million in grants to establish cryo-EM centers in New York, California, and Oregon, each with a few Krios machines. As with synchrotrons, time on the machines at those centers is freely available to researchers. Users simply arrive with their frozen samples—or ship them—and cross their fingers for good data.

Yet many leave empty-handed, says Claudia López, co-director of the cryo-EM center at Oregon Health & Science University (OHSU). "You can have the best microscope in the world, but if your sample is no good, there's nothing you can do about it," she says.

It's not just the chanciness of sample preparation; it's also the lack of access that frustrates Claudio Grosman, a biophysicist at the University of Illinois, Urbana-Champaign, who studies a membrane protein in muscle cells and neurons that binds to nicotine and other drugs. He is waiting to see whether he will get time at the Simons Center, which has added four NIH-funded Krios machines to three existing ones. "I am used to being limited by my own skills in my lab, by my own capacity to read and understand the literature," he says. "It's hard to compete with these labs that have a cryo-EM machine in their basement."

Even with seven machines, Carragher says queues at the Simons Center are long and fretful. "Everyone wants more time," she says, "everybody thinks they're not getting enough."

IT'S EASY TO SEE WHY. At LMB in July 2019, postdoc Byung-Gil Lee finally had time on a Krios after weeks of waiting—and weeks of purifying a protein until hundreds of trillions of copies were concentrated in a single drop. The object of his desire: cohesin, a protein involved in cell division that is defective in many cancers. Only bits of its structure have been solved with crystallography, and Lee had just 24 hours to try for a better picture using cryo-EM.

With a pipette, he beaded the solution onto 10 sample holders, each the size of a flea. Then he froze them in a flash, as a robot plunged them into a bath of liquid ethane. He held a hair dryer up to a dentist's palette of tiny tools and blasted away residual water, which could contaminate the samples. With tweezers, he slotted the sample holders into a rack and doused it in hissing liquid nitrogen to keep the samples frozen. He loaded the rack into the humming Krios towering behind him. An unseen mechanism slipped it into a bright beam of electrons accelerated to three-quarters the speed of light.

Swiveling to a computer command station, Lee fiddled with knobs and a joystick to zoom in. Splotches and blank spots on the sample holder, or grid, indicated that the thin films of ice didn't completely form. "I don't think this grid has good ice," he says. "This is common."

Lee ultimately found that just three of his 10 sample holders were worth inspecting. He didn't get enough snapshots for the computer to build up a 3D picture of the protein. He had to get back in the queue and wait for more microscope time.

How best to tackle such problems? One remedy is to be thriftier with samples of proteins, which are time-consuming and expensive to make and then casually wasted. Typically, researchers use pipettes to apply microliters of the protein solution to sample

CREDITS: (GRAPHIC) V. ALTOUNIAN/*SCIENCE*; (DATA) THE PROTEIN DATA BANK

Corrected 24 January 2020. See full text.

holders and blot away the excess with filter paper. In 2016, Carragher and colleagues reported developing a dispenser that sprays the solution like an inkjet printer, releasing picoliters instead of microliters—only one-millionth as much.

Naismith wants to look further upstream and reduce the amount of protein solution made in the first place. Cryo-EM needs fewer copies of a protein molecule than x-ray crystallography, and yet purifying techniques haven't changed significantly in 20 years, he says. Once the Rosalind Franklin Institute opens its doors, he says, its researchers will work on techniques that could supplant chromatographic columns—the tall tubes that separate and purify proteins.

Sample preparation has a more fundamental problem: The air-water interfaces at the top and bottom of the thin films are perilous for proteins. Drifting proteins that happen to encounter an edge in the moments before flash-freezing—which is likely given the films' thinness—tend to stick to the surface. "This peels them open and really destroys them," Russo says.

Russo and colleagues found a way to protect them, by undergirding the thin film with graphene, a one-atom-thick layer of pure carbon that is transparent to the electron beam. The researchers patterned the graphene with "functional groups"—chemical studs such as carboxyl or amine groups. Those groups jut into the solution and stick to passing proteins, preventing them from wandering into the air-water interface above. "Instead of having two dangerous surfaces, we're replacing one of them with one we can control," Russo says.

Other researchers are trying to wring efficiencies at the back end of the process. New K3 detectors released in 2017 by the company Gatan are not only accurate, but fast. They can vacuum up 6000 pictures in 1 day, several times the rate of a few years ago, Carragher says, allowing more proteins to be tested.

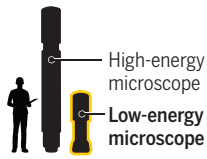
BUT HENDERSON SAYS the major bottleneck is the cost of the machines, which keeps them scarce.

A lack of competition helps explain the cost. Hitachi and JEOL make cryo-EM microscopes, but they have not put a dent in Thermo Fisher's commanding market share. Gatan has a similar hold on the market for detectors, which are sometimes sold separately from the microscopes at roughly $1 million a pop.

The monopolistic pressures nearly got worse. In June 2018, Thermo Fisher made a $925 million bid to buy Gatan and form a cryo-EM superpower. But in April 2019,

A costly view

Cryo–electron microscopy (cryo-EM) reveals the structure of proteins by probing a flash-frozen solution with a beam of electrons, and then combining 2D images of individual molecules into a 3D picture. The cost of high-energy microscopes has limited the method's adoption, but researchers are pushing for cheaper, smaller, lower-energy machines (right).



Electron source

NOW
A field emission gun generates a beam of electrons that is typically accelerated to energies of 300 kiloelectronvolts (keV).

FUTURE
At lower energies of 100 keV, a microscope wouldn't need an expensive spark-snuffing system.

Bending the beam

NOW
High-energy machines use heavy magnets to focus the electron beam. Thick lead shielding soaks up stray radiation.

FUTURE
A low-energy machine would be lighter and wouldn't require as much shielding.

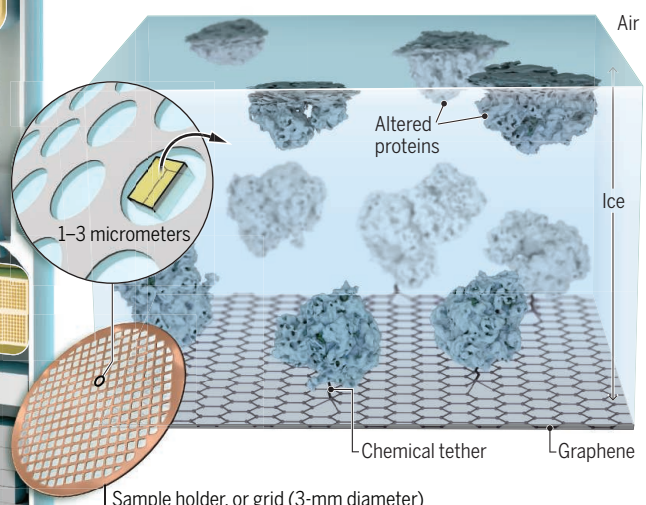
Protein sample dock

Imaging system

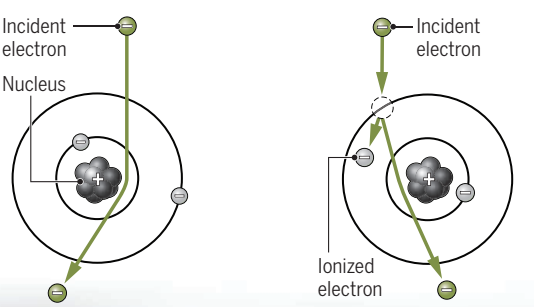
NOW
Electron detectors, typically made of silicon, are costly.

FUTURE
A low-energy machine would use a cheap detector customized for 100-keV electrons.

Deep freeze
Protein solutions are flash-frozen in baths of liquid ethane to create a clear thin film of ice. But the process is a crapshoot: Proteins are often drawn to, and altered by, the air-water interfaces. Layers of graphene, studded with sticky chemical tethers, could keep proteins from straying too close.



Close encounters
Electrons that are deflected by a protein's nuclei (below left) instead of colliding with its electrons (below right) are key to building a cryo-EM image.

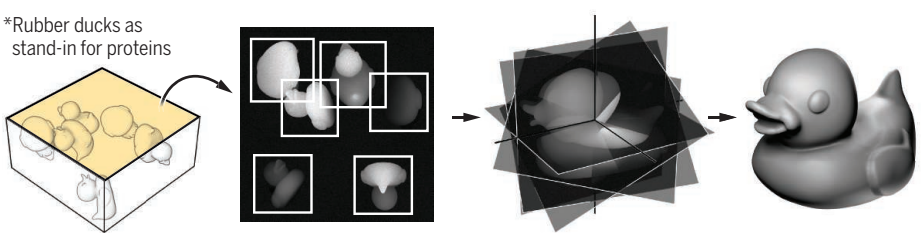


Elastic scattering
A low-energy electron beam yields more of these deflected electrons, which produce cryo-EM images.

Inelastic scattering
Electrons that ricochet off an atom's own electrons can create ions that destroy the proteins.

Building a 3D picture

Ideally, individual proteins freeze in random orientations. The microscope generates 2D images of each orientation. A computer identifies the 2D projections and uses them to calculate the 3D structure.



GRAPHIC: V. ALTOUNIAN/*SCIENCE*

Corrected 24 January 2020. See full text.

the U.K. Competition and Markets Authority found that such a merger would reduce competition and lead to even higher prices. By June 2019, Thermo Fisher had scrapped the deal. "We need competition," Carragher says. "Unless somebody's chasing you, why would you innovate?"

Steve Reyntjens, Thermo Fisher's director of product marketing for cryo-EM, says the company is not resting on its laurels. Scientists' productivity per Krios has risen as the company has improved the microscope's detectors and data collection software, he says. "We have a track record of innovating."

And a Krios is expensive for legitimate reasons, including the gear that accelerates electrons to energies of 300 kiloelectronvolts (keV). Operating at voltages of 300 kilovolts, more than 2500 times stronger than electricity from a U.S. wall socket, the machine requires a bulky transformer and thick, heavily insulated cabling. Massive, costly magnets are needed to focus the high-energy electrons into beams. "Everything needs to be tightly controlled and minutely aligned," Reyntjens says.

The high energies also require expensive safety measures. As the powerful electric fields accelerate the electrons, arcs—little lightning bolts—can form. To prevent arcing, the field emission guns that produce the electrons are suffused in sulfur hexafluoride (SF_6) gas, which snuffs out sparks. But the gas destroys ozone, makes toxic byproducts, and must be recycled. X-rays generated when stray 300-keV electrons hit metal are another hazard, requiring several centimeters of lead shielding around the microscope.

Krios machines are too tall for the average ceiling heights in most countries, and labs have to build special footings to handle the load, between 1000 and 2000 kilograms. López says OHSU spent $20 million to renovate space for its new microscopes.

Yet researchers widely believe the high energies are indispensable. The goal is to send electrons close enough to atomic nuclei in a protein to feel an attractive electrostatic force. Their paths are bent, or scattered elastically—meaning they lose no energy. The pattern of those elastically scattered electrons can be used to build up a picture.

Low-energy electrons are prone to getting too close. They collide either with the nucleus or, more often, with one of the protein's many electrons, scattering inelastically. That causes two problems: The electrons deposit energy into the specimen, eventually destroying it, and they ricochet at lower speeds and odd angles that don't contribute to a picture. At high energies, electrons zip through faster, with less time to "feel" the electrostatic forces, resulting in fewer of the good elastic scattering events.

In a trade-off between gathering information and avoiding specimen damage, 300 keV is seen as cryo-EM's sweet spot. But no one had actually tested that assumption. "Amazing, huh?" Russo muses.

When Henderson, Russo, and a colleague did so, they found a surprise: At 100 keV, they did indeed get more bad events, but also many more good ones—enough to make the trade-off worthwhile. For thin specimens, Henderson says, 100-keV microscopes should actually be better.

To prove the point, the LMB team cobbled together a 100-keV machine from spare parts and, in October 2019, published results in the crystallography journal *IUCrJ* showing they could resolve atoms on five well-known proteins. Such a microscope doesn't need thick lead shielding or an expensive SF_6 system, and the scope can be much smaller and lighter. "It's actually a huge difference," Naismith says.

The team hopes now to entice a company to build a 100-keV machine. At an August 2019 microscopy meeting in Portland, Oregon, Henderson and Russo met privately with Thermo Fisher's CEO and officials from other manufacturers to coax them into making something new and selling it for less. "They listen politely," Henderson says, but so far "we're on our own."

Reyntjens points to Glacios, a 200-keV cryo-EM microscope that Thermo Fisher sells for half as much as a Krios, and says, "Our long-term strategy is to support the democratization of cryo-EM."

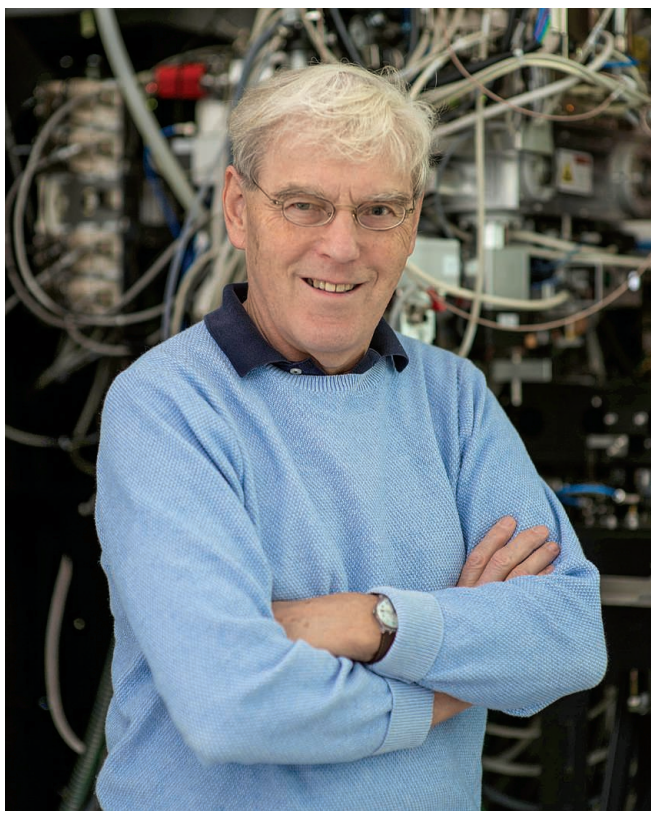
Henderson set about finding smaller companies to take up the 100-keV mantle. He persuaded a U.K. company, York Probe Sources, to build a cheap field emission gun. "I started from a clean slate and designed it for 100 kilovolts," says founder and electronics engineer Mohamed El Gomati.

Henderson also enlisted the help of Swiss company Dectris, which says it can build detectors for less than £150,000. Sacha De Carlo, Dectris's EM business development manager, expects the campaign by Henderson and colleagues will create demand for 100-keV detectors. "He's extremely influential," De Carlo says.

And next month, JEOL will deliver to LMB the first of three cheap 100-keV microscopes—a basic model to which Henderson plans to add the field emission gun and detector.

If he can wow fellow scientists with good results—and show that the price tag is well under $1 million—he's confident that Thermo Fisher will snap to attention. "Either through shame or blackmail or bribery, we'll get them to do it in the end," Henderson jokes.

Henderson already knows where he will put 100-keV machines. At LMB, he has been given space in a room formerly devoted to x-ray crystallography—a metaphor not lost on him. The small cryo-EM machines will fit easily in the room, with no need for renovations. One of the in-house x-ray sources has already been cleared away. A few technicians work on a declining library of crystals while Henderson squints to visualize the tools he's trying to muscle into existence. "X-rays are on their way out, really," he says. Cryo-EM—cheap cryo-EM—is on its way in. ■



Richard Henderson wants to enable more researchers to use cryo–electron microscopy, a technique he helped develop and shared a Nobel Prize for.

PHOTO: MRC LABORATORY OF MOLECULAR BIOLOGY

Corrected 24 January 2020. See full text.

Published by AAAS

POLICY FORUM

AGRICULTURE AND ENVIRONMENT

Overhaul environmental risk assessment for pesticides

Align regulation with environmental reality and policy

By **C. J. Topping[1], A. Aldrich[2], P. Berny[3]**

Among aspects of agricultural intensification that have been criticized for negative impacts on biodiversity (*1, 2*), pesticides have been linked to declines in insects, birds, and biodiversity in aquatic systems (*3–5*). If pesticide use is to blame, even partially, then this raises questions both about pesticide use and the regulatory procedures that are used to protect the environment (*4*). Environmental risk assessment (ERA) of pesticides does not account for many stressors that have intensified in recent years, such as climate change, habitat destruction, and increasing landscape homogeneity, the combination of which can aggravate effects of pesticides in nature. We describe how several assumptions underlying ERA may not hold in modern intensive agricultural landscapes, and the interaction among assumption violations may account for observed declines in biodiversity. Using European contexts to exemplify these global concerns, we review how regulatory ERA for pesticides has fallen out of step with scientific knowledge (*4*) and societal demands for sustainable food production and suggest systematic and recently feasible changes for regulation.

OUTDATED ASSUMPTIONS

The aim of ERA is to determine whether use of a pesticide can be made safe for a receiving environment. The current regulatory procedure evaluates each product separately for each agronomic use—a single-product, single-crop assessment, resulting in authorization or nonauthorization of the product for that use. This approach makes the regulatory scheme tractable, and if the risk is managed appropriately and the product applied accordingly, then pesticides should not pose an unacceptable threat.

Under the European Union (EU) pesticide Regulation No 1107/2009, the ERA uses

[1]Department of Bioscience, Aarhus University, Kalø, Denmark. [2]Department of Ecotoxicology, Agroscope, Wädenswil, Switzerland. [3]Department of Toxicology, Vetagro Sup Campus Vétérinaire, Marcy l'étoile, France. Email: cjt@bios.au.dk

PHOTO: WESTEND61/GETTY IMAGES

Pesticide is applied in a vineyard in Haro, La Rioja, Spain. It is unlikely that an organism in a given field would face a single-product scenario.

guidance documents that were developed according to the science of the early 1990s. A similar situation prevails in the United States, with some guidance being more than 20 years old. Subsequent guidance has attempted to improve flaws, largely by increasing complexity of assessments, maintaining the stepwise (tiered) approach. This relies on the assumption that managing risks through single-product, single-crop assessments provides sufficient ecosystem protection, and where harm is unavoidable, such as insecticide applications, that the ecosystem is sufficiently robust to recover. However, as monitoring studies now show, this is not the case. In our view, the binary ERA result of safe or unsafe does not consider the landscape by balancing environmental harm with ecological recovery because it does not address real spatial and temporal exposure.

Single product, single crop

Risk assessments are based on the use of a single pesticide in a specific crop. Yet, the number of mixtures and sequential treatments with pesticides in the landscape can be very high (*6*) and is the norm across Europe (*7*). Similar situations are described in the United States, Australia, and elsewhere. Any organism spending time in a single field is unlikely to face a single-product scenario, necessitating the consideration of the application sequence in the ERA. The current approach also ignores scale of use; hence, once approved for a crop, the area over which that crop is grown is not considered.

Misrepresented dynamics

Under EU pesticide Regulation No 1107/2009, short-term side effects of pesticides are allowed. Where direct exposure to nontarget organisms makes harm unavoidable (for example, insecticides applied to kill insect pests may kill nontarget insects), the ERA requires recovery experiments to demonstrate recovery potential. For most species, recovery in the contaminated area is by emigration from source habitats. However, this assumes a balanced source-sink dynamic that does not reflect intensive modern agriculture. The experiment, conducted for single fields or plots within an untreated area, does not represent the real ratio of source and sink habitat, something constantly changing because of continued agricultural intensification. The consequence is an underestimation of the risk in the long term (*8*) as source areas become depleted owing to dispersal of organisms into the sinks contaminated with pesticides.

In addition, in a mosaic of pesticide applications, wide-ranging animals (such as honey bees) are exposed to a cocktail of pesticides even more diverse than that applied to a single field. Current ERA also does not consider temporal dynamics. A declining population will be less resilient to future stressors, and thus, a spiral of decline may ensue. ERA ensures that assumptions regarding population health will be incorrect because multiple (regulated and nonregulated) stressors are ignored entirely.

Ecotoxicology

Some shortcomings of the ecotoxicological aspects of ERA are well known. The sensitivity of a few surrogate species is assumed to reflect the sensitivity of all organisms. Often, the choice of surrogate species is based on which species can be reared in the laboratory. Uncertainty in estimating impacts on one species on the basis of impacts on another is addressed with assessment factors (multipliers for the toxicity/exposure ratio, defined according to expert judgment in the uniform principles of Regulation No 1107/2009). For example, a no-effect dose for a surrogate species is assumed protective for all species if the ratio to the estimated exposure is greater than the assessment factor. The value of the assessment factors (typically between 2 and 100) varies with nontarget group and the amount of available data because the factors should address all ERA uncertainties.

Ecological interactions

Trophic interactions are not part of ERA today, even though indirect effects have been recognized as being important (for example, in EU pesticide Regulation No 1107/2009). However, EU and U.S. ERA schemes currently do not consider indirect effects (the partial exception being some aquatic mesocosm studies). This is a major factor when considering whether ERA is actually protective and has potentially profound knock-on effects in ecosystems.

Tiered and deterministic

The current approach to regulatory ERA is based on a tiered and mostly deterministic approach in order to rapidly identify substances of low concern, using assessment factors to cover for all uncertainties in the approach. Pesticides too toxic to pass the first tier (for example, toxicity/exposure ratio below an assessment value of 10 for acute toxicity in birds) enter a refinement approach aimed at reducing the perceived risk for assessed species. The assessment factors decrease with each step of the tiered approach as more data are available, implying that uncertainty in the assessments decreases with each step. However, increased specificity increases the risk of missing effects on other species. For example, the risk to other birds in cereal fields could be refined on the basis of exposure of skylarks feeding on seedlings (as the focal species in the refined assessment) but then miss effects on earlier- or later-breeding species. Hence, risk managers accept that use of pesticides that pass a refined risk assessment may potentially pose a high concern, yet no formal safety loop is in place to take timely action to fix potential oversights. Once pesticides are considered safe, they are placed on the market for 10 years, with restrictions based only on single use and single crop, not on scale of use. Over time, it has become apparent that certain aspects were overlooked or simply unknown, and adverse effects occur at larger spatiotemporal scales.

AN INTEGRATED SYSTEMS APPROACH

The regulatory community is recognizing increasing scientific knowledge and the shortfalls it highlights in the ERA. Recent scientific opinions by the European Food Safety Authority (EFSA) address scientific progress by increasing ERA complexity in a context-specific way within the tiered and deterministic approach. However, continually patching and expanding a conceptually flawed ERA is not seen by many risk assessors as a solution to address the real use of pesticides (*9*). Although the concept of balancing the risks of harm per pesticide per use seems sensible, the actual use of pesticides in the real world renders it scientifically naïve. Calls for pesticidovigilance (*10*) highlight that the "hypothesis" of there being no long-term harm to the environment generated by the application of ERA needs to be evaluated. Whereas monitoring may catch mistakes eventually, an improved prospective ERA would reduce the damage occurring in the first place.

Recently, recommendations from the Scientific Advisory Mechanism (SAM) to the EU call for a critical evaluation of pesticide regulatory assessment (*11*). Similarly, regulators in Germany have called for dramatic change (*9*). The overall picture is of a need to move to a more holistic (systems) view, which integrates far more than current ERA. The first-tier data requirements for toxicological and fate parameters should be maintained for a comparative risk assessment. But the process should then feed into a landscape analysis based on modeling, which would make the tiered, refined risk assessment under current ERA redundant.

> **"The overall picture is of a need to move to a more holistic (systems) view, which integrates far more than current environmental risk assessment."**

The systems approach can be considered at administrative and social-ecological levels. At the administrative level, there is currently a unidirectional flow of information from risk assessor to risk manager to farmer, which is separate from monitoring (if any exists). An integrated-systems approach based on a multidirectional flow of information and involvement of stakeholders could change focus from assessing single products to developing strategies for landscape management. The approach needs to integrate modeling and monitoring to cross-validate assessments and mechanistic understanding. Landscape management would provide an opportunity to tailor mitigation strategies to local contexts and link these to other agricultural management policy instruments (such as subsidy claims). Regular review and changes to authorization would be driven post-market by pesticidovigilance, taking agronomic as well as environmental monitoring into account. This would require focusing on cross-compliance and coherence between different, sometimes conflicting directives affecting the agroecosystem.

To be scientifically relevant and provide the rich spectrum of data for transparent risk management, the systems approach needs to be social-ecological, incorporating agronomic and ecological impacts and better integrating stakeholders in the process. For example, the farmer plays a pivotal role in the system but is currently largely ignored. By contrast, under a new approach, they would become part of the system, informing monitoring by communicating their pesticide use and giving feedback about the agroeconomic impact of the management decisions of pesticide authorization. In turn, the uncertainties in ERA and monitoring outputs should be communicated in a comprehensible way so that the farmer is apprised of the potential risks. Rather than "perfecting" the ERA by addressing all possible toxicological effects and exposure routes (increase in complexity), the new approach needs to be focused on the relevant system aspects.

New ERA tools are available

Achieving systems ERA is a tall order, but new data, approaches, and technology are available. Detailed landscape simulation models exist for a substantial and increasing part of the EU. These are able to represent current farming practices and pesticide use and represent terrestrial environmental

fate at landscape-scale [for example, (*8*)]. Based on Common Agricultural Policy (CAP) subsidy data to provide farm, field, and crop mapping, as well as topographic and satellite data available in the EU, they create a detailed dynamic simulation of farming and pesticide use and fate as well as simulate the distribution and abundance of a range of focal species. Environmental context is directly incorporated, including spatiotemporal interactions. These tools are suggested in EFSA's scientific opinions for landscape scale assessments [for example, (*12*, *13*)] and ongoing EFSA work on multiple stressors and bees and have recently seen use to support an evaluation of the Dutch crop protection policy (*14*).

Mixture effects are starting to be addressed in human RA by using cumulative assessment groups. The same basic principles would work for ERA. Effects of combined toxicity can be elucidated and quantified by using a model deviation ratio (measuring deviations from the assumption of additivity) (*15*). Current EFSA work focuses on combining this approach with quantitative structure-activity relationship (QSAR) modeling to classify pesticides into similar groups and predict their combined toxic effects. This addresses the issues of predicting effects of pesticide mixtures by using a system that is easily expandable and adjustable as more toxicological information becomes available. In order to classify pesticides, we need to base the selection of focal species and their toxicology on our current knowledge and experience of ecologically relevant and sensitive species.

Combining such simulation and mixture approaches would mean that a pesticide would be evaluated in terms of how it would be incorporated into an application schedule. Authorization would relate to scale of use and the application of mitigation measures, which would be incorporated into local landscape conditions. Thus, a balance between harm and recovery could be achieved. In the EU, this could be policed through the existing CAP subsidy scheme. It would be facilitated by a classificatory approach to pesticides that would allow control of use of groups of pesticides rather than individual products. The current toxicity approach would still be used to screen for extreme cases in a comparative assessment. ERA should identify hot spots for monitoring, and monitoring should provide a reality check of the ERA tools. Thus, in response to monitoring, the level of authorized use might be adjusted considering pesticide impacts and changes in the real world (for example, shifts in cropping patterns). This ERA process would provide more information on the local context for the national risk managers but entails tighter controls on post-authorization use.

Goals and structure must change

Policy goals for ERA are changing, yet ERA practice does not reflect this. In the EU, the pesticide Regulation No 1107/2009 is due for refit, but even the current regulation is ahead of current ERA in ecological realism, being targeted at protection of ecosystems. The EU has also launched the Green Deal, including reference to sustainable food production with fewer pesticides. The refit will, however, not make the paradigm shift from a single-product, single-crop ERA to an ecologically and agronomically realistic systems approach. This would require wider changes to authorization and protection goals, and therefore regulations. Protection goals would need to be considered in terms of overall system impacts. Authorization would no longer be a binary condition of safe or unsafe but instead would transparently communicate the accepted risk and regulate use in a sustainable way, similar to efforts to improve how antibiotics are used.

Yet, administrative systems are often very resistant to change. In the case of the EU regulatory system, few ERA guidance documents are currently available, despite numerous scientific opinions and proposed guidelines by EFSA, because no single body has the responsibility to initiate the guidance process. Guidance may take many years to develop once initiated, which becomes more critical because many non-EU countries (such as Brazil) use procedures adapted from the European scheme.

To alleviate these issues, SAM strongly suggests giving risk assessors "complete autonomy to determine all working procedures, methods, data requirements…" (*11*), which should help speed up the working process by removing the requirement that such mandates must be issued by the European Commission (EC). However, change of this magnitude cannot happen overnight. Therefore, we suggest several steps. Lessons learned from previous ERA and monitoring studies should be used to group pesticides and focus the ERA to avoid under- or overestimation of risk. Inclusion of spatiotemporal effects into the single-product, single-crop ERA should be expedited by including landscape-scale and year-on-year effects, as already suggested by EFSA (*12*, *13*). This can be done reasonably quickly under the current regulatory framework if an EC or European Parliament mandate is forthcoming. A multiple stressor approach (landscape modeling and pesticide mixture approaches) should be implemented for nontarget arthropods and bees, those being the most urgently needed, and expanded to further groups subsequently. Development should begin on pesticidovigilance systems linked to EU-wide prospective ERA, incorporating the systems view of impacts of use and restrictions of pesticides. A standing working group of scientists and stakeholders should be created and empowered to implement and maintain the systems ERA, considering new knowledge. Meanwhile, the policy, legal, and advocacy foundations should be prepared to implement the systems approach and pesticidovigilance in future regulations. This should include the EU collation and provision of pesticide use data collected under the Sustainability Directive.

Overall, the risks must be communicated to the public, whose choices will ultimately determine the future of agricultural production and landscapes. Recent studies about the plight of biodiversity have already alerted the public, which is now putting pressure on industry, farmers, and regulators. The time is therefore right. This future view of pesticide ERA and regulation cannot address all concerns nor remove all uncertainties. However, we should not let the perfect be the enemy of good. We can retarget ERA to address better the key environmental questions. This will always be a compromise that needs to be balanced with input from all stakeholders. However, it is feasible, and the need for action is urgent. ■

REFERENCES AND NOTES

1. F. Sánchez-Bayo, K. A. G. Wyckhuys, *Biol. Conserv.* **232**, 8 (2019).
2. C. A. Hallmann *et al.*, *PLOS ONE* **12**, e0185809 (2017).
3. M. A. Beketov, B. J. Kefford, R. B. Schäfer, M. Liess, *Proc. Natl. Acad. Sci. U.S.A.* **110**, 11039 (2013).
4. C. A. Brühl, J. G. Zaller, *Front. Environ. Sci.* **7**, 177 (2019).
5. P. Mineau, M. Whiteside, *PLOS ONE* **8**, e57457 (2013).
6. S. Fryday, H. Thompson, D. Garthwaite, "Background information for considering risk of exposure to multiple pesticides," Department of Environment Food and Rural Affairs (DEFRA) project PS2354 (DEFRA, 2011).
7. D. Garthwaite *et al. EFSA Support. Pub.* **12**, 846E (2015).
8. C. J. Topping *et al.*, *Environ. Toxicol. Chem.* **33**, 1499 (2014).
9. T. Frische, S. Egerer, S. Matezki, C. Pickl, J. Wogram, *Environ. Sci. Eur.* **30**, 8 (2018).
10. A. M. Milner, I. L. Boyd, *Science* **357**, 1232 (2017).
11. Group of Chief Scientific Advisors, "EU Authorisation Processes of Plant Protection Products From a Scientific Point of View," Scientific Advisory Mechanism Report (EU Publications, 2018).
12. EFSA Panel on Plant Protection Products and their Residues (PPR) *et al.*, *EFSA J.* **16**, e05125 (2018).
13. EFSA Panel on PPR, *EFSA J.* **13**, 3996 (2015).
14. A. Tiktak *et al.*, *Geïntegreerde Gewasbescherming nader beschouwd. Tussenevaluatie van de nota Gezonde Groei, Duurzame Oogst* (Planbureau voor de Leefomgeving, 2019).
15. E. Carnesecchi *et al.*, *Environ. Int.* **133** (Part B), 105256 (2019).

ACKNOWLEDGMENTS

We thank S. More and J. H. Williams for comments and R. Luttik for initial discussions. All authors are members of the EFSA Panel on Plant Protection Products and their Residues. All views expressed herein are the authors' alone and do not necessarily represent positions of EFSA or of the authors' employers.

10.1126/science.aay1144

PERSPECTIVES

BIOPHYSICS

Can phase separation buffer cellular noise?

By suppressing concentration fluctuations, condensation may stabilize cellular processes

By **Joshua A. Riback**[1] *and* **Clifford P. Brangwynne**[1,2]

Scientists have long marveled at the ability of complex patterns to emerge from seemingly chaotic origins. In biology, each cell must circumvent the stochastic nature, or noise, inherent to chemical reactions and molecular diffusion. This noise introduces large fluctuations in messenger RNA (mRNA) and protein concentrations, which might be deleterious to processes such as biosynthesis, macroscale organization, and self-replication. Noise may be decreased through compartmentalization, including liquid-liquid phase separation (LLPS), a process underlying the formation of membraneless compartments (*1*, *2*). On page 464 of this issue, Klosin *et al.* (*3*) provide evidence that LLPS can buffer noise in cells and may play a role in stabilizing various biological circuits.

LLPS appears to be a fundamental mechanism driving the assembly of dozens of membraneless organelles, also known as biomolecular condensates (*1*, *2*). Condensates are the location of various processes, including transcription and RNA processing, and are thought to facilitate these processes by concentrating key biomolecular factors (*4*, *5*). Although some data suggest that protein clustering can buffer noise (*6*), evidence for buffering by LLPS has been largely absent.

Compartmentalization of the cellular interior, whether by LLPS or other means, could provide a passive filter—that is, an energy-independent suppression of noise (*7*). For example, segregating factors into different compartments may avoid undesirable interactions that would occur under concentration fluctuations. Additionally, compartmentalization imposes a rate-limiting step, such as nuclear export, which effectively buffers fluctuations (*8*, *9*). The prospect that LLPS can filter noise is particularly attractive because it is grounded in the underlying physics of phase transitions.

LLPS occurs when a system becomes supersaturated with "sticky" molecules, including proteins that contain weakly self-associating, intrinsically disordered regions (IDRs), such that they condense into liquid droplets (thus separating molecules from the surrounding liquid phase). This is represented by the "phase diagram": For a single solute component, its concentration outside of the droplet will be fixed at the saturation concentration (C_{sat}), the maximum quantity of a substance that can dissolve (see the figure). This implies that fluctuations in total concentration primarily change the size (volume fraction) of each of the phases, not the concentrations of the component in each phase. This provides a potential mechanism for passive filtering of fluctuations in the availability of molecules for reactions and signaling.

Klosin *et al.* analyzed a physical framework to understand the non-equilibrium fluctuations imparted by LLPS of a transiently expressed protein, and demonstrated that phase separation lowers the concentration fluctuations in the cytoplasm and nucleoplasm versus those of the total expressed protein. These data provide a strong proof of principle that LLPS can act as a passive filter. Future studies should deduce the importance of such buffering in the robustness of cellular phenomena. This also demonstrates a connection between noise buffering and quantitative determination of phase diagrams—particularly the value of C_{sat}—in living cells.

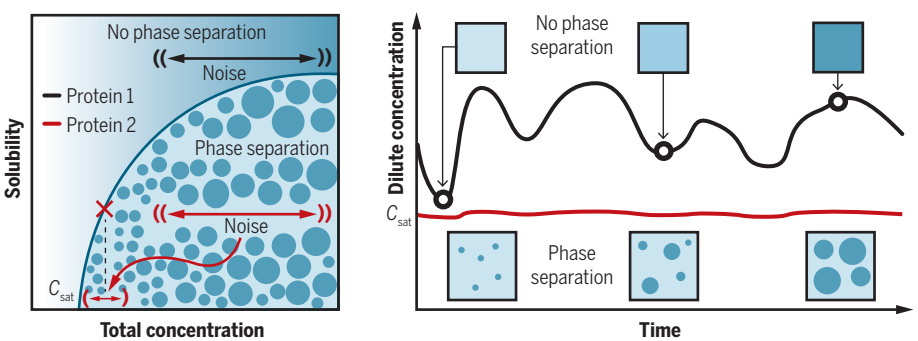
Although many assert that a hallmark of LLPS is a single fixed C_{sat} (*1*, *2*, *10*), quantifying such phase behavior has rarely been achieved for endogenous intracellular condensates. In the few cases where the phase diagrams of model IDR-containing proteins have been mapped in living cells (*11*), the idea of a fixed C_{sat} seems to be valid, implying that such condensates buffer concentration. However, in emerging results for endogenous systems, including for key components of the nucleolus, Cajal bodies, and processing (P) bodies, C_{sat} was found not to be fixed (*12*), consistent with theoretical considerations (*13*).

What is the key distinction between endogenous condensates and model IDR-driven droplets that may complicate the simple "fixed C_{sat}" picture? In endogenous condensates, multiple heterotypic interactions likely occur, often involving self-associating IDRs, as well as proteins with folded domains and nucleic acids. For such multicomponent systems, the phase behavior becomes more complex, and the concentration of a single component only represents part of a high-dimensional phase diagram (*14*).

Given that intracellular LLPS is not governed by fixed concentrations, can endogenous condensates buffer noise? Although Klosin *et al.* do not directly address multicomponent LLPS, they discuss fluctuations

A phase diagram constrains protein fluctuations

"Sticky" molecules (with low solubility) can undergo phase separation, which occurs only above the saturation concentration C_{sat} (left, defined by red X). Fluctuations in the total concentration of sticky proteins (protein 2) may be suppressed by phase separation (right, red). For proteins not undergoing phase separation (protein 1), less noise buffering occurs (right, black).



GRAPHIC: C. BICKEL/*SCIENCE*

[1]Department of Chemical and Biological Engineering, Princeton University, Princeton, NJ, USA. [2]Howard Hughes Medical Institute, Princeton University, Princeton, NJ, USA.
Email: cbrangwy@princeton.edu

in other factors that reduce but do not eliminate the buffering by LLPS. Critically, they provide evidence for noise buffering of an endogenously tagged protein, nucleophosmin (NPM1), which forms condensates in nucleoli. Previous in vitro experiments showed that NPM1 LLPS is stabilized by heterotypic interactions with numerous components, including RNA (*15*), consistent with emerging in vivo findings of a nonfixed C_{sat} (*12*).

A simple argument illustrates how the picture for multicomponent phase separation becomes more complex, even when there is a fixed C_{sat}. Consider the simplest system, which comprises N identical yet independent components. Without phase separation, expression noise manifests through each of these N degrees of freedom. However, upon phase separation, the system loses one degree of freedom, such that the added noise for each component would be $[1 - (1/N)]$ of their expression noise. This implies an interesting balance for multicomponent LLPS, as more components can be buffered, yet with each one being buffered to a lesser extent.

The study of Klosin *et al.* represents an important set of findings that open the door for further studies to delineate potential locations where LLPS may play a role in noise buffering. For example, could feedback through transcriptional condensation (*4*, *5*) be lowering the noise from stochastic mRNA production? Additionally, Cajal bodies and nuclear speckles, condensates relevant for mRNA processing, might have mechanisms to buffer processed mRNA availability. Cytoplasmic bodies—many of which contain mRNAs under various conditions, various stages of development, and in specific tissues—may contribute to cellular robustness by removing expression noise in translation. It is increasingly clear that LLPS must be considered to establish a complete description of noise buffering in living systems. ■

REFERENCES AND NOTES

1. S. F. Banani *et al.*, *Nat. Rev. Mol. Cell Biol.* **18**, 285 (2017).
2. Y. Shin, C. P. Brangwynne, *Science* **357**, eaaf4382 (2017).
3. A. Klosin *et al.*, *Science* **367**, 464 (2020).
4. D. Hnisz *et al.*, *Cell* **169**, 13 (2017).
5. J. Berry *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **112**, E5237 (2015).
6. T. E. Saunders *et al.*, *Dev. Cell* **22**, 558 (2012).
7. T. Stoeger *et al.*, *Cell* **164**, 1151 (2016).
8. K. B. Halpern *et al.*, *Cell Rep.* **13**, 2653 (2015).
9. N. Battich *et al.*, *Cell* **163**, 1596 (2015).
10. D. T. McSwiggen *et al.*, *Genes Dev.* **33**, 1619 (2019).
11. D. Bracha *et al.*, *Cell* **175**, 1467 (2018).
12. J. A. Riback *et al.*, bioRxiv 809210 (2019).
13. J. M. Choi *et al.*, *PLOS Comput. Biol.* **15**, e1007028 (2019).
14. W. M. Jacobs, D. Frenkel, *Biophys. J.* **112**, 683 (2017).
15. D. M. Mitrea *et al.*, *Nat. Commun.* **9**, 842 (2018).

ACKNOWLEDGMENTS

We thank R. Kriwacki, L. Pelkmans, and W. Jacobs for discussions on the manuscript.

10.1126/science.aba0446

AGING

Support cells in the brain promote longevity

Glial cells in the brain use neuropeptides to communicate stress responses and longevity

By **Jason Wayne Miklas**[1] *and* **Anne Brunet**[1,2]

Aging is a multifaceted process that results in organismal decay. At the cellular level, protein homeostasis is a key system that becomes dysregulated with age, causing the accumulation of aberrant or unfolded proteins. In a youthful individual, unfolded proteins normally trigger the unfolded protein response (UPR), which upregulates the protein clearance machinery and returns cells to a homeostatic state. The UPR is typically induced in a cell-autonomous manner. But some cells communicate protein folding stress to distal cells. For example, neurons communicate activation of the UPR to peripheral tissues to promote longevity in the worm *Caenorhabditis elegans* (*1*). On page 436 of this issue, Frakes *et al.* (*2*) show that support cells in the brain called glial cells (*3*) can also initiate long-range activation of the endoplasmic reticulum UPR (UPR^{ER}) in distal cells to coordinate stress resistance and longevity in *C. elegans* and that this occurs through neuropeptide secretion.

"...it will be interesting to determine whether similar neuropeptides are produced... in the human brain...."

A key component of the UPR^{ER} is the conserved transcription factor X-box–binding protein 1 (XBP-1), which coordinates a stress response program. Frakes *et al.* show that overexpressing a constitutively active form of XBP-1, *xbp-1s*, in glia is sufficient to extend life span in *C. elegans*. The authors identify four astrocyte-like cephalic sheath (CEPsh) glial cells as the specific subpopulation of glia that controls UPR^{ER} activation in distal intestinal cells, promoting life-span extension. XBP-1 expression in glia selectively triggers the UPR^{ER} but not other stress responses (such as mitochondrial UPR) in intestinal cells.

How do glial cells communicate with distal intestinal cells? In a previous study, neurons expressing *xbp-1s* induce the UPR^{ER} in a non–cell-autonomous manner by releasing small clear synaptic vesicles containing neurotransmitters that could in turn, directly or indirectly, affect intestinal cells (*1*). Frakes *et al.* show that unlike neurons, glia do not use the machinery involved in the release of small clear synaptic vesicles to regulate signaling with distal intestinal cells. The authors reasoned that the distance a signal from CEPsh glial cells would need to travel to intestinal cells (~300 μm in *C. elegans*) might require long-range-acting neuropeptides, which are secreted from neurons, neuroendocrine cells, and glia. There are 119 neuropeptide precursor genes in *C. elegans*, and their peptide products regulate key physiological processes, including cell-to-cell communication (*4*). Neuropeptides go through a series of processing steps before they are packaged in dense-core vesicles and transported out of the cell. Frakes *et al.* show that disruption of dense-core vesicle export and neuropeptide processing in glial cells suppresses UPR^{ER} activation in intestinal cells. Thus, neuropeptide secretion mediates the effect of glial cells on the periphery (see the figure).

Many interesting questions remain. The specific neuropeptide(s) being secreted are not known, nor are their downstream targets and mode of action. In mammals, several neuropeptides and neurohormones (such as growth hormone–releasing hormone) are secreted by neurons or neuroendocrine cells in the hypothalamo-pituitary axis and exert effects on energy metabolism in peripheral tissues (*5*). Conversely, other peptide hormones are produced by peripheral tissues—such as leptin (adipose), ghrelin (stomach), and insulin (pancreas)—and act in various regions of the brain and other organs (*6*). Some neuropeptides are conserved between *C. elegans* and humans (*7*). In *C. elegans*, the 119 neu-

[1]Department of Genetics, Stanford University, Stanford, CA, USA. [2]Glenn Laboratories for the Biology of Aging, Stanford University, Stanford, CA, USA. Email: anne.brunet@stanford.edu

ropeptide precursor genes can be divided into three categories: 31 FMRFamide (Phe-Met-Arg-Phe-amide)–like peptides and 40 insulin-like peptides, both of which share some homology with neuropeptide precursor genes in mammals, and 48 neuropeptide-like protein genes (*4*). Leveraging single-cell RNA-sequencing data and targeted screens could identify the key neuropeptides expressed by CEPsh glia. Because *C. elegans* CEPsh glia share similarities with mammalian glia (*8*)—notably, astrocytes—it will be interesting to determine whether similar neuropeptides are produced by astrocytes in the human brain and whether they could mediate long-range signaling.

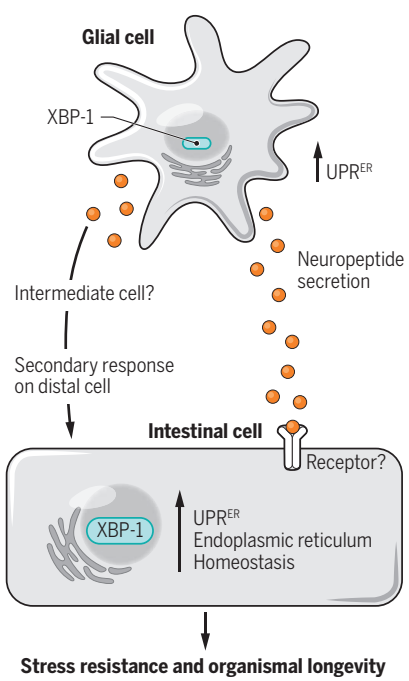
Once a specific neuropeptide or group of neuropeptides are identified, a key step will be to understand the mechanisms by which they influence life span. Neuropeptides could act directly on distal target cells or indirectly by affecting neuronal function. Identifying receptors for glial neuropeptides will be essential to decipher their mechanisms of action in distal cells. In previous studies, *xbp-1s* expression in neurons resulted in lipid metabolism remodeling and lysosomal activity increase in the intestine, mediating life-span extension (*9, 10*). It would be interesting to explore whether the mode of action of neuropeptides secreted by glia reveals unknown longevity pathways in target cells.

Glia have unexpected roles in guiding nervous system development (*3*), in facilitating neurotransmission at synapses (*3*), and in whole-organism functions such as susceptibility to obesity (*11*). Glial cells (astrocytes, oligodendrocytes, Schwann cells, and microglia) can be equal or even greater in number than neurons in mammalian brains (*12*). A key question is whether glia in mammals also express specific neuropeptides that can signal at a distance and whether this is triggered by environmental stimuli. Glia can secrete some peptides, such as neuropeptide Y (*13*), which is involved in feeding behavior, circadian rhythm (body clock), learning, and memory (*14*). Discovering how different types of glial cells interpret their environment and what other cells they modulate will be essential for understanding the response to loss of protein homeostasis and how it evolved.

As Frakes *et al.* and others have shown, changes in protein homeostasis can be communicated in a non–cell-autonomous manner. This may be advantageous to generate a rapid physiological response to remove unwanted proteins and return homeostasis to the organism. Using *C. elegans* as a model to investigate these communication networks at a molecular level will reveal how biological systems communicate with one another and ideally uncover mechanisms of action that are conserved in humans. Understanding how glial cells respond to stress and what neuropeptides they secrete may help identify specific therapeutic interventions to maintain or rebalance these pathways during aging and age-related diseases. ■

Distant longevity signaling

Glial cells can communicate stress resistance to distal cells through neuropeptide secretion. Activating the endoplasmic reticulum unfolded protein response (UPR^{ER}), a key component of which is XBP-1 (X-box–binding protein 1), in glial cells results in a non–cell-autonomous UPR^{ER} activation in distal intestinal cells, which leads to increased stress resistance and longevity in *Caenorhabditis elegans*. Neuropeptide secretion by glial cells is critical for this communication.



REFERENCES AND NOTES

1. R. C. Taylor, A. Dillin, *Cell* **153**, 1435 (2013).
2. A. E. Frakes *et al.*, *Science* **367**, 436 (2020).
3. J. B. Zuchero, B. A. Barres, *Development* **142**, 3805 (2015).
4. L. Frooninckx *et al.*, *Front. Endocrinol.* **3**, 167 (2012).
5. M. Lu *et al.*, *Signal Transduct. Target. Ther.* **4**, 3 (2019).
6. A. N. van den Pol, *Neuron* **76**, 98 (2012).
7. E. Van Sinay *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **114**, E4065 (2017).
8. G. Oikonomou, S. Shaham, *Glia* **59**, 1253 (2011).
9. S. Imanikia *et al.*, *Cell Rep.* **28**, 581 (2019).
10. S. Imanikia *et al.*, *Curr. Biol.* **29**, 2322 (2019).
11. M. Valdearcos *et al.*, *Cell Metab.* **26**, 185 (2017).
12. B. S. Khakh, M. V. Sofroniew, *Nat. Neurosci.* **18**, 942 (2015).
13. I. Prada *et al.*, *J. Cell Biol.* **193**, 537 (2011).
14. M. Botelho, C. Cavadas, *Trends Neurosci.* **38**, 701 (2015).

10.1126/science.aba4474

NEUROSCIENCE

The stillness of sleep

A key neuron in the basal ganglia commands both sleep and immobility

By **William Wisden** *and* **Nicholas P. Franks**

When animals fall asleep, skeletal muscle movement largely ceases. The lack of movement during sleep is an actively controlled process, just like sleep itself. There are specialized sleep-inducing neurons that mostly reside in the brainstem and hypothalamus (*1*). Until now, active repression of movement during sleep was thought to mainly apply to rapid eye movement (REM) sleep, which is when the neocortex exhibits a wake-like activity and dreaming is vivid. Conversely, for the first stage of sleep, non-REM (NREM) sleep, when activity of neurons in the neocortex synchronize at 0.5 to 4 Hz (called delta waves), it was unknown whether movement was actively repressed. On page 440 of this issue, Liu *et al.* (*2*) find that entering NREM sleep and stopping movement are wired together in mice. This is controlled by a brain region called the substantia nigra pars reticulata (SNr), which was thought to control motor actions only when mice are awake.

Liu *et al.* studied an inhibitory neuronal subtype in the SNr of mice, marked by the expression of the gene glutamic acid decarboxylase 2 (*Gad2*), which encodes a protein that synthesizes the inhibitory neurotransmitter molecule γ-aminobutyric acid (GABA). They discovered that these neurons send their axons to areas of the brain that simultaneously induce NREM sleep and inhibit movement (see the figure). For example, to inhibit movement, the *Gad2*$^+$ SNr neurons connect to the motor thalamus and other motor areas of the brain. But to induce sleep, they also inhibit arousal-inducing centers such as the locus ceruleus and dorsal raphe. Because of these connections, a specific circuitry now explains how movement is repressed during NREM sleep, as well as during REM sleep.

The neural circuitry that suppresses

Department of Life Science and UK Dementia Research Institute at Imperial College London, London, UK.
Email: w.wisden@imperial.ac.uk; n.franks@imperial.ac.uk

GRAPHIC: N. DESAI/*SCIENCE*

movement during REM sleep seems entirely different from the circuit that suppresses movement in NREM sleep. During REM sleep, brain-stem circuits actively suppress motor neurons in the spinal cord, which control skeletal muscle contraction. This means that skeletal muscles have no tone (atonia) during REM sleep (*3*). The two different circuits that suppress movement in NREM and REM sleep further emphasize the differences between these types of sleep; indeed, sleep researchers are still puzzled as to why two states of sleep exist.

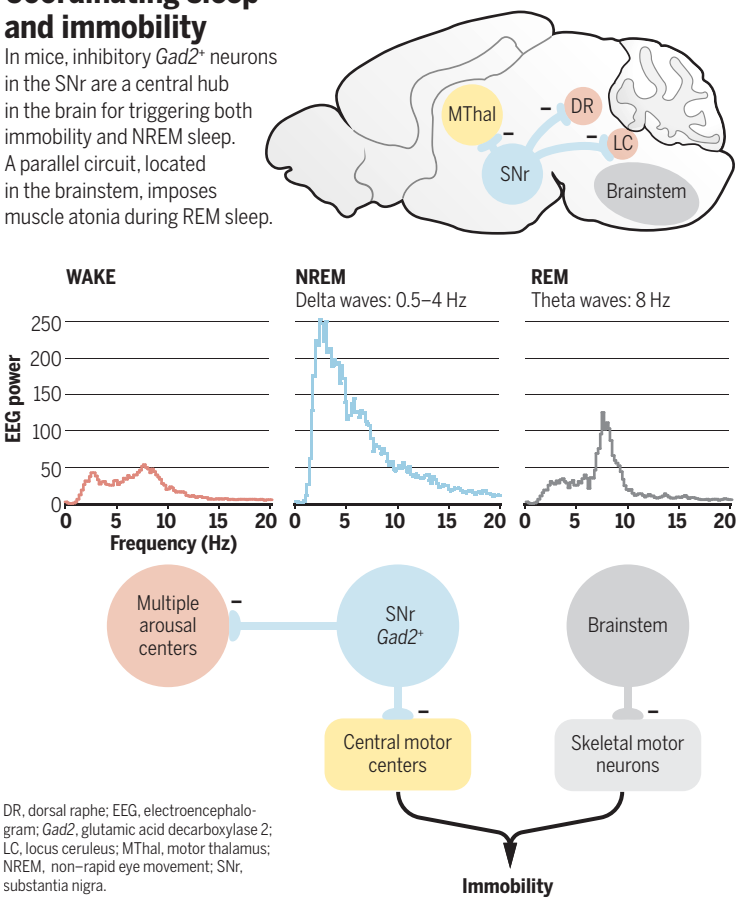
However, falling asleep is not like throwing a switch. Animals are not just awake, and then suddenly asleep (unless suffering from narcolepsy). Indeed, preparing to sleep is a complex and slow behavior. Humans perform specific sleep-preparatory behaviors, such as putting on bed clothes, getting into bed, and adopting particular postures. After feeling drowsy, humans slip into NREM sleep. There seem to be intriguing hints that, at least in mice, specific circuitry controls sleep-preparatory behavior: Nesting behavior in mice prior to sleep is initiated by inhibiting dopamine neurons in the ventral tegmental area (VTA), which is a brain region closely related to the SNr that controls goal-directed behaviors (*4*). Once in the nest, sleep-preparatory postures promote skin warming which, via the preoptic hypothalamus, induces sleep (*5*). Liu *et al.* discovered that the *Gad2*$^+$ SNr neurons might be contributing to the "drowsiness factor" of sleep-preparatory behavior. The authors used machine learning to make nuanced assessments of motor behavior and sleep entry based on electroencephalogram (EEG) readings from mice, which detect brain activity wave patterns. They correlated EEG activity of the *Gad2*$^+$ SNr neurons and found that they function to bias mice to enter NREM sleep by both promoting delta waves (during drowsiness) and inhibiting movement.

The study of Liu *et al.* adds to recently identified long-range neuronal connections that induce NREM sleep from brain areas, collectively known as the basal ganglia, that are classically considered to promote aspects of wakefulness such as motivation, movement, and motor planning. Nearly all of these basal ganglia and closely associated regions also have neurons that strongly promote NREM sleep, including the nucleus accumbens and caudate putamen (*6*), globus pallidus externa (*7*), and now the SNr (*2*), as well as the VTA (*8*). On the basis of these collective findings, the distributed nuclei of the basal ganglia seem to play a major role in regulating NREM sleep (*6*). It seems to be no coincidence that deep brain stimulation of the basal ganglia nuclei to reduce tremor in Parkinson's disease patients often reduces the sleep disturbances these patients suffer and improves their sleep (*7*). The added and distinctive feature of the new findings is that SNr neurons actively innervate and inhibit motor-generating centers as well as promote sleep, whereas presumably other NREM sleep–promoting neurons inhibit movement indirectly.

"It has been suggested that immobility serves the restorative function(s) of sleep."

Coordinating sleep and immobility

In mice, inhibitory *Gad2*$^+$ neurons in the SNr are a central hub in the brain for triggering both immobility and NREM sleep. A parallel circuit, located in the brainstem, imposes muscle atonia during REM sleep.



DR, dorsal raphe; EEG, electroencephalogram; *Gad2*, glutamic acid decarboxylase 2; LC, locus ceruleus; MThal, motor thalamus; NREM, non–rapid eye movement; SNr, substantia nigra.

Muscle activity is not completely suppressed during healthy sleep. Breathing continues, the heart beats, and in REM sleep, eye muscles are highly active, causing the eyes to flicker behind closed lids. But overall, one of the big mysteries of sleep is why it is necessary to go offline and be unconscious (*9*). Why has such a circuit evolved so that sleep and immobility are linked? It has been suggested that immobility serves the restorative function(s) of sleep (*10*). If the animal has to be offline and unconscious to carry out whatever restorative process is being served by sleep, then this may be the reason it needs to stay immobile. For example, it is widely supposed that paralysis during REM sleep stops dreams from being acted out (a process that fails during REM sleep behavioral disorder in humans). But there are also dreams during NREM sleep, so the *Gad2*$^+$ SNr-imposed immobility discovered by Liu *et al.* could serve a parallel function.

What happens to the *Gad2*$^+$ SNr neurons in animals (e.g., dolphins, and some migrating birds) that keep moving while they are partially asleep (*11*)? In these cases it might be expected that a modified circuitry exists, such that the sleep-promoting neurons in the SNr do not inhibit movement. *Gad2*$^+$ SNr neurons could control two desirable features of anesthesia: immobility and unconsciousness (*12*). It would be interesting to know whether stimulating *Gad2*$^+$ SNr neurons would aid entry into anesthesia, and if sedatives could selectively target this type of neuron (*12*). ■

GRAPHIC: V. ALTOUNIAN/*SCIENCE*

REFERENCES AND NOTES

1. T. E. Scammell, E. Arrigoni, J. O. Lipton, *Neuron* **93**, 747 (2017).
2. D. Liu *et al.*, *Science* **367**, 440 (2020).
3. J. Peever, P. M. Fuller, *Curr. Biol.* **27**, R1237 (2017).
4. A. Eban-Rothschild *et al.*, *Nat. Neurosci.* **19**, 1356 (2016).
5. E. C. Harding *et al.*, *Curr. Biol.* **28**, 2263 (2018).
6. Y. Oishi, M. Lazarus, *Neurosci. Res.* **118**, 66 (2017).
7. M. H. Qiu *et al.*, *Neuroscience* **322**, 115 (2016).
8. X. Yu *et al.*, *Nat. Neurosci.* **22**, 106 (2019).
9. M. G. Frank, H. C. Heller, *Handb. Exp. Pharmacol.* **253**, 3 (2019).
10. D. Liu, Y. Dan, *Annu. Rev. Neurosci.* **42**, 27 (2019).
11. N. C. Rattenborg *et al.*, *Nat. Commun.* **7**, 12468 (2016).
12. N. P. Franks, *Nat. Rev. Neurosci.* **9**, 370 (2008).

ACKNOWLEDGMENTS

The authors are supported by the Wellcome Trust and the UK Dementia Research Institute.

10.1126/science.aba4485

RESEARCH ARTICLE SUMMARY

NEUROGENETICS

Chromatin accessibility dynamics in a model of human forebrain development

Alexandro E. Trevino*, Nasa Sinnott-Armstrong*, Jimena Andersen*, Se-Jin Yoon, Nina Huber, Jonathan K. Pritchard, Howard Y. Chang, William J. Greenleaf†, Sergiu P. Pașca†

INTRODUCTION: The cerebral cortex is responsible for higher-order functions in the nervous system and has undergone substantial expansion in size in primates. The development of the forebrain, including the assembly of the expanded human cerebral cortex, is a lengthy process that involves the diversification and expansion of neural progenitors, the generation and positioning of layer-specific glutamatergic neurons, the cellular migration of γ-aminobutyric acid (GABA)–ergic neurons, and the formation and maturation of glial cells. Disruption of these cellular events by either genetic or environmental factors can lead to neurodevelopmental disease, including autism spectrum disorders and intellectual disability.

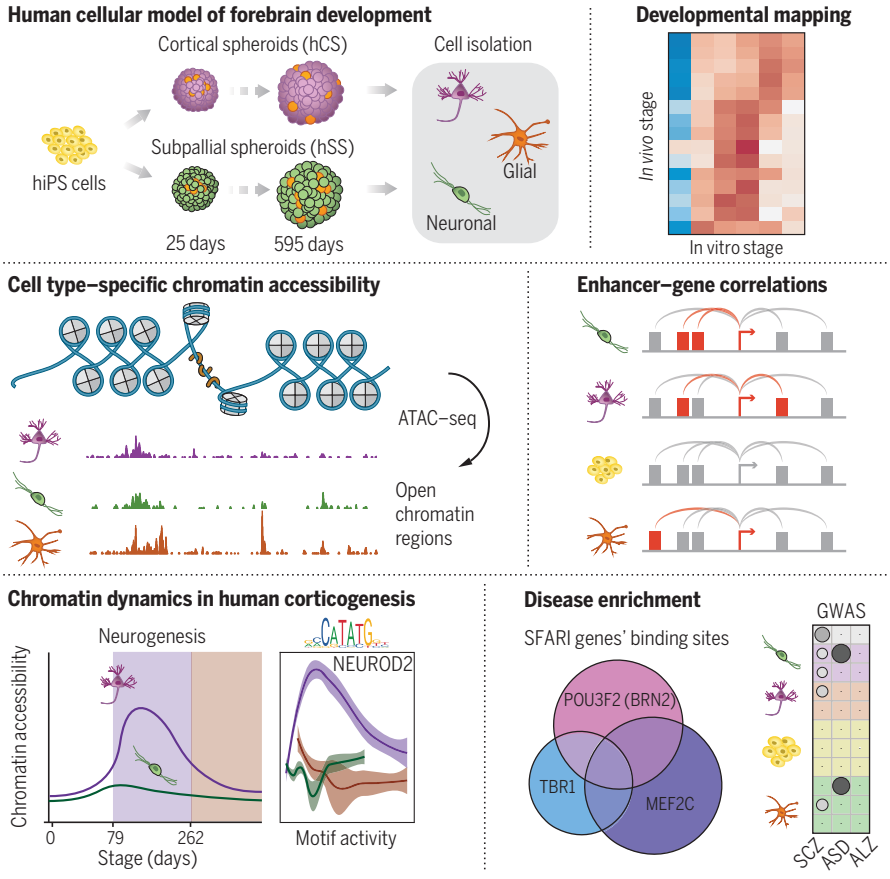
RATIONALE: Human forebrain development is, to a large extent, inaccessible for cellular-level study, direct functional investigation, or manipulation. The lack of availability of primary brain tissue samples—in particular, at later stages—as well as the limitations of conventional in vitro cellular models have precluded a detailed mechanistic understanding of corticogenesis in healthy and disease states. Therefore, tracking epigenetic changes in specific forebrain cell lineages over long time periods, has the potential to unravel the molecular programs that underlie cell specification in the human cerebral cortex and, by temporally mapping disease risk onto these changes, to identify cell types and periods of increased disease susceptibility.

RESULTS: We used three-dimensional (3D) directed differentiation of human pluripotent stem cells into dorsal and ventral forebrain domains and applied the assay for transposase-accessible chromatin with high-throughput sequencing (ATAC-seq) in combination with RNA-sequencing (RNA-seq) to map the epigenetic and gene expression signatures of neuronal and glial cell lineages over 20 months in vitro. We show, through direct comparison with primary brain tissue from our study and several epigenetic datasets, that human stem cell–derived 3D forebrain organoids recapitulated in vivo chromatin accessibility patterns over time. We then integrated these data to discover putative enhancer-gene linkages and lineage-specific transcription factor regulators, including a diverse repertoire of factors that may control cortical specification. We validated protein expression of some of these transcription factors using immunofluorescence, confirming cellular and temporal dynamics in both primary tissue and forebrain organoids. Next, we used this resource to map genes and genetic variants associated with schizophrenia and autism spectrum disorders to distinct accessibility patterns to reveal cell types and periods of susceptibility. Last, we identified a wave of chromatin remodeling during cortical neurogenesis, during which a quarter of regulatory regions are active, and then highlighted transcription factors that may drive these developmental changes.

ON OUR WEBSITE

Read the full article at http://dx.doi.org/10.1126/science.aay1645

CONCLUSION: Using long-term 3D neural differentiation of stem cells as well as primary brain tissue samples, we found that organoids intrinsically undergo chromatin state transitions in vitro that are closely related to human forebrain development in vivo. Leveraging this platform, we identified epigenetic alterations putatively driven by specific transcription factors and discovered a dynamic period of chromatin remodeling during human cortical neurogenesis. Finally, we nominated several key transcription factors that may coordinate over time to drive these changes and mapped cell type–specific disease-associated variation over time and in specific cell types. ■



Developing a human cellular model of forebrain development to study chromatin dynamics. ATAC-seq and RNA-seq studies over long-term differentiation of human pluripotent stem cells into forebrain organoids and in primary brain tissue samples reveal dynamic changes during human corticogenesis, including a wave associated with neurogenesis, and identify disease-susceptible cell types and stages.

The list of author affiliations is available in the full article online.
*These authors contributed equally to this work.
†Corresponding author. Email: wjg@stanford.edu (W.J.G.); spasca@stanford.edu (S.P.P.)
Cite this article as: A. E. Trevino *et al.*, *Science* **367**, eaay1645 (2020).

RESEARCH ARTICLE

NEUROGENETICS

Chromatin accessibility dynamics in a model of human forebrain development

Alexandro E. Trevino[1,2]*, Nasa Sinnott-Armstrong[3]*, Jimena Andersen[2,4]*, Se-Jin Yoon[2,4], Nina Huber[4], Jonathan K. Pritchard[3,5,6], Howard Y. Chang[3,6,7,8], William J. Greenleaf[3,9,10]†, Sergiu P. Pașca[4]†

Forebrain development is characterized by highly synchronized cellular processes, which, if perturbed, can cause disease. To chart the regulatory activity underlying these events, we generated a map of accessible chromatin in human three-dimensional forebrain organoids. To capture corticogenesis, we sampled glial and neuronal lineages from dorsal or ventral forebrain organoids over 20 months in vitro. Active chromatin regions identified in human primary brain tissue were observed in organoids at different developmental stages. We used this resource to map genetic risk for disease and to explore evolutionary conservation. Moreover, we integrated chromatin accessibility with transcriptomics to identify putative enhancer-gene linkages and transcription factors that regulate human corticogenesis. Overall, this platform brings insights into gene-regulatory dynamics at previously inaccessible stages of human forebrain development, including signatures of neuropsychiatric disorders.

The assembly of the human cerebral cortex is a dynamic and lengthy process that begins during early gestation and continues postnatally (*1*). It is characterized by the inside-out generation of neurons in the pallium from progenitors, followed by the generation and maturation of astrocytes and other glial cells. γ-aminobutyric acid (GABA)–ergic interneurons, born in the subpallium, migrate into the cerebral cortex, where they integrate to form circuits (*2*, *3*). Although genetic and environmental perturbations of corticogenesis can cause severe neurodevelopmental disease (*4*), a detailed understanding of the molecular programs that govern these cellular events remains elusive (*1*).

Epigenetic gene regulation plays a crucial role in controlling developmental transitions and cellular differentiation (*5*, *6*). Gene-distal enhancer elements are dynamic throughout development, exhibiting only brief activity in restricted cell populations, yet are enriched for disease-associated genetic variants (*7*, *8*). Moreover, evolutionary divergence of regulatory elements can affect phenotypes in the absence of protein coding changes (*9*). Chromatin accessibility has emerged as an accurate proxy for regulatory potential (*10*). Therefore, defining accessibility across human brain cell lineages and time could provide understanding of gene-regulatory principles and disease signatures in the human forebrain.

Previous studies have begun to define the transcriptome and epigenome of the developing human forebrain, including characterization of spatiotemporal gene expression in the cortex (*11*–*14*), the molecular signature of cortical progenitors (*15*, *16*), epigenetics of early brain development (*17*), and the impact of recently evolved enhancers on gene expression (*18*–*21*). However, brain tissue remains difficult to study, precluding the tracking of epigenetic dynamics in specific cellular lineages over long time periods and, consequently, a mechanistic understanding of forebrain development.

Human induced pluripotent stem cell (hiPS) cell–derived three-dimensional (3D) organoids provide a distinct opportunity to study brain development in vitro and have provided insight into early stages of brain development (*17*, *22*). To capture later stages of development, we developed methods to generate 3D brain region–specific organoids that resemble the dorsal forebrain, called cortical spheroids (hCSs) (*23*, *24*) and the ventral forebrain, called human subpallial spheroids (hSSs) (*25*). Differentiations into regionalized organoids are highly reliable (*26*) and demonstrate a transition from fetal to early postnatal stages at ~280 days in vitro (*24*).

Results

Characterization of accessibility in hCSs and hSSs

We differentiated hCSs and hSSs from hiPS cells for ~20 months in vitro (Fig. 1A). As described (*23*, *24*, *26*, *27*), hCSs recapitulate key features of cortical development in vitro, including the emergence of radial glial progenitors (Fig. 1B and fig. S1A), followed by the progressive generation of deep and superficial glutamatergic neurons (Fig. 1C and fig. S1B), and last, the generation and maturation of astrocytes (Fig. 1D and fig. S1, C and D). Ventral forebrain hSSs generate glutamic acid decarboxylase 67^+ ($GAD67^+$) and $GABA^+$ interneurons (Fig. 1E and fig. S1E) (*25*).

We collected 117 ATAC-seq (assay for transposase-accessible chromatin with high-throughput sequencing) samples and 54 RNA-sequencing (RNA-seq) samples from hCSs and hSSs derived from 7 hiPS cell lines (tables S1 and S2). At early stages, we isolated purified nuclei from whole, undissociated spheroids (*28*). After day 79, we purified glial and neuronal cell lineages by either surface marker immunopanning or fluorescence-activated cell sorting (FACS). For immunopanning, glial lineage cells were isolated by using an antibody against HepaCAM, and neuronal lineage cells were isolated by using an antibody against Thy1 (CD90) (*24*, *29*). For FACS, cells were labeled with viral reporters driven by the human glial fibrillary acidic protein (GFAP) promoter or the synapsin 1 (SYN1) promoter [pLV-GFAP::enhanced green fluorescent protein (eGFP) and adeno-associated virus (AAV)–SYN1::mCherry, respectively] (Fig. 1A) (*30*). We confirmed the specificity of these markers using a single-cell RNA-seq dataset (fig. S2). We refer to $GFAP^+$ or $HepaCAM^+$ cells as glial lineage cells, which may encompass radial glia, outer radial glia, and mature astrocytes, depending on the differentiation stage. We refer to $SYN1^+$ and $Thy1^+$ cells as neuronal lineage cells, which may encompass hCS (glutamatergic) or hSS (GABAergic) neurons. We also performed ATAC-seq in cerebral cortex isolated from human fetal tissue (HFT) at post-conception week (PCW) 20 and PCW21, including samples isolated by immunopanning (fig. S1, F and G).

To evaluate the quality of ATAC-seq libraries generated, we used several metrics, including enrichment at transcription start sites (TSSs) and insert size distributions (table S3) (*30*). hCS and hSS TSS enrichments averaged 25.3-fold. We identified 703,306 reproducible peaks across the entire dataset, representing 244,044 contiguous open chromatin regions (*30*). ATAC-seq technical and biological replicates were highly correlated within peak regions [mean Pearson's correlation coefficient (r) = 0.97 and 0.94, respectively]. To confirm reproducibility, we compared longitudinal ATAC-seq data from hCSs and hSSs derived from two human hiPS cell lines, generated from two individuals, and found an overall mean correlation between lines of $r = 0.94$ (fig. S3A) (*30*). We did not detect changes in apoptosis, necrosis, or unfolded protein response–related gene sets over time [Spearman's rank correlation

[1]Department of Bioengineering, Stanford University, Stanford, CA, USA. [2]Stanford Human Brain Organogenesis Program, Stanford, CA, USA. [3]Department of Genetics, Stanford University, Stanford, CA, USA. [4]Department of Psychiatry and Behavioral Sciences, Stanford University, Stanford, CA, USA. [5]Department of Biology, Stanford University, Stanford, CA, USA. [6]Howard Hughes Medical Institute, Chevy Chase, MD, USA. [7]Department of Dermatology, Stanford University, Stanford, CA, USA. [8]Center for Personal Dynamic Regulomes, Stanford University, Stanford, CA, USA. [9]Department of Applied Physics, Stanford University, Stanford, CA, USA. [10]Chan Zuckerberg Biohub, San Francisco, CA, USA.
*These authors contributed equally to this work.
†Corresponding author. Email: wjg@stanford.edu (W.J.G.); spasca@stanford.edu (S.P.P.)

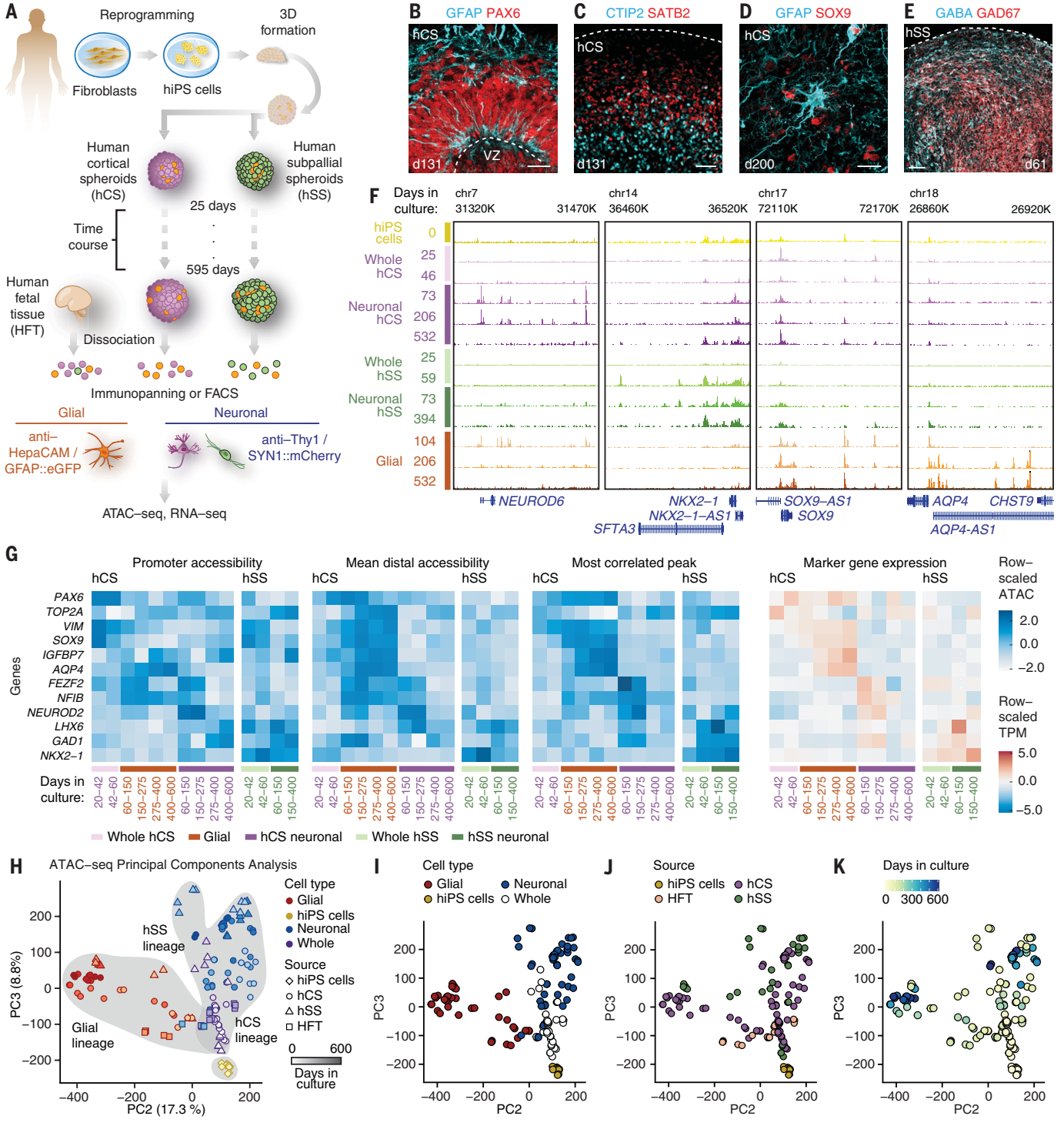


Fig. 1. Forebrain lineage markers and ATAC-seq clustering in hCSs and hSSs. (**A**) Generation of hCSs and hSSs from hiPS cells. Neuronal and glial cells were collected by immunopanning or FACS for ATAC-seq and RNA-seq. Day 20 to 60 spheroids were collected intact. Cells from HFT at PCW20 and PCW21 were collected. (**B** and **C**) Immunohistochemistry in hCSs at day 131 (d131) showing expression of GFAP and PAX6 in a ventricular zone (VZ)–like region, and the layer-specific markers CTIP2 and SATB2. (**D** and **E**) Immunohistochemistry for the astrocyte markers GFAP and SOX9 in hCS day 200 (d200) and GABA and GAD67 in hSS day 61 (d61). (**F**) Genome browser plots of marker gene accessibility for hiPS cells, hCS neurons, hSS neurons, and glial cells. (**G**) Promoter and distal accessibility levels and expression. For the left three panels, heatmap color indicates the scaled accessibility level at the promoter (left), all distal elements averaged (middle), or the distal element with activity most correlated to gene expression (right). The rightmost panel shows RNA expression over time [in transcripts per million (TPM)]. Expression of genes with multiple RefSeq annotations was averaged. (**H**) PCA of all ATAC-seq samples. hiPS cells (yellow), whole hCSs or hSSs (green), glial cells (red), and neuronal cells (blue) are shown. Shapes indicate hCSs (circles), hSSs (triangles), hiPS cells (diamonds), and HFT (squares). Stage is represented by a gradient. hiPS cells represent day 0. (**I**) PCA of all ATAC-seq samples. Whole samples refer to both hCSs and hSSs. (**J**) PCA of all ATAC-seq samples by source. (**K**) PCA of all ATAC-seq samples showing differentiation timing. HFT at PCW20 to PCW21 were labeled as d140. Scale bars, 20 μm (B) and (D), 50 μm (C) and (E).

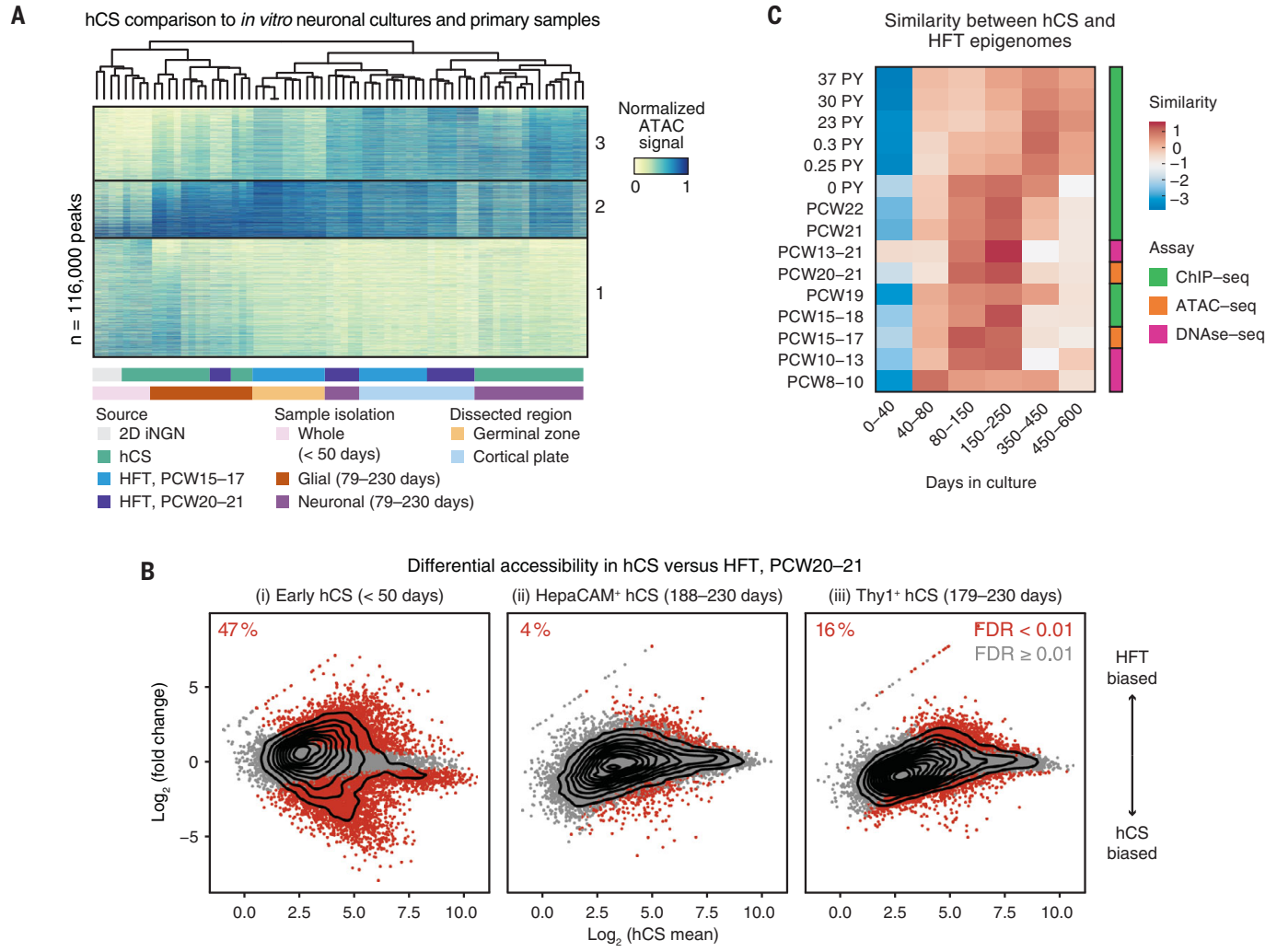


Fig. 2. Comparison of hCS and HFT chromatin accessibility landscapes. (**A**) Comparison of hCSs to HFT at the peak level by hierarchical and *K*-means clustering. Sample set includes iNGN, early hCSs, isolated hCS lineages (79 to 230 days), and PCW20 HFT and microdissected HFT from GZ and CP [from (*20*)]. (**B**) Differentially accessible peaks between early spheroids and glial and neuronal lineages (DESeq2). Differential peaks at FDR < 0.01 are in red; the three panels correspond to three differential tests. Panels compare (i) hCSs (<50 days) to HFT; (ii) HepaCAM$^+$ cells from hCSs (179 to 230 days) and HFT; and (iii) Thy1$^+$ cells from hCSs (179 to 230 days) and HFT. (**C**) Similarity of HFT samples to hCSs over time. Scaled Jaccard indices are plotted. Each row denotes the merged neuronal and glial hCS samples that are most similar to the HFT sample.

coefficient (ρ) = −0.08, 0.05, -0.01 and P = 0.96, 0.57, 0.74, respectively] (fig. S3, B to D, and table S4) (*30*).

ATAC-seq data revealed lineage- and time-specific accessibility differences near marker genes, including the glial marker *SOX9*, the cortical neuron marker *NEUROD6*, the subpallial progenitor marker *NKX2-1*, and the mature astrocyte marker *AQP4* (Fig. 1F). Because enhancer activity and gene expression can be coordinated (*31*), we explored relationships between gene expression and local chromatin accessibility patterns (Fig. 1G and fig. S4, A and B) (*30*). We found that average distal enhancer accessibility, defined as the mean ATAC-seq signal in peaks within 500 kb of the TSS, correlated more strongly with gene expression than promoter-proximal chromatin accessibility (Pearson's r = 0.52, $P = 9.9 \times 10^{-05}$ and r = 0.69, $P = 2.7 \times 10^{-6}$, respectively). We also visualized the best-correlated distal ATAC-seq peak for each gene across cell type and time (Pearson's r = 0.88, $P = 6.2 \times 10^{-15}$), emphasizing the concordance between the expression of well-established lineage markers and associated chromatin accessibility patterns. Conversely, nonforebrain markers were expressed at low levels and did not display lineage-specific patterns, and regulatory elements at these loci were poorly correlated to expression (Pearson's r = 0.23, P = 0.10; r = 0.11, P = 0.56, and r = 0.55, $P = 1.5 \times 10^{-3}$, respectively) (fig. S4, C and D) (*30*).

To examine the global structure of the ATAC-seq data, we performed principal components analysis (PCA) (Fig. 1, H to K). The first three principal components explained 48% of the variance in accessibility (fig. S5A) (*30*). Samples grouped into co-accessible cell populations: hiPS cells, whole hCSs and hSSs at 25 to 59 days in culture, early hCS-derived neurons at 79 to 230 days in culture, hSS-derived neurons, glial progenitors, and mature glia. Generally, we saw that hSS-derived neurons grouped with late-stage hCS-derived neurons (after 230 days in culture). Overall, PCA captured the differentiation of hiPS cells into neuronal and glial lineages (Figs. S5, B to D, and S6, A to E) (*30*).

Direct comparison of hCS lineages to primary human tissue

To examine the fidelity of our in vitro chromatin landscapes to those in vivo, we performed *K*-means clustering on ATAC-seq data

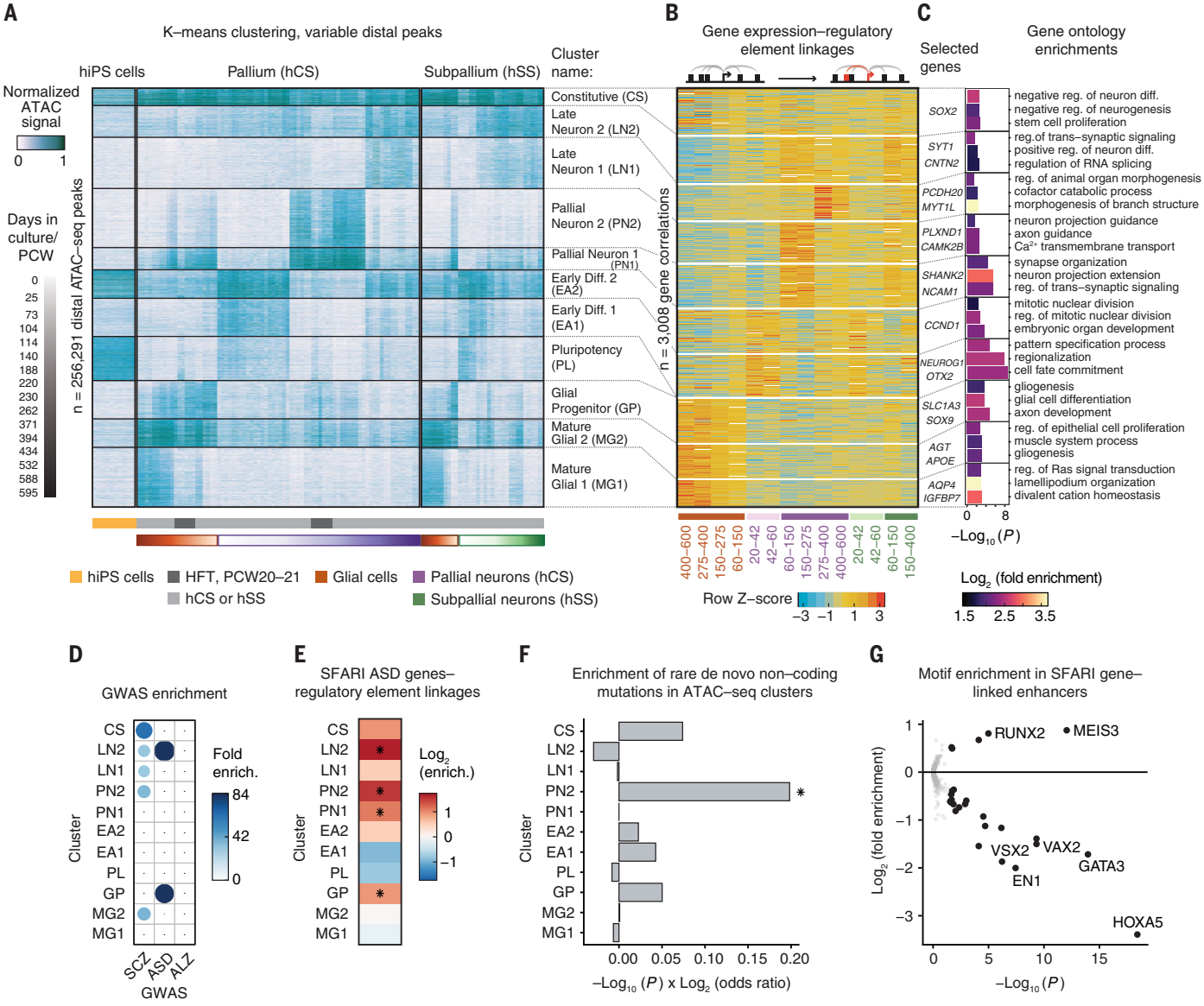


Fig. 3. Gene-regulatory dynamics in hCSs, hSSs, and HFT. (**A**) *K*-means clustering of variable, distal accessible peaks. Activity is represented as normalized ATAC signal without additional row-scaling. Samples were sorted by stage and cell lineage. HFT samples were included in the *K*-means algorithm as d140. Stage or days in culture is represented by a gradient. (**B**) Heatmap showing expression of genes correlated to accessible elements from each *K*-means cluster within 500 kb of the TSS. RNA-seq clusters correspond to ATAC-seq clusters. Cluster PL was omitted because of lack of RNA-seq. (Top) Schematic of the correlation approach, displayed as row-standardized TPM. (**C**) Representative genes from each *K*-means cluster along with the GO enrichment within each cluster. *P* values derive from the hypergeometric test, and the color indicates the fold enrichment. (**D**) Enrichment of GWAS signal for schizophrenia, ASD, and Alzheimer's disease in clusters, estimated using LD score regression. The size and color indicate fold enrichment; data points without significance are shown with a black dot. (**E**) Enrichment of SFARI ASD genes in the linked gene expression clusters. Colors correspond to enrichment in the cluster, and stars indicate significance ($P < 0.05$). (**F**) Enrichment of de novo noncoding ASD mutations. Enrichment scores [$-\log_{10}(P)$ * $\log_2$(enrichment)]. (**G**) Motif enrichment in SFARI genes relative to enhancers of other linked genes displayed as a volcano plot [$\log_2$(enrichment) versus $-\log_{10}(P)$].

from early hCSs (25 to 46 days in vitro), 2D neurons derived by overexpression of NGN2 (iNGN; 27 to 41 days in vitro), hCS-derived glial and neuronal lineages (79 to 230 days in vitro), PCW20 HFT, and microdissected germinal zone (GZ) and cortical plate (CP) at PCW15 to PCW17 (*20*). As a feature selection step aimed at reducing the number of potentially noisy peaks with weak chromatin dynamics, we used the top 20% of peaks ranked by standard deviation across this sample subset. We found three patterns of chromatin accessibility, which is consistent with hierarchical clustering of the samples ($K = 3$, $n = 116{,}000$ peaks) (Fig. 2A) (*30*). Accessible elements in the first cluster were most active in iNGN and hCSs at 25 to 46 days in vitro, whereas elements in the second cluster had the highest activity in hCSs and HFT glial lineage cells and in GZ, which is consistent with glial progenitors representing a larger proportion of the GZ. Elements in the third cluster demonstrated specific activity in neuronal lineage cells isolated from hCSs and HFT, and CP, but not in early hCSs or iNGN. Next, we computed differential accessibility between isolated hCS and HFT lineages directly. We found that 47% of regions were significantly differentially accessible between whole HFT samples and whole hCSs at 25 to 59 days [DEseq2 false

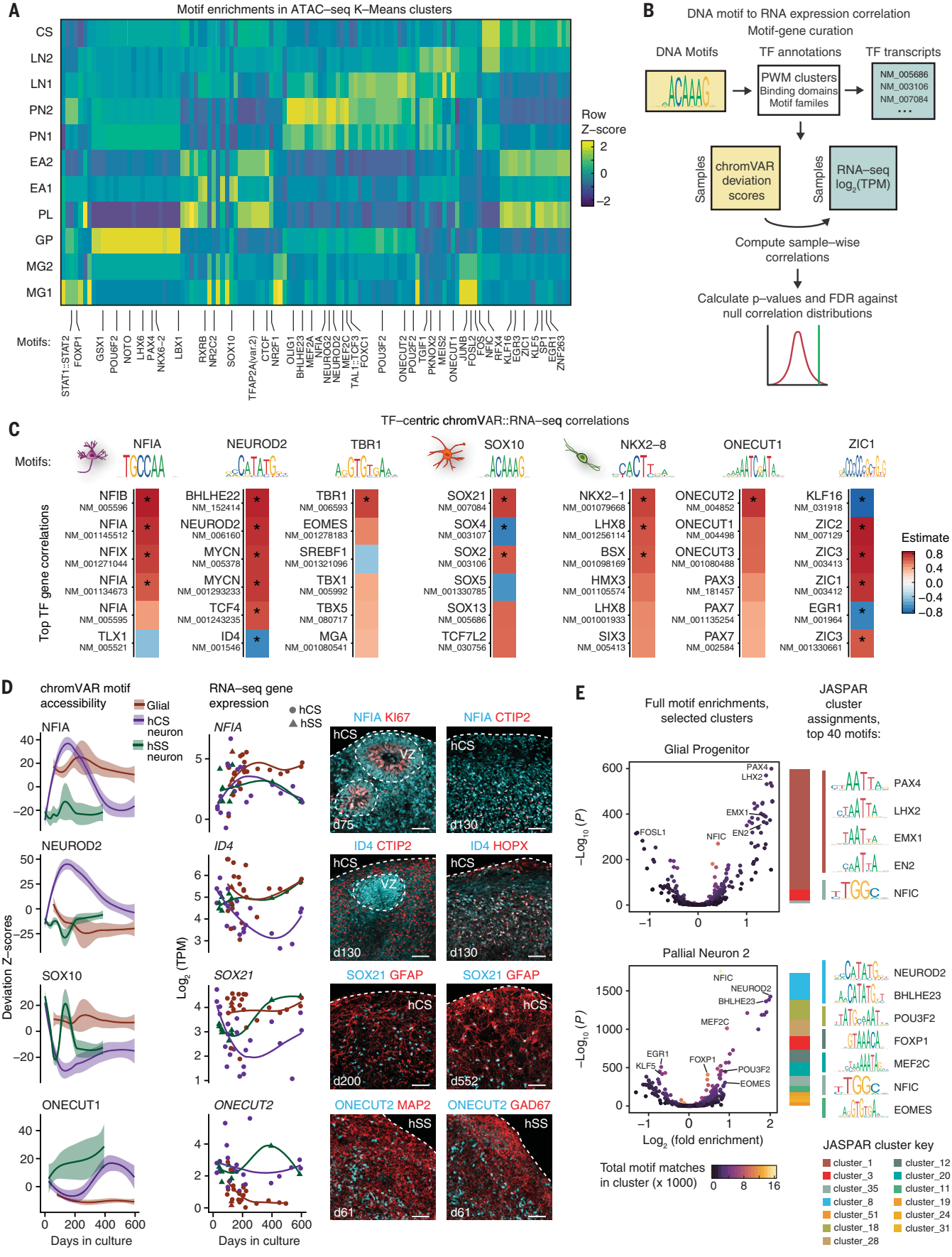

A
Motif enrichments in ATAC–seq K–Means clusters
CS
LN2
LN1
PN2
PN1
EA2
EA1
PL
GP
MG2
MG1
Row Z–score
2
0
–2
Motifs:
STAT1::STAT2
FOXP1
GSX1
POU6F2
NOTO
LHX6
PAX4
NKX6–2
LBX1
RXRB
NR2C2
SOX10
TFAP2A(var.2)
CTCF
NR2F1
OLIG1
BHLHE23
MEF2A
NFIA
NEUROD2
NEUROD2
MEF2C
TAL1::TCF3
FOXC1
POU3F2
ONECUT2
POU2F2
TGIF1
PKNOX2
MEIS2
ONECUT1
JUNB
FOSL2
FOS
NFIC
RFX4
KLF16
EGR3
ZIC1
KLF5
SP1
EGR1
ZNF263
B
DNA motif to RNA expression correlation
Motif-gene curation
DNA Motifs
TF annotations
TF transcripts
PWM clusters
Binding domains
Motif familes
NM_005686
NM_003106
NM_007084
...
Samples
chromVAR deviation scores
Samples
RNA–seq log2(TPM)
Compute sample–wise correlations
Calculate p–values and FDR against null correlation distributions
C
TF–centric chromVAR::RNA–seq correlations
Motifs:
NFIA
NEUROD2
TBR1
SOX10
NKX2–8
ONECUT1
ZIC1
Top TF gene correlations
NFIB NM_005596
NFIA NM_001145512
NFIX NM_001271044
NFIA NM_001134673
NFIA NM_005595
TLX1 NM_005521
BHLHE22 NM_152414
NEUROD2 NM_006160
MYCN NM_005378
MYCN NM_001293233
TCF4 NM_001243235
ID4 NM_001546
TBR1 NM_006593
EOMES NM_001278183
SREBF1 NM_001321096
TBX1 NM_005992
TBX5 NM_080717
MGA NM_001080541
SOX21 NM_007084
SOX4 NM_003107
SOX2 NM_003106
SOX5 NM_001330785
SOX13 NM_005686
TCF7L2 NM_030756
NKX2–1 NM_001079668
LHX8 NM_001256114
BSX NM_001098169
HMX3 NM_001105574
LHX8 NM_001001933
SIX3 NM_005413
ONECUT2 NM_004852
ONECUT1 NM_004498
ONECUT3 NM_001080488
PAX3 NM_181457
PAX7 NM_001135254
PAX7 NM_002584
KLF16 NM_031918
ZIC2 NM_007129
ZIC3 NM_003413
ZIC1 NM_003412
EGR1 NM_001964
ZIC3 NM_001330661
Estimate
0.8
0.4
0.0
–0.4
–0.8
D
chromVAR motif accessibility
Glial
hCS neuron
hSS neuron
Deviation Z–scores
NFIA
NEUROD2
SOX10
ONECUT1
Days in culture
RNA–seq gene expression
hCS
hSS
Log2 (TPM)
NFIA
ID4
SOX21
ONECUT2
Days in culture
NFIA KI67
NFIA CTIP2
ID4 CTIP2
ID4 HOPX
SOX21 GFAP
SOX21 GFAP
ONECUT2 MAP2
ONECUT2 GAD67
hCS
hSS
VZ
d75
d130
d200
d552
d61
E
Full motif enrichments, selected clusters
JASPAR cluster assignments, top 40 motifs:
Glial Progenitor
–Log10 (P)
PAX4
LHX2
EMX1
EN2
NFIC
FOSL1
Pallial Neuron 2
NFIC
NEUROD2
BHLHE23
MEF2C
EGR1
KLF5
FOXP1
POU3F2
EOMES
Log2 (fold enrichment)
Total motif matches in cluster (x 1000)
0
8
16
PAX4
LHX2
EMX1
EN2
NFIC
NEUROD2
BHLHE23
POU3F2
FOXP1
MEF2C
NFIC
EOMES
JASPAR cluster key
cluster_1
cluster_3
cluster_35
cluster_8
cluster_51
cluster_18
cluster_28
cluster_12
cluster_20
cluster_11
cluster_19
cluster_24
cluster_31

Fig. 4. Transcription factor activity in the forebrain. (**A**) Heatmap summarizing motif enrichments in ATAC-seq *K*-means clusters. The top 15 most-enriched motifs per cluster are displayed, and color indicates scaled $\log_2$(fold enrichment) versus other clusters. Select motifs are shown. (**B**) Schematic of chromVAR–RNA-seq expression correlation approach. Motifs are linked to TF genes that share a position weight matrix (PWM) cluster, family annotation, or binding domain by using available databases. For each motif, the correlation of chromVAR motif deviations to the expression values of each eligible TF gene is compared against a background set of correlations. A *P* value and FDR are computed against the null to determine significance. (**C**) ChromVAR-expression correlations for a subset of enriched motifs. Color indicates Pearson correlations for a TF motif–TF gene pair. The top six correlations are displayed. Stars indicate FDR-adjusted $P < 0.05$. Cell icons indicate the lineages with the greatest enrichments for a given motif. (**D**) ChromVAR motif accessibility (deviation *Z* scores), RNA expression ($\log_2$ TPM), and immunohistochemistry for candidate lineage-specific TFs identified by means of paired ATAC-seq and RNA-seq. For chromVAR, values are indicated with a smoothed line. RNA-seq values are shown by sample. Immunostainings show NFIA expression with KI67 and CTIP2 in hCSs (days 75 and 130); ID4 with HOPX and CTIP2 in hCS (day 130); SOX21 with GFAP in hCS (days 200 and 552); and ONECUT2 with MAP2 and GAD67 in hSS (day 61). (**E**) Motif volcano plots ($\log_2$ fold enrichment of motif in ATAC-seq cluster versus other clusters by $-\log_{10} P$ values) for GP and PN2. The color of each circle indicates the number of sequence matches in the foreground set for each motif. Cluster names are derived from the JASPAR database. The size of each colored box indicates the TF number from a given family that are enriched in the cluster. Scale bars, 50 μm (D).

discovery rate (FDR) < 0.01]. However, at later stages, only 4% of regions between HFT- and hCS-derived glia and 16% of regions between HFT- and hCS-derived neurons were differentially accessible (Fig. 2B) (*30*).

To explore the relationship of cortical development to our in vitro cultures over time, we integrated epigenetic datasets from PsychENCODE, ENCODE, and a recent study, representing human cortical tissue samples from early fetal to postnatal stages (*12*, *17*, *20*, *32*). We computed the Jaccard similarity index (*J*) between peaks from each primary tissue sample and hCSs at different stages of in vitro differentiation. To account for differences between studies, we standardized the scores for each primary sample (Fig. 2C). We found that HFT at PCW8 to PCW10 was most similar to hCSs at 40 to 80 days. Mid– to late–fetal developmental stages up to birth [PCW10 to 0 postnatal years (PY)] resembled cultures at 80 to 250 days, whereas postnatal samples were most similar to hCSs after 350 days. We observed similar trends when displaying unscaled values of *J* (fig. S7, A to C) (*30*). The mapping of early hCS samples (<50 days in vitro) to primary tissue samples was weaker likely because they represent earlier stages of development for which no in vivo data are currently available. Nonetheless, we chose to keep these samples in downstream analyses because they may inform future in vitro models of development and disease. Although changes in tissue cell composition and limitations of our in vitro model make comparisons to later stages challenging, these data suggest that forebrain organoids undergo progressive chromatin remodeling commensurate with that observed in primary tissue. Despite limited samples for comparison, these trends were consistent across multiple epigenomic data types and studies and across long time scales, from early fetal to postnatal development.

Gene-regulatory landscapes in forebrain differentiation

To explore patterns of chromatin accessibility in cell lineages isolated from HFT and 3D forebrain cultures, we performed *K*-means clustering on the top 40% of peaks ranked by standard deviation (hereafter, dynamic peaks; $n = 256{,}291$) (*30*). Inspection of these clusters revealed that they corresponded to six groups: (i) peaks with maximal accessibility in hiPS cells ["pluripotency" (PL)]; (ii) peaks specific to hCSs or hSSs between 25 and 59 days of differentiation ["early 1 and 2" (EA1 and EA2)]; (iii) glial lineage peaks from hCS, hSS, and HFT samples, which were subdivided into early and late-stages ["glial progenitor" (GP) and "mature glial 1 and 2" (MG1 and MG2)]; (iv) peaks specific to cortical neurogenesis between days 79 and 230, including HFT ["pallial neuron 1 and 2" (PN1 and PN2)]; (v) peaks specific to neurons from hCS after 230 days of differentiation and hSS ["late neuron 1 and 2" (LN1 and LN2)]; and (vi) peaks that were highly active across samples ["constitutive" (CS)] (Fig. 3A and table S1). We found that although hSS neurons were distinct from hCS neurons using a supervised approach (fig. S8A) (*30*), they did not form a separate cluster. Differential hSS-specific peaks comprised 41, 46, and 34% of the LN1, LN2, and EA2 clusters, respectively (fig. S8B) (*30*).

To verify the overlap of *K*-means clusters with epigenomes from various primary brain tissues and in vitro models, we computed the enrichment of reference datasets within the clusters (fig. S9, A to C) (*30*). Across studies and assays [including chromatin immunoprecipitation–sequencing (ChIP-seq), deoxyribonuclease-sequencing (DNase-seq), and ATAC-seq], we found that in vitro models were enriched in the early stage clusters EA1 and EA2, as well as the LN2 and CS clusters. Mid-fetal HFT (PCW14 to PCW22) datasets were strongly enriched in clusters PN1 and PN2, whereas glial maturation was well captured by a transition of enrichment from the GP to MG2 to MG1 clusters. Last, late fetal and early postnatal samples were enriched in the late neuronal LN2 cluster. Adult samples were not enriched in LN2, which could reflect time- or activity-dependent maturation not currently captured by our in vitro model. Using this collection of primary brain tissue– and cell culture–derived epigenomes, we computed the fraction of peaks from each dataset overlapping our clusters, revealing that enrichments were driven by large fractions of overlaps (fig. S9D) (*30*).

We next aimed to define gene expression programs associated with this developmental accessibility landscape. We adapted an approach to discover putative enhancer-gene linkages on the basis of the coordination of accessibility and gene expression across cell lineages (*33*). For each gene, we then computed correlations across samples between gene expression and accessible peaks within a 500-kb window, taking into account the distribution of spurious correlations. This method identified 28,940 links corresponding to 8294 genes. For each cluster, we selected the top 400 strongest enhancer-gene correlations in the cluster for further analysis (omitting PL) (Fig. 3B). These groups included established lineage-specific markers, such as *AQP4* and *IGFBP7* in MG1 and *NFIA* and *SHANK2* in PN1. Each group was enriched for specific gene ontology (GO) terms, including cell fate commitment (EA1), synaptic signaling (PN1), neuron projection and axon guidance (PN2), and cation homeostasis and lamellipodium organization (MG1) (Fig. 3C and table S5).

Disease signatures in forebrain chromatin landscapes

Noncoding regions contain the majority of common variants that influence human disease (*7*), and we reasoned that genetic risk for brain disorders might localize to chromatin accessibility in specific cell lineages. We mapped single-nucleotide polymorphisms, shown to confer disease risk by genome-wide association studies (GWASs), onto our accessibility landscape. Using linkage disequilibrium (LD) score regression (*34*, *35*), we measured enrichment of risk variants within each cluster (Fig. 3D). Variants associated with Alzheimer's disease, major depressive disorder, and epilepsy were not enriched in any clusters. Schizophrenia risk (*36*) was enriched across the MG2, PN2, LN1, LN2, and CS clusters, with PN2 cells exhibiting the most significant enrichment,

and considerable signal in the late neuronal clusters, which encompass hSS interneurons and late-stage hCS neurons (PN2 heritability enrichment = 37.7, block jackknife χ^2 test, $P = 4.5 \times 10^{-6}$; LN1 heritability enrichment = 30.8, block jackknife χ^2 test, $P = 5.2 \times 10^{-5}$). Meanwhile, the GP and LN2 clusters were enriched for autism spectrum disorder (ASD) risk (heritability enrichments = 84.7- and 99.3-fold, respectively; block jackknife χ^2 test, $P = 2.4\times 10^{-3}$ and 7.1×10^{-3}). Despite a paucity of genome-wide significant ASD-associated variants to date, we found that GP was consistently and strongly enriched for risk across multiple GWASs. hSS peaks were enriched for genetic risk for schizophrenia, autism, and attention deficit hyperactivity disorder (fig. S10A) (*30*).

The Simons Foundation Autism Research Initiative (SFARI) curates a database of ASD-associated genes (http://gene.sfari.org). Although recent studies have examined expression patterns of these genes during fetal development or postnatally (*12*, *37*), relatively little is known about the specific lineages that may be involved in ASD at different stages. Across our model, we found that 81% of all SFARI genes exhibited variable expression in our transcriptomic data ($n = 851$ out of 1054 genes, Fisher's exact test for enrichment $P = 4.19 \times 10^{-4}$, odds ratio = 1.18), and we identified significant enhancer-gene linkages for 54% of them ($n = 570$, $P = 3.51 \times 10^{-5}$, odds ratio = 1.25) (fig. S10B) (*30*). Consistent with our analysis of GWAS data, SFARI genes were significantly enriched for enhancer linkages in clusters GP and LN2 (hypergeometric test, $P = 2.24 \times 10^{-2}$ and 4.36×10^{-5}). In addition, we uncovered an enrichment in pallial neurons (PN1 and PN2 hypergeometric test, $P = 8.82 \times 10^{-3}$ and 4.36×10^{-4}) (Fig. 3E). The SFARI database includes genes with evidence from not only common variants but also de novo and rare syndromic and nonsyndromic mutations. This association with pallial neurons is further supported by the enrichment of recently identified de novo noncoding ASD mutations (*38*) within the PN2 cluster ($P = 0.027$) (Fig. 3F and fig. S10C). Last, we found that SFARI gene-linked enhancers were significantly enriched for MEIS3 homeobox transcription factor (TF) motifs, both relative to other gene linkages (Fig. 3G) and to unlinked enhancers at SFARI loci (fig. S10D).

Evolutionary conservation of epigenetic information in human forebrain

We investigated how evolutionary conservation might vary across our data, and we found that PhyloP conservation scores were not evenly distributed in our clusters, with PL and MG1 having the lowest PhyloP scores (indicating less conservation) and CS having the highest (fig. S11A) (*30*). Low conservation in MG1 could reflect morphological and phenotypic differences that have been observed between astrocytes in humans compared with other vertebrates (*39*). Peaks with an enhancer-gene linkage in the GP, PN1, and LN1 clusters demonstrated significantly reduced conservation scores relative to peaks without linkage, potentially reflecting a degree of evolutionary adaptation at these forebrain-specific loci contributing to gene expression (Mann-Whitney-Wilcoxon test with Bonferroni correction, $P_{adj} = 1.9 \times 10^{-2}$, 8.5×10^{-7}, and 1.6×10^{-3}, respectively). Although these differences suggest that a degree of selection may be acting on particular brain functions, they were quantitatively small. Therefore, we looked for evidence of process-specific selection, which might support the functional relevance of these observations. To do so, we defined gene-linked peaks in the top and bottom 5% of conservation scores as highly conserved and poorly conserved sets, respectively. We detected enrichment for forebrain differentiation, brain development, and transcriptional regulation GO terms among genes with highly conserved peak links relative to a background of genes with poorly conserved links (fig. S11B) (*30*). Last, we filtered genes to those with multiple highly conserved or poorly conserved linked enhancers and measured the distributions of these enhancers across our *K*-means clusters, relative to a permuted background (*30*). Poorly conserved enhancers were strongly enriched in cluster PL, and well-conserved enhancers were enriched in clusters PN1 and PN2 (fig. S11C).

Conservation can also be measured experimentally in vivo with enhancer-driven reporters. We quantified the overlap of our data with the VISTA mouse enhancer dataset (*40*) and observed that the VISTA enhancers that were accessible in our dynamic peaks were enriched for forebrain activity (odds ratio = 1.37, $P = 7.83 \times 10^{-6}$) (fig. S11D). Clusters MG2, GP, EA2, PN1, and CS were significantly enriched for forebrain-only activity (Benjamini-Hochberg adjusted $P = 7.13 \times 10^{-5}$, 6.20×10^{-4}, 4.78×10^{-2}, 1.64×10^{-3}, and 1.76×10^{-9}) (fig. S11E). Selected VISTA enhancers from PN1 and LN1—which include hSS enhancers—displayed activity in dorsal and ventral forebrain, respectively (fig. S11F) (*30*).

Last, we were interested in exploring how sequences that may have recently evolved in the human lineage are distributed during corticogenesis. To do this, we considered the overlap between our forebrain chromatin accessibility dataset and human accelerated regions (hARs), which are noncoding genomic loci, previously implicated in brain development, that exhibit high vertebrate conservation and increased substitution rates in humans compared with other great apes (*41*, *42*). We observed that 35% of all hARs ($n = 965$ out of 2734) overlapped with our ATAC-seq data, corresponding to a rate of 3.8 hARs per thousand dynamic peaks (fig. S11G and table S6) (*30*). These overlaps were broadly distributed across clusters, with the lowest enrichments in PL, and the highest in LN2, corresponding to late stages of differentiation (fig. S11H). Subpallial neuron peaks also exhibited overlap with hARs (4.05 to 4.34 hARs per megabase), including loci we linked to *GAD1*, *DLX2*, and the axon guidance ephrin receptor *EPHA5*. Recent work has provided evidence that a number of these loci may have functional consequences for developmental timing in humans (*19*). Together, our analyses uncovered lineage-specific patterns of functional conservation and acceleration across our chromatin accessibility landscape.

Transcription factor motifs in human forebrain development

We next explored the sequence features that underlie ATAC-seq clusters. We calculated the enrichment of DNA sequence motifs in each cluster versus all the other clusters (Fig. 4A and table S7) (*30*). We found that ZIC1, SP1, KLF5, and CTCF motifs were enriched in clusters PL and EA2, which is consistent with DNA binding domains shown to regulate pluripotency (*43*). Cluster GP was enriched for SOX10 and NFIC motifs, as well as a group of AT-rich sequence motifs bound by homeobox factors such as LHX and PAX, which play key roles in early cortical specification (*44*). The MG1 and MG2 clusters exhibited strong enrichment for STAT1::STAT2, SOX10, and the FOS family of motifs, which may play roles in astrocyte differentiation and maturation (*45*).

Sequence-based TF motif analysis does not provide a one-to-one correspondence between binding motifs and particular TF proteins (*46*, *47*). To nominate specific TFs that may drive chromatin changes in an unbiased way, we used chromVAR to quantify the accessibility motifs across all samples and dynamic peaks (*48*) then computed sample-wise Pearson correlations between chromVAR motif deviations and log-transformed gene expression values of TFs from RNA-seq (Fig. 4B and table S8) (*30*). We highlighted TF genes with expression that were significantly correlated or anticorrelated to dynamic cluster-specific motif deviations (Fig. 4C). For each motif, we then plotted the chromVAR deviations alongside the RNA-seq profiles of each putative TF. Last, we verified the expression and localization of some of these factors in hCSs and HFT using immunohistochemistry (Fig. 4D).

In the glial lineage, we found that nuclear factor I A (NFIA) motif accessibility was strongly correlated with the expression of the *NFIA* gene (Pearson's $r = 0.77$). NFIA protein expression was widespread in hCSs (Fig. 4D, top row, and fig. S12, A to D) and HFT (fig. S12E), which is consistent with its RNA-seq expression and motif accessibility patterns. More specifically, NFIA was expressed in progenitors, including

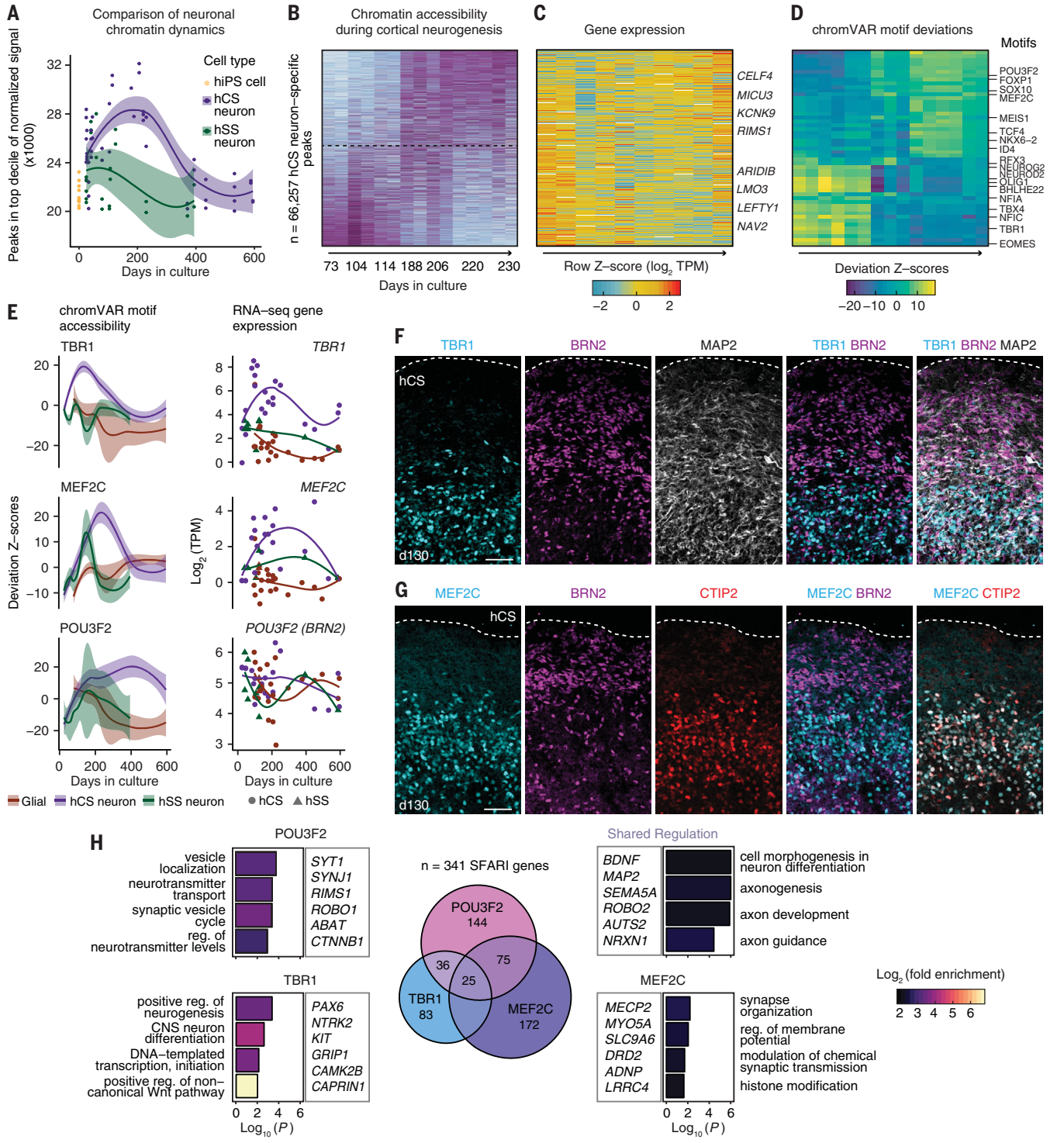


Fig. 5. A wave of chromatin accessibility and coordinated gene regulation during cortical neurogenesis. (**A**) Comparison of highly active open chromatin regions in neuronal lineages. The number of peaks in the top decile of ATAC-seq signal is shown. (**B**) Heatmap of hCS neuron peaks (DESeq2, FDR-adjusted $P < 0.01$) over 79 to 230 days. Peaks are sorted by the day-weighted mean of accessibility across these time points. Signal is displayed as row-scaled normalized accessibility. (**C**) Gene expression patterns of genes significantly linked to peaks in (B), displayed as row-scaled $\log_2$ (TPM). Select genes are shown. (**D**) ChromVAR deviations in hCSs. Row-scaled deviation *Z*-scores for each motif are plotted and sorted by the day-weighted mean of the deviations. (**E**) ChromVAR deviations and RNA-seq expression values for TBR1, MEF2C, and POU3F2 (BRN2). (**F**) Immunohistochemistry of hCS in cryosections at day 130 showing expression of the layer-specific markers TBR1 and BRN2 (or POU3F2) with MAP2. (**G**) Immunohistochemistry hCS at day 130 showing expression of MEF2C, BRN2, and CTIP2. (**H**) For each TF motif, the number of genes from the SFARI database with a linked enhancer having that motif is shown. Euler diagram indicates genes with binding motifs for more than one of the indicated TFs in linked enhancers. GO enrichments for groups of these SFARI genes are shown. "Shared regulation" indicates genes with two or more motifs. Scale bars, 50 μm (F) and (G).

HOPX$^+$ and SOX9$^+$ cells in hCSs and HFT, as well as in neurons, as shown by colocalization with CTIP2. The NFIA motif was also correlated to *NFIB*, which had similar protein expression patterns (fig. S12, F to H) (*30*). NEUROD2, which was strongly enriched in PN1 and PN2 clusters, was anticorrelated with *ID4* (r = –0.66), which also has a basic helix-loop-helix (bHLH) motif. Both *ID3* and *ID4* RNA expression increased over time in glial cells, and ID4 protein was localized to the nuclei of HOPX$^+$ cells but not CTIP2$^+$ neurons in hCSs (Fig. 4D, second row). In HFT, ID4 protein expression colocalized with HOPX$^+$ nuclei in the outer subventricular zone but not with CTIP2$^+$ cells (fig. S12I). ID-family TFs have been shown to dimerize with and repress other bHLH-family proteins (*49*), and Id4$^{-/-}$ mice displayed altered neural and glial differentiation (*50*). Thus, ID4 may act to maintain the radial glia population. The SOX10 motif was correlated with *SOX2*, *SOX9*, and most strongly *SOX21* expression (r = 0.73), and the latter increased in glial cells over time. We found protein expression of SOX21 in GFAP-expressing radial glia in both hCSs and HFT (fig. S13, A and B). Consistent with our genomic data, we found nuclear SOX21 expression in mature astrocytes at later stages in hCSs (200 and 552 days) (Fig. 4D, third row). In examining motifs specific to subpallial neurons (fig. S8A), we found that *ONECUT2*, *NKX2-1*, and *RFX4* were highly correlated to ONECUT1, NKX2-8, and RFX2 motif deviations (Fig. 4C). ONECUT2 was expressed in the nucleus of a subset of hSS cells but not hCS cells (Fig. 4D, bottom row, and fig. S14, A and B). ZIC1 was similarly expressed in the nuclei of hSS cells, localized with GABA, and was expressed to a much lesser degree in hCS cells (fig. S14, C to E). Last, RFX4 expression was also higher in hSSs and localized with GAD67 expression (fig. S14, F to I) (*30*). Together, these examples illustrate how this data can be interrogated to identify TFs that may regulate lineage specification in the forebrain.

To explore the diversity of sequence motifs in each of our dynamic clusters, we grouped enriched motifs ($\log_2$ fold enrichment > 0.6) according to their JASPAR family assignments and computed the Shannon entropy of these labels by using the total number of enriched motifs as a prior estimate (fig. S15). The PL cluster had the highest diversity by this metric, followed by the PN2 cluster (H_{PL} = 2.61 and H_{PN2} = 2.39), whereas MG2 and GP had the lowest (H_{MG2} = 0.47 and H_{GP} = 0.66). As described above, GP accessibility was primarily associated with homeobox TFs. Meanwhile, the PN2 cluster exhibited a number of diverse yet highly enriched families, including bHLH, T-box, natural killer (NK)–related, and MADS box factors (Fig. 4E). Although this analysis does not capture accessibility contributions by weakly enriched factors, it suggests that the period of cortical neurogenesis corresponding to 79 to 230 days is associated with a greater diversity of active TFs, prompting us to explore this stage further.

Gene-regulatory dynamics during human cortical neurogenesis

Upon closer inspection of the neuronal lineages, we observed more extensive chromatin remodeling in hCS neuronal lineages when compared with hSS interneurons, particularly between 79 and 230 days (Fig. 5A and table S9) (*30*). This could reflect the distinct trajectory of cortical interneurons, which require not just fate specification but also migration and integration into circuits (*51*). Migration and functional integration stages are not captured in our model without assembly. A differential test revealed that 26% of all dynamic peaks (n = 66,257) were specific to hCS and HFT neurons in this time window (FDR-adjusted P < 0.01) and that this wave of differential neuronal accessibility overlapped as expected with clusters PN1 and PN2 (J = 0.53). We then sorted hCS neuron–specific peaks by their ATAC-seq signal over time and observed dynamic changes (Fig. 5B).

Corticogenesis involves the sequential generation of layer-specific neurons over ~20 weeks in utero. To investigate the gene regulation underlying this process, we applied the analytical tools described above. Using our linked enhancer framework, we selected the strongest associations with the chromatin remodeling wave between 79 and 230 days and similarly sorted and plotted their expression in neurons over time (Fig. 5C); *NAV2*, *LEFTY1*, and the transcriptional cofactors *LMO3* and *ARID1B* were among the genes expressed at earlier stages, whereas *KCNK9*, *MICU3*, and *RIMS1* were found at later stages. We next selected TF motifs that were significantly enriched in clusters PN1 or PN2 and sorted their chromVAR deviations by the same metric, revealing concomitant progressions in motif accessibility over time (Fig. 5D); EOMES, TBR1, and NEUROD2 motifs represented the earliest phase; TCF4 and MEF2C represented a middle phase; and FOXP1, CUX2, and POU3F2 (BRN2) motifs represented the latest stages. The chromVAR deviations for many of these motifs also had strong correlations to the expression of at least one cognate TF (Fig. 5D).

We next identified chromatin accessibility linked to *TBR1* (TBR1 motif), which is known to be expressed in early-born deep-layer neurons, and accessibility linked to *POU3F2* (POU3F2 motif), which is expressed in late-born, superficial-layer neurons (*52*) and which has previously been shown to regulate gene expression networks related to the pathogenesis of schizophrenia and bipolar disorder (*53*). The expression of these transcripts peaked in hCS neurons between 93 and 114 days for *TBR1* and 220 and 250 days for *BRN2*, and immunohistochemistry in hCSs and HFT confirmed this pattern of protein expression (Fig. 5, E and F). Our analysis highlighted MEF2C as a key TF motif associated with accessibility at the middle stage of the corticogenesis accessibility wave (with r = 0.83 correlation with the *MEF2C* gene). *MEF2C* has been previously associated with ASD (*54*), but its gene regulatory activity in the context of human cortical development has not been described. Consistent with our genomic measurements, we found that MEF2C was not expressed in progenitor cells or CTIP2$^+$ neurons at 75 days (fig. S16A). However, MEF2C partially colocalized with CTIP2 and RORB in hCSs after 130 days (Fig. 5G and fig. S16, B and C) as well as in HFT (fig. S16, D and E). These convergent lines of evidence suggest an important role for MEF2C in midcortical neurogenesis.

Because the *MEF2C* gene as well as the *TBR1* and *BRN2* genes have been linked to ASD (*55*, *56*), we next explored predicted downstream pathways and cell functions related to this association in human corticogenesis. Approximately 32% (n = 341 genes) of SFARI ASD genes had a linkage to an element containing one of these three motifs. More specifically, TBR1-specific gene linkages included the TF *PAX6* and calcium/calmodulin–dependent kinase *CAMK2B*, and overall, these genes were enriched for processes related to transcription regulation and cell signaling (relative to a background of variable genes in our forebrain data); MEF2C-specific gene targets included the Rett syndrome–related *MECP2* and the dopamine receptor *DRD2* and were enriched for synaptic organization as well as histone modification; BRN2-specific genes included *SYT1* and *SYTJ1* and were enriched for synaptic activity (Fig. 5H). Last, when we looked at the genes targeted by two or more of these three motifs, we discovered that they converged on a signature of axonogenesis and chemotaxis that included *BDNF*, *SEMA5A*, and *ROBO2* (Fig. 5H). Overall, these examples illustrate how this platform can be interrogated to identify factors that regulate development and disease risk during periods of human forebrain development that are inaccessible.

Discussion

The assembly of the human forebrain is a lengthy developmental process that, in primates, extends into postnatal life (*1*, *2*, *57*). Development can be disrupted by genetic and environmental factors, leading to disease. Our molecular understanding of forebrain development comes mostly from animal models. Recently, transcriptomic and epigenetic methods have been applied in human brain tissue (*18*, *20*, *21*, *58*). However, critical developmental time periods corresponding to mid- and

late gestation are still poorly characterized owing to tissue availability. We have used 3D organoids that can be maintained over long time periods to study chromatin accessibility dynamics during human forebrain development in specific cell lineages. Direct comparisons with human tissue validate that in vivo forebrain regulatory programs broadly map onto those in long-term cultures. Our accessibility landscape extends across cell lineages, forebrain domains, and time and reveals lineage-specificity of disease risk and conservation. We identified TFs associated with astrocyte maturation and interneuron specification and discovered a protracted period of chromatin remodeling during human cortical neurogenesis. Last, we have described how several key TFs may coordinate over time to regulate cellular functions as pallial neurons develop.

In contrast to expressed genes, regulatory elements are more numerous and cell lineage–restricted, and their sequences can provide clues about the gene programs that drive specification (*6*). By combining epigenetic information with transcriptomic and protein measurements, we nominate TFs with substantial regulatory potential in forebrain cell lineages. We have generated examples of how the data can be used to elucidate the regulation of forebrain lineages. We anticipate that future studies could leverage our data to experimentally demonstrate the detailed mechanisms that govern the production of pallial neurons versus glial cells, the specification of interneurons, and cortical cellular maturation.

Efforts to link disease risk to epigenetic or transcriptomic trajectories have primarily used dissected brain tissue at early stages of development or later postnatal stages (*17*, *20*, *21*, *37*, *59*, *60*). By capturing both human neurogenesis and gliogenesis in vitro, we have further mapped genetic ASD risk to glial progenitor cells as well as mid- and late-stage neurons. Moreover, we have defined a set of ASD genes enriched for chemotaxis and axonogenesis that appear to be regulated by key TFs. Single-cell epigenomic assays, which have recently become tractable and scalable, will be especially suited to resolving how TFs coordinate across cells to achieve cortical specification and disease gene regulation.

The brain region–specific organoid platform holds the promise of flexibly modeling the development of other diverse brain regions, which are also currently inaccessible to molecular study, and modeling patient genetic backgrounds. Patient-derived organoids will allow for genetic and pharmacological manipulation of phenotypes to identify disease mechanisms, and we anticipate that comparison of epigenetic trajectories will provide further insights into pathophysiology.

Last, our data indicate that organoids intrinsically undergo chromatin state transitions in vitro that are closely related to human forebrain development in vivo. However, a full understanding of brain development will require the study of not only these intrinsic programs but also extrinsic cues that influence development and disease. The assembly of organoids from multiple brain regions, in combination with genetic or optogenetic tools, could be used to comprehensively assess the impact of connectivity and activity on development and maturation.

Methods

Detailed materials and methods can be found in the supplementary materials.

Generation of hCS and hSS from hiPS cells

Intact hiPS cell colonies maintained on mouse embryonic fibroblast feeders or feeder-free cells were enzymatically lifted before being transferred into ultralow-attachment plates in hiPS cell medium supplemented with the SMAD inhibitors dorsomorphin (5 μM) and SB-431542 (10 μM), as described (*25*–*27*). On the 6th day in suspension, the medium was switched to neural medium supplemented with epidermal growth factor (EGF) (20 ng/ml) and fibroblast growth factor 2 (FGF-2) (20 ng/ml). The neural medium was changed every day until day 17 and then every other day until day 24. For the generation of hSSs, the medium was additionally supplemented with inhibitor of Wnt production 2 (IWP-2) (5 μM) on days 4 to 24, smoothened agonist (SAG) (100 nM) on days 12 to 24, retinoic acid (RA) (100 nM) on days 12 to 15, and allopregnanolone (100 nM) on days 15 to 24. From day 25 to day 43, the neural medium for both hCSs and hSSs was changed every other day and supplemented with brain-derived neurotrophic factor (BDNF) (20 ng/ml) and neutrophin-3 (NT3) (20 ng/ml). From day 45 onwards, hCSs and hSSs were maintained in neural medium without growth factors with medium changes every 4 to 6 days.

Human tissue

Human brain tissue was obtained under a protocol approved by the Research Compliance Office at Stanford University. PCW20 and PCW21 forebrain tissue was delivered overnight on ice and immediately processed after arrival.

Dissociation, immunopanning, and cell sorting

Dissociation of neural spheroids and human tissue into single cells was performed as described (*24*–*26*, *29*). Tissue was chopped and incubated in 40 U/ml papain enzyme solution at 37°C for 90 min then gently triturated to achieve a single-cell suspension. Single-cell suspensions were then either FACS-sorted or immunopanned to achieve astrocyte- or neuron-enriched populations. For FACS-sorting, cells separated on the basis of their expression of either GFP (pLV-GFAP::eGFP glia) or mCherry (AAV-DJ-hSYN1::mCherry neurons). For immunopanning, the single-cell suspension was added to a series of plastic petri dishes precoated with either antibody to against HepaCAM (glia) or antibody against Thy1 (neurons) and incubated for 10 to 30 min at room temperature. Bound cells were incubated in a trypsin solution (for glia) or Accutase (for neurons) at 37°C for 3 to 5 min then gently pipetted off the plates and processed for RNA-seq and ATAC-seq.

ATAC-seq processing

ATAC-seq was performed by resuspending cells in lysis buffer with Tween-20, followed by centrifugation at 4°C for 10 min and incubation with transposase at 37°C for 30 min. Sequencing was performed on an Illumina NextSeq. Data were processed against the GRCh38/hg38 reference genome. ATAC-seq data were aligned by using Bowtie2. Peaks were called by using MACS2, filtered to q value < 0.01, then merged within biological replicate groups and tiled to a width of 500 base pairs. Tiles were evaluated for variability and signal strength across samples and removed if they failed these metrics. Last, counts were generated for each peak set by using multiBamCov, cleaned with edgeR, and quantile normalized.

RNA-seq processing

RNA was isolated from single-cell suspensions and homogenized tissue by use of TRIzol, purified with direct-zol RNA MicroPrep kits, and processed with an Ovation NuGen RNA-seq kit. The quality of the libraries was assessed with BioAnalyzer before sequencing on an Illumina NextSeq. Reads were trimmed by using Skewer then pseudo-aligned to the RefSeq transcriptome with Kallisto to obtain per-transcript TPM values, as well as count estimates. Read depth, alignment rate, and RNA integrity number (RIN) score were evaluated, and samples were kept with RIN > 7.5. To correct for batch effects across samples collected at different times, we used the package "limma" on the matrix of log-transformed TPM values: "removeBatchEffect(log2(TPM + 1))".

REFERENCES AND NOTES

1. J. C. Silbereis, S. Pochareddy, Y. Zhu, M. Li, N. Sestan, The cellular and molecular landscapes of the developing human central nervous system. *Neuron* **89**, 248–268 (2016). doi: 10.1016/j.neuron.2015.12.008; pmid: 26796689
2. J. L. R. Rubenstein, Annual Research Review: Development of the cerebral cortex: implications for neurodevelopmental disorders. *J. Child Psychol. Psychiatry* **52**, 339–355 (2011). doi: 10.1111/j.1469-7610.2010.02307.x; pmid: 20735793
3. C. P. Wonders, S. A. Anderson, The origin and specification of cortical interneurons. *Nat. Rev. Neurosci.* **7**, 687–696 (2006). doi: 10.1038/nrn1954; pmid: 16883309
4. G. Juric-Sekhar, R. F. Hevner, Malformations of cerebral cortex development: Molecules and mechanisms. *Annu. Rev. Pathol.*

14, 293–318 (2019). doi: 10.1146/annurev-pathmechdis-012418-012927; pmid: 30677308
5. J. D. Buenrostro *et al*., Integrated single-cell analysis maps the continuous regulatory landscape of human hematopoietic differentiation. *Cell* **173**, 1535–1548.e16 (2018). doi: 10.1016/j.cell.2018.03.074; pmid: 29706549
6. A. B. Stergachis *et al*., Developmental fate and cellular maturity encoded in human regulatory DNA landscapes. *Cell* **154**, 888–903 (2013). doi: 10.1016/j.cell.2013.07.020; pmid: 23953118
7. ENCODE Project Consortium, An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**, 57–74 (2012). doi: 10.1038/nature11247; pmid: 22955616
8. F. Grubert *et al*., Genetic control of chromatin states in humans involves local and distal chromosomal interactions. *Cell* **162**, 1051–1065 (2015). doi: 10.1016/j.cell.2015.07.048; pmid: 26300125
9. E. Z. Kvon *et al*., Progressive loss of function in a limb enhancer during snake evolution. *Cell* **167**, 633–642.e11 (2016). doi: 10.1016/j.cell.2016.09.028; pmid: 27768887
10. S. L. Klemm, Z. Shipony, W. J. Greenleaf, Chromatin accessibility and the regulatory epigenome. *Nat. Rev. Genet.* **20**, 207–220 (2019). doi: 10.1038/s41576-018-0089-8; pmid: 30675018
11. H. J. Kang *et al*., Spatio-temporal transcriptome of the human brain. *Nature* **478**, 483–489 (2011). doi: 10.1038/nature10523; pmid: 22031440
12. M. Li *et al*., Integrative functional genomic analysis of human brain development and neuropsychiatric risks. *Science* **362**, eaat7615 (2018). doi: 10.1126/science.aat7615; pmid: 30545854
13. J. A. Miller *et al*., Transcriptional landscape of the prenatal human brain. *Nature* **508**, 199–206 (2014). doi: 10.1038/nature13185; pmid: 24695229
14. T. J. Nowakowski *et al*., Spatiotemporal gene expression trajectories reveal developmental hierarchies of the human cortex. *Science* **358**, 1318–1323 (2017). doi: 10.1126/science.aap8809; pmid: 29217575
15. M. B. Johnson *et al*., Single-cell analysis reveals transcriptional heterogeneity of neural progenitors in human cortex. *Nat. Neurosci.* **18**, 637–646 (2015). doi: 10.1038/nn.3980; pmid: 25734491
16. A. A. Pollen *et al*., Molecular identity of human outer radial glia during cortical development. *Cell* **163**, 55–67 (2015). doi: 10.1016/j.cell.2015.09.004; pmid: 26406371
17. A. Amiri *et al*., Transcriptome and epigenome landscape of human cortical development modeled in organoids. *Science* **362**, eaat6720 (2018). doi: 10.1126/science.aat6720; pmid: 30545853
18. A. E. Jaffe *et al*., Mapping DNA methylation across development, genotype and schizophrenia in the human frontal cortex. *Nat. Neurosci.* **19**, 40–47 (2016). doi: 10.1038/nn.4181; pmid: 26619358
19. S. Kanton *et al*., Organoid single-cell genomic atlas uncovers human-specific features of brain development. *Nature* **574**, 418–422 (2019). doi: 10.1038/s41586-019-1654-9; pmid: 31619793
20. L. de la Torre-Ubieta *et al*., The dynamic landscape of open chromatin during human cortical neurogenesis. *Cell* **172**, 289–304.e18 (2018). doi: 10.1016/j.cell.2017.12.014; pmid: 29307494
21. H. Won *et al*., Chromosome conformation elucidates regulatory relationships in developing human brain. *Nature* **538**, 523–527 (2016). doi: 10.1038/nature19847; pmid: 27760116
22. S. P. Pașca, The rise of three-dimensional human brain cultures. *Nature* **553**, 437–445 (2018). doi: 10.1038/nature25032; pmid: 29364288
23. A. M. Pașca *et al*., Functional cortical neurons and astrocytes from human pluripotent stem cells in 3D culture. *Nat. Methods* **12**, 671–678 (2015). doi: 10.1038/nmeth.3415; pmid: 26005811
24. S. A. Sloan *et al*., Human astrocyte maturation captured in 3D cerebral cortical spheroids derived from pluripotent stem cells. *Neuron* **95**, 779–790.e6 (2017). doi: 10.1016/j.neuron.2017.07.035; pmid: 28817799
25. F. Birey *et al*., Assembly of functionally integrated human forebrain spheroids. *Nature* **545**, 54–59 (2017). doi: 10.1038/nature22330; pmid: 28445465
26. S.-J. Yoon *et al*., Reliability of human cortical organoid generation. *Nat. Methods* **16**, 75–78 (2019). doi: 10.1038/s41592-018-0255-0; pmid: 30573846
27. S. A. Sloan, J. Andersen, A. M. Pașca, F. Birey, S. P. Pașca, Generation and assembly of human brain region-specific three-dimensional cultures. *Nat. Protoc.* **13**, 2062–2085 (2018). doi: 10.1038/s41596-018-0032-7; pmid: 30202107
28. M. R. Corces *et al*., An improved ATAC-seq protocol reduces background and enables interrogation of frozen tissues. *Nat. Methods* **14**, 959–962 (2017). doi: 10.1038/nmeth.4396; pmid: 28846090
29. Y. Zhang *et al*., Purification and characterization of progenitor and mature human astrocytes reveals transcriptional and functional differences with mouse. *Neuron* **89**, 37–53 (2016). doi: 10.1016/j.neuron.2015.11.013; pmid: 26687838
30. Materials and methods are available as supplementary materials.
31. H. A. Pliner *et al*., Cicero predicts cis-regulatory dna interactions from single-cell chromatin accessibility data. *Mol. Cell* **71**, 858–871.e8 (2018). doi: 10.1016/j.molcel.2018.06.044; pmid: 30078726
32. A. Kundaje *et al*., Integrative analysis of 111 reference human epigenomes. *Nature* **518**, 317–330 (2015). doi: 10.1038/nature14248; pmid: 25693563
33. M. R. Corces *et al*., The chromatin accessibility landscape of primary human cancers. *Science* **362**, eaav1898 (2018). doi: 10.1126/science.aav1898; pmid: 30361341
34. B. Bulik-Sullivan *et al*., An atlas of genetic correlations across human diseases and traits. *Nat. Genet.* **47**, 1236–1241 (2015). doi: 10.1038/ng.3406; pmid: 26414676
35. H. K. Finucane *et al*., Partitioning heritability by functional annotation using genome-wide association summary statistics. *Nat. Genet.* **47**, 1228–1235 (2015). doi: 10.1038/ng.3404; pmid: 26414678
36. Schizophrenia Working Group of the Psychiatric Genomics Consortium, Biological insights from 108 schizophrenia-associated genetic loci. *Nature* **511**, 421–427 (2014). doi: 10.1038/nature13595; pmid: 25056061
37. N. N. Parikshak *et al*., Integrative functional genomic analyses implicate specific molecular pathways and circuits in autism. *Cell* **155**, 1008–1021 (2013). doi: 10.1016/j.cell.2013.10.031; pmid: 24267887
38. J. Zhou *et al*., Whole-genome deep-learning analysis identifies contribution of noncoding mutations to autism risk. *Nat. Genet.* **51**, 973–980 (2019). doi: 10.1038/s41588-019-0420-0; pmid: 31133750
39. N. A. Oberheim *et al*., Uniquely hominid features of adult human astrocytes. *J. Neurosci.* **29**, 3276–3287 (2009). doi: 10.1523/JNEUROSCI.4707-08.2009; pmid: 19279265
40. A. Visel, S. Minovitsky, I. Dubchak, L. A. Pennacchio, VISTA Enhancer Browser—A database of tissue-specific human enhancers. *Nucleic Acids Res.* **35** (Database), D88–D92 (2007). doi: 10.1093/nar/gkl822; pmid: 17130149
41. J. A. Capra, G. D. Erwin, G. McKinsey, J. L. R. Rubenstein, K. S. Pollard, Many human accelerated regions are developmental enhancers. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **368**, 20130025 (2013). doi: 10.1098/rstb.2013.0025; pmid: 24218637
42. K. S. Pollard *et al*., Forces shaping the fastest evolving regions in the human genome. *PLOS Genet.* **2**, e168 (2006). doi: 10.1371/journal.pgen.0020168; pmid: 17040131
43. S. K. Balakrishnan, M. Witcher, T. W. Berggren, B. M. Emerson, Functional and molecular characterization of the role of CTCF in human embryonic stem cell biology. *PLOS ONE* **7**, e42424 (2012). doi: 10.1371/journal.pone.0042424; pmid: 22879976
44. V. S. Mangale *et al*., Lhx2 selector activity specifies cortical identity and suppresses hippocampal organizer fate. *Science* **319**, 304–309 (2008). doi: 10.1126/science.1151695; pmid: 18202285
45. K. Hisanaga, S. M. Sagar, K. J. Hicks, R. A. Swanson, F. R. Sharp, c-fos proto-oncogene expression in astrocytes associated with differentiation or proliferation but not depolarization. *Brain Res. Mol. Brain Res.* **8**, 69–75 (1990). doi: 10.1016/0169-328X(90)90011-2; pmid: 2166202
46. A. Jolma *et al*., DNA-binding specificities of human transcription factors. *Cell* **152**, 327–339 (2013). doi: 10.1016/j.cell.2012.12.009; pmid: 23332764
47. A. Jolma *et al*., DNA-dependent formation of transcription factor pairs alters their binding specificity. *Nature* **527**, 384–388 (2015). doi: 10.1038/nature15518; pmid: 26550823
48. A. N. Schep, B. Wu, J. D. Buenrostro, W. J. Greenleaf, chromVAR: Inferring transcription-factor-associated accessibility from single-cell epigenomic data. *Nat. Methods* **14**, 975–978 (2017). doi: 10.1038/nmeth.4401; pmid: 28825706
49. H. A. Sikder, M. K. Devlin, S. Dunlap, B. Ryu, R. M. Alani, Id proteins in cell growth and tumorigenesis. *Cancer Cell* **3**, 525–530 (2003). doi: 10.1016/S1535-6108(03)00141-7; pmid: 12842081
50. L. Bedford *et al*., Id4 is required for the correct timing of neural differentiation. *Dev. Biol.* **280**, 386–395 (2005). doi: 10.1016/j.ydbio.2005.02.001; pmid: 15882580
51. A. Kepecs, G. Fishell, Interneuron cell types are fit to function. *Nature* **505**, 318–326 (2014). doi: 10.1038/nature12983; pmid: 24429630
52. H. Zeng *et al*., Large-scale cellular-resolution gene profiling in human neocortex reveals species-specific molecular signatures. *Cell* **149**, 483–496 (2012). doi: 10.1016/j.cell.2012.02.052; pmid: 22500809
53. C. Chen *et al*., The transcription factor POU3F2 regulates a gene coexpression network in brain tissue from patients with psychiatric disorders. *Sci. Transl. Med.* **10**, eaat8178 (2018). doi: 10.1126/scitranslmed.aat8178; pmid: 30545964
54. H. Li *et al*., Transcription factor MEF2C influences neural stem/progenitor cell differentiation and maturation in vivo. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 9397–9402 (2008). doi: 10.1073/pnas.0802876105; pmid: 18599437
55. J. H. Notwell *et al*., TBR1 regulates autism risk genes in the developing neocortex. *Genome Res.* **26**, 1013–1022 (2016). doi: 10.1101/gr.203612.115; pmid: 27325115
56. N.-P. Tsai *et al*., Multiple autism-linked genes mediate synapse elimination via proteasomal degradation of a synaptic scaffold PSD-95. *Cell* **151**, 1581–1594 (2012). doi: 10.1016/j.cell.2012.11.040; pmid: 23260144
57. B. Martynoga, D. Drechsel, F. Guillemot, Molecular control of neurogenesis: A view from the mammalian cerebral cortex. *Cold Spring Harb. Perspect. Biol.* **4**, a008359 (2012). doi: 10.1101/cshperspect.a008359; pmid: 23028117
58. S. Preissl *et al*., Single-nucleus analysis of accessible chromatin in developing mouse forebrain reveals cell-type-specific transcriptional regulation. *Nat. Neurosci.* **21**, 432–439 (2018). doi: 10.1038/s41593-018-0079-3; pmid: 29434377
59. D. Velmeshev *et al*., Single-cell genomics identifies cell type-specific molecular changes in autism. *Science* **364**, 685–689 (2019). doi: 10.1126/science.aav8130; pmid: 31097668
60. A. J. Willsey *et al*., Coexpression networks implicate human midfetal deep cortical projection neurons in the pathogenesis of autism. *Cell* **155**, 997–1007 (2013). doi: 10.1016/j.cell.2013.10.020; pmid: 24267886

ACKNOWLEDGMENTS

We thank members of the Greenleaf, Pasca, Pritchard, and Chang laboratories for discussion and advice, especially J. Trombetta, M. Mumbach, J. Granja, D. Calderon, F. Birey, R. M. Agoglia, T. Khan, and A. M. Pasca for analytical and technical assistance. We thank K. Kovary for designing the interactive webpage. We thank the Stanford Research Computing Center for computational resources. **Funding:** This work was supported by NIH grants P50HG007735 and UM1HG009442 (to H.Y.C. and W.J.G.) and U19AI057266 (to W.J.G.), BRAINS Award MH107800 and U01 MH115745 as part of the National Institute of Mental Health Convergent Neuroscience Consortium (to S.P.P.), the Rita Allen Foundation (W.J.G.), the Baxter Faculty Scholar (W.J.G. and S.P.P.), S. Coates and the V.J Coates Foundation (S.P.P.), the Human Frontiers Science RGY0065 (W.J.G), the Stanford Wu Tsai Neurosciences Rejuvenation Project and Human Brain Organogenesis Program (S.P.P), and the Kwan Fund (S.P.P). W.J.G. is a Chan Zuckerberg Biohub investigator and acknowledges grants 2017-174468 and 2018-182817 from the Chan Zuckerberg Initiative. S.P.P. is a New York Stem Cell Foundation Robertson Stem Cell Investigator and a Chan Zuckerberg Ben Barres Investigator. J.K.P. and H.Y.C. are investigators of the Howard Hughes Medical Institute. Fellowship support was provided by the NSF Graduate Research Fellowship Program, the Siebel Scholars, the Enhancing Diversity in Graduate Education Program, and the Weiland Family Fellowship (A.E.T.); the National Defense Science and Engineering Graduate Fellowship and the Stanford Graduate Fellowship (N.S-A.); the MCHRI Fellowship (N.H.); the Walter V. and Idun Berry Fellowship (J.A.); and the Stanford Dean's Fellowship (J.A. and N.H.). **Author contributions:** A.E.T., N.S-A., J.A., N.H., W.J.G., and S.P.P. conceived the project and designed experiments. A.E.T., W.J.G. and S.P.P. wrote the manuscript with input from authors. J.A. generated and maintained forebrain 3D long-term cultures and performed validations. S.-J.Y. and N.H. performed cell isolation from 3D cultures or primary tissue. A.E.T. and N.S-A. performed ATAC-seq and RNA-seq experiments and analyses with input from J.K.P. and H.Y.C.; W.J.G. and S.P.P. supervised all aspects of the work. **Competing interests:** Stanford University has filed a provisional patent application that covers the generation of region-specific brain organoids and their assembly (U.S. patent application number 62/477,858 and 15/158,408). **Data availability:** Data used for the analyses presented in this work are available under GEO accession no. GSE132403. Please see the supplementary materials for detailed availability of the provided tables and data. A website associated with the manuscript, including an interactive data browser, is available at http://brainchromatin.stanford.edu.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/367/6476/eaay1645/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S16
Tables S1 to S10
References (*61*–*76*)

13 June 2019; accepted 16 December 2019
10.1126/science.aay1645

NANOPHOTONICS

Coherent scanning tunneling microscopy

Ultrafast phase-stabilized optical pulses add temporal resolution to high spatial resolution

By **Clarice D. Aiello**

The ultimate goal of any nanotechnology is to resolve and control elementary processes in matter. In general, although many spectroscopies can achieve high temporal resolution and many microscopies can achieve high spatial resolution, achieving both is difficult. On page 411 of this issue, Garg and Kern (*1*) improve the limits of concomitant spatial and temporal resolutions by combining scanning tunneling microscopy (STM) (*2*) with an ingenious phase-locking train of ultrafast optical pulses. The authors go on to demonstrate that their instrument can deconvolve femtosecond electron dynamics with nanoscale resolution. These feats are only made more interesting because the phase-locking scheme enables an imprint of the laser's (coherent) phase onto the tunneling electrons. This demonstration of "light-wave electronics" at atomic scales is unprecedented.

Coherence and coherent control have widely varying meanings across fields, so it is important to be specific in how the terms are meant in these studies (*3*). Electron tunneling through a barrier is not a process that physicists usually think of as "coherent." For example, when a voltage is applied onto a STM tip near a metal surface, the stimulated tunneling current lacks coherence because electrons are affected by the randomly evolving phase of the applied voltage.

By contrast, the experiments of Garg and Kern exhibit coherent tunneling, defined here as an electronic current with a well-defined phase at the nanoscale. The carrier-envelope phase φ_{CEP} is defined as the phase between the carrier wave and the maximum of its intensity envelope (*4*). The approach to control electronics with CEP-stabilized optical pulses is known as light-wave electronics (*5*). When a STM tip is excited with CEP-stabilized pulses, a very strong electric field of ~1 V/Å is created that can lower the tunneling barrier (see the figure). The induced tunneling current will be coherent, provided that the pulses' phase is indeed stable for the duration of the tunneling process (which is estimated to be on the femtosecond time scale).

In such cases, the tunneling current intensity (I) is predicted to oscillate with φ_{CEP}, that is, I is proportional to $\exp(i\varphi_{CEP})$. It is in this sense that the pulses' phase is imprinted onto the tunneling electrons. In the experimental setup of Garg and Kern, φ_{CEP} could be chosen and tuned at will. This capability, which was enabled by the CEP feed-forward lock they developed, will be of relevance for a broad class of instrument builders.

As a proof-of-principle experiment, a CEP-stabilized train of ultrafast optical pulses was impinged on a STM tip near a gold surface. By varying the CEP, Garg and Kern could indeed measure modulations of tunneling current intensity. This demonstrated one essential ingredient of nanoelectronics, namely, control over atomic-level currents. The other essential ingredient is manipulation speed. Such optical pulses, which are on the several-femtosecond scale, correspond to frequencies of hundreds of terahertz, which compares favorably to current electronics that operate in the gigahertz to terahertz regimes (*6*).

Finally, Garg and Kern use the same experimental method to study fast carrier-decay dynamics in nanostructures. A STM tip was placed in the close vicinity of gold nanorods and excited by CEP-stabilized optical pulses in a type of pump-probe configuration. The tunneling current varies with the delay between pulse and probe pulses with a dominant frequency component matching a plasmon resonance in the nanorods. This observation suggests that the laser is effectively exciting a plasmon resonance that decays within a time scale (~40 fs) that was well resolved by the method.

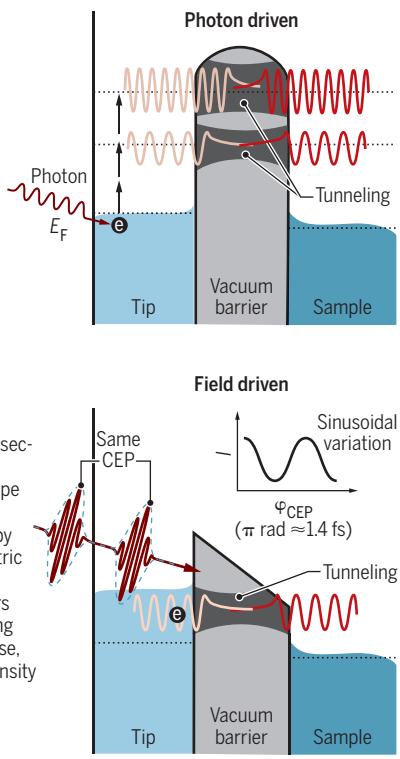
More broadly, the experimental tool kit advanced here can be put to use to investigate charge-transport phenomena in nanostructures and in single molecules adsorbed onto a surface, including molecules of biological relevance. The reported experiments should encourage research into unexplored avenues. For example, manipulating the spins of the tunneling electrons with similar techniques would allow investigations of the chiral-induced spin selectivity effect (*7*) at the single-molecule level. Additionally, quantum computing protocols might harness the coherent tunneling phase. In any case, Garg and Kern's present experimental tour de force should already entice the reader to be on the lookout for nanomicroscopy and nanoelectronics techniques that use the combined power of STMs and CEP-stabilized ultrafast optical pulses. ■

Coherently through the barrier

Garg and Kern irradiated a scanning tunneling microscope tip with phase-stabilized femtosecond optical pulses. The resulting strong electric fields produce in-phase electrons (e) whose atomic-scale flux can be tuned within the same time frame. E_F, Fermi energy.

Photon-promoted tunneling
Low fluxes of femtosecond pulses promote the tunneling of electrons by exciting them to higher energy states. Tunneling electrons are not in phase.

Field-induced tunneling
High fluxes of femtosecond pulses with the same carrier-envelope phase (CEP) φ_{CEP} promote tunneling by creating a high electric field that changes the shape and lowers the barrier. Tunneling electrons are in phase, and the current intensity I is proportional to the experimentally tunable CEP.



GRAPHIC: A. KITTERMAN/*SCIENCE*

Quantum Biology Tech (QuBiT) Lab, Electrical and Computer Engineering Department, University of California, Los Angeles, Los Angeles, CA, USA. Email: cla@ucla.edu

REFERENCES AND NOTES

1. M. Garg, K. Kern, *Science* **367**, 411 (2020).
2. B. Voigtländer, *Scanning Probe Microscopy: Atomic Force Microscopy and Scanning Tunneling Microscopy* (NanoScience and Technology Series, Springer, 2015).
3. T. Baumgratz *et al.*, *Phys. Rev. Lett.* **113**, 140401 (2014).
4. K. Yoshioko *et al.*, *Nat. Photonics* **10**, 762 (2016).
5. E. Goulielmakis *et al.*, *Science* **317**, 769 (2007).
6. X. Mei *et al.*, *IEEE Electron Device Lett.* **36**, 327 (2015).
7. R. Naaman, D. H. Waldeck, *J. Phys. Chem. Lett.* **3**, 2178 (2012).

10.1126/science.aba0714

SYNTHETIC CHEMISTRY

Rethinking chemistry for a circular economy

Chemical complexity complicates product recycling and manufacturing sustainability

By **Klaus Kümmerer[1,2], James H. Clark[3], Vânia G. Zuin[3,4]**

Earth is running out of resources needed for manufacturing materials such as chemicals, minerals, and petroleum. Thus, these components are available only at increasing economic and environmental costs. As an important contribution to a sustainable future, chemistry and its products must be adapted to a circular economy (CE)—a system aimed at eliminating waste, circulating and recycling products, and saving resources and the environment (*1*).

Nearly 140,000 industrial chemicals are marketed worldwide, and new chemicals are becoming more complex (e.g., stereochemistry, functional groups) (*2*). Products of the chemical and allied industries contain mixtures of elements and molecules, and these products and their constituents are found everywhere, including in waste products, soils, water, air, plants, food, animals, and the human body.

SHAPING A CIRCULAR ECONOMY

In a CE, products should be used as long as possible until the end of their lives. Modern circularity thinking includes the design of products with adapted lifetime, reusability, ease of repair, and recycling ability—all made with renewable resources (*3*). The development and implementation of CE approaches in China (*4*), the United States (*5*), the European Union (*1*), and other countries (*6*) are supported by international organizations such as the World Health Organization (WHO) and the United Nations (UN) (*2*). These efforts will help address Earth's resource and waste challenges and contribute to sustainable development. However, greater success will have to come from changes at the product-design level, led by scientists who strive to decipher, at the atomic and molecular levels, how chemical products and their underpinning synthetic chemistry fit into a CE. This includes identifying suitable starting materials and modifying them to the desired grade, structure, and function for an application. With today's diverse and interconnected chemical, material, and product flows, manufacturers must learn what can and should be circulated and recycled and what can and should not.

LIMITS OF CIRCULATION AND RECYCLING

Scientists expect metal demands to increase substantially for use in industrial chemistry and for products and processes that reduce CO_2 emissions or increase digitalization and (electro)mobility or enhanced communication. A sustainable future depends on the availability of high-grade metals—both common (e.g., aluminum and copper) and specialty (noble, rare earth) ones—sourced from virgin ores or from recycling.

Copper, for example, is indispensable for many products, such as wiring, wind turbines, electric motors, information technology, generators, sensors, and electronic devices. Since the beginning of the 20th century, copper production has grown more than 3000%, with a predicted market deficit of 600,000 metric tons by 2021 (*7*). New copper ores frequently fall short of the desired quality; thus, upgrading requires more energy and resources. Extraction of ores from deep mines creates a large environmental footprint, as this process generates more waste, mobilizes more toxic elements (e.g., arsenic), and is more likely to require access to protected lands of indigenous peoples, risking social disturbance.

Metal recycling can be advantageous given the energy-intensive processes of mining, grinding, and extraction of ores. However, many materials need specific grades and mixtures of metals and other elements, and each of these demands separation and purification at the atomic and molecular levels in the recycling process. Some ingredients cannot be separated and are lost from further use by dissipation.

By 2012, humankind had mined ~560 million metric tons of copper (19 million in 2010 alone) (*7*, *8*), and about half of this is still in use. But where is the other half, and what can researchers learn about the CE from this loss? Some irrevocable losses result from low recollection rates. Other losses occur when recycled materials can only be reused by mixing with virgin metals so as to meet the desired quality (i.e., high grades of steel or aluminum). Some products have a long lifetime, such as those used in the human-made environment (the "built world"), e.g., metals and polymers. As societal demand for products increases, manufacturers will need more virgin resources, because, for the foreseeable future, demand cannot be met entirely by a CE. Scientists have begun to exploit the vast quantities of accumulated wastes, including those that are metal-rich (e.g., from electronics and steel). But along with these wastes come other unwanted wastes and added energy demands.

Unlike metals, organic compounds can be synthesized. However, their recycling can require added chemicals and be energy-intensive and expensive. Plastics, for example, often consist of one or more (co) polymers along with several additives, such as plasticizers, flame retardants, coloring agents, ultraviolet-light stabilizers, and antioxidants, which are often hazardous and difficult to separate in recycling. Even when recycling pure materials, such as polyethylene terephthalate from plastic water bottles, more than 5% is lost in the process. Often recycling is only possible at the atomic or molecular level (e.g., depolymerization of plastics or extraction of metals and other elements from electronic materials), which results in the loss of macroscopic form, molecular structure, and desirable properties of the products. In some cases, recycling is impossible. Such materials and products must be avoided in a CE.

The use of renewable bioresources for organic molecules has a long history and is increasingly practiced in industrial chemistry. Agro-industrial or forestry waste products, for example, are complex and difficult to separate. Thus, the challenge lies in transforming waste into specific products with defined properties and degrees of complexity. Additional complications arise from the buildup of unwanted chemicals in products during use, natural aging, and the recycling process itself, as well as the connected flows of materials and products. For example, mechanical recycling (remolding) results in polluted and lower-quality recyclates, thus preventing reuse in the same application.

In fact, recycling unavoidably leads to the downgrading of materials and dissipative losses of utilizable elements and energy. The complexity and diversity of today's products and their components (on atomic, molecular,

[1]Institute of Sustainable and Environmental Chemistry, Leuphana University Lüneburg, Lüneburg, Germany. [2]International Sustainable Chemistry Collaborative Centre (ISC3), (ISC3), Bonn, Germany. [3]Green Chemistry Centre of Excellence, University of York, York, UK. [4]Department of Chemistry, Federal University of São Carlos, São Paulo, Brazil. Email: klaus.kuemmerer@uni.leuphana.de

material, and building-block levels) enlarge this dissipation of metals and other elements, pollutants, and energy.

These losses are apparent in an observed increase of thermal and material entropy. Whereas Earth can export thermal entropy to a certain degree by emission into space, it cannot get rid of the material entropy that comes from the mixing and diluting of materials, including lowering the concentrations of the elements therein. These insurmountable limitations (according to the laws of thermodynamics) must be taken into account and considered for chemistry, including metallurgy and materials science and its products necessary for a CE.

Products with open-environmental applications, such as pesticides, cosmetics, biocides, and pharmaceuticals, can be neither circulated nor recycled, because low concentrations and high dispersion on application make recollection impossible. For liquid waste and waste water, even advanced treatment removes product components poorly (*9*). In addition, nearly 80% of the world's effluent is not treated at all (*9*). Furthermore, environmental wastes and pollutants often are the sources of new, sometimes even more toxic, molecules. Therefore, scientists must design new molecules and materials for such applications, to permit fast and complete environmental mineralization while retaining their desired functions (*10*).

SIMPLIFYING COMPLEXITY

A series of chemistry keystones lie at the center of a CE and must be introduced into education, legislation, and industry (see the box). Most of today's chemical products are synthetic, based on nonrenewable resources, and formed into complex articles such as plastics. Recovering molecular value from these will require a considerable investment in funding and energy. Future products must constrain the levels of complexity of their constituent resources and not change them in recycling.

Keeping the chemical structure of the final product and its building blocks (form, composition, stereochemistry, and functional groups) as similar as possible to the starting material is desirable for both virgin and recycled materials. This will aid subsequent recycling and reduce the number of downstream chemical processes and unit operations, as these cause waste production, entropy increase, and resource consumption, including energy (*11*). If materials and molecules must be changed in form and composition, manufacturers should aim to further reduce complexity. Final products should be as simple in composition as possible, minimizing additives and avoiding toxic components and elements not easily separated for recovery.

Integrating chemistry into a circular economy

- **Keep molecular complexity to the minimum** required for the desired performance, including end of life (complex molecules require more synthesis steps, may have additional undesirable properties, and can be more difficult to recycle).
- **Design products for recycling,** including all additives and other components of the product.
- **Reduce and simplify diversity and dynamics** of substance, material, and product flows; e.g., use fewer chemicals overall (both number and quantity), design for less resource intensity, and adapt innovation speed of products to adaptation speed of recycling.
- **Avoid complex products** (e.g., multiple components, materials).
- **Minimize use of product components** that cannot easily be separated and recycled (e.g., solvents, metals).
- **Design products** not suitable for capture and recycling for complete fast mineralization at the end of their lives (e.g., pharmaceuticals, pesticides, personal care and cleaning products).
- **Prevent raw materials** from becoming critical through reduced use and efficient recovery and recycling (e.g., many metals).
- **Avoid entropic losses and transfers** (e.g., dissipation of metals, energy).
- **Avoid rebound effects** (e.g., using less carbon often means higher demand for metals).
- **Be responsible for/develop ownership of your product** throughout its complete life cycle, including recycling.
- **Ensure traceability** and consider use of product digital passports (e.g., composition of products, components, and processes).
- **Develop and apply circular metrics** (e.g., giving credit to the use of by-products).
- **Change traditional chemical practices** based on "bigger-faster" into "optimal adapted-better-safer" and change ownership to rent, lease, and share business models.
- **Keep processes as simple as possible** with a minimum number of steps, auxiliaries, energy, and unit operations (e.g., separations, purification).
- **Design processes for optimal material recovery** of auxiliaries, unused substrates, and unintended by-products (based on quality and quantity).

To avoid mixing varying constituents and increasing unnecessary complexity, flows of products and their constituents must be kept as separate as possible at all stages of their life cycle, from resource extraction and synthesis to use and recycling. If the flows are highly variable, recycling will be hindered. Similarly, local flows are preferable to global ones, and more complete knowledge of the composition of materials and products will help with speedy recycling. Of course, it is necessary to locate waste streams and decipher how to collect them, or they might be lost forever.

If speed of innovation on the product side is higher than on the recycling side, then the problem will outpace the solution. Society needs a greater investment in cutting-edge recycling technologies. Industry and academia must give these advances the same recognition as that received by new molecules, materials, and product designs. Making companies as responsible for recycling of their products as they are for product performance should drive industry participation in research and development (R&D) of recycling technologies and in product design for a CE. The high value attracted by some recycled products and the increasing costs of virgin resources should transform recycling and product-design research from a burden to a business opportunity.

Integration of the waste and chemical sectors, including industries that produce large volumes of chemical or metal-rich wastes (such as agri-food, textiles, electronics, plastics, metals, and alloys), allows for a better understanding of chemistry and its products and links sources with sinks. Demanding only a "green" CE without reflecting on the systemic role of chemistry will not contribute much to sustainability. The chemistry necessary for a CE will come to fruition only through a new attitude toward chemistry education, chemical research and engineering, and product design (*12*, *13*). ■

REFERENCES AND NOTES

1. European Commission, Implementation of the Circular Economy Action Plan (2019).
2. U.N. Environment Programme, Global Chemicals Outlook II (2019).
3. J. H. Clark, *Green Chem.* **21**, 1168 (2019).
4. J. A. Mathews, H. Tan, *Nature* **531**, 440 (2016).
5. J. de Boer, H. M. Stapleton, *Science* **364**, 231 (2019).
6. V. G. Zuin, *Curr. Opin. Green Sustain. Chem* **2**, 40 (2016).
7. R. K. Valenta, D. Kemp, J. R. Owen, G. D. Corder, É. Lèbre, *J. Clean. Prod.* **220**, 816 (2019).
8. B. W. Schipper *et al*., *Resour. Conserv. Recycling* **132**, 28 (2018).
9. K. Kümmerer, D. D. Dionysiou, O. Olsson, D. Fatta-Kassinos, *Science* **361**, 222 (2018).
10. K. Kümmerer, *Sustain. Chem. Pharm.* **12**, 100136 (2019).
11. J. H. Clark, *Nat. Chem.* **1**, 12 (2009).
12. P. G. Mahaffy, S. A. Matlin, T. A. Holme, J. MacKellar, *Nat. Sustain.* **2**, 362 (2019).
13. K. Kümmerer, *Angew. Chem. Int. Ed.* **56**, 16420 (2017).

10.1126/science.aba4979

Forensic investigators compare a photograph of Mengele with a model of a skull in April 1986.

PHOTO: DAVID G. MARWELL

BOOKS *et al.*

HISTORY OF MEDICINE

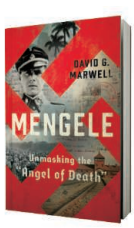
A notorious Nazi, revealed

A U.S. Department of Justice insider's biography reveals new details about Josef Mengele

By **Patricia Heberer Rice**

It is perhaps the most darkly iconic image of the Holocaust. A transport of Jewish prisoners arrives at Auschwitz and is ordered to divide by gender and form ranks of five. Amid the prisoners stands a man in a white coat, directing victims either to forced labor or to the gas chamber. This physician is Josef Mengele, nicknamed the "Angel of Death" for his cold demeanor on the ramp. Since his disappearance during the Cold War, he has come to personify the "Nazi doctor," a manifestation of evil that emerged as the surviving symbol of Nazi genocide.

Drawing from new sources and scholarship, historian David Marwell has written a compelling work that dispels many of the myths obscuring the identity of the infamous physician. *Mengele: Unmasking the "Angel of Death"* is at once a compact biography of the notorious war criminal, a detailed account of Mengele's flight to South America, and an absorbing narrative of the quest to bring him to justice. Marwell is at his best and most fascinating when he is separating the historical Mengele from the conjecture and half-truths that surround him.



Mengele: Unmasking the "Angel of Death"
David G. Marwell
Norton, 2020. 456 pp.

Josef Mengele was born in 1911 in Günzburg, Bavaria, the son of a prosperous manufacturer. In 1937, he received an appointment as a physician and the following year defended his dissertation in medicine. Mengele later painted himself as a virulent anti-Semite whose engagement with Nazi ideology began quite early. However, Marwell argues convincingly that Mengele's upbringing and early political orientation were conservative, Catholic, and imbued not with the Nazis' racialist anti-Semitism but with the latent cultural anti-Semitism of his milieu. He joined the Nazi Party in 1937 and the SS the next year. While in medical school, he became a true believer in the constructs of racial hygiene prevalent in Nazi Germany.

In 1942, Mengele served as a medical officer with the 5th SS Panzer Division "Wiking." Two important details emerge here from Marwell's research. First, we learn that Mengele's unit engaged in vicious atrocities against Jews in Ukraine. Although it is uncertain whether Mengele participated in these activities, it is clear that he witnessed extreme violence against civilians and was inured to it. Second, Marwell convincingly argues that it was unlikely that Mengele was wounded in January 1943. It is frequently suggested that an injured Mengele came to Auschwitz in May 1943 in a routine transfer. However, Marwell surmises that Mengele applied for the position, encouraged by his mentor Otmar von Verschuer, a leading scientist known for his genetic research with twins.

Mengele began his career at Auschwitz as the medical officer responsible for Birkenau's Gypsy camp and, following its liquidation in November 1943, undertook a new position as chief camp physician of Auschwitz II (Birkenau). "If Auschwitz … stands as a symbol of the Holocaust," writes Marwell, "then Mengele, as perpetrator, has come to serve a similar role for the death camp itself."

Stories of Mengele's inhumane medical experimentation are legion, but his propensity for twin research is often misinterpreted. Although common wisdom suggests that Mengele's preoccupation with twins reflected a desire to increase the German birth rate, Marwell shows that twin research was considered, at the time, the gold standard for the exploration of the genetics of disease. Mengele, Marwell maintains, was not a "mad scientist," as he is often depicted. Despite the immoral and often lethal nature of his experiments, his methodology was consistent with that of other researchers in the scientific establishment, and Mengele himself was firmly entrenched in Germany's mainstream medical community.

Of course, Mengele's chief crime lay in his work on the ramp. He is associated more closely with "selection duty" than any other medical officer at Auschwitz, although he performed this task no more often than did any of his colleagues. The pervasive image of Mengele at the ramp in so many survivor accounts has to do not only with his postwar notoriety but also with the fact that Mengele often appeared at selections while off duty, searching for twins.

After the war, Mengele evaded capture and, with the aid of his prosperous family, fled to South America. In February 1979, he suffered a stroke and drowned while swimming near Bertioga, Brazil, and was buried under the fictive name Wolfgang Gerhard. The last chapters of Marwell's book provide a page-turning account of the hunt for Mengele, ending in the exhumation of his body in 1985. As a U.S. Department of Justice official, Marwell describes his own participation in the forensic examination, which confirmed that the infamous Mengele had at last been found. ■

The reviewer is at the Jack, Joseph and Morton Mandel Center for Advanced Holocaust Studies, United States Holocaust Memorial Museum, Washington, DC 20024, USA. Email: pheberer-rice@ushmm.org

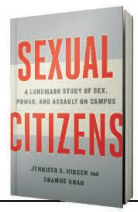
10.1126/science.aba0950

PUBLIC HEALTH

Confronting campus sexual assault

Inadequate sex education and socialization collide in built spaces that stymie consent

Sexual Citizens: A Landmark Study of Sex, Power, and Assault on Campus
Jennifer S. Hirsch and Shamus Khan
Norton, 2020. 432 pp.



By **Claire M. Renzetti**

Although there are variations in format and content, there are also remarkable similarities in orientation programs designed to introduce freshmen to college life. All programs instruct students on practical matters such as course scheduling, academic advising, and choosing a major. But most programs also have sessions on navigating campus life outside the classroom, with at least one session addressing issues related to healthy relationships.

According to Jennifer Hirsch and Shamus Khan, the authors of *Sexual Citizens*, these sessions help students become "acutely attuned to the importance of consent." Nevertheless, as they convincingly demonstrate in this profoundly eye-opening book, in their everyday social interactions, many students rely on a variety of tacit signals and contextual cues to discern consent. In short, despite students' ability to recite verbatim the legal standard of affirmative consent, their behavior belies their knowledge.

Sexual Citizens is one of the products of a 5-year study of undergraduates at Columbia University called the Sexual Health Initiative to Foster Transformation (SHIFT). The goal of SHIFT was to identify the social causes of campus sexual assault in order to develop more effective prevention strategies. This mixed-methods project, involving nearly 30 researchers from multiple disciplines, included a population-based survey, a 60-day daily diary study, 17 focus groups, multihour interviews, and hundreds of hours of ethnographic field observation of students socializing in various settings.

Throughout the book, we hear from a broad range of student participants. Some of their stories are harrowing, others are heartbreaking. But Hirsch and Khan succeed in relating all of the students' accounts with empathy rather than moralistic judgment, and in doing so, they allow readers to recognize similarities to their own students, or their children, or perhaps even their younger selves.

Hirsch and Khan present a novel model for explaining and responding to campus sexual assault. At its crux are three concepts: sexual projects, sexual citizenship, and sexual geographies.

Sexual projects are the reasons why people seek particular sexual experiences. As Hirsch and Khan point out, many of the students studied were "figuring out their sexual projects through trial and error," primarily because no one had talked to them much about sex, except to condemn it as bad or to scare them into not having it. The negative consequences for these students, they argue, derive largely from their lack of comprehensive sex education and the unwillingness of most parents to have conversations with their children that "conveyed that sex would be an important and potentially joyful part of their life, and so they should think about what they wanted from sex, and how to realize those desires with other people in a respectful way."

Dorm rooms in which beds are the only seating option are part of the problem.

The concept of sexual citizenship refers to one's own and others' equivalent right to sexual self-determination. In discussing this concept, Hirsch and Khan explore the reasons why "some people feel entitled to others' bodies, and others do not feel entitled to their own bodies." Again, students' lack of comprehensive sex education and sexual socialization is a critical explanatory factor: "...sexual illiteracy reflects a denial of young people's sexual citizenship, with a climate of shame and silence that are part of the social context of campus sexual assault."

Hirsch and Khan use the concept of sexual geographies to integrate the role that built environments play in students' behavior and interactions. Think, for example, of the typical place on campus where a student couple can hang out with a modicum of privacy: a cramped dorm room where the most comfortable place to sit is the twin bed. As Hirsch and Khan demonstrate throughout *Sexual Citizens*, "These spatial dynamics—control, access, feeling at ease—are major players in sexual assault."

Hirsch and Khan present the familiar statistics on campus sexual assault and also spend considerable time addressing one of the well-known risk factors: alcohol consumption, particularly binge drinking. But the authors' three-pronged explanatory model is far more valuable in that it redirects our attention from individual bad actors to focus instead on the social roots of campus sexual assault. In the concluding chapter of *Sexual Citizens*, Hirsch and Kahn deftly draw on several recent public health campaigns to underline the possibility of successful interventions through collective action. Here, they make recommendations to parents, politicians, policymakers, religious leaders, educators, university administrators, and community members more broadly.

If we wish to raise children for whom the risk of sexual assault is significantly reduced, we must all share responsibility for eliminating confusion around sexual projects, bringing clarity to the right to sexual citizenship, and creating sexual geographies that reduce power inequalities. ■

The reviewer is at the Center for Research on Violence Against Women and the Department of Sociology, University of Kentucky, Lexington, KY 40506, USA, and editor of the journal *Violence Against Women* (Sage Publications). Email: claire.renzetti@uky.edu

10.1126/science.aba1150

Giant panda Bei Bei spends a final day in Washington, D.C., before departing for China.

PHOTO: XINHUA/ALAMY STOCK PHOTO

Edited by **Jennifer Sills**

Time to update China's panda loan terms

In November, 2019, 4-year-old panda Bei Bei, born on American soil to parents on loan from China, departed the National Zoo in Washington, D.C. (*1*). China's loan arrangements specify that all panda cubs born abroad are the property of China and must be sent back to China for scientific and reproduction purposes (*2*). The 10-year loan agreements, which cost the host country a rental fee of US$1 million, have been in place since the mid-1980s and currently cover about 50 pandas in captivity outside China (*3*). However, the current terms of the agreements are not in the best interests of the pandas or panda research.

When the loans were originally created in 1982, China's panda population, then at about 1100 individuals (*4*), was dwindling. At that time, it was important to augment the population with new panda cubs. However, the population has since grown, recently reaching 2300 individuals (*5*). Because the local panda gene pool is now large enough to enable breeding, it is no longer necessary to send all panda cubs back to China.

If returned panda cubs were released to their original wild habitat in China and scientifically monitored remotely, then their ecological value might justify uprooting them from their place of birth. However, upon their return to China, the cubs remain in captivity (*6*). Given the challenges that the repatriated cubs face, including difficult adjustments to a new climate, feeding regimen, and breeding culture (*7*–*9*), returning them to China may not be to their benefit.

The current loan arrangements should be updated to ensure that they are furthering the goals of species protection, scientific research, and cultural exchanges. The length of the loan period should be more flexible, with extensions available. Cubs should be returned to China only if it is in their best interest. Overseas hosts should have the opportunity to conduct observation and breeding projects for cubs born on their soil. This would require effective cooperation between Chinese panda researchers and their overseas counterparts at the host zoos. The financial revenue collected through the loans—which sometimes seems to serve as a commercial benefit for Chinese zoo and attractions developers (*10*) rather than a self-sustaining investment in species conservation—should be better monitored and allocated instead to support and research programs for the returned pandas. Finally, China's wildlife authorities should develop a more supportive regulatory framework of panda protection and research at a global scale, as well as promote international cooperation and long-term eco-cultural exchanges through the loan arrangements.

Qi Yan[1]*, Yunhong Hu[2], Haobin Ben Ye[3]
[1]School of Tourism and Social Management, Nanjing Xiaozhuang University, Nanjing 210017, China. [2]College of Landscape Architecture, Nanjing Forestry University, Nanjing 210046, China. [3]School of Tourism Management, South China Normal University, Guangzhou 511400, China.
*Corresponding author. Email: yorknjcn@msn.com

REFERENCES AND NOTES

1. C. Reid, "Bei Bei the giant panda leaving the National Zoo for China," *CBS News* (2019).
2. T. Waldrop, "Bye bye, Bei Bei: Beloved giant panda headed from DC to China in a private jet" *CNN* (2019).
3. X. Xi, "Pandas are expensive to loan and difficult to raise" *Sina.com* (2013); http://view.163.com/special/reviews/giantpanda1017.html [in Chinese].
4. J. Hu, Z. Zhang, R. Wei, *Acta Theriologica Sinica* **31**, 10 (2011) [in Chinese].
5. Z. Jiang *et al.*, *Science* **353**, 657 (2016).
6. C. J. Loucks *et al.*, *Science* **294**, 1465 (2001).
7. J. Liu *et al.*, *Science* **317**, 1513 (2007).
8. B. Yang *et al.*, *Science* **339**, 521 (2003).
9. S. Zhao *et al.*, *Nat. Genet.* **45**, 67 (2013).
10. F. Wei, Z. Zhang, J. Hu, *Acta Theriologica Sinica* **31**, 412 (2011) [in Chinese].

10.1126/science.aba3948

China's dams isolate Asian elephants

The Asian elephant (*Elephas maximus* L.) is designated as a grade-I protected species in China and listed as an endangered species by the International Union for Conservation of Nature (IUCN) (*1*). There are only about 300 wild individuals remaining in China (*2*). Despite the species' recognized vulnerability, China has exacerbated the threats to its survival by failing to consider the cascading effects of dam construction.

The Mekong upstream district (named Lancang River in China) (*3*), a major Asian elephant habitat (*4*), is also abundant in water resources and well suited to the construction of hydropower stations (*5*). Accordingly, China began construction on the Jinghong Hydropower Station in 2003 as part of a renewable energy plan (*6*). The environmental impact assessment of the project did not comprehensively describe how it would affect Asian elephants, and the station went into operation in 2008 on schedule (*7*).

Landscape connectivity among habitats and protected areas is crucial for conservation of wildlife, especially endangered flagship species such as Asian elephants, which require a large home range covering a variety of ecosystems (*8*). After the Jinghong Hydropower Station dam was completed, water levels rose and widened, making the mud banks more wet and slippery on both sides (*7*). Although Asian elephants could get down the banks to the river and swim across, their flat soles prevented them from climbing back out (*9*), stranding them in the water. As a result, no elephants have crossed the Mekong in the past decade, and movement routes and gene flow of Asian elephants living on either side of the Mekong have been blocked by the reservoir (*7*).

The restricted habitat has been particularly challenging for the Menghai-Lancang elephant population, which includes 18 of China's 300 elephants. This small subpopulation, which once lived in the Xishuangbanna National Nature Reserve located on the east bank of the Mekong, migrated to the west bank before the dam was completed and now can't return (*10*). Trapped on the west bank, they move freely among villages, farmlands, and fragmented forests, causing serious human-elephant conflicts. Between 2011 and 2019, these elephants killed 27 people and injured more than 50 (*11*). In addition, the herd has ruined crops and destroyed houses, causing US$3 million in economic losses (*11*).

The negative effects of elephants isolated by dams are enormous and long-lasting. The

Chinese government should take appropriate measures to minimize these impacts, such as establishing corridors for Asian elephants to cross the Mekong and increasing habitat connectivity and opportunities for gene exchange. To protect the isolated Menghai-Lancang Asian elephants and prevent future human-elephant conflicts, the forests and scrub-grasslands in Lancang county and Menghai county, which are suitable as elephant habitats, should be planned and constructed as conservation areas.

Zhihong Wang, Zhengling Li, Yongjing Tang, Chongxue Yao, Yu Liu, Guilian Jiang, Fang Wang, Liang Liang, Wenlan Zhao, Gaofan Zhu, Mingyong Chen*
Yunnan Asian Elephant Field Scientific Observation and Research Station of the Ministry of Education, Kunming 650091, China and Asian Elephant Research Center of Yunnan University, Kunming 650091, China.
*Corresponding author.
E-mail: mychen1108@ynu.edu.cn

REFERENCES AND NOTES

1. W. W. Li *et al.*, *Glob. Ecol. Conserv.* **19**, e00664 (2019).
2. Y. Zhao *et al.*, *World For. Res.* **31**, 2 (2018) [in Chinese].
3. Z. F. Xu *et al.*, *Acta Phytoecologica Sinica* **28**, 585 (2004).
4. L. Lin *et al.*, *Acta Theriologica Sinica* **35**, 1 (2015) [in Chinese].
5. J. T. Liu *et al.*, *Water Power* **34**, 28 (2008) [in Chinese].
6. M. J. Tian *et al.*, *Hongshui River* **28**, 10 (2009) [in Chinese].
7. Z. B. Yang, Ed., "Assessment to the impacts on terrestrial ecosystems in the reservoir area caused by Lancangjiang-Jinghong Hydropower Station during the construction phase" (Yunnan Science and Technology Press, Kunming, 5, 2013) [in Chinese].
8. W. Suksavate *et al.*, *Glob. Ecol. Conserv.* **19**, e00685 (2019).
9. M. Y. Chen, Ed., "Wild elephant in China" (Yunnan Science and Technology Press, Kunming, 7, 2008) [in Chinese].
10. Q. Y. Wang *et al.*, *For. Inventory Plan.* **43**, 30 (2018) [in Chinese].
11. G. Z. Sun *et al.*, *For. Construct.* **210**, 1 (2019) [in Chinese].

10.1126/science.aba5991

Conservationists must address meat and dairy

In October 2019, more than 300 conservation experts at the International Union for Conservation of Nature Species Survival Commission Leaders' Meeting in Abu Dhabi published an urgent call for action to address the impacts of human activities on wildlife (*1*). However, the letter did not explicitly address meat and dairy consumption. The meat-eating habits of humans drive climate change, deforestation, and pollution of both terrestrial and aquatic ecosystems (*2*–*5*). Conservationists, as authorities on environmental science, have credibility with the general public (*6*). It is our responsibility to use our expertise and influence to explicitly address the pressing issue of meat and dairy production.

We plead with conservationists to engage in the anti-meat movement, today carried out primarily by vegan activists and animal rights organizations. All conservation projects should include an educational program that raises awareness about the irreversible impacts of meat and dairy consumption on biodiversity. Our consumption behaviors must change if we are to save what is left of wildlife. Conservationists should set a positive example and lead the debate.

Luis Fernando Marin da Fonte[1]*, Ana Paula Foletto Marin[2]
[1]Trier University, Trier, Germany. [2]London, UK.
*Corresponding author. Email: pulchella@gmail.com

REFERENCES AND NOTES

1. IUCN, "The Abu Dhabi call for global species conservation action" (2019). www.iucn.org/species/about/species-survival-commission/ssc-leadership-and-steering-committee/ssc-leaders-meeting-2019/abu-dhabi-call-global-species-conservation-action.
2. B. Machovina, K. J. Feeley, W. J. Ripple, *Sci. Total Environ.* **536**, 419 (2015).
3. J. Poore, T. Nemecek, *Science* **360**, 987 (2018).
4. H. C. J. Godfray *et al.*, Science **361**, eaam5324 (2018).
5. M. Springmann *et al.*, *Nature* **562**, 519 (2018).
6. C. Funk, M. Hefferon, B. Kennedy, C. Johnson, "Trust and mistrust in Americans' views of scientific experts" (2019); www.pewresearch.org/science/2019/08/02/trust-and-mistrust-in-americans-views-of-scientific-experts/.

10.1126/science.aba5867

Science Webinars

AAAS

Science Webinars help you keep pace with emerging scientific fields!

Stay informed about scientific breakthroughs and discoveries.

Gain insights into current research from top scientists.

Take the opportunity to ask questions during live broadcasts.

Get alerts about upcoming free webinars.

Sign up at: webinar.sciencemag.org/stayinformed

CALL FOR PAPERS

SPJ
SCIENCE PARTNER JOURNALS

SIBET CAS

spj.sciencemag.org/bmef

BME Frontiers

OPEN ACCESS

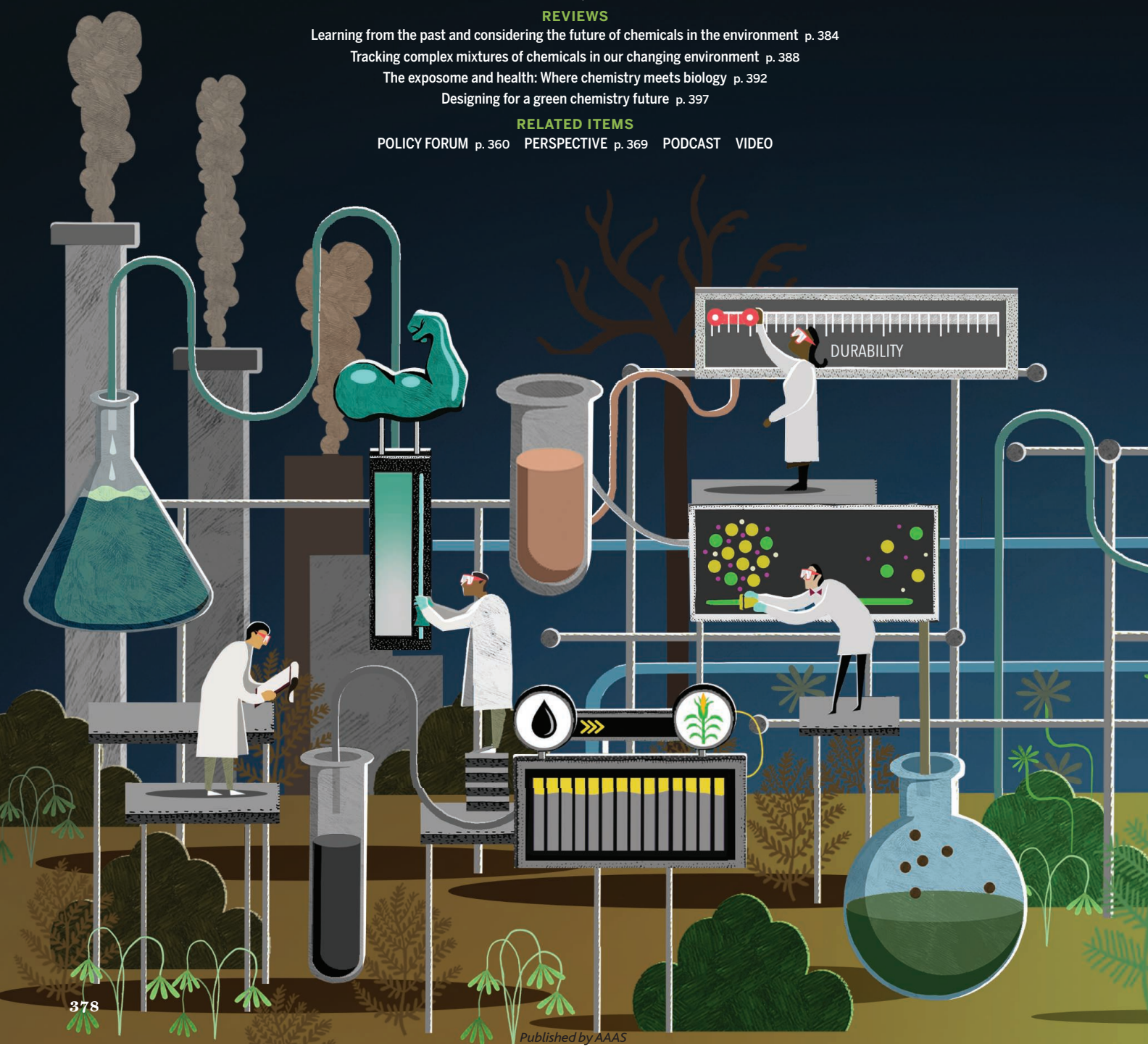
Biomedical Engineering (BME) Frontiers is a **Science Partner Journal** distributed by the **American Association for the Advancement of Science** (AAAS) in collaboration with the **Suzhou Institute of Biomedical Engineering and Technology, Chinese Academy of Sciences** (SIBET CAS). *BME Frontiers* aims to serve as an effective platform for the multidisciplinary community of biomedical engineering. The journal will publish breakthrough research in the fields of pathogenic mechanisms as well as disease prevention, diagnosis, treatment, and assessment.

The Science Partner Journals (SPJ) program was established by the American Association for the Advancement of Science (AAAS), the nonprofit publisher of the *Science* family of journals. The SPJ program features high-quality, online-only, open access publications produced in collaboration with international research institutions, foundations, funders and societies. Through these collaborations, AAAS expands its efforts to communicate science broadly and for the benefit of all people by providing top-tier international research organizations with the technology, visibility and publishing expertise that AAAS is uniquely positioned to **offer as the world's largest general science membership society**.

Submit your research to *Biomedical Engineering Frontiers* today!
Learn more at: **spj.sciencemag.org/bmef**

ARTICLE PROCESSING CHARGES WAIVED UNTIL 2021

A CLEANER, GREENER FUTURE FOR CHEMICALS



Published by AAAS

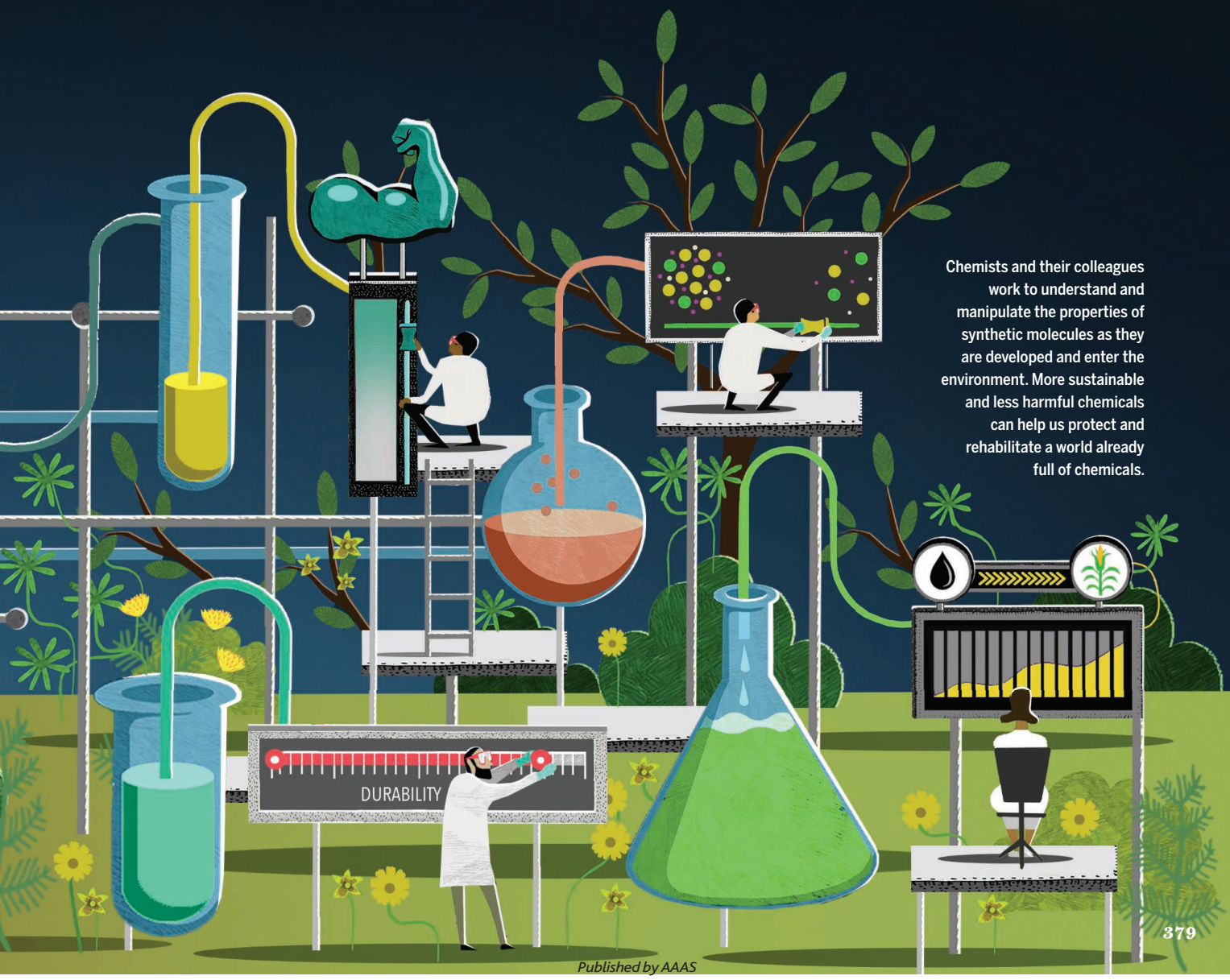
By **Michael Funk** *and* **Caroline Ash**

Since the Industrial Revolution, developments in chemistry have transformed entire sectors of the global economy, often providing great benefits to society and quality of life. But the production, mass distribution, and disposal of increasingly complex and persistent chemical products have resulted in many cases of ecological and environmental damage. Harmful effects for people are often concentrated in those communities least able to avoid exposure, and concern over unfamiliar chemicals in our food, water, and homes is widespread. How can we continue to develop molecules that address today's challenges while ensuring that we understand the effects of complex and ubiquitous chemicals on our health and the environment?

On a global level, reducing greenhouse gas emissions and fossil-derived raw inputs is imperative to achieve a sustainable future. New chemical transformations are necessary to supplement or supplant many of those we rely on currently. We must also reckon with the fate of the myriad complex chemicals used in agriculture, consumer products, drugs, and materials. We now have many sophisticated tools to understand what happens to these molecules as they are released into the environment. Our health, and that of ecosystems around the world, depends on our commitment to gathering this information and taking action accordingly.

Julia Fahrenkamp-Uppenbrink conceived this special issue.

ILLUSTRATION: ADAM SIMPSON



Chemists and their colleagues work to understand and manipulate the properties of synthetic molecules as they are developed and enter the environment. More sustainable and less harmful chemicals can help us protect and rehabilitate a world already full of chemicals.

Published by AAAS

CAN DO

To forge a safer lining for food containers, a chemical company teams up with unlikely allies

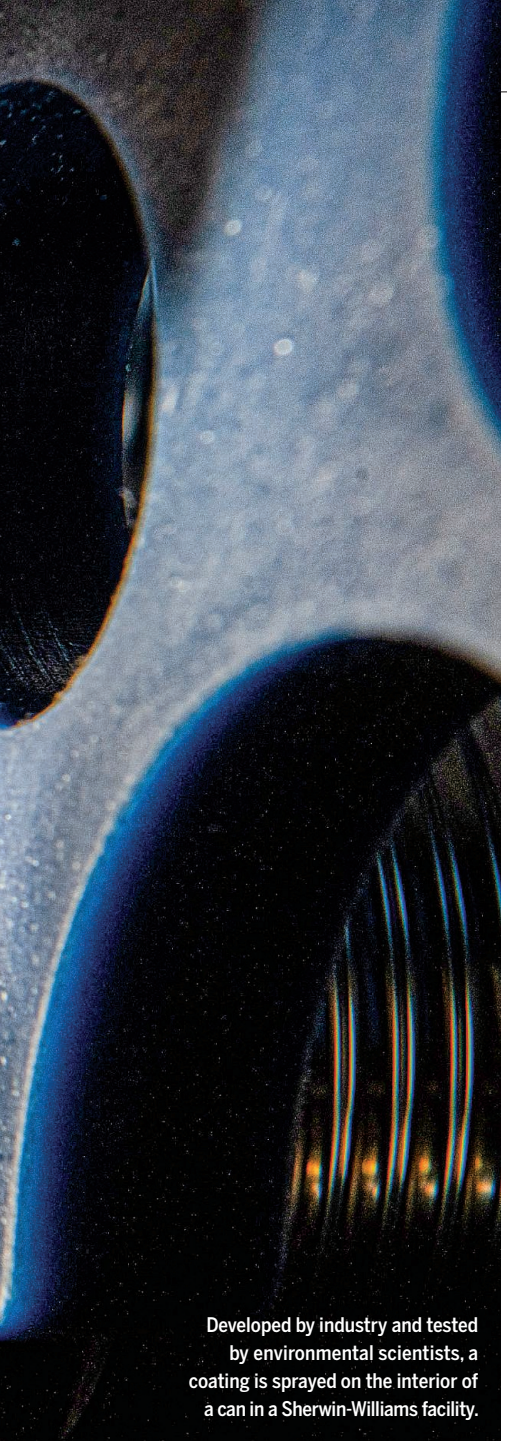
By **Warren Cornwall**, *in Pittsburgh; Photography by* **Rob Larson**

A soda can might seem unworthy of scientific research. But just ask chemist Tom Mallen. With his back to a laboratory table, he holds up a silvery disk and points to the spot where a metal pull tab is pinned to the middle. That spot is one of the toughest tests for the thin plastic layer coating the insides of cans. Enormous strain is placed on the lining as a machine pounds the metal lid to make a small bump where the pull tab attaches, and then squashes the bump to pin the tab in place. That makes it "the most difficult fabrication in this whole universe," he declares.

Designing coatings that can withstand this rough treatment—and meet a host of other requirements—has defined much of Mallen's 31-year career. Now, he and other employees of Sherwin-Williams, best known for selling paint, are trying a new way to develop such industrial chemicals.

The plastic lining of a can is just 2 micrometers thick, less than one-eighth the thickness of a human hair. A ubiquitous yet invisible part of the modern world, can coatings must stand up to the rigors of manufacturing and then last for years in baths as acidic as lemon juice while preserving a seamless barrier between the food or drink and the metal. Any crack in that layer can mean corrosion or a weird metallic taste. Good performance ensures a 3-year-old can of Coke will have the same fizz and flavor as one bought yesterday.

Those properties can come at a cost, however. The coating is often made from bisphenol A (BPA), a chemical that has gained notoriety because of evidence that it can disrupt the dance of hormones that influence growth and development. But replacing BPA isn't simple. Companies produced more than 6 million tons in 2018, making it one of the world's most common synthetic chemicals. It is cheap, durable, and flexible—



Developed by industry and tested by environmental scientists, a coating is sprayed on the interior of a can in a Sherwin-Williams facility.

PHOTO: ROB LARSON

a key building block not just in can linings, but also in products as varied as automobile dashboards and sales receipts.

A decade ago, in a nondescript two-story building on a dead-end street in a scruffy industrial neighborhood here, Mallen and colleagues embarked on a quest to find a molecule that could do the near-miraculous things BPA can do without the downsides. In place of the usual corporate research initiative shrouded in secrecy, the company tried something almost unheard of among chemical companies: It sought scrutiny from some of the industry's fiercest critics.

"I haven't seen another company take that approach," says Tom Neltner, chemicals policy director for the Environmental Defense Fund, who met twice with company officials.

As a result of the effort, the manufacturer and prominent university scientists known for criticizing BPA have joined forces to check for possible health effects of a candidate chemical. The endeavor has produced a promising molecule while earning tentative praise from environmental advocacy groups, and some point to it as a model for how companies might uncover safer chemicals. But although Sherwin-Williams's molecule is already in production, it must also win over price-sensitive can manufacturers and exacting food companies. And no one knows whether it will prove truly benign.

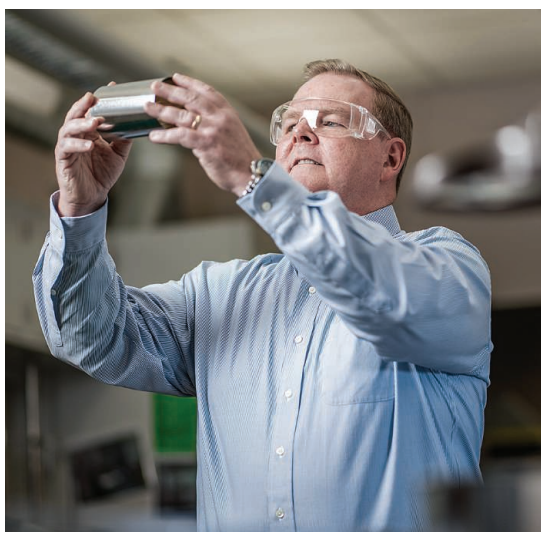
MALLEN HARDLY HAS THE MAKINGS of an iconoclast. The 59-year-old has spent his career at the paint and chemical company Valspar, which Sherwin-Williams bought in 2017. Starting as a laboratory chemist, he rose to become a company vice president in charge of regulatory affairs for the division that makes can coatings. Baby-faced, with neatly trimmed blond hair, he speaks with the measured, matter-of-fact manner of the upstate New York farm boy that he is. But when company chemists approached him in 2009 with a plan to find a new can coating, he decided the company would need to break from tradition.

Starting in the late 1990s, scientists had sounded the alarm about BPA, particularly its ability to mimic the hormone estrogen. In animal and epidemiological studies, university researchers linked BPA to breast and prostate cancer, reduced fertility, diabetes, genital defects, and altered behavior. The Centers for Disease Control and Prevention found the chemical in the urine of 93% of U.S. adults.

Although the chemical industry and federal regulators have argued that the evidence doesn't prove the compound poses health risks at the levels found in people, many consumers now shun BPA. Today, it is banned from baby bottles and toddler cups in Canada, Europe, and the United States. In 2015, France banned BPA in all food containers. California now requires warning labels or signs for food and drink cans with BPA. And makers have scrambled for alternatives, which quickly made their way into products sporting "BPA free" labels.

But the alternatives have flaws. Some substitutes—often related chemicals in the bisphenol family—appear to have similar hormone-mimicking properties. Can linings made from other plastics have a shorter shelf life, work only for particular foods and drinks, or require more coatings in the factory—a major drawback on assembly lines churning out 2000 cans per minute.

Valspar was in the thick of the controversy. The company's can coating division relied heavily on BPA and struggled to find satisfactory replacements. Although many companies were running away from bisphenols, those compounds still promised to outperform other chemicals such as polyesters or acrylics. So Jeff Niederst, a chemist heading Valspar's efforts to develop linings, posed a question: What if Valspar could find a bisphenol that wasn't an endocrine disrupter?



Tom Mallen has helped lead Sherwin-Williams's efforts to replace the ubiquitous chemical bisphenol A.

The suggestion set off a blizzard of meetings. Valspar faced a multimillion-dollar, multiyear research investment marked with uncertainty. And if the company stuck with bisphenols, it would need to prove itself to a public skittish about that class of chemicals. To succeed, Mallen recommended the team open its work to outside inspection. "This whole idea of embracing transparency," Mallen recalls, "we were going to have to get very comfortable with that."

IN SEARCHING FOR A COATING that wouldn't disturb the endocrine system, Valspar found itself in an emerging area of toxicology. Existing chemical safety tests used by companies and government regulators often rely on decades-old methods that look for glaring effects such as changes in organ weight. That situation has prompted researchers to look for faster, more sensitive ways to flag hazardous chemicals and pinpoint safer ones. U.S. federal agencies, including the National Institutes of Health's National Toxicology Program and the Environmental Protection Agency (EPA), have

spent the past decade developing ways to quickly check thousands of chemicals by dosing cells cultured in petri dishes, a project called Tox21. The European Commission, meanwhile, is pouring €50 million into research to improve ways to identify endocrine disrupters.

But much of that work is still experimental, and none has produced a definitive set of endocrine tests. "I'm not sure we'll ever get there," says Bernard Robaire, a pharmacologist at McGill University who is part of a Canadian chemical-screening initiative.

So Valspar decided to go beyond the usual regulatory testing, cobbling together its own screening on the advice of scientists inside and outside the company. It hired Mark Maier, a toxicologist who had worked for a drug company and a pesticide industry group. He used a computer program to search hundreds of bisphenol compounds for molecular structures that appeared unlikely to bind with estrogen receptors.

Some bisphenols were prohibitively expensive. Others didn't react well to form a polymer, a long chain of molecules, that would make up the can lining. The top contenders were sent to a private lab for testing on yeast cells engineered to glow in response to chemicals that acted like estrogen or testosterone. The most promising was tetramethyl bisphenol F (TMBPF), a rarely used chemical previously tested as an insulator in electronic circuit boards.

Although it cost 50% more than BPA, TMBPF was comparable in price to other BPA substitutes, company officials say. And it shared some of BPA's toughness. TMBPF had another apparent advantage. Manufacturers use two chemical steps to forge BPA into can linings. In the second stage, some BPA molecules don't get incorporated into the polymer chains. Those strays are thought to account for most of the BPA that leaches from linings. Valspar chemists found a way to use TMBPF just once, early in the process, stanching the release of the unlinked molecules. The Food and Drug Administration and EPA gave Valspar the green light to use the new chemical in cans in 2014, after routine tests.

BEFORE VALSPAR put the chemical on the market, executives wanted to be sure it would also pass muster with people sounding the alarm about BPA. "We couldn't afford to spend a lot of time working on a wrong material," Mallen says. The company decided "to go to the expert critics" so that they, too, could scrutinize it.

With the help of a public relations firm, Burson-Marsteller (now BCW), Valspar crafted a plan to ask prominent BPA skeptics to study the chemical for endocrine effects. If the results were promising, Valspar would court environmental activists and journalists with a message of transparency.

Maier began by contacting Maricel Maffini, a Frederick, Maryland–based biologist who consulted for environmental groups on toxic chemical issues. The two had met in 2012 on a bus in Parma, Italy, at a food safety meeting. Maffini's reputation preceded her: "I had heard all these horrible stories," Maier recalls. "She was vilified by industry."

Maier, however, was impressed with her thoughtfulness, so he asked her to look at data on TMBPF. In 2014, the company hired her as a consultant. Maffini says she wrestled with how colleagues in the environmental world might view her work for industry. "I thought, 'OK, these are people that are trying to do something different. Just screaming into the wind is not going to lead us very far.'"

Maffini became a bridge. She had been a research assistant professor at the Tufts University lab of Ana Soto, a reproductive endocrinologist and leading scientist raising concerns about BPA. In 2015, Valspar and Maffini approached Soto about analyzing the chemical, and she agreed.

The company was hands-off. Rather than sign a contract with Soto, Valspar made a donation to Tufts with no implied obligation: She would study the chemical however she saw fit. "They didn't try to twist my arm," Soto says.

She put the molecule through tests. It didn't seep from coatings when soaked in acid or alcohol baths meant to simulate various foods. More important, it did not activate estrogen receptors in engineered breast cells or spur growth of breast cancer cells, she, Maier, Maffini, and colleagues reported in February 2017 in *Environmental Science & Technology*. The molecule also didn't affect several genes that respond to estrogen. "As far as estrogenicity, we are pretty sure that there is no activity, and we are happy with that," Soto says.

For the company, the finding was a coup—a well-known independent researcher had declared that the molecule showed no signs of sharing one of BPA's biggest flaws. When Mallen learned the results, he says, "I had quite a few scotches that night."

Estrogen mimicking wasn't the only concern, so Valspar sought out other experts. Thomas Zoeller, an endocrinologist at the University of Massachusetts, Amherst, and an authority on how BPA affects the thyroid, reviewed the company's test data and organ samples. The results suggest "it doesn't seem to affect the thyroid system," Zoeller says.

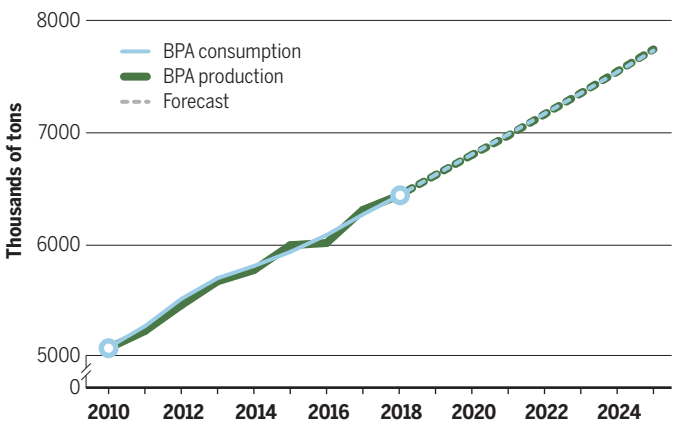
Outside researchers didn't deliver uniformly good news, however. Valspar approached scientists at Baylor Medical School, who found that TMBPF blunted estrogen's effect on test cells and that a polymer made from the molecule had a similar effect on testosterone, according to a 2017 paper in *PLOS ONE*. Adam Szafran, a molecular biologist who helped lead the research, says the findings weren't conclusive and could be specific to the prostate cells they tested.

Mallen acknowledges that those results raise questions about the compound. But he says company-sponsored research showed that changes in test cells don't translate into effects on an entire organism. That study, published online in *Food and Chemical Toxicology* in October 2019, showed no endocrine-related effects on rats fed TMBPF for 3 months.

THE CAUTIOUS PRAISE Sherwin-Williams is earning from chemical safety advocates suggests to Maier that other companies might benefit from emulating its approach. He points to the agribusiness giant Monsanto, now owned by Bayer, which has been hit with hundreds of millions of dollars in legal judgments over claims that its glyphosate herbicide caused cancers. "Monsanto wouldn't be in the trouble they

Defying the doubts

Despite concerns, worldwide production and consumption of bisphenol A (BPA) has increased, although separate figures for North America show a small decline.



CREDITS: (GRAPHIC) N. DESAI/*SCIENCE*; (DATA) WILLIAMS & MARSHALL STRATEGY

are now if they had taken a transparent, proactive, collaborative approach," says Maier, now working as a consultant in Albuquerque, New Mexico.

But some companies see downsides to the public spotlight, says Paul Anastas, a founder of the green chemistry movement at Yale University. Businesses quietly working to find safer replacement chemicals worry that attention could bring unwanted questions about the safety of current products. "The nail that sticks out gets the hammer," he says.

Nor have critics wholeheartedly embraced Sherwin-Williams's new coating. "It's a step forward," Neltner says. But, he adds, "We can't make a statement that this product is absolutely safe."

And Soto hesitates to go beyond her estrogen-related findings. "I can't tell you there isn't something else there because I haven't tested for it."

That statement highlights a challenge for Sherwin-Williams: How far does it need to go to prove itself? The endocrine system's complexity stymies easy assessment. The variety of hormones, their fluctuations throughout life, and their effects on different parts of the body create a kaleidoscope of ways a chemical might trigger effects. "It's very complicated," says Terry Collins, a chemist at Carnegie Mellon University. "You are really asking what God knows about how to build life."

Collins is part of a group of scientists who proposed a five-step approach in 2013 to look for endocrine disruption. The process starts with computer simulations and ends with extensive testing on rodents. TMBPF hasn't passed two of those tests yet: Will minuscule amounts cause problems, even though the much bigger doses used in standard toxicology tests don't? And will exposure for one generation of lab animals and their fetuses translate into illness for offspring?

To satisfy European regulators, Sherwin-Williams is planning tests that "will definitely answer at least 95% of the remaining questions," Mallen says. The tests will expose rats to TMBPF starting in utero, and potential U.S. and European academic collaborators could recommend additional tests piggybacking on the study. That new round of experiments will be "what a lot of people are waiting for," Mallen says.

Despite the remaining questions, the company has started to sell its coating in the United States, branded as valPure V70. It has been used in 22 billion cans since 2017. That's a modest fraction of the estimated 350 billion aluminum beverage cans and 100 billion steel food cans produced each year worldwide.

A large reaction vessel in a Sherwin-Williams laboratory contains the company's newly developed coating.

PHOTO: ROB LARSON

Winning more of the market is daunting. A fraction of a penny per container could sway canmakers, and BPA is cheaper. Though now used in just 10% of steel cans in the United States, BPA is still in roughly half of all aluminum cans, Mallen says.

And consumers show little sign of demanding any particular non-BPA product. Cans often look alike under the gloss of brand labels. Even Mallen can't tell which cans are treated with his company's product. At a lunch break in a company conference room, he turns a Diet Coke in his hand, looking for the tiny insignia of the company that made the can. "I don't think they're using V70," he says. "Maybe they are."

Sherwin-Williams hopes its molecule might become a universal can liner—like BPA without the drawbacks. Getting there will take more than chemical safety testing. While Mallen scrutinizes his soda can, researchers in nearby rooms run elaborate torture tests to perfect formulas that will work with as many foods and drinks as possible—olives, alcoholic ciders, sauerkraut, dog food. All that for the humble can. ■

REVIEW

Learning from the past and considering the future of chemicals in the environment

Andrew C. Johnson[1]*, Xiaowei Jin[2], Norihide Nakada[3], John P. Sumpter[4]

Knowledge of the hazards and associated risks from chemicals discharged to the environment has grown considerably over the past 40 years. This improving awareness stems from advances in our ability to measure chemicals at low environmental concentrations, recognition of a range of effects on organisms, and a worldwide growth in expertise. Environmental scientists and companies have learned from the experiences of the past; in theory, the next generation of chemicals will cause less acute toxicity and be less environmentally persistent and bioaccumulative. However, researchers still struggle to establish whether the nonlethal effects associated with some modern chemicals and substances will have serious consequences for wildlife. Obtaining the resources to address issues associated with chemicals in the environment remains a challenge.

Synthetic chemicals have enabled marked improvements in food production and living standards (*1*). Although concerns exist about the many hundreds of chemicals in the environment, there are only a few, albeit notable, examples of chemicals actually harming wildlife populations (Fig. 1). These examples demonstrate that hydrophobic (lipophilic) chemicals can both persist in the environment and bioconcentrate, meaning that the highest exposures manifest in the longest-lived top predators. In addition, tests of acute toxicity on a limited range of laboratory-friendly species are not predictive for all species and effects, and chronic tests on a wider range of organisms are needed. Knowledge gained from such disasters should make the use of chemicals increasingly safer. However, our past failures suggest that we must be prepared for more surprises in the future.

Proportion of chemicals for which adequate environmental information is known

In places where data are accessible, such as the United States and Europe, the number of chemicals and substances on the market is believed to be around 75,000 to 140,000 (*2*, *3*). However, empirical data on persistence are available for ~0.2%, bioconcentration data for 1%, and aquatic toxicity for 11% of chemicals registered in the European Union (*4*, *5*), and similar data have been reported for the United States (*2*). In the absence of such substantive information for the majority of chemicals, computational predictive methods can provide some help in terms of risk assessment (*2*, *4*). Nevertheless, the task is complicated by the formation of breakdown products in the environment, for which we have little to no information. An additional challenge to our efforts to assess risk from these many chemicals entering the environment is the potential for mixture effects. These effects may lead to higher impacts on organisms than would have been predicted on the basis of individual chemical-based risk assessments (*6*). Today's research funding model tends to encourage widening and deepening studies on the current chemical, or group of chemicals, perceived to be of most concern, rather than supporting research on a larger proportion of the chemicals being discharged and considered potentially problematic (*7*).

Chemical risks are not equal, nor is exposure

Given the vast array of chemicals contaminating our natural environment, where should we focus our greatest attention? For instance, the risk of copper harming wildlife is reported to be five orders of magnitude higher than the risk from the drug atenolol (*8*) when comparing median exposure with median toxicity values for rivers in the United Kingdom. In other words, the risk of harm from atenolol is only 0.001% of the risk from copper. In fact, metals dominate the top 10 of 71 chemicals of concern studied in the United Kingdom (*8*) (Fig. 2) and are similarly highly ranked in China (*9*).

Chemical exposure from wastewater, which can be expressed as the extent to which the wastewater generated by an individual will be diluted by the natural river flow (*10*), is not evenly spread around the world. Depending on landmass, population size, and rainfall amount, some countries will face constant and widespread elevated exposure to chemicals in wastewater, whereas other nations will experience much less exposure (Fig. 3).

New chemicals and new places of concern

In modern society, chemicals are expected to serve a variety of functions; examples include

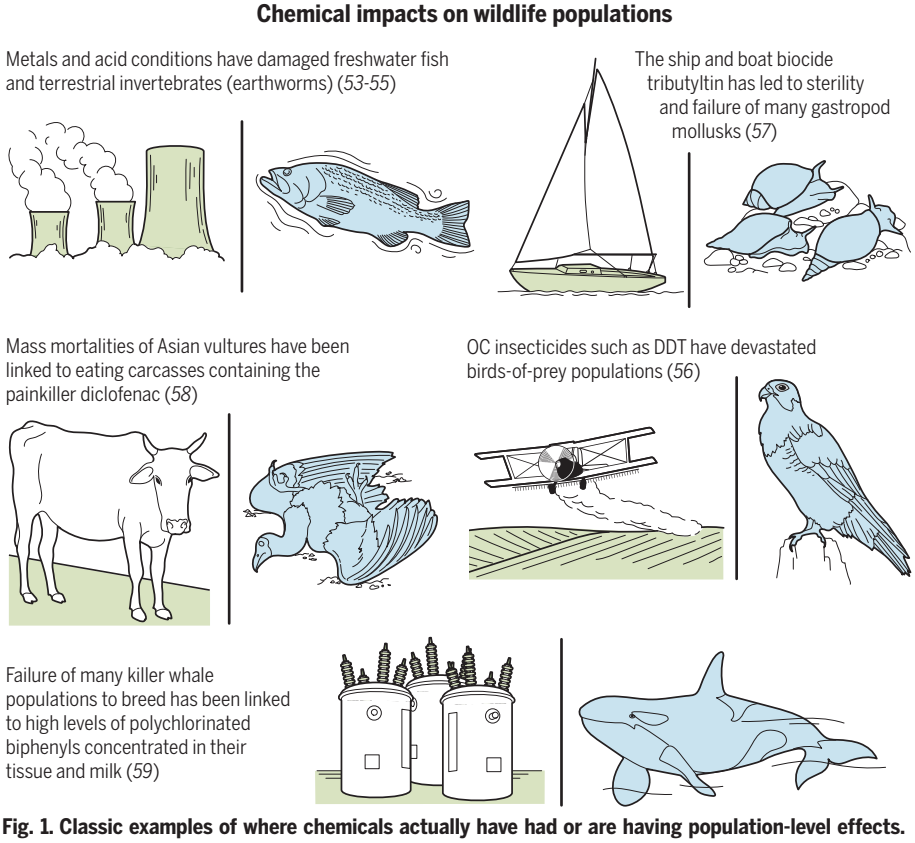


Fig. 1. Classic examples of where chemicals actually have had or are having population-level effects. See (*53*–*59*). OC, organochlorine.

[1]Centre for Ecology and Hydrology, Benson Lane, Crowmarsh Gifford, Wallingford, Oxfordshire OX10 8BB, UK. [2]China National Environment Monitoring Centre, Anwai Dayangfang No. 8, Chaoyang District, Beijing, China. [3]Research Center for Environmental Quality Management, Kyoto University, 1-2 Yumihama, Otsu, Shiga, 520-0811, Japan. [4]Institute for the Environment, Health and Societies, Brunel University London, Uxbridge, Middlesex UB8 3PH, UK.
*Corresponding author. Email: ajo@ceh.ac.uk

medicines, flame retardants, and pesticides. We now recognize that the very properties that can make these chemicals work effectively can simultaneously be deleterious for the wider environment. For instance, as medical knowledge grows, the expectation for new pharmaceutical-based treatments for diverse health conditions will continue. A current example is the incentive for drug companies to devise more effective compounds to treat a range of age-related conditions (*11*). Additionally, ethinylestradiol has been a very effective oral contraceptive, but the combination of its potency and persistence has made it an endocrine disrupter in wild fish downstream of wastewater effluent (*12*). If some of the new pharmaceuticals act as agonists or antagonists on the endocrine system, then the estrogen-based disruption may expand to a wider variety of fish (*13*). On the subject of flame retardants, problems with the persistence and toxicity of polybrominated diphenyl ethers have led to a wider range of replacement candidate substances, including nonhalogenated organic or metal compounds with phosphate groups, hydroxide, or stannate groups (*14*). Finally, concerns over pesticide mobility, nontarget toxicity, and persistence have markedly reduced the number of products for sale. The pest-control approaches of tomorrow are likely to be more precisely targeted to affect RNA interference, pheromones, and sterility. New flame retardants and insecticides should be much safer than older ones, but we must be alert to unexpected consequences, as have been observed for neonicotinoids (insect-specific postsynaptic agonists). These compounds, once considered sustainable, are now known to cause population decline in wild bees (*15*).

The modern economy has been transformed by globalization. As a result, much chemical production has been transferred to Asia (*16*), where chemical sales are now 168% of those in the United States and Europe combined (Fig. 4). However, in some cases, weak regulation or uneven local enforcement has led to severe pollution hotspots. Examples include atmospheric contamination with chlorofluorocarbons coming from the Shandong and Hebei provinces of China (*17*), gross perfluorooctanoic acid pollution from a vast Chinese manufacturing site (*18*), and water contamination with antibiotics from a manufacturing plant in India (*19*). Unfortunately, successful management of industrial waste, and pollution more generally, is far from straightforward. Setting water quality targets is a good step, but such benchmarks are successful only where independent regulators take consistent, high-quality measurements and are supported by an independent judiciary, on both the local and national scale. The degree to which environmental protection is improved by centralization or when it is devolved to local

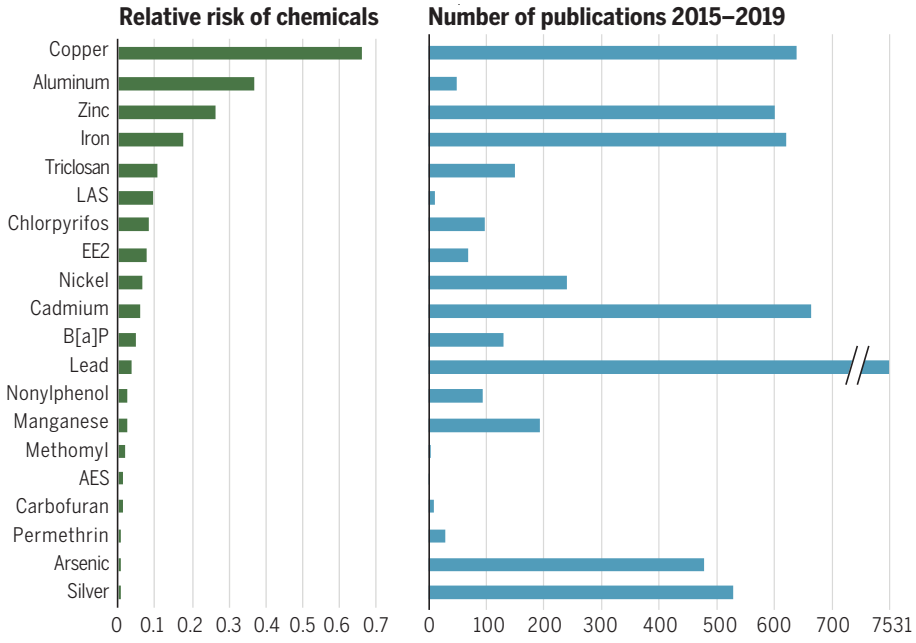


Fig. 2. The 20 highest-ranked chemicals from a pool of 71 common chemicals found in British rivers. Chemicals are ranked according to the ratio of median river concentration versus the fifth percentile of aquatic ecotoxicity data. Relative risk numbers are presented as ratios. Data are from (*8*). Also shown is the number of publications found on Web of Science in September 2019 under the search "chemical AND environment AND risk" for the period 2015–2019 for the chemicals listed at left. LAS, linear alkylbenzene sulfonates; EE2, ethinylestradiol; B[a]P, benzo[a]pyrene; AES, alcohol ethoxysulfates.

administrations is debatable (*20*). In the case of local governance in China, there is evidence for uneven application of regulations (*21*, *22*). Protection is also boosted by a national commitment to transparency, in which scrutiny by the public, environmental nongovernmental organizations, and journalists is accepted. Nevertheless, this approach has not been adopted globally (*23*, *24*).

Reasons for optimism

Progress in regulation and management of chemicals in the environment

Chemical regulations in the 1960s and 1970s concentrated on remediating past pollution and controlling the emission of a limited number of pollutants. The approach today is becoming forward-looking to ensure that new chemicals poised to enter the market will conform to minimum human safety and environmental standards. Examples include the Toxic Substances Control Act (TSCA) in 1976 (Public Law 94-469) in the United States, as well as Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH) (EC 19072006) in the European Union. Still, because many chemicals entered the market before these laws were enacted, a retrospective authorization process is trying to catch up. Although not perfect, the establishment of regulations such as TSCA and REACH set an important precedent: The onus to demonstrate that a chemical is safe for humans and the environment should lie with the manufacturer. In Europe, the phrase used to describe this concept is "no data, no market" (*25*).

Analytical developments, knowledge of undesirable chemical characteristics, and alternatives to animal testing

Developments in analytical chemistry continue to drive down limits of detection. With the use of nontargeted screening (NTS) methods, it is becoming possible to search for and tentatively identify all molecules present in a sample, both known and unknown (*26*). NTS has been applied to reveal the range of compounds in urban runoff water (*27*), to investigate unusual pollution incidents, and to identify the industrial premises responsible for such incidents (*28*). Recently, historic analytical raw data from previous studies have been used to retrospectively analyze "new" pollutants that were not originally targeted in these studies (*29*). These new approaches will help make the environment more transparent with respect to chemical contaminants.

There is now much shared knowledge on the undesirability of properties such as hydrophobicity and persistence in chemicals intended for discharge to the environment. In the consumer goods industry, recognition of poor biodegradability has led to the replacement of branched alkylbenzene sulfonates by linear forms, long-chain dialkyl quaternary surfactants by ester-based quaternaries, nonylphenol ethoxylates (which also have toxicity concerns) by alcohol ethoxylates, and musk xylene by macrocyclic musks. Although not driven primarily by environmental

concerns, an increasing proportion of newly registered pharmaceuticals are the so-called "biologics." For example, 12 of the 30 new drugs registered for the German market in 2016 (noted by the German Pharma Association) and 75 of the 200 recent top-selling retail drugs in the United States (*30*) are made from biological materials such as proteins, genes, allergens, and cells. These substances are not considered to pose the persistence issues of small synthetic molecules.

Understandably, ethical concerns have arisen about subjecting large numbers of animals to laboratory toxicity tests for the many thousands of chemicals yet to be registered, and these concerns have encouraged the development of toxicity and exposure models (*2*). Computer models have been used to help predict which chemicals will be of greatest concern (in silico risk assessment)—in other words, those that will be persistent, bioaccumulative, and toxic (PBT). In a survey of 95,000 chemicals, a model predicted that only 3 to 5% were likely to be PBT (*4*).

Better wastewater treatment and international chemical initiatives

Shifting from primary wastewater treatment (settling) to secondary treatment (biological) and increasing biological treatment time in secondary treatment from simple methods such as trickling filters to activated sludge (*31*, *32*) have considerable benefits for general water quality and chemicals reduction. The widespread adoption of the activated sludge process (ASP) in towns and cities around the world, with a biological treatment time of 8 hours or more, has done a great deal to protect rivers from the worst consequences of high chemical exposures. In China, it is now reported that the water distributed to 94% of urban population receives wastewater treatment, with 81% undergoing advanced processes such as ASP (*33*). Introduction of these methods can substantially improve water quality and, hence, biodiversity as compared with previous, less efficient treatments (*31*). As a society, we now have the capacity to introduce stringent tertiary treatment to eliminate almost all organics from wastewater effluent, as is being done in some parts of Switzerland (*34*).

Developed and developing countries share many of the same chemical challenges. This is particularly true with regard to many persistent pollutants, which know no boundaries. It is encouraging to see international agreements on persistent organic pollutants (Stockholm Convention), mercury (Minamata Convention), hazardous waste disposal (Basel Convention), and certain hazardous chemicals and pesticides (Rotterdam Convention). Sensible advice on managing chemicals, with respect to legal, economic, technical, and voluntary instruments and the adoption of safer alternatives, is now available to all countries (*35*).

Reasons for pessimism

Continuing uncertainty over the importance of nonlethal effects

Once we move away from apical end points (lethal or end points that disrupt reproduction or growth), it remains a matter of speculation as to whether the response to a chemical observed in the laboratory really translates to harm for individuals or populations in the wild. In theory, the detailed mechanistic detection of an adverse outcome pathway (AOP) predicts harmful effects ranging from the molecular level up to the population level (*36*). AOPs have been used to confidently predict population effects on fish from endocrine disrupters (*37*), yet these effects have not been observed in the field (*38*). Whether the development of AOPs will aid in the environmental risk assessment of chemicals is presently unclear. Similarly, the question of whether gene, protein, or metabolite expression studies can, on their own, predict actual impacts on wildlife populations or food webs (*39*) remains to be answered.

Data quality and the relevance of research topics

It is now widely accepted that a high proportion of published research is not reproducible, a situation sometimes called the reproducibility crisis (*40*–*42*). Reasons may include perverse incentives on scientists to publish "exciting" research and a general lack of training for researchers (*43*). Two common associated problems are poor experimental design and bias (*44*). In ecotoxicology, many scientists conduct their research on animals that are not routinely used in regulatory tests and that other researchers rarely use.

The focus of public concern over chemicals is unpredictable. This can lead to sudden demands for information, which can overwhelm other research areas. Inevitably, many fundable topics will have to be dropped so that resources can be concentrated in an area of new concern. One area of marked growth has been the study of nanoparticles and the environment: A search for this topic on Web of Science revealed an increase from 36 papers per year in 2000 to 4200 per year in 2017. Yet many studies appear to show a modest relative risk, at least for common metal-based nanoparticles (*8*, *45*). Another example may be the study of bisphenol A (BPA), an additive used in many plastic items, which has been shown to exhibit weak estrogen activity. Many hundreds of studies on BPA's presence and possible harm to the environment have been published (a September 2019 search of Web of Science with the terms "BPA," "effect," and "environment" revealed 630 papers). Yet the evidence that BPA is adversely affecting wildlife is essentially nonexistent (*46*). On the other hand, there are many thyroid activity, cardiovascular, antiepileptic, and muscle relaxant drugs for which few, if any, studies of possible effects on aquatic wildlife have been carried out.

Perhaps surprisingly, the focus of research into chemicals in the environment is not necessarily linked to their relative risk. For the top 20 highest-risk chemicals in British rivers (Fig. 2), publications related to their environmental risk varied between 7531 for lead to only 2 for the anionic surfactant alcohol ethoxysulfates in the period 2015–2019 (Fig. 2).

This area of science is prone to the "bandwagon" effect, by which many papers only demonstrate what we already know. Did we need ≥250 papers to tell us that ethinylestradiol poses a risk to fish? Everything we need to know to protect the environment was communicated in the first half a dozen papers. A current trend is this desire to search for increasingly more subtle effects, such as the expression of one or a few genes being altered, when the consequences of those effects are entirely unknown.

Risk assessments are falling further behind, and scientists tend to stay in their silos

Thorough risk assessment is costly and may require decades of research. Given

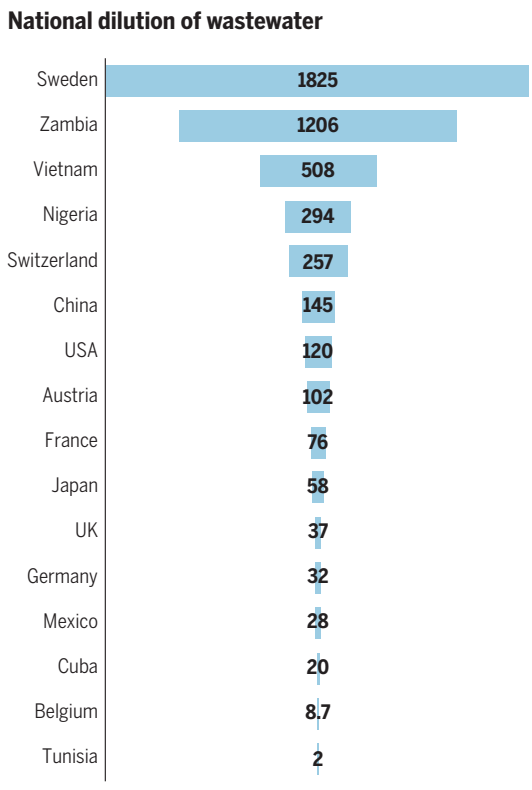


Fig. 3. Examples of the relative dilution of an individual's wastewater, based on national median annual natural flow divided by the national annual inland wastewater volume. Data are from (*10*).

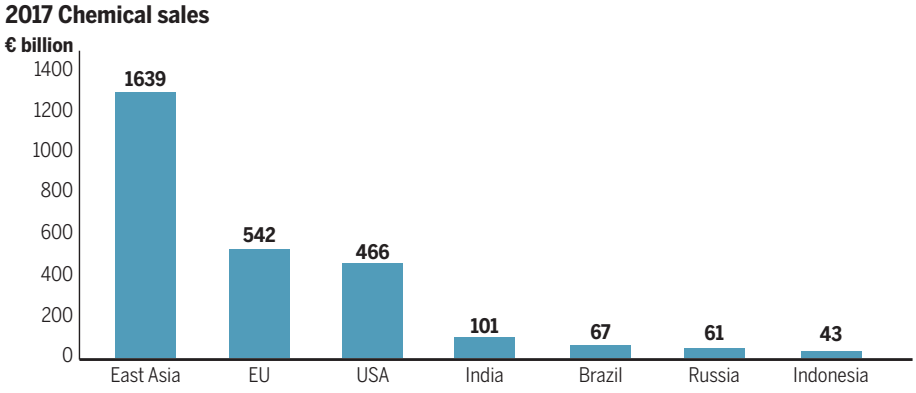


Fig. 4. Level of chemical sales in 2017. East Asia includes China, Japan, and South Korea (*16*). EU, European Union; €, euros.

the range of species and number of end points that could be examined, it seems certain that we will never catch up by using traditional approaches (*47*). If this assertion is correct, then persevering with the current testing strategy does not seem appropriate. Ethical objections to the use of animals, particularly vertebrates, in tests are increasing, yet we continue to add more tests to the Organisation of Economic Cooperation and Development (OECD) battery of accepted (eco)toxicity tests. Efforts to rethink how the environmental risks of a chemical can be assessed, with an expanded role for predictive modeling of harmful properties, are ongoing, but regulators remain cautious about relying on such information (*48*).

The study of chemicals in the environment appears to revolve largely around the two disciplines of ecotoxicology and environmental chemistry. In their publications, ecotoxicologists commonly state that "effects were observed at environmentally relevant concentrations," whereas environmental chemists are often tempted to assert that their "highest measured concentrations exceeded reported effect (toxic) concentrations" (*49*). Such statements imply that chemicals are harming the environment, possibly to a serious extent, on a daily basis. However, it is unclear, based on the evidence of ecotoxicology and environmental chemistry alone, whether we are exaggerating the dangers and thus overregulating or, alternatively, underestimating risks (as has been proposed from mixture effects) and thus failing to protect (*47*). Additionally, a third community of scientists—ecologists—has much to offer, in theory, in assessing chemical impacts on wildlife. The presence of long-term wildlife monitoring is vital for such research, but we see surprisingly few examples of collaboration between ecologists, ecotoxicologists, and environmental chemists. Ecologists have highlighted alarming declines in some wildlife populations (*50*, *51*), and, despite many confounding variables, long-term ecological data can be extremely compelling toward establishing a link between competing arguments, such as those concerning neonicotinoids and bees (*15*, *39*). To determine the true harm of chemicals, these different scientists will need to collaborate closely (*52*).

Outlook

Adapting to the immensely difficult societal and environmental challenges of tomorrow will undoubtedly require new chemicals and chemical solutions. The production of chemicals, their diversity, and their use around the world has never been greater. Our ability to manage the risks is finely balanced, with reasons to be both pessimistic and optimistic. Unfortunately, the sheer volume of chemicals on the market, and presumably also entering the environment, currently outpaces our ability to assess the risks. Although there are no guarantees, our past knowledge combined with in silico modeling of hazards will be beneficial in gauging relative risk. Provided that long-term wildlife monitoring efforts are maintained, particularly in areas with the greatest chemical exposure, we may have some confidence that our use of chemicals is sustainable.

"Our ability to manage the risks is finely balanced, with reasons to be both pessimistic and optimistic."

REFERENCES AND NOTES

1. J. E. Casida, G. B. Quistad, *Annu. Rev. Entomol.* **43**, 1–16 (1998).
2. P. P. Egeghy *et al.*, *Sci. Total Environ.* **414**, 159–166 (2012).
3. R. Judson *et al.*, *Environ. Health Perspect.* **117**, 685–695 (2009).
4. S. Strempel, M. Scheringer, C. A. Ng, K. Hungerbühler, *Environ. Sci. Technol.* **46**, 5680–5687 (2012).
5. L. Posthuma, J. van Gils, M. C. Zijp, D. van de Meent, D. de Zwart, *Environ. Toxicol. Chem.* **38**, 905–917 (2019).
6. T. J. Thrupp *et al.*, *Sci. Total Environ.* **619–620**, 1482–1492 (2018).
7. C. G. Daughton, *Sci. Total Environ.* **466–467**, 315–325 (2014).
8. A. C. Johnson *et al.*, *Sci. Total Environ.* **599-600**, 1372–1381 (2017).
9. A. C. Johnson *et al.*, *Environ. Toxicol. Chem.* **37**, 1115–1121 (2018).
10. V. D. J. Keller, R. J. Williams, C. Lofthouse, A. C. Johnson, *Environ. Toxicol. Chem.* **33**, 447–452 (2014).
11. D. Bunke *et al.*, *Environ. Sci. Eur.* **31**, 32 (2019).
12. K. L. Thorpe *et al.*, *Environ. Sci. Technol.* **37**, 1142–1149 (2003).
13. J. P. Sumpter, A. C. Johnson, *Environ. Sci. Technol.* **39**, 4321–4332 (2005).
14. S. L. Waaijers *et al.*, *Sci. Total Environ.* **463–464**, 1042–1048 (2013).
15. B. A. Woodcock *et al.*, *Nat. Commun.* **7**, 12459 (2016).
16. CEFIC, "Facts and figures of the European chemical industry" (The European Chemical Industry Council, 2018); http://old.cefic.org/Documents/RESOURCES/Reports-and-Brochure/Cefic_FactsAnd_Figures_2018_Industrial_BROCHURE_TRADE.pdf.
17. M. Rigby *et al.*, *Nature* **569**, 546–550 (2019).
18. P. Wang *et al.*, *Environ. Pollut.* **218**, 1234–1244 (2016).
19. J. Fick *et al.*, *Environ. Toxicol. Chem.* **28**, 2522–2527 (2009).
20. H. Y. Zhao, R. Percival, *Transnatl. Environ. Law* **6**, 531–549 (2017).
21. S. H. Guo, J. Q. Lu, *J. Clean. Prod.* **212**, 1054–1061 (2019).
22. T. Hong, N. N. Yu, Z. G. Mao, *J. Clean. Prod.* **231**, 649–659 (2019).
23. G. Li, Q. He, S. Shao, J. Cao, *J. Environ. Manage.* **206**, 1296–1307 (2018).
24. S. B. Kedzior, *J. Environ. Dev.* **26**, 272–296 (2017).
25. European Commission, Introduction to REACH regulation (2019); https://ec.europa.eu/environment/chemicals/reach/reach_en.htm.
26. E. L. Schymanski *et al.*, *Anal. Bioanal. Chem.* **407**, 6237–6255 (2015).
27. B. Du *et al.*, *Environ. Sci. Process. Impacts* **19**, 1185–1196 (2017).
28. J. Hollender, E. L. Schymanski, H. P. Singer, P. L. Ferguson, *Environ. Sci. Technol.* **51**, 11505–11512 (2017).
29. M. C. Campos-Mañas, I. Ferrer, E. M. Thurman, J. A. Sánchez Pérez, A. Agüera, *Sci. Total Environ.* **664**, 874–884 (2019).
30. N. A. McGrath, M. Brichacek, J. T. Njardarson, *J. Chem. Educ.* **87**, 1348–1349 (2010).
31. A. C. Johnson *et al.*, *Environ. Toxicol. Chem.* **38**, 1820–1832 (2019).
32. M. Gardner *et al.*, *Sci. Total Environ.* **456–457**, 359–369 (2013).
33. Q. H. Zhang *et al.*, *Environ. Int.* **92–93**, 11–22 (2016).
34. M. Bourgin *et al.*, *Water Res.* **129**, 486–498 (2018).
35. UNEP, "Strategic Approach to International Chemicals Management. SAICM texts and resolutions of the International Conference on Chemicals Management" (United Nations Environment Programme, 2007).
36. G. T. Ankley *et al.*, *Environ. Toxicol. Chem.* **29**, 730–741 (2010).
37. G. T. Ankley *et al.*, *Aquat. Toxicol.* **92**, 168–178 (2009).
38. A. C. Johnson, Y. Chen, *Sci. Total Environ.* **589**, 89–96 (2017).
39. M. Yamamuro *et al.*, *Science* **366**, 620–623 (2019).
40. M. Hanson, L. Baxter, J. Anderson, K. Solomon, R. Brain, *Sci. Total Environ.* **685**, 1221–1239 (2019).
41. M. L. Hanson *et al.*, *Sci. Total Environ.* **578**, 228–235 (2017).
42. E. Loken, A. Gelman, *Science* **355**, 584–585 (2017).
43. C. A. Mebane *et al.*, *Integr. Environ. Assess. Manag.* **15**, 320–344 (2019).
44. C. A. Harris *et al.*, *Environ. Sci. Technol.* **48**, 3100–3111 (2014).
45. D. A. Notter, D. M. Mitrano, B. Nowack, *Environ. Toxicol. Chem.* **33**, 2733–2739 (2014).
46. E. Mihaich *et al.*, *Environ. Toxicol. Chem.* **31**, 2525–2535 (2012).
47. A. C. Johnson, J. P. Sumpter, *Environ. Toxicol. Chem.* **35**, 1609–1616 (2016).
48. ECHA, "The use of alternatives to testing on animals for the REACH Regulation" (European Chemicals Agency, 2017); https://echa.europa.eu/documents/10162/13639/alternatives_test_animals_2017_en.pdf.
49. L. Weltje, J. P. Sumpter, *Environ. Sci. Technol.* **51**, 11520–11521 (2017).
50. F. Sánchez-Bayo, K. A. G. Wyckhuys, *Biol. Conserv.* **232**, 8–27 (2019).
51. C. A. Hallmann *et al.*, *PLOS ONE* **12**, e0185809 (2017).
52. M. O. Gessner, A. Tlili, *Freshw. Biol.* **61**, 1991–2001 (2016).
53. C. A. Mebane, R. J. Eakins, B. G. Fraser, W. J. Adams, *Elem. Sci. Anth.* **3**, 000042 (2015).
54. D. J. Spurgeon, S. P. Hopkin, D. T. Jones, *Environ. Pollut.* **84**, 123–130 (1994).
55. J. Herrmann *et al.*, *Ambio* **22**, 298–307 (1993).
56. D. A. Ratcliffe, *J. Appl. Ecol.* **7**, 67 (1970).
57. C. D. Sayer *et al.*, *Environ. Sci. Technol.* **40**, 5269–5275 (2006).
58. J. L. Oaks *et al.*, *Nature* **427**, 630–633 (2004).
59. J. P. Desforges *et al.*, *Science* **361**, 1373–1376 (2018).

ACKNOWLEDGMENTS

We thank M. Jürgens (CEH) for technical support.
Funding: A.C.J. and J.P.S. are grateful to NERC for grant NE/S000100/1 supporting the ChemPop project.
Competing interests: A.C.J. and J.P.S. are currently members of the Defra (UK) Hazardous Substances Advisory Committee. No other competing interests are known.

10.1126/science.aay6637

REVIEW

Tracking complex mixtures of chemicals in our changing environment

Beate I. Escher[1,2]*, Heather M. Stapleton[3], Emma L. Schymanski[4]

Chemicals have improved our quality of life, but the resulting environmental pollution has the potential to cause detrimental effects on humans and the environment. People and biota are chronically exposed to thousands of chemicals from various environmental sources through multiple pathways. Environmental chemists and toxicologists have moved beyond detecting and quantifying single chemicals to characterizing complex mixtures of chemicals in indoor and outdoor environments and biological matrices. We highlight analytical and bioanalytical approaches to isolating, characterizing, and tracking groups of chemicals of concern in complex matrices. Techniques that combine chemical analysis and bioassays have the potential to facilitate the identification of mixtures of chemicals that pose a combined risk.

Chemicals are the basis of life, but some anthropogenic organic chemicals pose inherent dangers. Pesticides, industrial chemicals, pharmaceuticals, and other synthetic chemicals can enter the environment and the food chain, causing unwanted effects and disease. Medical research indicates that as much as two-thirds of chronic human disease risk cannot be explained by genetics alone and may result from the environment or gene-environment interactions (*1*). Furthermore, the *Lancet* Commission on Pollution and Health has estimated that 16% of global premature deaths are linked to pollution (*2*). These statistics highlight the need for research to elucidate the complex links among exposure to chemicals, environmental quality, and health.

Concentrations of many legacy chemicals are decreasing after national and international actions led to near-global phase-out of these chemicals (*3*). However, the number of new chemicals is rising, with the Chemical Abstract Service Registry growing from 20 million to 156 million chemicals between 2002 and 2019. Regulation of problematic chemicals can take decades; once enacted, such rules can lead to chemical substitutions that are less well characterized. There have been several cases in which the replacement chemical had properties, including toxicity, similar to those of the chemical it was intended to replace. Notable examples include plasticizers, flame retardants, chlorinated paraffins, and polyfluoroalkyl substances.

More recent industrial and agricultural chemicals, pharmaceuticals, and personal care products are not generally persistent, but they are ubiquitous as a result of their continuous use and global sources. When degraded, the resulting transformation products may be more persistent and may occur at higher concentrations than their parent compounds (*4*). Generally, degradation leads to transformation products that are more water-soluble and less toxic; however, some transformation products are more toxic than their parent (*5*).

The mixture challenge

Historically, chemical pollution was often attributed to a defined group of industrial chemicals. Today, awareness is increasing that we are exposed to a true cocktail of chemicals, only a fraction of which have been identified. There is no equity in the global distribution of these pollutants; developing countries may be at highest risk, given that large-scale production is moving to these countries and adding to their challenges in developing chemical regulations and infrastructure such as wastewater treatment (*2*). Awareness of the need to deal with complex chemical mixtures has increased with the introduction of the exposome concept, which integrates all human exposure from chemical and nonchemical stressors in relation to adverse health effects in humans (*6*) and can be expanded to any biota.

Chemicals can contribute to toxicity in a complex mixture even if they are present below their own effect threshold and/or analytical detection limit (*7*). Chemicals with the same modes of toxic action tend to follow the mixture concept of "concentration addition," whereas those with different modes of action act according to "independent action" (*8*). At low effect levels and low, environmentally realistic concentrations, concentration-effect curves are linear and the two models converge. Synergy is generally limited to mixtures with a small number of components at high concentrations (*9*) but becomes less relevant for low doses (*10*). Because synergy rarely leads to more than a factor of 10 increase in effect for any synergistic combination and only a few components will interact, it is of lower priority for environmental mixtures. It is safe to assume that most environmental mixtures with tens of thousands of chemicals of diverse modes of action at low concentrations will act according to the simple additive model, but the big unknown is the contribution of unidentified chemicals to the effect of environmental mixtures. Therefore, we discuss bioanalytical tools as a quantitative measure of mixture effects in monitoring studies.

Although this review focuses on anthropogenic organic chemicals, the relevance of mixtures is exacerbated by metals, inorganics, and particles (e.g., plastics, nanomaterials). Moreover, chemical mixtures can act jointly with multiple stressors caused by external factors such as oxygen levels, increasing temperature, and ocean acidification.

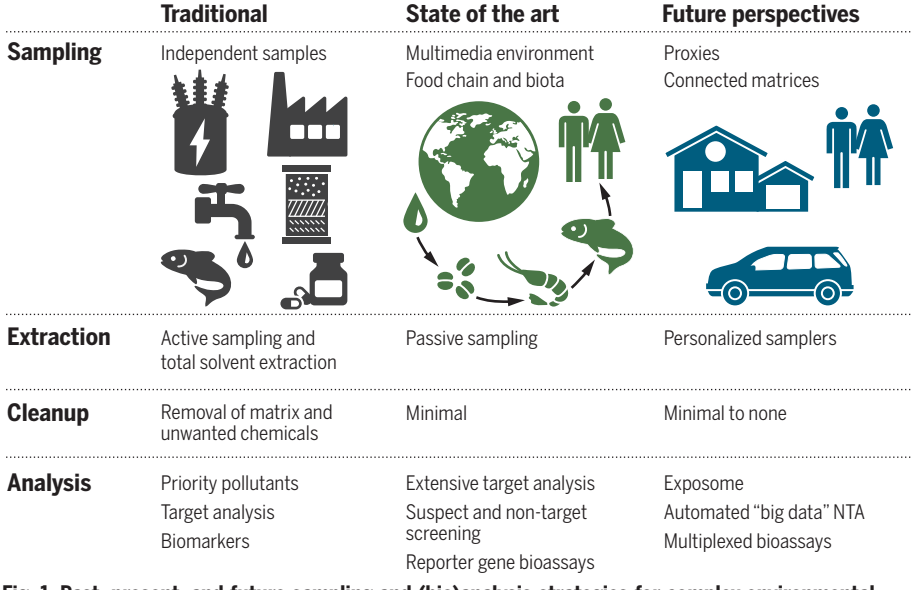


Fig. 1. Past, present, and future sampling and (bio)analysis strategies for complex environmental mixtures in the environment and in humans. NTA, non-target analysis.

CREDIT: KELLIE HOLOSKI/SCIENCE

[1]Department of Cell Toxicology, Helmholtz Centre for Environmental Research–UFZ, DE-04318 Leipzig, Germany. [2]Environmental Toxicology, Center for Applied Geoscience, Eberhard Karls University Tübingen, DE-72074 Tübingen, Germany. [3]Nicholas School of the Environment, Duke University, Durham, NC 27708, USA. [4]Luxembourg Centre for Systems Biomedicine (LCSB), University of Luxembourg, 4367 Belvaux, Luxembourg.
*Corresponding author. Email: beate.escher@ufz.de

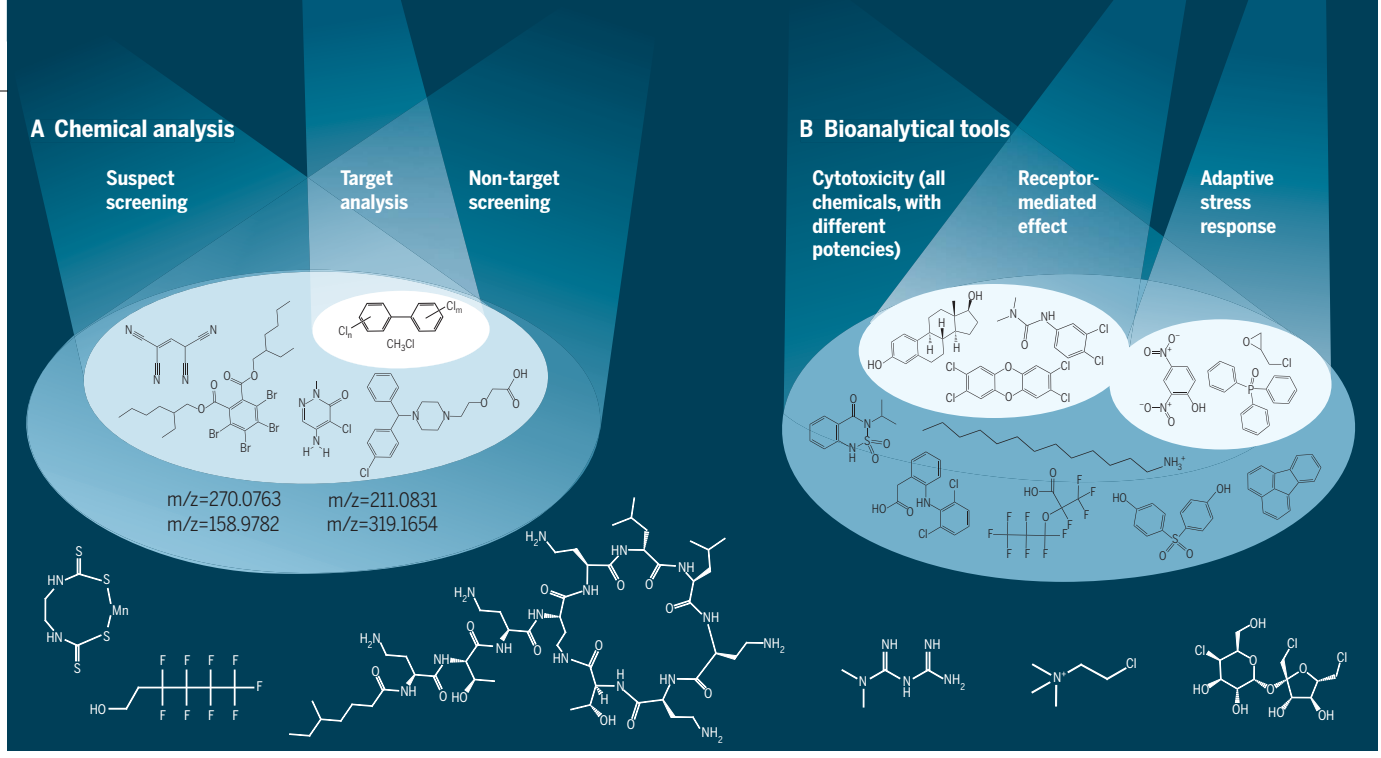


Fig. 2. Spotlight on the chemical universe. (**A**) Chemical analysis. (**B**) Bioanalytical tools. The overlap that exists between the chemicals "seen" by chemical analysis and bioanalysis is not depicted. The chemicals depicted are illustrative examples; selection is not comprehensive.

Monitoring chemicals

Early studies focused on hydrophobic persistent organic pollutants, which accumulate in soils, sediments, or lipid-rich tissues of organisms (Fig. 1A). Research has evolved to include a focus on polar contaminants as well, particularly in surface and ground waters used as drinking water sources (*11*).

Today, we recognize that there is a strong interconnectivity of chemicals in different environmental compartments (Fig. 1B). Some chemicals may preferentially accumulate in one environmental compartment over another and have different degradability in different compartments, but once emitted into the environment, they travel between all environmental compartments and along the food chain to humans. Simple exposure models and assessment of physicochemical properties of the chemicals of interest may make it easier to design sampling strategies that direct monitoring efforts toward relevant matrices.

Sampling and sample preparation

To identify chemical mixtures posing the greatest risk, sampling and monitoring plans must be developed to address variability in space, time, and composition, and to determine whether sampling should be continuous or discrete (*12*). Passive sampling techniques are alternative sampling methods applied to air, water, sediments, and biota in field or laboratory studies (*13*). Passive samplers are typically polymers (e.g., low-density polyethylene, polyurethane, or silicone) and offer time-integrative or equilibrium-based sampling.

Most environmental matrices and tissues cannot be directly analyzed because endogenous chemicals would disturb analysis. Pollutants are present in very low concentrations. Therefore, extraction is required to isolate and enrich pollutants, and cleanup steps are required to remove coextracted matrix (Fig. 1). Care must be taken not to alter the mixture composition during processing and to quantitatively track the enrichment. Recovery standards (e.g., isotope-labeled analogs) can be applied prior to chemical analysis, but they cannot be used for bioanalysis, for which independent recovery experiments must be conducted (*14*). Persistent organic pollutants can tolerate harsh chemical treatments to remove matrix components before analysis (Fig. 1A), but recent moves to focus on more labile contaminants require more directed sample preparation approaches. Because any cleanup is time-consuming and will lead to the loss of a fraction of chemicals, recent developments strive for minimal cleanup or none at all (Fig. 1C).

Water is an important source and sink of pollutants from wastewater treatment plants, urban runoff, and agricultural applications (*15*). If organic matter content is not too high, water can be directly injected into liquid chromatography instruments. This approach will likely grow in popularity as the sensitivity of analytical instruments increases. Solid-phase extraction is a versatile method for enrichment of chemicals from water samples that captures a large fraction of organic chemicals with generally good extraction efficacy in water, apart from limitations for small or ionized chemicals (*14*).

Chemicals in more complex matrices such as soil, sediment, plants, biota, and human tissue are often extracted with organic solvents. This approach may cause co-extraction of lipids and matrix components and requires tedious cleanups, including acid digests and gel permeation chromatography or silica gel columns, to remove the matrix and isolate the chemicals of interest (*16*). Certain chemicals (such as hormones, pyrethroid pesticides, and dioxins) pose a risk at very low concentrations, such that extensive cleanup and highly sensitive, dedicated target analyses are required to differentiate them from co-occurring chemicals present at higher concentrations. Thousands of chemicals can be detected in human blood, but xenobiotic chemicals are found at levels that are orders of magnitude lower than biological markers of human metabolism (*17*). Even with the current high dynamic range and high-resolution analytical methods, not all chemicals will be quantifiable because phenomena such as matrix suppression will interfere with analysis.

Because sampling of tissues from living organisms, particularly humans, can be logistically challenging and may pose ethical concerns, proxies are needed to evaluate exposure to mixtures (Fig. 1C). Rather than blood or tissue, noninvasive samples (e.g., hair, nails, urine, deciduous teeth) can be used as proxies for exposure, although knowledge of uptake rates and toxicokinetics is required (e.g., to relate urine concentrations to ambient exposure levels). Hand wipes were validated as a measure of personal exposure via correlations with biomarkers of exposure (*18*), but they only measure recent exposures and can be confounded by behavior such as hand-washing. Silicone wristbands have gained popularity as passive personal samplers (Fig. 1C) that can be worn for several days to weeks; these measure average integrated exposures by air

and possibly dermal and inadvertent dust ingestion pathways (*19*). Wristbands have been validated for several classes of semivolatile organic contaminants ubiquitous in indoor environments (*20*). House dust is often contaminated with a complex mixture of organic and inorganic contaminants that have been released from various building materials and consumer products into the home environment. Although increasing attention is given to the characterization of house dust using non-target analysis (NTA) to characterize indoor sources of human exposure, the use of wristbands may have the advantage of estimating exposure in multiple microenvironments in addition to the home. However, wristbands are limited to assessments of neutral organics and do not effectively characterize metal and dietary exposures.

Chemical analysis

The growth, evolution, and accessibility of high-resolution mass spectrometry (HR-MS) in environmental laboratories (*21*) has opened a Pandora's box of chemical complexity. HR-MS, often coupled with either gas or liquid chromatography, can detect tens of thousands of "features" (accurate molecular masses associated with unknown chemicals) in biological and environmental samples (*22*). Fragmentation information from tandem mass spectrometry (HR-MS/MS) is required to gain more information on structural features and assign more confident identifications (*23*). Methods for introducing analytes into a mass spectrometer strongly influence the range of chemical classes that can be detected in a sample; although different methods are complementary, they do not necessarily overlap. Although target analysis remains an essential component of chemical risk assessment, this approach illuminates only a narrow portion of chemical exposures and offers no information on unknown or previously unexpected contaminants that fall outside the targeted method (Fig. 2A). Suspect screening—using HR-MS data and lists of known chemicals of interest ("suspect lists") to identify contaminants without prior knowledge of their presence—widens the contaminant space and is currently a popular approach for providing semiquantitative analysis of complex mixtures. Even with the best NTA methods, some chemicals remain outside the spotlight, such as those that elute too early or late from the column, are poorly ionized by existing ionization methods, or are not yet interpreted correctly with current knowledge.

HR-MS/MS offers the possibility of performing routine target analysis, suspect screening, and discovery-based NTA in an all-in-one approach (Fig. 2A). With NTA, the limited scope of priority pollutants has been left behind, as even unknown masses can now be tracked in the environment (*21*), giving environmental chemists unprecedented power to detect new and emerging contamination and thereby covering a much larger chemical space (Fig. 2A). HR-MS/MS is now used for studies as diverse as daily monitoring of river water, tracing historical sediment contamination (*21*), characterizing indoor dust compositions (*24*), or performing retrospective screening of emerging contaminants (*25*).

Through the evolution of computational workflows based on exact mass and fragment matching, HR-MS/MS data can now be archived and used for plots showing the distribution of chemicals across time, space, or various matrices (Fig. 3) (*25*). The increasing popularity of NTA has seen a rapid increase of suspect lists to help find chemicals of interest, such as the CompTox Chemicals Dashboard (*26*) and the NORMAN Network Suspect List Exchange (www.norman-network.com/nds/SLE/). The CompTox Chemicals Dashboard, alongside PubChem (*27*), forms an important data source with its predicted physicochemical properties, connections to toxicity information, and data structures allowing access to metadata of mixtures in NTA via "MS-ready" data mappings (*28*). NTA, especially coupled with suspect screening of classes of compounds, can complement targeted analytical techniques but can only supply part of the picture of the chemicals causing toxicity in complex samples (Fig. 2).

Bioanalytical tools to capture mixture effects

Traditionally, whole-organism in vivo bioassays were used to evaluate the toxicity of wastewater effluent and sediment, but such bioassays suffered from limited sample throughput and an inability to distinguish the effects of pollutants from matrix components, salinity, or pH. This has changed with the advancement of cell-based in vitro bioassays, which are animal-protective and are amenable to high-throughput robotics (*29*). For example, the ToxCast/Tox21 program (*30*) seeks to develop predictive models to reduce or eliminate future in vivo testing. The program heralded a paradigm shift in toxicity testing, with in vitro methods now included in human health risk assessments to elucidate mechanisms of toxicity and to prioritize chemicals for further testing. The application of high-throughput in vitro assays toward ecological risk assessment and for monitoring and complex environmental samples is only emerging but has great potential to accelerate risk assessment and reduce animal testing (*31*).

In principle, sample extracts with little or no cleanup can be tested using in vitro bioassays, but care must be taken to avoid matrix effects. Samples that contain natural organic matter can suppress effects when cell-free receptors or enzyme systems are used, whereas co-extracted lipids and organic matter may lead to a decrease in sensitivity of cell-based bioassays because they act as an additional partitioning phase in the assay, decreasing the bioavailable fraction of more hydrophobic chemicals. A good understanding of the dosing process (*32*) and stringent quality control is vital when testing environmental samples in bioassays. Despite these practical limitations, bioassays are essential to capture the effects of all chemicals in mixtures. Every chemical will contribute to cytotoxicity, even if present below the instrumental detection limit or below the effect threshold of the single chemical (Fig. 2B).

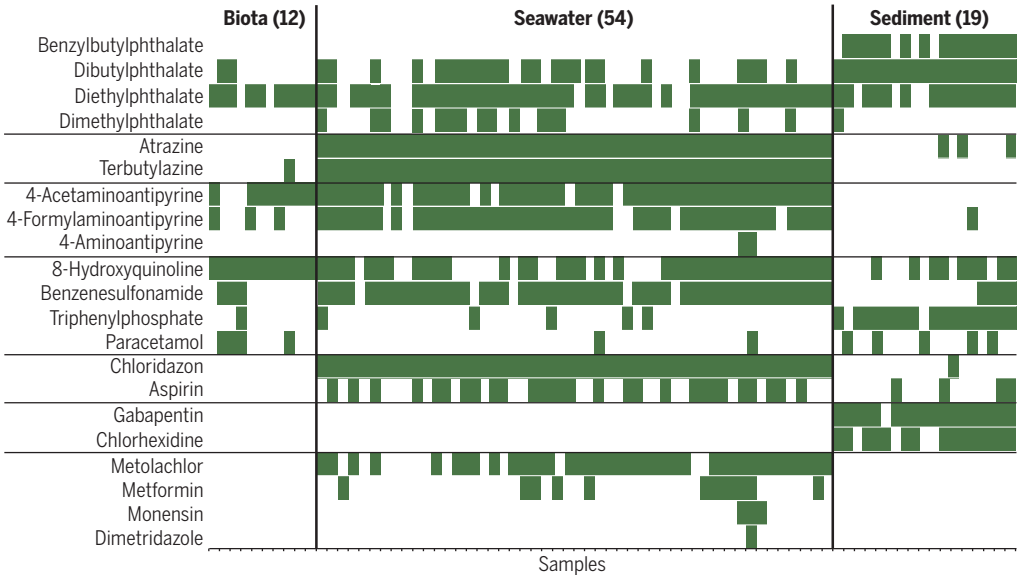


Fig. 3. Illustrative example of occurrence of chemicals in suspect screening of various matrices: biota (fish, mussels), seawater, and sediments. Green indicates presence with six or more matching HR-MS/MS fragments; white indicates absence or fewer than six HR-MS/MS fragments detected. [Created/modified from https://norman-data.eu/NORMAN-REACH/; details in (*25*)]

Reporter gene cell lines are popular in vitro assays and target one specific mode of action (MOA)—for example, binding to a hormone receptor, activation of metabolic enzymes through receptors such as the aryl-hydrocarbon receptor, or adaptive stress responses. This is accomplished by transient or stable transfection of a cell with a plasmid that contains multiple copies of the response element of the target receptor or transcription factor of interest, followed by multiple reporter genes. These encode reporter proteins that can be quantified easily, such as the enzyme luciferase. Cytotoxicity interference can mask specific effects, but this can be minimized by running a parallel cytotoxicity assay to avoid artifacts resulting from the so-called cytotoxicity burst (*33*). Concentration-effect curves are often linear at effect levels below 30%, which greatly simplifies mixture modeling and allows the calculation of bioanalytical equivalent concentrations (BEQ_{bio}) of complex mixtures in a sample (*34*). BEQ_{bio} relates the toxicity of a mixture to a well-characterized chemical for that MOA, and can be interpreted in terms of risk-scaled concentrations. These are easy to communicate to regulators and the public because they report mixture effects in easily understandable units—for example, "This water sample contains mixtures of chemicals that have the same effect as 6 ng of estradiol per liter of water," or "This fish has accumulated a chemical mixture that has the same effect as 3 pg of 2,3,7,8-tetrachlorodibenzodioxin per gram of fish." Effect-based trigger values based on acceptable BEQ_{bio} have been derived for diverse bioassays from drinking water guideline values and environmental quality standards and were proposed to be included in a future regulatory implementation of mixture effects in envrionmental monitoring (*35*).

Combining analytical and bioanalytical tools

Effect-directed analysis (EDA) or toxicity identification evaluation (TIE) can be used to identify risk drivers in complex mixtures and to separate bioactive chemicals that could otherwise be masked by matrix effects (*36*). A sample extract is separated by mass, hydrophobicity, or polarity by means of chromatography or physical separation into fractions, which are tested individually in bioassays. Each bioactive fraction is further fractionated until one or more bioactive subfractions are identified that contain chemicals that explain a majority of the observed effect(s). EDA has, for instance, helped to identify risk drivers in contaminated sites, industrial effluents, and sediments (*36*) or unknown chemicals that disturb thyroid function in blood of polar bears (*37*).

It is also possible to link measured concentrations and effects by modeling (*38*). The expected effects of quantified chemicals, expressed as bioanalytical equivalent concentrations from chemical analysis (BEQ_{chem}), can be predicted for all quantified and toxicologically characterized chemicals by multiplication of the detected concentration C_i with the relative effect potencies (REP_i) and adding the contributions from all chemicals in the mixture (Fig. 4). An equivalent concept is exposure-activity ratios of mixtures (EAR), where active concentrations are related to detected concentrations (*31*). These mass balance approaches apportion the effect of the known chemicals and identify chemicals that are mixture risk drivers among the known chemicals. For example, in Fig. 4A, the risk driver (light blue bar) would be chemical 2, because it has the highest relative effect potency (light gray bar) despite its low concentration (dark gray bar). If one compares the experimental BEQ_{bio} of the mixture with the BEQ_{chem} for matching bioassays, in many cases $BEQ_{bio} > BEQ_{chem}$, even if several hundreds of chemicals are quantified (*39*). The contribution of the unknown chemicals ($BEQ_{unknown}$) can be quantified using the simple effect balance equation $BEQ_{unknown} = BEQ_{bio} - BEQ_{chem}$. Certain receptor-mediated effects, such as activation of hormone receptors or inhibition of photosystem II, are triggered by a few chemicals that act in a highly specific manner. For these cases, it is often possible to close the mass balance and explain the entire biological effect of a sample (light green bar in Fig. 4B) by the predicted mixture effect of the known chemicals (dark blue bar). In contrast, for less specific effects, (nonspecific MOA in Fig. 4C), a large fraction of effect typically remains unidentified. For bioassays such as the fish embryo toxicity assay, activation of the oxidative stress response, or genotoxicity, the unknown effect may amount to as much as 90 to 99.99% of BEQ_{bio} (*39*). These mixture-modeling approaches historically suffered from a lack of toxicity data for individual chemicals, but with the large high-throughput screening databases becoming available and suspect screening lists expanding, more and more chemicals can be included in the BEQ_{chem} and EAR calculations and applied in monitoring studies (*40*).

Because typical chromatograms from environmental samples contain tens of thousands of features, we cannot expect that the small subset of features that are identified and quantified contain all drivers of the mixture risk. Measured BEQ_{bio} values are real even if their causative agents cannot be fully explained.

Outlook

The number of chemicals identified in environmental samples using sophisticated instrumental analysis is steadily increasing, and we have developed better tools to investigate their combined effects and mechanisms of toxicity. However, research is drowning in disconnected details instead of capturing the bigger picture. It is still difficult to elucidate the drivers of chemical stress in the environment; the links among environment, wildlife, and people can only be made by applying common monitoring approaches.

Tracking chemicals and their transformation products in the environment and in our bodies at ever-lower concentrations is an immense (bio)analytical challenge: Sampling, extraction, chemical detection, and data analysis all need to be tuned to each other to obtain robust information. Complementing high-resolution mass spectrometry with bioanalytical tools, especially in vitro bioassays, can now yield information on effects related to all chemicals in a sample, equating to risk-scaled total concentrations. A smart combination of these tools has the potential to revolutionize environmental monitoring. Over time, the storage of NTA data coupled with results and metadata from bioassays will allow "big data" methods to take over and help tease out the relationship between signals

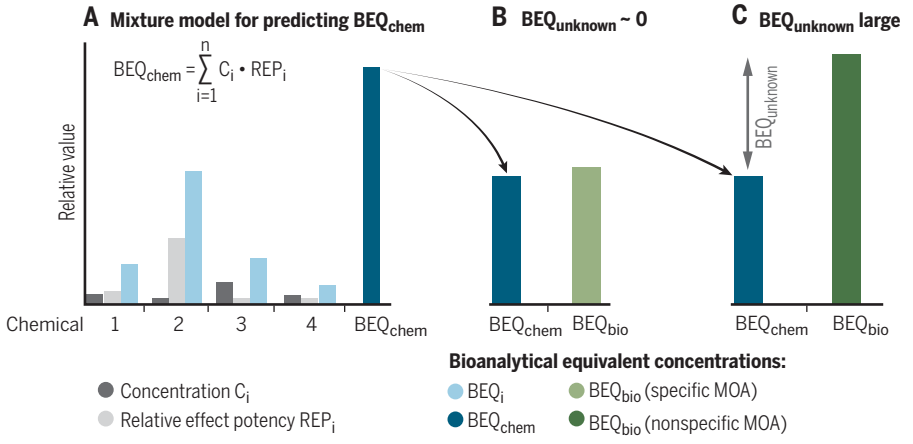


Fig. 4. Mixture toxicity models for linking detected concentrations with predicted and experimental mixture effects. (**A**) Example for calculation of BEQ_{chem} from concentrations C_i and relative effect potencies REP_i. (**B** and **C**) Comparison between the predicted mixture effects of the detected chemicals (BEQ_{chem}) and those directly measured in bioassays (BEQ_{bio}) allows derivation of the effect contribution of the unknown chemicals ($BEQ_{unknown}$). (B) For specific modes of action (MOA), BEQ_{chem} is often close to BEQ_{bio}. (C) For nonspecific effects, $BEQ_{unknown}$ is large.

and toxicity. Quantifying mixture effects is a way to capture all existing chemicals and their bioactive transformation products. Given the clear relevance of mixtures and the fact that thousands of chemicals are occurring in the environment and in our bodies, a shift in the existing regulatory paradigm toward mixture effects is urgently needed (*7*, *35*).

REFERENCES AND NOTES

1. S. M. Rappaport, M. T. Smith, *Science* **330**, 460–461 (2010).
2. P. J. Landrigan *et al*., *Lancet* **391**, 462–512 (2018).
3. R. A. Hites, *Environ. Sci. Technol.* **53**, 8585–8590 (2019).
4. K. Fenner, S. Canonica, L. P. Wackett, M. Elsner, *Science* **341**, 752–758 (2013).
5. B. I. Escher, K. Fenner, *Environ. Sci. Technol.* **45**, 3835–3847 (2011).
6. C. P. Wild, *Cancer Epidemiol. Biomark. Prev.* **14**, 1847–1850 (2005).
7. A. Kortenkamp, M. Faust, *Science* **361**, 224–226 (2018).
8. A. Kortenkamp, *Curr. Opin. Pharmacol.* **19**, 105–111 (2014).
9. N. Cedergreen, *PLOS ONE* **9**, e96580 (2014).
10. A. Boobis *et al*., *Crit. Rev. Toxicol.* **41**, 369–383 (2011).
11. L. M. Bexfield, P. L. Toccalino, K. Belitz, W. T. Foreman, E. T. Furlong, *Environ. Sci. Technol.* **53**, 2950–2960 (2019).
12. C. Ort, M. G. Lawrence, J. Rieckermann, A. Joss, *Environ. Sci. Technol.* **44**, 6024–6035 (2010).
13. B. Zabiegała, A. Kot-Wasik, M. Urbanowicz, J. Namieśnik, *Anal. Bioanal. Chem.* **396**, 273–296 (2010).
14. P. A. Neale *et al*., *Environ. Sci. Process. Impacts* **20**, 493–504 (2018).
15. B. Petrie, R. Barden, B. Kasprzyk-Hordern, *Water Res.* **72**, 3–27 (2015).
16. E. Simon *et al*., *Environ. Sci. Technol.* **45**, 7936–7944 (2011).
17. S. M. Rappaport, D. K. Barupal, D. Wishart, P. Vineis, A. Scalbert, *Environ. Health Perspect.* **122**, 769–774 (2014).
18. S. Poothong *et al*., *Environ. Sci. Technol.* **53**, 1985–1993 (2019).
19. S. G. O'Connell, L. D. Kincl, K. A. Anderson, *Environ. Sci. Technol.* **48**, 3327–3335 (2014).
20. S. C. Hammel, A. L. Phillips, K. Hoffman, H. M. Stapleton, *Environ. Sci. Technol.* **52**, 11875–11885 (2018).
21. J. Hollender, E. L. Schymanski, H. P. Singer, P. L. Ferguson, *Environ. Sci. Technol.* **51**, 11505–11512 (2017).
22. E. L. Schymanski *et al*., *Anal. Bioanal. Chem.* **407**, 6237–6255 (2015).
23. E. L. Schymanski *et al*., *Environ. Sci. Technol.* **48**, 2097–2098 (2014).
24. J. E. Rager *et al*., *Environ. Int.* **88**, 269–280 (2016).
25. N. A. Alygizakis *et al*., *Trends Anal. Chem.* **115**, 129–137 (2019).
26. A. J. Williams *et al*., *J. Cheminform.* **9**, 61 (2017).
27. S. Kim *et al*., *Nucleic Acids Res.* **47**, D1102–D1109 (2019).
28. A. D. McEachran *et al*., *J. Cheminform.* **10**, 45 (2018).
29. M. S. Attene-Ramos *et al*., *Drug Discov. Today* **18**, 716–723 (2013).
30. R. Kavlock *et al*., *Chem. Res. Toxicol.* **25**, 1287–1302 (2012).
31. D. L. Villeneuve *et al*., *Environ. Toxicol. Chem.* **38**, 12–26 (2019).
32. A. Jahnke *et al*., *Environ. Sci. Technol.* **50**, 5424–5431 (2016).
33. R. Judson *et al*., *Toxicol. Sci.* **152**, 323–339 (2016).
34. B. I. Escher, P. A. Neale, D. L. Villeneuve, *Environ. Toxicol. Chem.* **37**, 2273–2280 (2018).
35. W. Brack *et al*., *Environ. Sci. Eur.* **31**, 10 (2019).
36. W. Brack *et al*., *Sci. Total Environ.* **544**, 1073–1118 (2016).
37. E. Simon *et al*., *Environ. Sci. Technol.* **47**, 8902–8912 (2013).
38. Z. Tousova *et al*., *Sci. Total Environ.* **601–602**, 1849–1868 (2017).
39. P. A. Neale *et al*., *Water Res.* **123**, 734–750 (2017).
40. B. R. Blackwell *et al*., *Environ. Sci. Technol.* **53**, 973–983 (2019).

ACKNOWLEDGMENTS

We thank many colleagues for stimulating discussion and the following scientists for critical review of this manuscript: R. Altenburger (UFZ), R. Balling (LCSB), G. Getzinger (Duke), L. Henneberger (UFZ), J. Hollender (Eawag), M. Krauss (UFZ), F. Leusch (Griffith University), P. Neale (Griffith University), R. Schlichting (UFZ), R. Singh (LCSB), and C. Zarfl (EKUT). **Funding:** Supported by the Helmholtz Association under the recruiting initiative scheme, which is funded by the German Ministry of Education and Research (B.I.E.); Luxembourg National Research Fund grant A18/BM/12341006 (E.L.S.); and National Institute of Environmental Health Sciences grant R01 ES 016099 (H.M.S.). **Author contributions:** B.I.E. developed the outline; all authors contributed to the literature study, writing, and revising the manuscript. **Competing interests:** None declared.

10.1126/science.aay6636

REVIEW

The exposome and health: Where chemistry meets biology

Roel Vermeulen[1,2]*, Emma L. Schymanski[3], Albert-László Barabási[4,5,6], Gary W. Miller[7]*

Despite extensive evidence showing that exposure to specific chemicals can lead to disease, current research approaches and regulatory policies fail to address the chemical complexity of our world. To safeguard current and future generations from the increasing number of chemicals polluting our environment, a systematic and agnostic approach is needed. The "exposome" concept strives to capture the diversity and range of exposures to synthetic chemicals, dietary constituents, psychosocial stressors, and physical factors, as well as their corresponding biological responses. Technological advances such as high-resolution mass spectrometry and network science have allowed us to take the first steps toward a comprehensive assessment of the exposome. Given the increased recognition of the dominant role that nongenetic factors play in disease, an effort to characterize the exposome at a scale comparable to that of the human genome is warranted.

A basic tenet of biology is that the phenotype results from a combination of genes and environment. The field of genomics has provided an extraordinary level of genetic knowledge, aided by large-scale, unbiased genome-wide association studies (GWAS). A similar level of analysis, however, is still lacking for the environmental influences on the phenotype. The "exposome" concept was conceived by C. P. Wild in 2005 as a way to represent the environmental, i.e., nongenetic, drivers of health and disease (*1*). For these external forces to have an effect on health, they must alter our biology, suggesting that a detailed analysis of accessible biological samples at different molecular levels, coupled with information provide snapshots of both the internal (biological perturbations) and external contributors to the exposome. As Rappaport and Smith described in 2010, "toxic effects are mediated through chemicals that alter critical molecules, cells, and physiological processes inside the body … under this view, exposures are not restricted to chemicals (toxicants) entering the body from air, water, or food, for example, but also include chemicals produced by inflammation, oxidative stress, lipid peroxidation, infections, gut flora, and other natural processes" (*2*). The exceptional variety and dynamic nature of nongenetic factors (Fig. 1) presents us with an array of sampling and analytical challenges. Fifteen years after the exposome concept was introduced, this review discusses progress in assessing the chemical component of the exposome and its implications on human health.

"…9 million deaths per year…were attributed to air, water, and soil pollution alone."

From environment to genes

Mapping the human genome revolutionized our ability to explore the genetic origins of disease, but also revealed the limited predictive power of individual genetic variation for many common diseases. For example, genetics contributes to less than half of the risk for heart disease, the leading source of mortality in the United States and many other parts of the world (*3*). The health impact of environmental risk factors was highlighted by the Global Burden of Disease (GBD) project, which estimated the disease burden of 84 metabolic, environmental, occupational, and behavioral risk factors in 195 countries and territories, and found that these modifiable risks contribute to ~60% of deaths worldwide (*4*). Using established causal exposure–disease associations, 9 million deaths per year (16% of all deaths worldwide) were attributed to air, water, and soil pollution alone (*5*). However, the true impact of the environment is likely to be grossly underestimated by these studies, as many of the known chemicals of concern were not considered and less than half of the nongenetic risk burden was explained, suggesting the existence of missing exposome factors (*4*). These missing factors are analogous to the missing heritability challenge observed in genetic studies. Even with this incomplete inventory, the economic costs of chemical pollution are considerable, with healthcare and disability-related productivity loss estimated at $4.6 trillion U.S. dollars per year, representing 6.2% of global economic output (*5*). Reducing or preventing chemical pollution is a multifaceted problem that involves medical, legal, and regulatory input (see Box 1).

Measuring chemicals en masse

Several research efforts have pioneered different approaches for the systematic mapping of the exposome, taking advantage of developments in mass spectrometry, sensors, wearables, study design, biostatistics, and bioinformatics (*6*)—advances that now position us to pursue Dr. Wild's original vision of the exposome (*1*). A prime example is how high-resolution mass spectrometry (HRMS) has transformed our ability to measure multitudinous chemical species in a wide range of media, expanding our analytical window beyond targeted analysis of well-known metabolites and priority pollutants (*7*). HRMS provides the means to simultaneously measure a vast number of exogenous and endogenous compounds, offering a description of the system and its changes in response to exposure to environmental factors (*6*, *8*). As Fig. 2 (top panel) indicates, HRMS is capable of measuring thousands to tens of thousands of chemical features in a single analytical run, although most of these features remain unannotated. Although the systems biology approaches in metabolomics originally focused on detecting endogenous metabolites, HRMS methods can also detect exogenously derived small molecules such as pharmaceuticals, pesticides, plasticizers, flame retardants, preservatives, and microbial metabolites (*9*). Historically, these exogenous compounds were often viewed as noise and artifact but in reality they carry direct evidence of the complex environments to which living organisms are exposed.

Data resources relevant for HRMS-based exposomics range from specialized lists [e.g., (*10*)] to medium-sized databases containing tens to hundreds of thousands of chemicals, through to huge resources such as PubChem (*11*), which has 96 million entries (see Fig. 2). Of the >140,000 chemicals produced and used heavily since the 1950s, only ~5000 are estimated to be dispersed in the environment widely enough to pose a global threat to the human population, although many thousands more would be expected to affect individuals, local communities, or specific occupational settings (*5*). Specialized lists compiled by, for example, the U.S. Environmental Protection Agency (EPA) (*12*) and environmental communities such as the NORMAN Network (*13*) often contain additional information (e.g., exposure data and product information) to help annotate chemicals of interest in the study context. Medium-sized databases such as Human Metabolome Database (HMDB) (*14*)

[1]Institute for Risk Assessment Sciences, Division of Environmental Epidemiology, Utrecht University, Utrecht, the Netherlands. [2]Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht, the Netherlands. [3]Luxembourg Centre for Systems Biomedicine, University of Luxembourg, Belvaux, Luxembourg. [4]Network Science Institute, Northeastern University, Boston, MA, USA. [5]Department of Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA. [6]Department of Network and Data Science, Central European University, Budapest, Hungary. [7]Department of Environmental Health Sciences, Mailman School of Public Health, Columbia University, New York, NY, USA.
*Corresponding author. Email: R.C.H.Vermeulen@uu.nl (R.V.); gm2815@cumc.columbia.edu (G.W.M.)

are commonly used in approaches involving metabolic network analysis, offering typically one to a few possible chemicals per feature of exact mass detected by HRMS. Databases that contain spectral information (i.e., structural "fingerprints") can be used to increase the confidence of exact mass matching when experimental fragmentation information is available (*15*, *16*). Comprehensive chemical resources such as PubChem are so large that they often offer several thousand possible chemical candidates per exact mass. Despite the exceptional size of the chemical space, the knowledge and computational tools required to interrogate these data are increasingly available (*15*, *17*). For instance, incorporation of literature and patent information with in silico methods has greatly improved annotation rates (from <22% to >70%) for >1200 chemicals in HRMS experiments using PubChem (*17*).

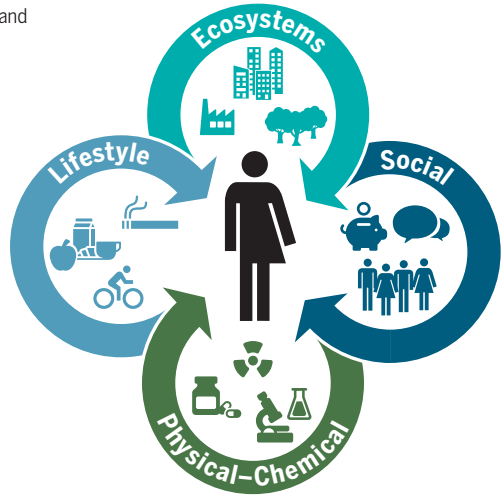
Chemicals are not static entities; they react in our bodies and the environment to form metabolites or transformation products. Computational tools exist to predict such metabolic and environmental transformations (*15*, *18*) but often produce many false-positive and false-negative candidates. Merely predicting first-order reactions of PubChem chemicals would result in billions of possibilities (Fig. 2, second row from bottom). As a result, few studies so far have been able to successfully capitalize on this information in high-throughput identification efforts. The dispersed nature of the essential chemical, metabolite, and spectral information across a wide range of resources with various formats and forms of accessibility (fully open, academic use only, commercial, etc.) is a major impediment to progress in the field.

Integrating chemical knowledge

The interconnected nature of the available chemical information indicates the need for an interdisciplinary and integrative approach to further define the exposome and the associated data science challenges. Literature mining of PubMed and mapping to discrete chemicals can be used to compile and synthesize the chemical information in the scientific literature (*10*, *19*). The expansion and automation of literature mining for more accurate chemical candidate retrieval during high-throughput identification, e.g., with MetFrag (*17*) or other in silico approaches (*15*), will be crucial for faster, more efficient annotation of the complex and highly varying datasets that characterize studies of chemical exposures and health.

Many of the chemicals of interest in exposomics come from the same or related sources (e.g., industrial processes, consumer goods, diet), meaning that such exposures exhibit a population structure (i.e., complex correlations and dynamic patterns) akin to observed correlations in complex biological systems. Thus, the reduction of dimensional complexity will be possible by grouping correlated exposures. Indeed, several reports have shown correlation patterns between different chemicals and chemical families within populations (*20*, *21*). These relationships between chemicals can be presented as networks of chemicals (i.e., exposure enrichment pathways) that reveal communities of exposures (*20*, *21*), which in turn can be used to explore the impact that they have on the biological system (see the following section).

Much of our current knowledge about the health effects of chemicals comes from epidemiological and toxicological studies in which a few pollutants are analyzed in relation to a specific phenotype, representing a hypothesis-driven path toward understanding exposure–disease relationships. However, our exposures are not a simple sum of a handful of chemicals. To overcome the limitations of traditional epidemiological studies, environment-wide association studies (EWAS) have been proposed for identifying new environmental factors in disease and disease-related phenotypes at scale. EWAS was inspired by the analytical procedures developed in GWAS (*22*) in which a panel of "exposures," analogous to genotype variants, is studied in relation to a phenotype of interest. For example, using the National Health and Nutrition Examination Survey dataset, an EWAS study explored the associations of 543 environmental attributes with type 2 diabetes, identifying five statistically



Ecosystems
Food outlets, alcohol outlets
Built environment and urban land uses
Population density
Walkability
Green/blue space

Lifestyle
Physical activity
Sleep behavior
Diet
Drug use
Smoking
Alcohol use

Social
Household income
Inequality
Social capital
Social networks
Cultural norms
Cultural capital
Psychological and mental stress

Physical–Chemical
Temperature/humidity
Electromagnetic fields
Ambient light
Odor and noise
Point, line sources, e.g. factories, ports
Outdoor and indoor air pollution
Agricultural activities, livestock
Pollen/mold/fungus
Pesticides
Fragrance products
Flame retardants (PBDEs)
Persistent organic pollutants
Plastic and plasticizers
Food contaminants
Soil contaminants
Drinking water contamination
Groundwater contamination
Surface water contamination
Occupational exposures

Fig. 1. The exposome concept. The exposome is an integrated function of exposure on our body, including what we eat and do, our experiences, and where we live and work. The chemical exposome is an important and integral part of the exposome concept. Examples of external stressors are adapted from (*39*). These stressors are reflected in internal biological perturbations (Fig. 3); therefore, exposures are not restricted to chemicals (toxicants) entering the body, but also include chemicals produced by biological and other natural processes.

Box 1. The exposome and regulation.

Many of the influential regulatory bodies in Europe and North America have been expanding their computational and high-throughput approaches to address the increasing number of chemicals to which humans are exposed, but there are still major challenges regarding prioritization. Networks such as NORMAN (*13*), which bridge scientists, regulators, and practitioners, are becoming increasingly valuable avenues of knowledge exchange. Large-scale exposome studies provide a systematic approach to prioritization, allowing regulatory bodies to focus on those chemicals that have the largest adverse effects on health. If systematic analysis reveals major adverse effects on human health from exposure to currently approved or potential replacement chemicals, then those compounds should be removed from the marketplace. Although thousands of compounds are classified as "generally recognized as safe," they have never been subjected to the scientifically rigorous testing systems currently in place. A data-driven exposome approach ignores historical decision-making and can help to evaluate the effects of classes of chemicals on specific biological pathways known to be perturbed, which will help in the design of new compounds with minimal impact on human health and the environment.

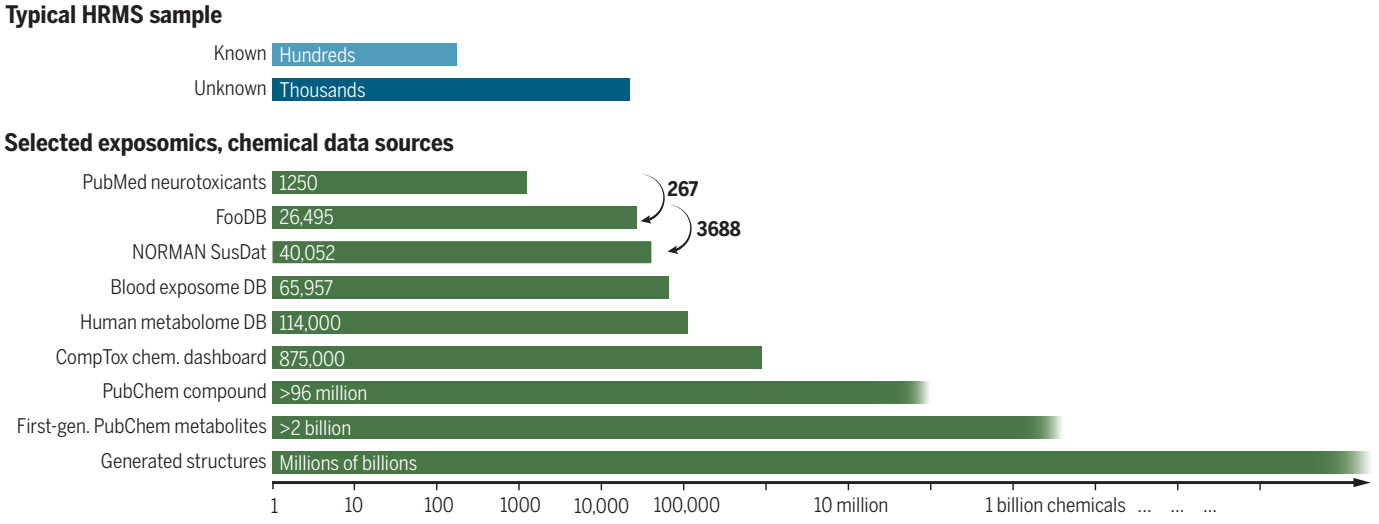


Fig. 2. Chemical complexity of HRMS and the exposome. Top: Known versus unknown features in a typical HRMS measurement [data from (*7*)]. Bottom: Selected data sources relevant to the chemical exposome (*10*–*14*, *19*). Arrows show the overlap of potential neurotoxicants in FooDB (http://foodb.ca/) and FooDB components in NORMAN SusDat (www.norman-network.com/nds/susdat/) (prioritized chemicals of environmental interest).

significant associations (including persistent organic pollutants and pesticides) validated across independent cohorts (*22*). However, by focusing on a predetermined list of chemicals, these initial EWAS studies likely suffer from the same limitations of candidate gene searches. Further, current EWAS approaches do not test for interactions and/or combinations of factors (mixtures). Recent efforts have been undertaken to develop statistical methods to screen for interactions and test the effects of mixtures or to apply frameworks such as aggregated exposure and adverse outcome pathways to study combinatorial effects (*9*).

As systematic exposomics moves forward to elucidate the impact of the constellation of chemical exposures on our health, increasingly rich and high-dimensional data must be captured (Fig. 3). In addition, defining the appropriate frameworks for establishing controls, as well as background and negative responses, is essential for enabling causal inference. To aid inference, more insights into the boundaries of what are "normal" responses are required and necessitate definitions of a reference exposome.

"Network science...offers a platform...to [understand]... the impact of multiple exposures."

Network science to address exposome complexity

The challenge in understanding the role of the exposome on our health lies not only in the large number of chemical exposures in our daily lives, but also in the complex ways that they interact with cells. A reductionist approach might isolate the role of a single variable, but it will inadequately capture the complexity of the exposome. Network science (*23*), which has well-developed applications in medicine and systems biology (*24*), offers a platform with which to achieve an understanding of the impact of multiple exposures. Each chemical will exert its effect through interactions with various cellular components supplying or perturbing cellular networks. To capture the diversity in these interactions, we must first catalog the sum of all physical interactions as a multilayer network (*25*) consisting of several distinct biological layers (Fig. 3). Although each of these networks will rely on different biological mechanisms, they are not independent; for example, protein production is governed by the regulatory network, and the catalysis of the metabolic reactions is in turn governed by the enzymes and protein complexes of the regulatory network (*26*).

To fully understand the role of the exposome, we must similarly develop a multilayer network–based framework capable of unveiling the role of chemicals, their combinations, and biological perturbations on our health. However, there are several data and methodological challenges. The first challenge is the paucity of systematic data on the various dimensions of exposure, from bioavailability to protein-binding information of the hundreds of thousands of exposome molecules. The U.S. National Toxicology Program, the EPA, and the European Molecular Biology Laboratory (EMBL) are developing platforms to generate, collate, and organize data on chemical–biological interactions, but there is a need for high-throughput approaches that offer greater coverage (*12*, *27*, *28*). The second challenge in developing a framework is that the current statistical toolset assumes that we are faced with a collection of random variables that are independent, identically distributed, and measured with equal precision. In a network environment, these assumptions are inherently false, as interactions couple the probability distribution of most network-based variables. Furthermore, most of the chemicals we are exposed to represent communities of exposures, so the effect of a chemical is rarely observed in isolation. Therefore, identifying meaningful associations from high-dimensional exposomic data poses major statistical and computational challenges that need to be addressed in parallel with experimental developments. The third challenge is that, beyond cataloging interactions, we must also understand the dynamics of the biochemical pathways (*29*) through which different elements of the exposome affect our health. Indeed, the human interactome, representing the sum of all physical interactions within a cell (Fig. 3), is often depicted as a static graph but is in reality a temporal network (*30*) with nodes and links that disappear and reemerge depending on factors ranging from the cell cycle to variability in environmental exposures across the life course. Modeling the fully temporal nature of these networks remains a challenge, as the kinetic constants underlying metabolic processes are not known and we currently lack systematic tools with which to identify them (*31*).

Informative exposome study designs

A systematic and unbiased assessment of the exposome that does not focus on a selected set of readily measured or priority chemicals requires access to biological samples that

reflect exposures, biological effects, and, preferably, the health phenotype of interest. This is challenging because it will be rare that the variability of exposures (E) aligns perfectly with the kinetics of the biological effects (B) or the etiological time window of the health phenotype, including developmental and transgenerational effects (P). Optimizing each step (E–B and B–P) in separate studies, however, has the disadvantage that overlapping patterns in each step restrict us from unveiling the true association between exposure and the health phenotype (E–P). The meet-in-the-middle (MITM) design attempts to address this challenge (*32*). In MITM, exposures can be assessed in individuals using HRMS or upstream estimates of external factors (Fig. 1) and are compared with downstream biological changes in persons who develop a specific health phenotype and those who do not.

The MITM approach using HRMS data has successfully identified single and combinatorial effects of chemicals (*33*–*36*). For example, the HELIX study explored the early-life exposome of population-based birth cohorts and identified several environmental chemicals that were associated with lung function in children (*35*). The EXPOsOMICs study showed how air pollution alters biological pathways, particularly linoleate metabolism, which predicted the occurrence of adult-onset asthma and cardiovascular disease (*36*).

Scaling up

By pooling studies, sample sizes for GWAS have increased from a few thousand to tens to hundreds of thousands of individuals over the past decade (*37*). However, enrollment in studies of nongenetic environmental exposures remains relatively low. The large-scale genomic consortia efforts allowed GWAS to detect many common genetic traits related to health phenotypes and, although the combined effects of the identified traits are still modest, they provide insights into the underlying biological pathways of disease. It is estimated that sample sizes of 500,000 to 2,000,000 are needed to explain a substantial portion of the projected genomic heritability of common chronic diseases (*38*). For the multitude of factors within the exposome, most of which likely exert small effects, similar or even greater sample sizes would be required for future environmental studies and EWAS (*22*). Scaling exposome research to these numbers will require a joint effort across multiple cohort consortia and research programs. Recently funded programs to work toward a human exposome project are a first step toward reaching tens of thousands of people with detailed environmental and biological analysis of exposures. Although these numbers are large enough to identify the most prevalent and strongest chemical risk factors, progressive increments in sample size will be needed for a systematic understanding of the impact of combinatorial exposome factors on specific and rare phenotypes. The systematic identification of the impact of nongenetic factors and chemical exposures would enable the establishment of an exposome risk score (ERS) akin to the polygenic risk score (PRS) (see Box 2).

Next steps for the exposome

The rate, volume, and variety of chemicals being introduced into our environment continue to expand. The importance of these chemical exposures on human health is exemplified by the large proportion of disease caused by as yet unknown exposome factors (*3*). Indeed, the nongenetic component exceeds known and missing heritability. We therefore need innovative ways to study these factors and translate our findings into policy. Currently, many of the regulatory agencies are expanding their computational and high-throughput approaches to account for the ever-increasing number of chemicals, but there are still major challenges regarding prioritization and new approaches are urgently needed (see Box 1). Open science efforts such as Global Natural Product Social Molecular Networking (GNPS), which allow users to archive huge amounts of raw data and in return offers computational mass spectrometry workflows coupled with open mass spectral libraries and continuous updates of new identifications, are beginning to be leveraged for large-scale studies (*20*). However, as discussed above, we must address several challenges to exploit the full potential of exposome research as it relates to improving our understanding of exposure to complex chemical mixtures. To address these challenges, we must: (i) improve our technology to screen for exogenous chemicals and their biological consequences at higher-throughput rates and lower costs; (ii) continue to develop

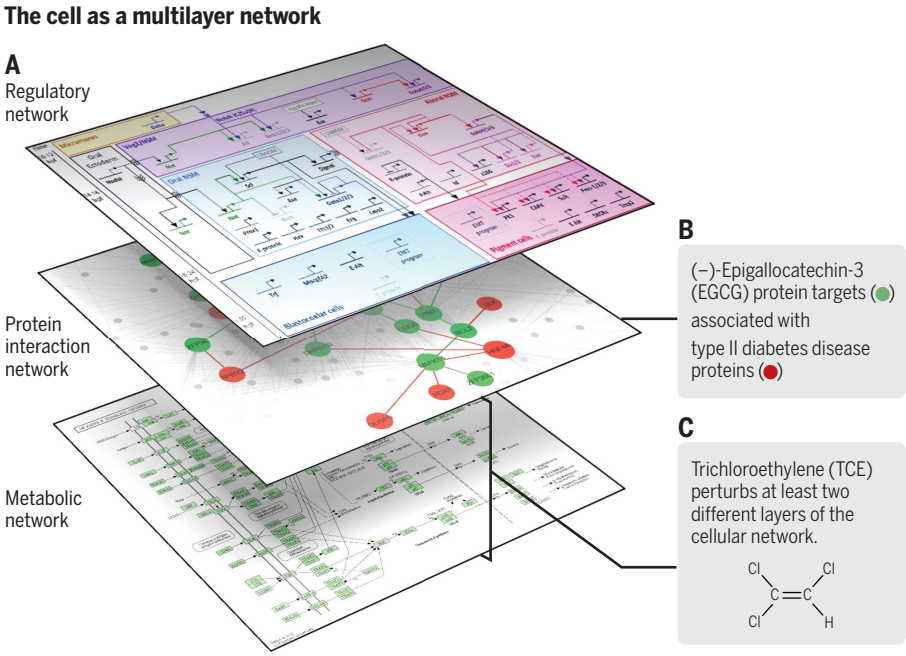


Fig. 3. Impact of the exposome on subcellular networks. (**A**) Network medicine views the cell as a multilayer network with three principal, interdependent layers: (i) a regulatory network capturing all interactions affecting RNA and protein expression, (ii) a protein interaction network that captures all binding interactions responsible for the formation of protein complexes and signaling, and (iii) a metabolic network representing all metabolic reactions, including those derived from the microbiome, a network of interacting bacteria linked through the exchange of metabolites. Exposome-related factors can affect each layer of this multilayer network. (**B**) For example, the polyphenol epigallocatechin gallate (EGCG), a biochemical compound in green tea with potential therapeutic effects on type 2 diabetes mellitus (T2D), binds to at least 52 proteins (*40*). Network-based metrics reveal a proximity between these targets and 83 proteins associated with T2D, suggesting multiple mechanistic pathways to potentially account for the relationship between green tea consumption and reduced risk of T2D. (**C**) As another example, trichloroethylene (TCE) is a volatile organic compound that was widely used in industrial settings and is now a widespread environmental contaminant present in drinking water, indoor environments, ambient air, groundwater, and soil. Multiple lines of evidence support a link between TCE exposure and kidney cancer and possibly non-Hodgkin's lymphoma (*33*). TCE perturbs at least two different layers of the cellular network: It covalently binds to proteins from the protein interaction network, altering their function, and affects the cellular metabolic network, eventually leading to adenosine triphosphate (ATP) depletion. Network-based tools could be used to explore the mechanistic role of many other exposome chemicals on our health and to build experimentally testable hypotheses.

the current chemical and spectral data resources needed to identify these signals in samples; (iii) increase the scale and scope of studies to a level that provides the necessary statistical power to precisely characterize the effects of the chemicals and their combinations; (iv) further develop and support cheminformatic and bioinformatic tools, including network theory and network medicine, to elucidate the constellation of the chemical environment and its biological consequences; and (v) ensure adequate protection for the generation of false-positive results by insisting on replication in independent studies and the use of methods to establish causation, such as Mendelian randomization, within-sibling comparisons, and exposure-negative and outcome-negative controls.

Box 2. Toward an ERS.

There has been substantial progress in the identification of genetic risk factors for chronic diseases. Analysis of high-risk mutations and estimation of PRS for these diseases are now becoming routine and can be included when developing individual-based (i.e., precision) prevention and treatment strategies. Similarly, the establishment of ERS would help to summarize relevant nongenetic risk factors, enabling the identification of hotspots of concern where multiple environmental factors come together, and would aid in the prioritization of risk factors on the basis of their population and individual impact. For example, an ERS could provide data on exposure to chemical toxicants that are based on the biological processes or organ systems that are most vulnerable and couple them with indices of associated biological injury or response. Such ERS, in contrast to PRS, would be time varying and dynamic through age-related exposures and susceptibilities.

Conclusion

A concerted and systematic effort to profile the nongenetic factors associated with disease and health outcomes is urgently needed because we lack important insights that might assist us in curtailing the ever-growing burden of chronic disease on society. Emerging exposome research frameworks are poised to enable the systematic analysis of nongenetic factors involved in disease. Technology has enabled the first generation of studies to evolve into the comprehensive study of combinatorial chemical exposures. Future efforts must ensure that analytical approaches and study designs are rigorous and validated. A coordinated and international effort to characterize the exposome, akin to the Human Genome Project, would provide rigorous data to allow exposome-based EWAS to be conducted at the scale of GWAS. By taking advantage of the nontargeted nature of HRMS, EWAS provide a true complement to GWAS. Consolidating knowledge garnered from GWAS and EWAS would allow us to map the gene and environment interface, which is where nature meets nurture and chemistry meets biology.

REFERENCES AND NOTES

1. C. P. Wild, *Cancer Epidemiol. Biomarkers Prev.* **14**, 1847–1850 (2005).
2. S. M. Rappaport, M. T. Smith, *Science* **330**, 460–461 (2010).
3. J. P. Ioannidis, E. Y. Loy, R. Poulton, K. S. Chia, *Sci. Transl. Med.* **1**, 7ps8 (2009).
4. GBD 2016 Risk Factors Collaborators, *Lancet* **390**, 1345–1422 (2017).
5. P. J. Landrigan *et al*., *Lancet* **391**, 462–512 (2018).
6. M. M. Niedzwiecki *et al*., *Annu. Rev. Pharmacol. Toxicol.* **59**, 107–127 (2019).
7. E. L. Schymanski *et al*., *Environ. Sci. Technol.* **48**, 1811–1818 (2014).
8. D. C. Sévin, A. Kuehne, N. Zamboni, U. Sauer, *Curr. Opin. Biotechnol.* **34**, 1–8 (2015).
9. B. I. Escher *et al*., *Environ. Int.* **99**, 97–106 (2017).
10. E. L. Schymanski *et al*., *Environ. Sci. Process. Impacts* **21**, 1426–1445 (2019).
11. S. Kim *et al*., *Nucleic Acids Res.* **47** (D1), D1102–D1109 (2019).
12. A. J. Williams *et al*., *J. Cheminform.* **9**, 61 (2017).
13. V. Dulio *et al*., *Environ. Sci. Eur.* **30**, 5 (2018).
14. D. S. Wishart *et al*., *Nucleic Acids Res.* **46** (D1), D608–D617 (2018).
15. I. Blaženović, T. Kind, J. Ji, O. Fiehn, *Metabolites* **8**, 31 (2018).
16. E. L. Schymanski *et al*., *Environ. Sci. Technol.* **48**, 2097–2098 (2014).
17. C. Ruttkies, E. L. Schymanski, S. Wolf, J. Hollender, S. Neumann, *J. Cheminform.* **8**, 3 (2016).
18. Y. Djoumbou-Feunang *et al*., *J. Cheminform.* **11**, 2 (2019).
19. D. K. Barupal, O. Fiehn, *Environ. Health Perspect.* **127**, 097008 (2019).
20. J. M. Gauglitz *et al*., *Food Chem.* **302**, 125290 (2020).
21. S. Li *et al*., *Reprod. Toxicol.* S0890-6238(18)30603-8 (2019).
22. C. J. Patel, J. Bhattacharya, A. J. Butte, *PLOS ONE* **5**, e10746 (2010).
23. A.-L. Barabási, *Network Science* (Cambridge Univ. Press, 2016).
24. J. Loscalzo, A.-L. Barabási, E. K. Silverman, *Network Medicine: Complex Systems in Human Disease and Therapeutics*. (Harvard Univ. Press, 2017).
25. G. Bianconi, *Multilayer Networks: Structure and Function* (Oxford Univ. Press, ed. 1, 2018).
26. A.-L. Barabási, Z. N. Oltvai, *Nat. Rev. Genet.* **5**, 101–113 (2004).
27. D. Mendez *et al*., *Nucleic Acids Res.* **47** (D1), D930–D940 (2019).
28. R. R. Tice, C. P. Austin, R. J. Kavlock, J. R. Bucher, *Environ. Health Perspect.* **121**, 756–765 (2013).
29. A. Barrat, M. Barthelemy, A. Vespignani, *Dynamical Processes on Complex Networks* (Cambridge Univ. Press, 2008).
30. P. Holme, J. Saramäki, "Temporal networks as a modeling framework," in *Understanding Complex Systems* (Springer, 2013), pp 1–14.
31. M. Santolini, A.-L. Barabási, *Proc. Natl. Acad. Sci. U.S.A.* **115**, E6375–E6383 (2018).
32. M. Chadeau-Hyam *et al*., *Environ. Mol. Mutagen.* **54**, 542–557 (2013).
33. D. I. Walker *et al*., *Int. J. Epidemiol.* **45**, 1517–1527 (2016).
34. B. Warth *et al*., *Anal. Chem.* **89**, 11505–11513 (2017).
35. L. Agier *et al*., *Lancet Planet. Health* **3**, e81–e92 (2019).
36. A. Jeong *et al*., *Environ. Int.* **119**, 334–345 (2018).
37. T. Beck, R. K. Hastings, S. Gollapudi, R. C. Free, A. J. Brookes, *Eur. J. Hum. Genet.* **22**, 949–952 (2014).
38. Y. Zhang, G. Qi, J. H. Park, N. Chatterjee, *Nat. Genet.* **50**, 1318–1326 (2018).
39. M. C. Turner *et al*., *Annu. Rev. Public Health* **38**, 215–239 (2017).
40. H. Iso, C. Date, K. Wakai, M. Fukui, A. Tamakoshi; JACC Study Group, *Ann. Intern. Med.* **144**, 554–562 (2006).

ACKNOWLEDGMENTS

We thank the following colleagues for critical review of this manuscript: R. Balling, M. Chadeau-Hyam, G. Downward, L. P. Fried, D. P. Jones, V. Kalia, V. Lenters, G. Menichetti, R. Singh, Í. Valle, B. van de Water, and J. Vlaanderen. **Funding:** R.V. is supported by an EU H2020-EXPANSE grant, the NWO Gravitation Program, and intramural funding from Utrecht University. E.L.S. is supported by the Luxembourg National Research Fund (FNR grant no. A18/BM/12341006). G.W.M. is supported by the National Institutes of Health (NIH grant nos. U2ESC030163 and RC2 DK118619). A.-L.B. is supported by the NIH (grant no. P01HL132825) and the American Heart Association (grant no. 151708). **Competing interests:** A.-L.B. is founder of Nomix, Foodome, and Scipher Medicine, companies that explore the role of networks in health. The other authors declare no competing interests. **Data and materials availability:** All data are available in the main text.

10.1126/science.aay3164

REVIEW

Designing for a green chemistry future

Julie B. Zimmerman[1,2,3]*, Paul T. Anastas[2,3,4], Hanno C. Erythropel[1,3], Walter Leitner[5,6]

The material basis of a sustainable society will depend on chemical products and processes that are designed following principles that make them conducive to life. Important inherent properties of molecules need to be considered from the earliest stage—the design stage—to address whether compounds and processes are depleting versus renewable, toxic versus benign, and persistent versus readily degradable. Products, feedstocks, and manufacturing processes will need to integrate the principles of green chemistry and green engineering under an expanded definition of performance that includes sustainability considerations. This transformation will require the best of the traditions of science and innovation coupled with new emerging systems thinking and systems design that begins at the molecular level and results in a positive impact on the global scale.

The scientific question facing the chemical sector when designing for the future Earth is not whether products of the chemical industry will be necessary, because they surely will be. Rather, the question is, what will be the character, nature, and production processes of synthetic chemicals needed for a sustainable civilization? Chemistry has a long history of inventing essential and beneficial products and processes with extraordinary performance; however, this technological progress has often been realized using a narrow definition of function, which does not account for adverse consequences.

Today's chemical sector follows a linear path (Fig. 1, left), in which feedstocks, mostly fossil and finite in nature, are pushed through a production chain that relies on reagents that are designed to be highly reactive but are often also unintentionally persistent and/or toxic, which is consequential for worker exposure as well as accidental or intentional release (e.g., methyl isocyanate release in Bhopal, India, and dioxin spills in Times Beach, Missouri, and Seveso, Italy). Many of these processes generate waste (often itself toxic, persistent, and bioaccumulating), at rates higher than the intended product, particularly as product complexity increases (e.g., 5 to 50 times for specialty chemicals and 25 to 100 times for pharmaceuticals) (*1*). Similarly, the resulting chemical products are often designed for their intended use while relying on circumstantial controls to limit exposures to hazards that have often not been assessed, potentially owing to the historic lack of tools and models, as evidenced by the multitude of unintended adverse consequences (*2*).

Given the need for the many functions provided by the products of the chemical industry, the question, as we look to the future, must include two goals: How do we (i) keep and greatly expand upon the advances in performance while (ii) limiting or eliminating the detrimental impacts that threaten the sustainability of human and planetary well-being? Answering this question is an important and urgent scientific challenge. There is a plethora of achievements in the fields of "green chemistry" (*3*, *4*) and "green engineering" (*5*) that have demonstrated that more performance and functionality from our chemical products and processes can be realized while decreasing adverse impacts. These successes need to be made systematic and not anecdotal. To succeed, not only do the conditions and circumstances by which we make and use chemical products need to be altered but the inherent nature of the chemical products and reagents themselves across the entire value chain from feedstock to application also needs to be changed (Fig. 1). This requires changing the nature of the very definition of "performance" from function alone to function and sustainability, which can only be realized through thoughtful design of the intrinsic properties of the molecules and their transformations.

Call for design and innovation in an integrated systems framework

Pursuing improved design for sustainability in complex systems is acutely challenging when using a traditional reductionist approach (*6*). In the chemical sector, a reductionist focus on function has resulted in pharmaceuticals that not only extend and enhance life and well-being but also persist in our water after their useful life, causing exposure to unintended populations (*7*); agricultural chemicals that increase crop yields while causing fish kills and degrading groundwater (*8*); and chemicals that make materials durable and impart long-term performance but bioaccumulate in our bodies and in the biosphere (*9*). Despite many examples of the failure modes of the reductionist approach, we often continue to use this framework to address sustainability challenges, as evidenced by the focus on isolated, individual metrics (e.g., greenhouse gas emissions, energy or water consumption, and ecotoxicity) instead of on reframing sustainability as a multidimensional problem of nested complex systems (*10*).

Incrementally improving current products and processes through efficiency measures is the focus of many sustainability efforts in the chemical industry, but this approach will not be sufficient. Instead, we need transformative, disruptive innovations to sustainably provide desired functions. Systems must be considered in their entirety to identify solutions that do not shift impacts or cause unintended consequences elsewhere. The traditional reductionist approach, then, must be coupled with integrative systems thinking to inform designs for a sustainable future (*11*). For example, knowledge about the functional performance of a molecule is only a minimal requirement; knowledge about the potential hazards of the molecule is also needed. Instead of solving a singular challenge in ways that may exacerbate other challenges (e.g., trading off bio-based fuels for land-use pressures or competition with food), there is now the opportunity for so-called "nexus solutions" (*12*), that is, solutions that synergistically advance multiple sustainability challenges—for example, an earth abundant metal catalyst that splits water using sunlight, yielding hydrogen for energy storage and producing water when that hydrogen is combusted for energy recovery (*13*). Designing future fuels that are produced in a carbon-neutral way to reduce local air pollution emissions and increase engine efficiency is another example (*14*). Although there are current discussions on cascading nonlinear problems (e.g., increased fossil energy generation→greater water stress→refugee migration→civil unrest and military conflict), systems thinking and design targeting nexus solutions allow for the possibility of cascading nonlinear solutions: discreet and specific actions that create benefits that are multiplied and magnified (e.g., CO_2 utilization shifts a waste to a feedstock→the use of toxic reagents, such as phosgene, is avoided→CO_2 emissions decrease→the rise of CO_2 levels slows→global climate change is mitigated).

Expanding the definition of performance beyond technical function to include sustainability

Realizing the necessary changes to the chemical sector requires a redefinition of how the concept of performance has historically been measured. Since commercial synthetic chemistry was started in the mid-19th century with

[1]Department of Chemical and Environmental Engineering, Yale University, New Haven, CT, USA. [2]School of Forestry and Environmental Studies, Yale University, New Haven, CT, USA. [3]Center for Green Chemistry and Green Engineering, Yale University, New Haven, CT, USA. [4]School of Public Health, Yale University, New Haven, CT, USA. [5]Max Planck Institute for Chemical Energy Conversion, 45470 Mülheim an der Ruhr, Germany. [6]Institut für Technische und Makromolekulare Chemie, RWTH Aachen University, 52074 Aachen, Germany.
*Corresponding author. Email: julie.zimmerman@yale.edu

the introduction of Perkin's mauve dye, the criterion by which chemical products were judged was "performance." Performance has been almost entirely defined as the ability to efficiently accomplish a narrowly defined function (e.g., color for dyes, adhesion for glues, or the ability to kill pests for pesticides). However, the unintended consequences of this singular focus demonstrate the imperative that our definition of performance must broaden to include all of the aspects that we care about beyond function, particularly sustainability.

This expanded definition of performance requires chemical product and process designers to know not only the mechanisms of technical functional performance but also the mechanisms of harm that the substances can induce. This expanded definition of performance implicitly requires that anyone designing, inventing, and intentionally making a chemical product or process must have a working knowledge of the molecular hazard, be it a global, physical, or toxicological hazard. After more than a century of incidents and accidents that have resulted in adverse consequences to human health and the environment, the training for chemists that integrates toxicology deeply into the curriculum remains the exception and not the rule. Enabling the consideration of chemical hazards in the same manner as chemical performance will require that our education programs, like our technologies, expand the definition of function to include sustainability attributes.

The redefinition of performance also directly affects the business model of the chemical sector, because one part of the strategy of diminishing adverse impact is to reduce the absolute quantity of material needed, thereby reducing harm potential across the life cycle. The "F-factor" (*15*) partially captures the concept of maximal functionality: The notion is to maximize function while minimizing chemical or material mass, which is akin to the Moore's Law challenge for computing power (*16*). Of course, the notion of minimizing mass is intended as a proxy for limiting the amount of feedstock, processing and transport energy, manufacturing waste, end-of-life material for management, and, subsequently, the associated harm. This approach can be applied to whatever service is required—for example, color, lubrication, or cleaning—and will shift the profit imperative from selling as much material as possible to selling as much function as possible with a concomitant decrease in hazard. Such a paradigm shift is in accordance with the United Nations Industrial Development Organization's emerging emphasis on "chemical leasing" (*17*), or selling the function, rather than a quantity, of the chemical. This reduces costs of material production and values efficiency and efficacy in ways that maximize profits in alignment with sustainability (Fig. 1, right).

The inherent nature of tomorrow's chemicals and materials

Tomorrow's chemical products should be designed to preserve efficacy of function while reducing or eliminating hazard (*3*). Here, hazard is defined broadly to consider physical (e.g., explosivity and corrosivity), global (e.g., greenhouse gases and ozone depletion), and toxicological (e.g., carcinogenicity and endocrine disruption) hazards. Chemical risk, defined as a function of both hazard and exposure, is often "managed" by solely focusing on limiting exposure to hazardous chemicals, for example, through the use of protective equipment or air emissions scrubbers. When exposure control mechanisms fail, the results can be catastrophic because the hazard part of the equation has not been addressed. The alternative approach of green chemistry is to shift the focus of risk reduction to reducing hazard. Notably, hazard is an inherent property of chemicals and is ultimately the result of a design choice. Accordingly, a shift to proactive design of chemicals and processes based on an understanding of the molecular mechanisms causing human and environmental toxicity as well as physical and global hazards more broadly is necessary (*18*). As such, an expanded definition of performance should include functionality as well as the inherent nature of the chemicals, including that they are nondepleting, nontoxic, and nonpersistent in the environment.

Nondepleting

Transitioning from fossil- to renewable-based chemistry must be thoughtfully designed within an integrated systems context to consider negative impacts that could be caused by land transformation, water use, or competition with food production. It is imperative that the important move to renewably sourced feedstocks is achieved using benign processes and includes a paradigm shift from linear to circular processes (Fig. 1). Thereby, materials that are currently regarded as low-value waste must be treated as a renewable feedstock in tomorrow's chemical sector (*19*). Examples of utilizing low-value "waste" include the conversion of lignin from papermill waste to produce vanillin and feedstocks such as catechol, phenol, or guaiacols (*20*) and the direct use of CO_2 to partially replace petroleum-based propylene oxide in the production of polyurethanes, which substantially reduces the carbon footprint and simultaneously improves other relevant environmental parameters (*21*, *22*). Going one step further, chemists need to consider "waste design": how synthesis routes can be adapted to avoid relying on treatment and disposal of by-products and instead include design decisions that render by-products useful as raw materials (Fig. 1) (*23*).

Nontoxic

The design of nontoxic chemicals can only be achieved through collaboration between

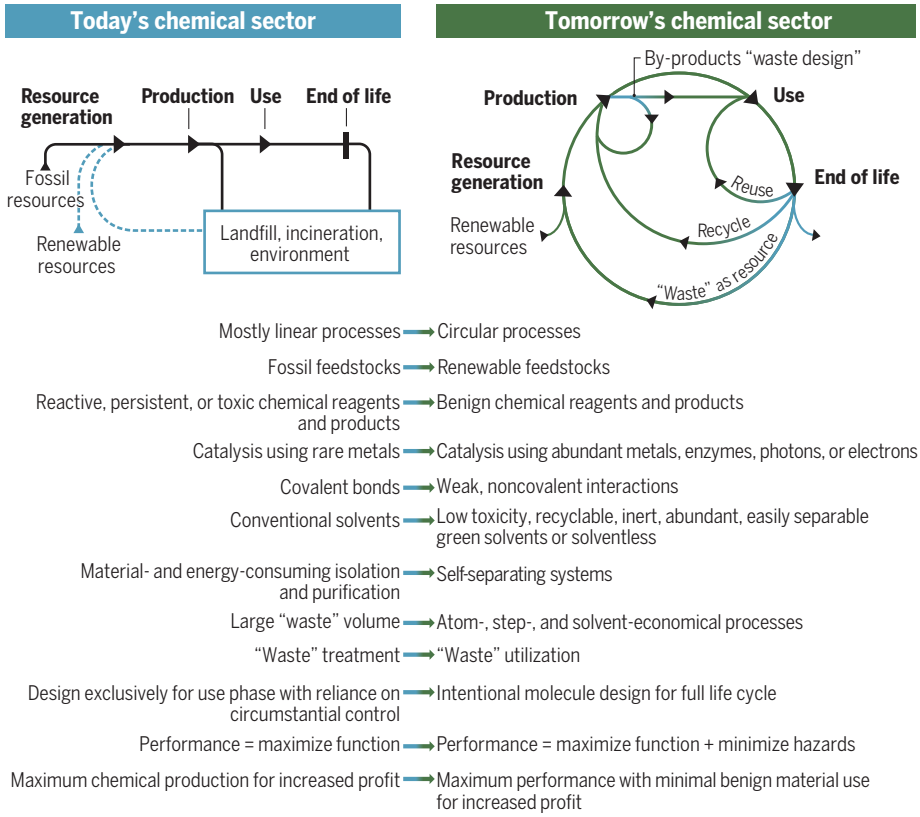


Fig. 1. Characteristics of today's and tomorrow's chemical sectors.

chemistry, toxicology, genomics, and other related fields. It is necessary to investigate and understand the underlying molecular mechanisms, including how molecules are absorbed, distributed, metabolized, and excreted from the body (ADME), and how physicochemical properties, including solubility, reactivity, and cell permeability, influence these events (*24*). Efforts to predict and model toxicity are ongoing, for example, based on molecular initiating events by examining structure-activity relationships [(quantitative) structure-activity relationships QSAR and SAR] (*24*). However, models rely on limited available toxicity data. Current efforts to generate this data include the ToxCast and Tox21 programs in the United States (*25*) and ToxRisk in the European Union (*26*).

Nonpersistent

Tomorrow's chemicals must be designed to degrade easily and rapidly to nonpersistent compounds that do not adversely affect the environment. Recent examples include a rapidly degradable insecticide with very low toxicity to mammals (*27*), chemical modification to improve biodegradability of an otherwise environmentally recalcitrant beta-blocker (*28*), and rapidly biodegradable, nontoxic, and partially renewably sourced succinic acid–based plasticizer additives for polyvinyl chloride (PVC) polymers (*29*). An understanding of the molecular features and environmental mechanisms that lead to persistence are required to build predictive QSAR and SAR models (e.g., the U.S. Environmental Protection Agency's EPI Suite). It is crucial to routinely assess the potential persistence of synthetic compounds for every (newly) designed compound that will be ultimately dispersed in the environment (e.g., pharmaceuticals and personal care products) (*30*).

However, persistence may be a desirable attribute when considering the amounts of energy and molecular complexity embedded in a chemical compound and its synthetic route. An evaluation of whether this "investment" can be harnessed beneficially in a value-added application rather than pursuing a design for a degradation pathway is an important alternative consideration (Fig. 2). For renewably sourced and benign molecules with high levels of complexity and embedded energy and, importantly, that are not intentionally distributed in the environment, the goal should be reintegration into the value chain through reuse or recycle loops. However, if the molecule is intentionally or ultimately distributed in the environment, degradability is imperative, regardless of its complexity or embedded energy.

Redesigning the chemical value chain with nonfossil feedstocks

Today's chemical production relies almost exclusively on oil, gas, and coal as the sources of carbon. The petrochemical value chain that emerged in the second half of the 20th century forms a highly integrated network, sometimes referred to as the "petroleum tree" (Fig. 3). Petrochemical refineries produce less than a dozen building blocks, in particular, short-chain olefins and aromatics. Together with synthesis gas, which is a mixture of CO and H_2, these few compounds form the stem of carbon that constitutes the huge molecular diversity in branches, twigs, and leaves of more than 100,000 chemical substances in final products.

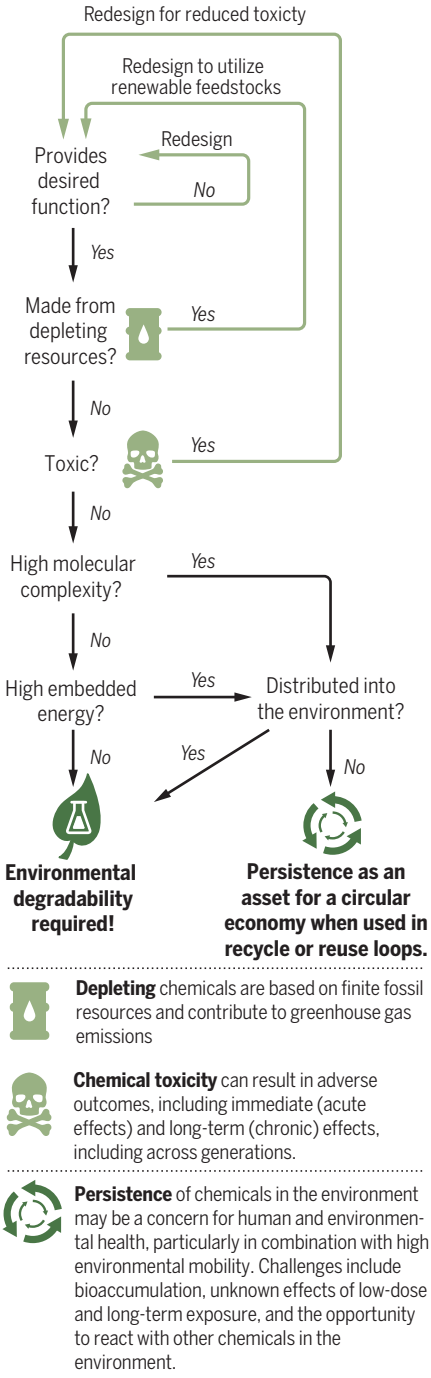


Fig. 2. Decision tree for chemical design.

The value generation of petrochemistry largely results from the powerful arsenal of synthetic methods for the introduction of functional groups into molecules. Consequently, the availability of starting materials and the desired product functions have a direct feedback connection with the development of chemical production routes and processes. Improvements in the corresponding synthetic methods will undoubtedly remain a major field of research with direct positive impacts on the environmental footprint of chemical products (Fig. 3, top, green arrows). Owing to the inherent problems of the depletion of a finite resource, connection to climate change, and inextricable links to toxic congeners, petroleum is not a sustainable option, and ultimately, the design of new value chains based on nonfossil carbon sources and energy from renewable resources will pave the way to closed carbon cycles. This paradigm shift, marking nothing less than the next industrial (r)evolution in chemistry, has already started (Fig. 3, bottom, green arrows).

Major scientific breakthroughs and innovations will be necessary to make the utilization of renewable carbon sources increasingly competitive. Advanced hydrogen technologies (*31*) and electrochemical processes (*32*) are actively researched to harness the required "decarbonized" energy. Among carbon sources, lignocellulosic biomass, one of the most abundantly available raw materials on Earth (*33*), and CO_2 (*34*) are available in sufficient quantities to achieve a "defossilization" of the chemical value chain. Recycled plastic material is another potential large-volume carbon source in the framework of a circular economy (*35*). In contrast to fossil resources made up from carbon and hydrogen only, these feedstocks comprise mostly highly oxidized and "overfunctionalized" molecules. Their transformation, therefore, requires innovative synthetic concepts and methodologies for the removal or reorganization of functional groups.

One option to deal with the challenge of complexity is to remove it—for example, through intermediate syngas production combined with the Fischer-Tropsch technology to produce a petroleum proxy as a drop-in replacement (*36*). Thereby, additional roots can be added to the petroleum tree from almost any nonfossil feedstock. "Green carbon" is then transported through all existing branches and leaves. Although this approach allows for repurposing established knowledge and infrastructure, it destroys and reassembles valuable functionality.

An alternative strategy is to take advantage of the inherent higher complexity of renewable feedstocks to open molecular shortcuts to the desired functional groups. For example, the production of important platform chemicals such as 1,3-propanediol or succinic acid by microbial fermentation of (waste) glycerol or sugars near room temperature in water is

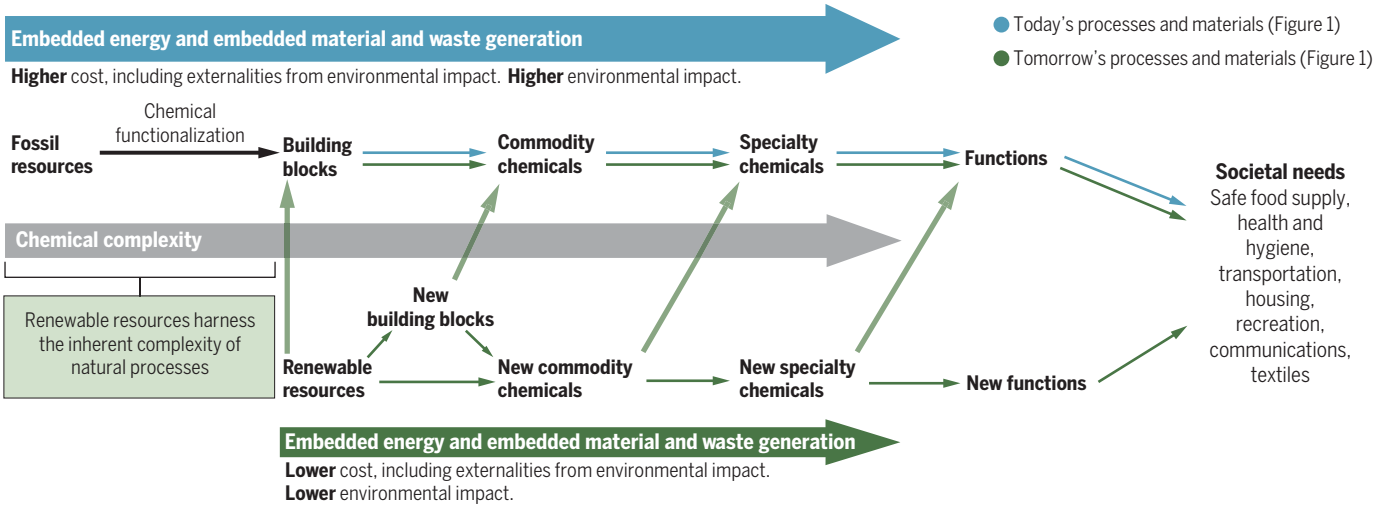


Fig. 3. Benefits of moving from fossil to renewable resources using greener transformation schemes and process chains in terms of embedded energy; embedded materials, including water; waste generation; and environmental and economic costs.

increasingly rivaling the multistep petrochemical routes (*37*). Similar considerations apply to the use of CO_2 that can be integrated into the value chain through existing large-volume base chemicals, methanol being one of the most promising options (*38*). Tailored chemo- and biocatalysts that can be adapted to variations in feedstock quality and fluctuations in energy supply, together with the development of highly integrated and energy-efficient purification processes, will be important scientific motors behind this development.

Perhaps to even greater advantage, renewable feedstocks provide entirely new chemical building blocks (*39*) that can drive the realization of improved functionality without the historic adverse impacts to human health and the environment. One molecule that has recently garnered interest is monosaccharide-derived furane dicarboxylic acid (FDCA), a possible building block for new polyester products such as containers for carbonated liquids (*40*). The incorporation of CO_2 directly into polymer chains of consumer products is already being industrialized (*21*). A growing number of new synthetic methodologies based on selective coupling of CO_2 and H_2 with other substrates illustrates their broad potential for construction of functional groups at late stages of product synthesis (*41*).

The necessary integration of the molecular and engineering sciences with the systems thinking on product performance is obvious in this area. This will enable integration of energy and material fluxes at the chemistry-energy nexus (*42*) and defines new opportunities for sector coupling of chemistry with, for example, agriculture, steel, or cement industries. Deconstructing a highly complex target molecule through "retrosynthetic analysis" to plan its assembly from existing building blocks and synthetic methods is a central pillar of today's synthetic organic chemistry (*43*). The same conceptual thinking can be transferred into a new design framework for synthetic pathways to improved target products from renewable feedstocks (*44*).

Conclusion

Einstein is famously quoted as saying "problems cannot be solved at the same level of awareness that created them" (*45*). The new tools and approaches we need include attaining mastery of weak-force interactions as a design tool, as we have realized with covalent bonds; designing complex, nonideal mixtures rather than synthesizing a single molecule to achieve function; understanding the molecular scale at the complete dynamic reality rather than as a simple, static snapshot; understanding and controlling long-range interactions for chemical reactivity at localized structures; and progressing from mathematical analysis of a series of experiments toward statistical mining of large and diverse datasets. As it is in nature, the concept of "waste" must disappear from our design frameworks, such that we instead think in terms of material and energy flows. Hazards to the inhabitants and systems of the biosphere by our chemical products and processes should be viewed as a critical design flaw and performance defined in terms of both primary functionality and sustainability.

REFERENCES AND NOTES

1. R. A. Sheldon, *Green Chem.* **19**, 18–43 (2017).
2. United Nations Environment Programme (UNEP), "Global chemicals outlook II: From legacies to innovative solutions" (UNEP, 2019).
3. P. T. Anastas, J. C. Warner, *Green Chemistry: Theory and Practice* (Oxford Univ. Press, 1998).
4. H. C. Erythropel *et al.*, *Green Chem.* **20**, 1929–1961 (2018).
5. P. T. Anastas, J. B. Zimmerman, *Environ. Sci. Technol.* **37**, 94A–101A (2003).
6. P. W. Anderson, *Science* **177**, 393–396 (1972).
7. K. Kümmerer, *Chemosphere* **75**, 417–434 (2009).
8. N. V. Fedoroff *et al.*, *Science* **327**, 833–834 (2010).
9. K. C. Jones, P. de Voogt, *Environ. Pollut.* **100**, 209–221 (1999).
10. P. T. Anastas, *Trends Chem.* **1**, 145–148 (2019).
11. S. A. Matlin, G. Mehta, H. Hopf, A. Krief, *Nat. Chem.* **8**, 393–398 (2016).
12. J. Liu *et al.*, *Science* **347**, 1258832 (2015).
13. A. J. Bloomfield, S. W. Sheehan, S. L. Collom, P. T. Anastas, *ACS Sustain. Chem.& Eng.* **3**, 1234–1240 (2015).
14. W. Leitner, J. Klankermayer, S. Pischinger, H. Pitsch, K. Kohse-Höinghaus, *Angew. Chem. Int. Ed.* **56**, 5412–5452 (2017).
15. J. Clark, R. Sheldon, C. Raston, M. Poliakoff, W. Leitner, *Green Chem.* **16**, 18–23 (2014).
16. M. Poliakoff, P. Licence, M. W. George, *Angew. Chem. Int. Ed.* **57**, 12590–12591 (2018).
17. P. Schwager, N. Decker, I. Kaltenegger, *Curr. Opin. Green Sustain. Chem.* **1**, 18–21 (2016).
18. P. Coish *et al.*, *Toxicol. Sci.* **161**, 241–248 (2018).
19. C. O. Tuck, E. Pérez, I. T. Horváth, R. A. Sheldon, M. Poliakoff, *Science* **337**, 695–699 (2012).
20. S. Gillet *et al.*, *Green Chem.* **19**, 4200–4233 (2017).
21. J. Langanke *et al.*, *Green Chem.* **16**, 1865–1870 (2014).
22. N. von der Assen, A. Bardow, *Green Chem.* **16**, 3272–3280 (2014).
23. J. B. Zimmerman, P. T. Anastas, in *Sustainability Science and Engineering*, M. A. Abraham, Ed. (Elsevier, 2006), vol. 1, chap. 10, pp. 201–221.
24. F. Melnikov, J. Kostal, A. Voutchkova-Kostal, J. B. Zimmerman, P. T. Anastas, *Green Chem.* **18**, 4432–4445 (2016).
25. A. M. Richard *et al.*, *Chem. Res. Toxicol.* **29**, 1225–1251 (2016).
26. M. Daneshian, H. Kamp, J. Hengstler, M. Leist, B. van de Water, *Arch. Toxicol.* **90**, 1021–1024 (2016).
27. G. D. Thompson, R. Dutton, T. C. Sparks, *Pest Manag. Sci.* **56**, 696–702 (2000).
28. T. Rastogi, C. Leder, K. Kümmerer, *RSC Advances* **5**, 27–32 (2015).
29. H. C. Erythropel *et al.*, *Polymer (Guildf.)* **89**, 18–27 (2016).
30. A. J. Ebele, M. A.-E. Abdallah, S. Harrad, *Emerg. Contam.* **3**, 1–16 (2017).
31. C. Niether *et al.*, *Nat. Energy* **3**, 476–483 (2018).
32. M. Wang *et al.*, *Nat. Commun.* **10**, 3602 (2019).
33. C. Somerville, H. Youngs, C. Taylor, S. C. Davis, S. P. Long, *Science* **329**, 790–792 (2010).
34. Q. Liu, L. Wu, R. Jackstell, M. Beller, *Nat. Commun.* **6**, 5933 (2015).
35. R. Geyer, J. R. Jambeck, K. L. Law, *Sci. Adv.* **3**, e1700782 (2017).
36. B. Kamm, *Angew. Chem. Int. Ed.* **46**, 5056–5058 (2007).
37. J. H. Ahn, Y.-S. Jang, S. Y. Lee, *Curr. Opin. Biotechnol.* **42**, 54–66 (2016).
38. J. Artz *et al.*, *Chem. Rev.* **118**, 434–504 (2018).
39. A. Corma, S. Iborra, A. Velty, *Chem. Rev.* **107**, 2411–2502 (2007).
40. I. Delidovich *et al.*, *Chem. Rev.* **116**, 1540–1599 (2016).
41. J. Klankermayer, S. Wesselbaum, K. Beydoun, W. Leitner, *Angew. Chem. Int. Ed.* **55**, 7296–7343 (2016).
42. R. Schlögl, *Angew. Chem. Int. Ed.* **50**, 6424–6426 (2011).
43. E. J. Corey, *Angew. Chem. Int. Ed.* **30**, 455–465 (1991).
44. J. Klankermayer, W. Leitner, *Science* **350**, 629–630 (2015).
45. A. Einstein, *New York Times*, 25 May 1946, p. 11.

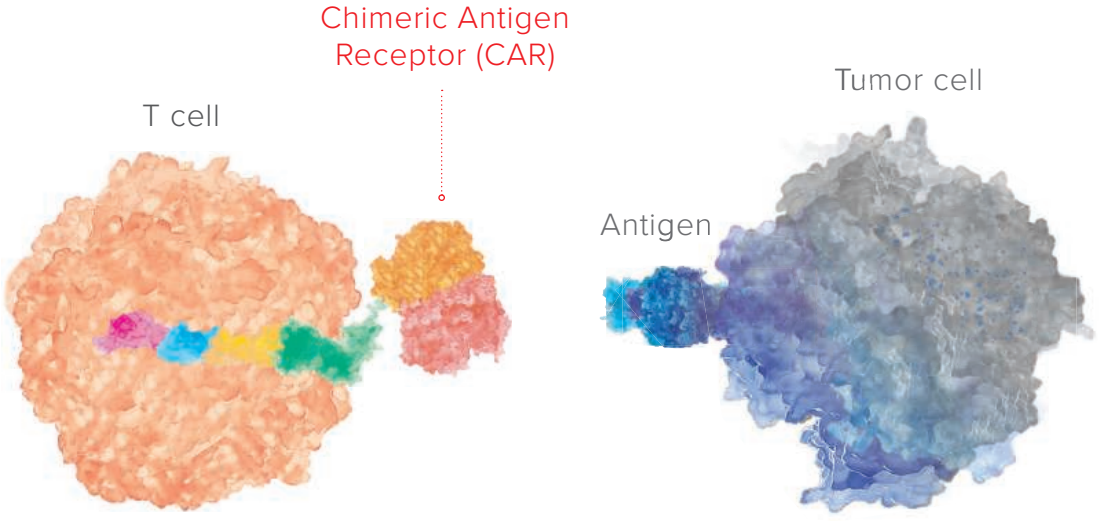
ACKNOWLEDGMENTS

Funding: J.B.Z. and P.T.A. acknowledge research funding by the Molecular Design Research Network (1339637), which is supported by the National Science Foundation. W.L. acknowledges research funding by the "Fuel Science Center" funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy–Exzellenzcluster 2186, the Fuel Science Center (ID 390919832), and the Kopernikus Project Power-to-X funded by the Bundesministerium für Bildung und Forschung (BMBF, ID 03SFK2A). **Competing interests:** The authors declare no competing interests.

10.1126/science.aay3060

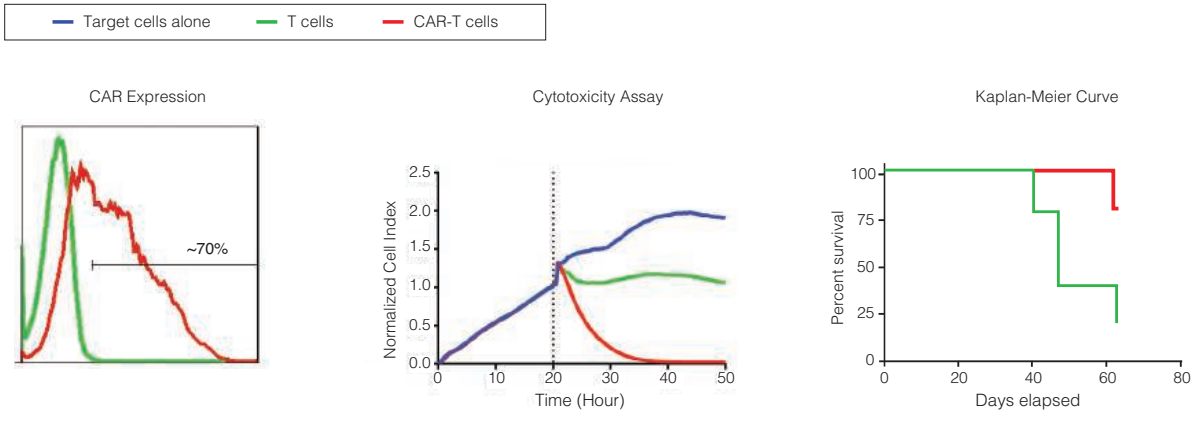
CAR-T/NK Cells

Ready-to-use CAR-T cell products to advance your research



CD19 • BCMA • Mesothelin • Her2 • EGFR

ProMab Biotechnologies is the world's first CRO for CAR-T and have established several patents in CAR-T technology to support immunotherapy research. With over 100 off-the-shelf CAR-T products and customizable development services, we are ready to help you advance your CAR-T discovery.



From left to right: CAR-T expression measured by FACS utilizing scFv specific antibody. Real-time cytotoxicity assay (RTCA) with CAR-T cells compared against target cells and T-cell negative control. Kaplan-Meier survival curve showing that CAR-T cells significantly prolonged survival compared to T cells in the NSG mice model.

All products are for research only

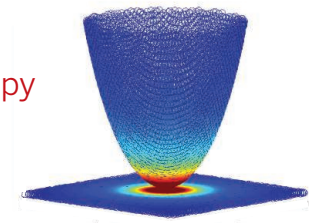
Discover more | www.ProMab.com | info@promab.com

ProMab
Biotechnologies, Inc.

RESEARCH

Ultrafast microscopy at atomic scale

Garg and Kern, p. 411



IN *SCIENCE* JOURNALS

Edited by **Michael Funk**

PHOTO: DANIELPRUDEK/ISTOCK.COM

PHYSIOLOGY

Breathing faster in thin air

When individuals who live at sea level ascend to higher altitudes, they automatically breathe faster to increase their oxygen uptake. This response is triggered by glomus cells, oxygen-sensitive neuron-like cells in the carotid body. Using genetically modified mice, Moreno-Domínguez *et al*. found that the gene expression pattern mediated by the oxygen-sensitive transcription factor HIF2α was necessary for this acute hypoxic response in glomus cells (see the Focus by Bishop and Ratcliffe). The authors propose that HIF2α may set the acute oxygen-sensing ability of other tissues as well. —WW

Sci. Signal. **13**, eaay9452, eaba1302 (2020).

A mountain climber en route to Mount Everest base camp in the Nepalese Himalayas

SUPERNOVAE

Modeling a superluminous supernova

Superluminous supernovae can be up to 100 times brighter than normal supernovae, but there is no consensus on how such bright transients are produced. Jerkstrand *et al*. identified emission lines of iron in the spectrum of a superluminous supernova that appeared more than a year after the explosion. The authors explored models of several possible mechanisms, finding that only one is consistent with all the observations: a normal Type Ia supernova that interacts with a dense shell of circumstellar material. The shell must have been ejected by the progenitor star less than a century before the explosion, perhaps owing to interaction with a binary partner. —KTS

Science, this issue p. 415

QUANTUM SENSING

Detecting a single magnetic excitation

Quantum-enhanced sensing is one of the near-term applications for the developing field of quantum technologies. A promising approach to quantum sensing relies on entangling a well-controlled system (the sensor) to the system of interest to detect quanta of excitations in the latter. By entangling a superconducting qubit with a ferrimagnetic crystal, Lachance-Quirion *et al*. demonstrate that they can detect a quantum of magnetic excitation (a magnon) within the sample. The demonstration of a high-efficiency single-magnon detector will be useful for quantum sensing as well as an active component of hybrid quantum systems and should find a wide range of applications in quantum technologies. —ISO

Science, this issue p. 425

NANOMATERIALS

Heterostructured nanorod libraries

The synthesis of nanostructures with well-defined interfaces between different materials can enable applications in areas such as catalysis and solar energy harvesting. Nanomaterials containing several different materials are usually synthesized through top-down approaches, such as surface growth or templating techniques, which usually produce small quantities of particles. Steimle *et al*. performed up to seven cation-exchange reactions (with ions such as Zn^{2+} and Co^{2+})

Published by AAAS

on solution-synthesized copper sulfide nanorods. In principle, more than 65,000 different variations of materials and their interfaces can be made, depending on the order and extent of cation exchange. —PDS

Science, this issue p. 418

CANCER IMMUNOTHERAPY

A one-two, CAR-T cell punch

Chimeric antigen receptor (CAR)–T cells have been clinically effective in killing certain hematological malignancies, but achieving long-term patient responses for solid tumors remains a challenge. Reinhard *et al*. describe a two-part "CARVac" strategy to overcome poor CAR-T cell stimulation and responses in vivo. They introduce the tight junction protein claudin 6 (CLDN6) as a new CAR-T cell target and designed a nanoparticulate RNA vaccine encoding a chimeric receptor directed toward CLDN6. This lipoplex RNA vaccine promotes CLDN6 expression on the surface of dendritic cells, which in turn stimulates and enhances the efficacy of CLDN6-CAR-T cells for improved tumor therapy. —PNK

Science, this issue p. 446

ORGANIC CHEMISTRY

A twisted small-molecule synthesis

Some molecules are easy to draw on paper, whereas others contain rings and contortions that require one to think in three dimensions. Reisberg *et al*. set out to synthesize the bicyclic small molecule tryptorubin A but found that their initial attempt produced a molecule with the right bonds but the wrong molecular shape, a form of noncanonical atropisomerism. The authors then devised a synthesis where they locked in the correct isomer before forming the second ring, which produced a product indistinguishable from the authentic natural product. Such structural isomers may be lurking when working with complex small molecules with constrained rotation. —MAF

Science, this issue p. 458

EMERGING INFECTIONS

Taking the bite out of diseases

Arthropods are the most abundant animals on Earth and can transmit diseases such as dengue or West Nile virus to humans. Bryden *et al*. tried to manipulate the immune reaction at the site of mosquito bites to restrict viral dissemination. They found that local Toll-like receptor 7 (TLR7) activation shortly after infection dampened replication of a model alphavirus in mice. This held true for clinically relevant arboviruses and in human skin explants. Viral restriction was due to activation of skin macrophages and heightened type I interferon production. Topical TLR7 activation after mosquito bites could be a broad-acting approach to abrogate arboviruses. —LP

Sci. Transl. Med. **12**, eaax2421 (2020).

DEVELOPMENTAL BIOLOGY

Shaping the early amniote embryo

Gastrulation is an essential step in development in which the internal tissues of the body are set apart. In birds and mammals, a similar cascade of molecular events is known to specify embryonic territories, but how they are physically remodeled has remained elusive. Working with avian embryos, Saadaoui *et al*. identified a cable that encircles the embryo as the engine of gastrulation and described the collective cell movements as similar to the motion of a fluid. One side of this contractile ring pulls more strongly than the other, entraining the large-scale tissue movements that shape the early body plan. The embryo margin, previously known to function in molecular regulation, thus emerges as a dual mechanical and molecular organizer of development. —BAP

Science, this issue p. 453

IN OTHER JOURNALS

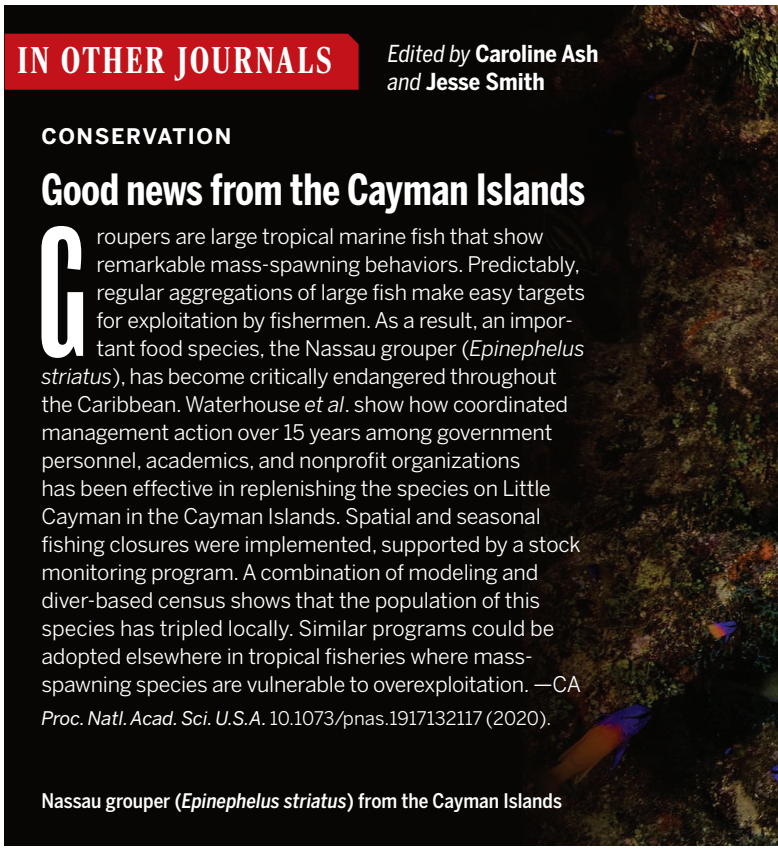
Edited by **Caroline Ash** *and* **Jesse Smith**

CONSERVATION

Good news from the Cayman Islands

Groupers are large tropical marine fish that show remarkable mass-spawning behaviors. Predictably, regular aggregations of large fish make easy targets for exploitation by fishermen. As a result, an important food species, the Nassau grouper (*Epinephelus striatus*), has become critically endangered throughout the Caribbean. Waterhouse *et al*. show how coordinated management action over 15 years among government personnel, academics, and nonprofit organizations has been effective in replenishing the species on Little Cayman in the Cayman Islands. Spatial and seasonal fishing closures were implemented, supported by a stock monitoring program. A combination of modeling and diver-based census shows that the population of this species has tripled locally. Similar programs could be adopted elsewhere in tropical fisheries where mass-spawning species are vulnerable to overexploitation. —CA

Proc. Natl. Acad. Sci. U.S.A. 10.1073/pnas.1917132117 (2020).



Nassau grouper (*Epinephelus striatus*) from the Cayman Islands

PHOTO: ALEX MUSTARD/MINDEN PICTURES

CANCER

Ethnicity reflected in tumor genomes

In the United States, African Americans are more likely to develop and succumb to lung cancer than European Americans. Several factors likely contribute to this racial disparity, including the possibility that disease biology differs between the two groups. Tumor genome sequencing can shed light on this hypothesis. Through targeted sequencing of 129 tumors, Mitchell *et al*. found somatic mutations in the *PTPRT* and *JAK2* genes in more than 30% of lung adenocarcinomas from African Americans versus 10% of tumors from European Americans. The proteins encoded by *PTPRT* and *JAK2* function in cellular signaling pathways implicated in cancer. Whether identification of these mutations will lead to new therapies is unclear, but the study broadly supports the idea that tumor biology may differ across racial groups. —PAK

Nat. Commun. **10**, 5735 (2019).

THIN FILMS

Synthesizing single-layer diamond

The carbon allotropes of diamond and graphene have different types of bonding that lead to their exceptional properties. Bakharev *et al*. pull off the impressive trick of making a monolayer carbon film that is diamond-like in its bonding. The authors accomplish this by attaching fluorine atoms to the carbon film, creating "F-diamane." Diamane is a long-sought-after, but challenging to make, material that should have useful properties. F-diamane may find use in a variety of applications, from microelectronics as a semiconductor to a seed material for growing single-crystal diamond films. —BG

Nat. Nanotechol. **15**, 59 (2020).

HUMAN GENETICS

Predicting transmission risk for de novo mutation

The reduction in sequencing costs and the increase

PHOTO: ZOLTAN BAGOSI / ALAMY STOCK PHOTO

in prediction accuracy make individual assessment of genetic risk from mutations more attractive and valuable. Examining mutation rates in blood and sperm, Breuss *et al*. surveyed families in which a child has been diagnosed with autism spectrum disorder. The authors found a small set of individuals in which a mosaic of potentially causative mutations was observed in the father's blood or sperm. Differential mutational processes seem to govern when the genetic variants arise. Screening for paternal mosaicism might help determine the risk of autism in future children of fathers that carry a de novo mutation. —LMZ

Nat. Med. **26**, 143 (2020).

PLANT SCIENCE

Some seeds store better than others

Seeds from species of the genus *Coffea*, although intolerant to complete drying, reflect a range of tolerance to desiccation. Stavrinides *et al*. studied seed transcriptomes and proteomes to investigate the divergent response to desiccation. In two *Coffea* species that produce seeds tolerant to desiccation, stress-related genes were up-regulated during seed development and mitochondrial physiology was down-regulated. In a *Coffea* species with seeds less tolerant to desiccation, genes involved in auxin production and response were up-regulated. The authors conclude that the seeds best positioned to withstand desiccation are those that can shut down respiration and tolerate oxidative stress. —PJH

J. Exp. Bot. 10.1093/jxb/erz508 (2019).

Raw, dried robusta coffee beans (*Coffea canephora*)

MACHINE LEARNING

Guiding solvothermal synthesis

Machine learning is rapidly revolutionizing computer-aided synthesis design, occasionally producing vivid use cases when the reaction parameters important for the synthesis are hidden in a complex chemical space. Xie *et al*. report a machine learning–assisted framework for the synthesis of metal-organic nanocapsules (MONCs), giant molecular building units potentially useful in different fields, based on predicting the crystallization propensity using experimental attempts as a training dataset. Machine-learning algorithms achieve prediction accuracies of more than 90%, considerably outperforming trained chemists, and the generated synthesis parameters direct solvothermal crystallization to new structures of MONCs. The proposed strategy for the discovery of new materials can be applied more broadly beyond MONCs. —YS

J. Am. Chem. Soc. 10.1021/jacs.9b11569 (2019).

OPTICS

Linking microwaves to photons

Communication networks seldom have all components operating at a single frequency. Components are individually designed to operate at optimum frequencies, with the signal then converted from one component to another by means of a transducer. Hybrid quantum systems, for example, will typically require links between microwaves, which operate the qubits, and photons that will transport the quantum information across the network. Shao *et al*. developed a thin-film acoustic resonator that allows the conversion of microwaves into optical signals. The acoustic resonators and optical waveguide circuits are patterned within a suspended layer of lithium niobate. Control of the coupling between the microwaves driving the resonators and the resulting modulation of the optical signal demonstrates a platform to develop hybrid signal-processing technologies. —ISO

Optica **6**, 1498 (2019).

ALSO IN *SCIENCE* JOURNALS

Edited by **Michael Funk**

NEUROGENETICS

Organoids recapitulate brain development

Gene expression changes and their control by accessible chromatin in the human brain during development is of great interest but limited accessibility. Trevino *et al*. avoided this problem by developing three-dimensional organoid models of human forebrain development and examining chromatin accessibility and gene expression at the single-cell level. From this analysis, they matched developmental profiles between the organoid and fetal samples, identified transcription factor binding profiles, and predicted how transcription factors are linked to cortical development. The researchers were able to correlate the expression of neurodevelopmental disease risk loci and genes with specific cell types during development. —LMZ

Science, this issue p. 404

RESEARCH METHODS

More diversity at the top

A detailed knowledge of cell differentiation hierarchies is important for understanding diverse biological processes such as organ development, tissue regeneration, and cancer. Single-cell RNA sequencing can help elucidate these hierarchies, but it requires reliable computational methods for predicting cell lineage trajectories. Gulati *et al*. developed CytoTRACE, a computational framework based on the simple observation that transcriptional diversity—the number of genes expressed in a cell—decreases during differentiation. CytoTRACE outperformed other methods in several test cases and was successfully applied to study cellular hierarchies in healthy and tumor tissue. —PAK

Science, this issue p. 405

NANOPHOTONICS

Shining a light on STM

Whereas ultrafast light pulses can provide a window into the dynamics of some of the fastest processes in condensed matter systems, scanning tunneling microscopy (STM) provides snapshots of atomic-scale spatial resolution of surfaces. By irradiating the tunnel junction with carrier-enveloped, phase-stabilized femtosecond light pulses, Garg and Kern effectively combined the two methods to demonstrate a technique capable of both high temporal and high spatial resolution of the elementary processes in matter (see the Perspective by Aiello). By tracking the decay dynamics of plasmonic excitations on the surface of gold nanorods, they illustrate the power of the technique for the tracking and control of ultrafast electronic processes on the atomic scale. —ISO

Science, this issue p. 411; see also p. 368

SURFACE MICROSCOPY

Imaging reactive surface water

Recent developments in transmission electron microscopy (TEM) have enabled imaging of single atoms, but adsorbed gas molecules have proven more challenging because of a lack of sufficient image contrast. Yuan *et al*. adsorbed water and carbon monoxide (CO) on a reconstructed nanocrystalline anatase titanium dioxide (TiO_2) surface that has protruding TiO_3 ridges every four unit cells, which provide regions of distinct contrast. Water adsorption on this surface during environmental TEM experiments led to the formation of twinned protrusions. These structures developed dynamic contrast as the water reacted with coexposed CO to form hydrogen and carbon dioxide. —PDS

Science, this issue p. 428

NEUROSCIENCE

Interneurons control brain arousal states

The underlying circuit mechanisms coordinating brain arousal and motor activity are poorly understood. Liu *et al*. found that glutamic acid decarboxylase 2 (GAD2)–expressing, but not parvalbumin-expressing, interneurons in a part of the brain known as the substantia nigra promote sleep (see the Perspective by Wisden and Franks). Parvalbuminergic neurons fire at higher rates in states of high motor activity, and their activation increases movement termination consistent with the function of the substantia nigra in suppressing unwanted movements during action selection. By contrast, GAD2 neurons are preferentially active in states of low motor activity. In addition to motor suppression, their activation powerfully enhances the transition from quiet wakefulness to sleep, which differ mainly in the arousal level rather than motor behavior. GAD2 interneurons thus provide general suppression of both motor activity and brain arousal to promote states of quiescence. —PRS

Science, this issue p. 440; see also p. 366

LIFE-SPAN EXTENSION

Taking the stress out of life

In the model organism *Caenorhabditis elegans*, a roundworm, it has been shown that neurons can communicate proteostasis to the periphery to affect aging. Frakes *et al*. have now identified astrocyte-like glial cells that also act as central regulators of systemic protein homeostasis and aging (see the Perspective by Miklas and Brunet). They found that the life span of *C. elegans* can be extended by expression of a constitutively active version of the transcription factor XBP-1s, which mediates the unfolded protein response of the endoplasmic reticulum (UPR^{ER}), in a specific subset of glial cells. Glial XBP-1s initiates induction of the UPR^{ER} in distal intestinal cells, which makes the worms more resistant to chronic ER stress. Neuropeptide signaling was required for glial-mediated longevity and induction of the peripheral UPR^{ER}, suggesting a distinct mechanism from that initiated by neuronal XBP-1s. Thus, in this animal model of aging, a mere four cells can control organismal physiology and aging —SMH

Science, this issue p. 436; see also p. 365

PHASE SEPARATION

Keeping the noise down

Protein concentrations in a cell fluctuate considerably because of stochasticity in gene expression and variations in the cell's microenvironment. How cells cope with concentration fluctuations when precision is important is unclear. Klosin *et al*. used a combination of theoretical and experimental work to demonstrate that phase-separated compartments can effectively reduce protein concentration noise in cells (see the Perspective by Riback and Brangwynne). The results suggest that phase separation provides a mechanism to enhance the robustness of biological systems. —SMH

Science, this issue p. 464; see also p. 364

ASTROBIOLOGY

Understanding the boundaries of life

Life on Earth has several universal chemical characteristics, including the presence of a lipid bilayer cell membrane that separates the aqueous cell contents from the environment. If life has evolved on other planets, it may use entirely different chemical

components and principles. Polarity-inverted membrane structures composed of acrylonitrile, termed azotosomes, have previously been proposed as possible membrane alternatives under conditions that resemble those found on Saturn's moon Titan. Sandström and Rahm now argue, on the basis of molecular dynamics simulations, that azotosomes made of acrylonitrile in methane would be thermodynamically unstable and therefore unable to self-assemble in a manner similar to lipid bilayers. Such modeling efforts may be useful for constraining future speculation about what chemical assemblies are possible in unfamiliar environments. —JJdP

Sci. Adv. 10.1126/sciadv.aax0272 (2020).

PLANT SCIENCE

Filling in the gaps

In a plant seed, the embryo lies dormant surrounded by nutritive endosperm while awaiting suitable conditions to germinate. A hydrophobic cuticle around the embryo protects it from catastrophic water loss during the early days of growth. Doll *et al*. identified a back-and-forth signaling pathway that ensures an intact cuticle. The precursor of a signaling peptide is made in the embryo and transferred to the endosperm, where it is processed into an active form. The activated peptide diffuses back into the embryo to activate receptor-like kinases that drive cuticle development. Serve and return continues until all leaks in the cuticle are filled in and the peptide can no longer cross the barrier. —PJH

Science, this issue p. 431

RESEARCH ARTICLE

RESEARCH METHODS

Single-cell transcriptional diversity is a hallmark of developmental potential

Gunsagar S. Gulati[1]*, Shaheen S. Sikandar[1]*, Daniel J. Wesche[1], Anoop Manjunath[1], Anjan Bharadwaj[1], Mark J. Berger[2]†, Francisco Ilagan[1], Angera H. Kuo[1], Robert W. Hsieh[1], Shang Cai[3], Maider Zabala[1]‡, Ferenc A. Scheeren[4], Neethan A. Lobo[1]‡, Dalong Qian[1], Feiqiao B. Yu[5], Frederick M. Dirbas[6], Michael F. Clarke[1,7], Aaron M. Newman[1,8]§

Single-cell RNA sequencing (scRNA-seq) is a powerful approach for reconstructing cellular differentiation trajectories. However, inferring both the state and direction of differentiation is challenging. Here, we demonstrate a simple, yet robust, determinant of developmental potential—the number of expressed genes per cell—and leverage this measure of transcriptional diversity to develop a computational framework (CytoTRACE) for predicting differentiation states from scRNA-seq data. When applied to diverse tissue types and organisms, CytoTRACE outperformed previous methods and nearly 19,000 annotated gene sets for resolving 52 experimentally determined developmental trajectories. Additionally, it facilitated the identification of quiescent stem cells and revealed genes that contribute to breast tumorigenesis. This study thus establishes a key RNA-based feature of developmental potential and a platform for delineation of cellular hierarchies.

In multicellular organisms, tissues are hierarchically organized into distinct cell types and cellular states with intrinsic differences in function and developmental potential (*1*). Common methods for studying cellular differentiation hierarchies, such as lineage tracing and functional transplantation assays, have revealed detailed roadmaps of cellular ontogeny at scales ranging from tissues and organs to entire model organisms (*2*–*4*). Though powerful, these technologies cannot be applied to human tissues in vivo and generally require prior knowledge of cell type–specific genetic markers (*2*). These limitations have made it difficult to study the developmental organization of primary human tissues under physiological and pathological conditions.

Single-cell RNA-sequencing (scRNA-seq) has emerged as a promising approach to study cellular differentiation trajectories at high resolution in primary tissue specimens (*5*). Although a large number of computational methods for predicting lineage trajectories have been described, they generally rely upon (i) a priori knowledge of the starting point (and thus, direction) of the inferred biological process (*6*, *7*) and (ii) the presence of intermediate cell states to reconstruct the trajectory (*8*, *9*). These requirements can be challenging to satisfy in certain contexts, such as human cancer development (*10*). Moreover, with existing in silico approaches, it is difficult to distinguish quiescent (noncycling) adult stem cells that have long-term regenerative potential from more specialized cells. Although gene expression–based models can potentially overcome these limitations [e.g., transcriptional entropy (*11*–*13*), pluripotency-associated gene sets (*14*), and machine learning strategies (*15*)], their utility across diverse developmental systems and single-cell sequencing technologies is still unclear.

Here, we systematically evaluated RNA-based features, including nearly 19,000 annotated gene sets, to identify factors that accurately predict cellular differentiation status independently of tissue type, species, and platform. We then leveraged our findings to develop an unsupervised framework for predicting relative differentiation states from single-cell transcriptomes. We validated our approach through comparison to leading methods and explored its utility for identifying key genes associated with stem cells and differentiation in both healthy tissues and human cancer.

Results

RNA-based correlates of single-cell differentiation states

Our initial goal was to identify robust, RNA-based determinants of developmental potential without the need for a priori knowledge of developmental direction or intermediate cell states marking cell fate transitions. Using scRNA-seq data, we evaluated ~19,000 potential correlates of cell potency, including all available gene sets in the Molecular Signatures Database (n = 17,810) (*16*), 896 gene sets covering transcription factor binding sites from ENCODE (*17*) and ChEA (*18*), an mRNA expression–derived stemness index (mRNAsi) (*15*), and three computational techniques that infer stemness as a measure of transcriptional entropy [StemID, SCENT, and SLICE (*11*–*13*)]. We also explored the utility of "gene counts," or the number of detectably expressed genes per cell. Although anecdotally observed to correlate with differentiation status in a limited number of settings [alveolar development in mouse and thrombocyte development in zebrafish (*19*, *20*)], the reliability of this association and whether it reflects a general property of cellular ontogeny are unknown.

To assess these RNA-based features, we compiled a training cohort consisting of nine gold standard scRNA-seq datasets with experimentally confirmed differentiation trajectories. These datasets were selected to prioritize commonly used benchmarking datasets from earlier studies and to ensure a broad sampling of developmental states from the mammalian zygote to terminally differentiated cells (table S1). Overall, the training cohort encompassed 3174 single cells spanning 49 phenotypes, six biological systems, and three scRNA-seq platforms (fig. S1A and table S1). To evaluate performance, we used Spearman correlation to compare each RNA-based feature, averaged by phenotype, against known differentiation states (Fig. 1A). We then averaged the results across the nine training datasets to yield a final score and rank for every feature (table S2).

This systematic screen revealed many known and unexpected correlates of differentiation status (Fig. 1B, fig. S1B, and table S2). However, one feature in particular showed notable performance: the number of detectably expressed genes per cell (gene counts). Appearing in the top 1% of the ranked list (104 of 18,711), this data-driven feature compared favorably to well-established stem cell signatures, including cell cycle and pluripotency genes (*14*, *15*), yet also showed evidence of distinctive biology and broader applicability. For example, regardless of whether we examined cycling cells, noncycling cells, or published data from the earliest stages of human embryogenesis prior to the up-regulation of pluripotency factors (*21*), gene counts generally decreased with successive stages of differentiation (fig. S2, A and B, left). Pluripotency genes, by contrast, showed an arc-like pattern early in human embryogenesis that was characterized by progressively increasing expression until the emergence of

[1]Institute for Stem Cell Biology and Regenerative Medicine, Stanford University, Stanford, CA 94305, USA. [2]Department of Computer Science, Stanford University, Stanford, CA 94305, USA. [3]School of Life Sciences, Westlake University, Zhejiang Province, China. [4]Department of Medical Oncology, Leiden University Medical Center, 2333 ZA Leiden, Netherlands. [5]Chan Zuckerberg Biohub, San Francisco, CA 94305, USA. [6]Department of Surgery, Stanford Cancer Institute, Stanford University, Stanford, CA 94305, USA. [7]Department of Medicine, Stanford University, Stanford, CA 94305, USA. [8]Department of Biomedical Data Science, Stanford University, Stanford, CA 94305, USA.
*These authors contributed equally to this work. †Present address: Color Genomics Inc., Burlingame, CA 94010, USA. ‡Present address: Onena Medicine S.L., Donostia-San Sebastián, Guipúzcoa 20009, Spain.
§Corresponding author. Email: amnewman@stanford.edu

embryonic stem cells, followed by decreasing expression (fig. S2B, right).

These findings suggested that gene counts might extend beyond isolated experimental systems to recapitulate the full spectrum of developmental potential. To test this possibility, we compiled, remapped, and normalized a set of in vivo mouse lineage trajectories based on five plate-based scRNA-seq experiments encompassing 5059 cells and 30 phenotypes that together spanned all major potency levels (*22*) (table S3 and materials and methods). Indeed, when averaged according to known phenotypes and assessed across independent studies, the association between gene counts and differentiation was maintained ($R^2 = 0.89$, $P = 1.8 \times 10^{-8}$) (Fig. 1C and materials and methods). Notably, this relationship was also observed in other organisms, including *Caenorhabditis elegans* (Fig. 1D) and zebrafish (table S4), suggesting that it is a general feature of cellular ontogeny.

Because the transcriptional output of a cell is associated with its genome-wide chromatin profile, we next tested whether single-cell gene counts are ultimately a surrogate for global chromatin accessibility, which has been shown to decrease with differentiation in certain contexts (*23*–*25*). To do this, we compared single-cell gene counts derived from scRNA-seq data with paired bulk ATAC-seq (assay for transposase-accessible chromatin sequencing) profiles obtained from a recent study of in vitro mesodermal differentiation from human embryonic stem cells (hESCs) (*26*). Indeed, genome-wide chromatin accessibility was observed to progressively decrease with differentiation of hESCs into paraxial mesoderm and lateral mesoderm lineages (Fig. 1E and fig. S3, A and B). We observed strong concordance between the number of accessible peaks and the mean number of detectably expressed genes per phenotype (fig. S3A). Similar patterns were observed for

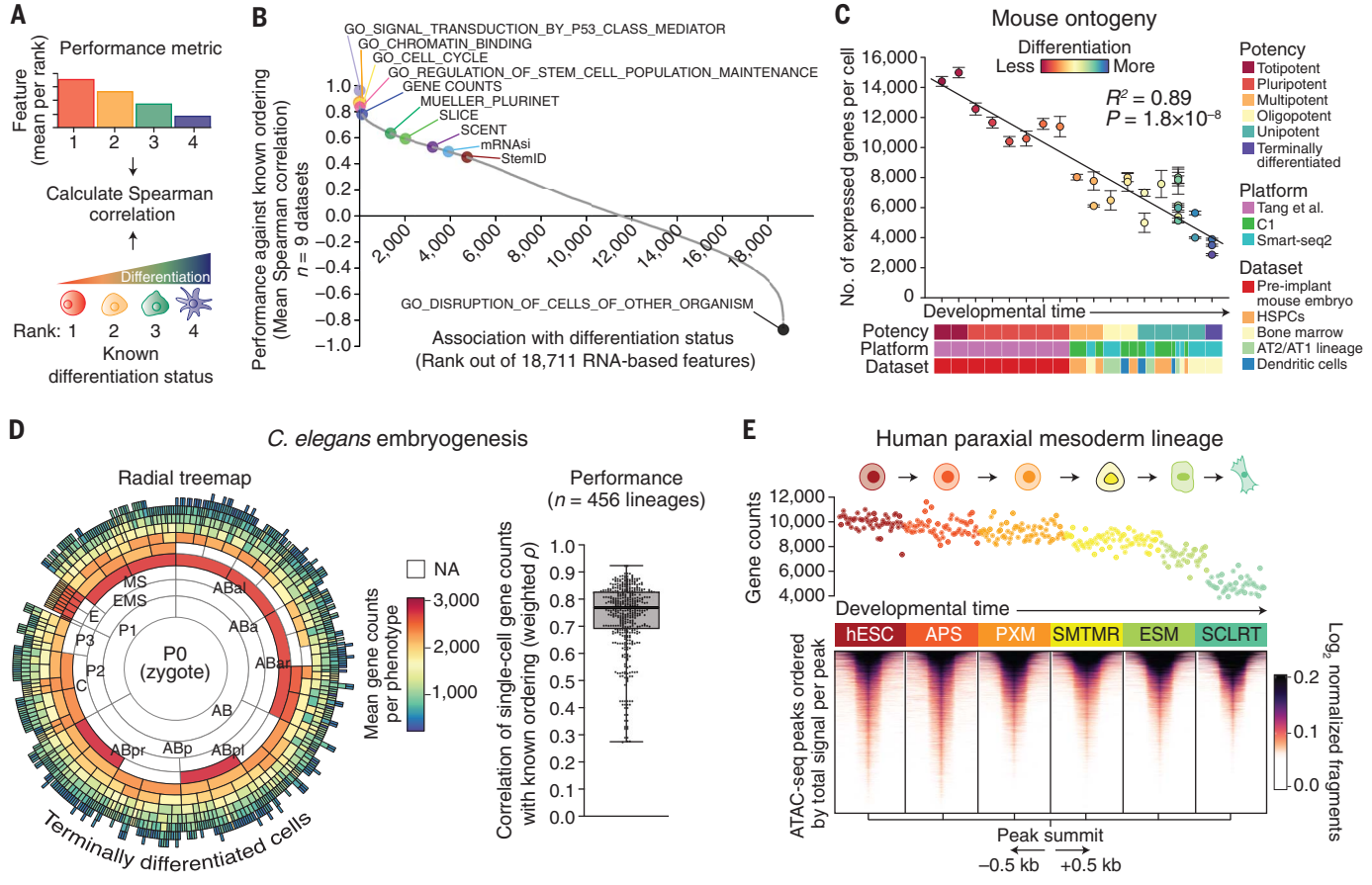


Fig. 1. RNA-based determinants of developmental potential. (**A** and **B**) In silico screen for correlates of cellular differentiation status in scRNA-seq data. (A) Depiction of the scoring scheme. Each phenotype was assigned a rank on the basis of its known differentiation status (less differentiated = lower rank), and the values of each RNA-based feature (fig. S1A) were mean-aggregated by rank for each dataset (higher value = lower rank). Performance was calculated as the mean Spearman correlation between known and predicted ranks across all nine training datasets (table S1). (B) Performance of all evaluated RNA-based features for predicting differentiation states in the training cohort, ordered by mean Spearman correlation (fig. S1 and table S2). (**C**) The developmental ordering of 30 mouse cell phenotypes across 17 developmental stages shown as a function of single-cell gene counts (table S3). Data are expressed as means ± 95% confidence intervals. The linear regression line and coefficient of determination (R^2) are shown. (**D**) Performance of gene counts for ordering *C. elegans* embryogenesis. (Left) Radial tree map showing gene counts for each cell type with available scRNA-seq data from a recent study (*48*). NA, not available. Embryogenesis originates at the center of the plot [P0 (zygote)] and moves outwards towards terminally differentiated cells, with concentric rings representing sequential cell divisions. (Right) Boxplot showing weighted Spearman correlations between single-cell gene counts and developmental lineages with available transcriptomic data ($n = 456$). (**E**) Association between single-cell gene counts and chromatin accessibility in cells from an in vitro differentiation series of purified phenotypes from the human paraxial mesoderm lineage [Mesoderm (C1) dataset; table S1]. (Top) Association of single-cell gene counts with differentiation. Each point represents a cell colored by known phenotype (below). (Bottom) Heat map showing chromatin accessibility profiles for the same phenotypes as above. Peaks are centered by their summit, defined as the base with maximum coverage, shown within a window of 1 kb (±0.5 kb), and ordered top to bottom within each phenotype by decreasing total signal per peak. Cell type abbreviations are defined in materials and methods.

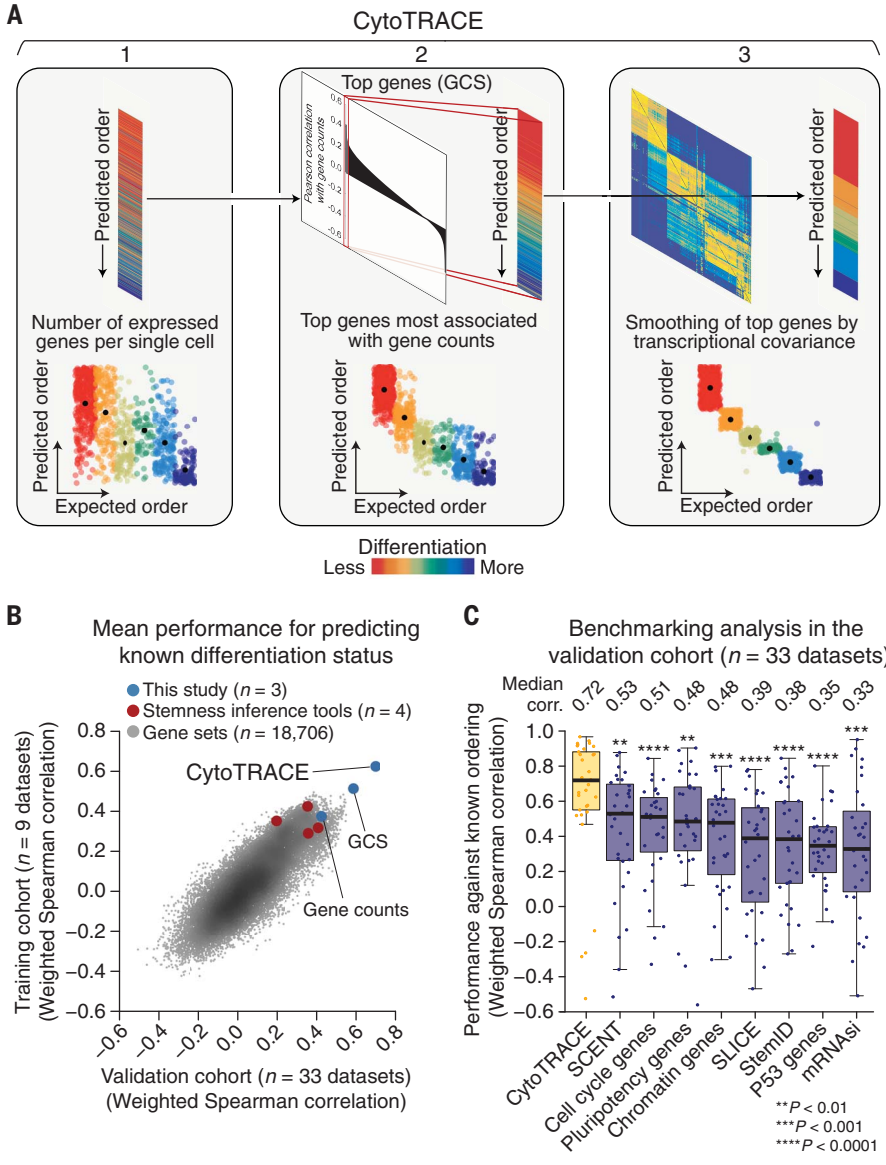


Fig. 2. Development and validation of CytoTRACE. (**A**) Schematic overview of the CytoTRACE framework applied to the hESC in vitro differentiation (C1) dataset (materials and methods and table S1). (**B**) Scatterplot comparing the average performance of 18,706 annotated gene sets, four stemness inference methods, gene counts, GCS, and CytoTRACE in the training and validation cohorts (table S2). (**C**) Boxplots showing the single-cell performance of CytoTRACE against RNA-based features and methods in the validation cohort (n = 33 datasets; table S2). Each point represents the Spearman correlation, weighted by number of cells per phenotype, between predicted and known differentiation states for a given dataset, calculated as described in materials and methods. Statistical significance was assessed by a one-sided paired Wilcoxon signed-rank test against CytoTRACE (table S4).

lung adenocarcinoma cells jointly profiled by ATAC-seq and RNA-seq (sci-CAR) and for human hematopoiesis, supporting the generality of this result (fig. S3, C to E).

Development of CytoTRACE

The number of expressed genes per cell generally showed consistent performance with respect to key technical parameters and was generally correlated with mRNA content (figs. S4 to S7 and supplementary text). However, in some datasets, such as that for in vitro differentiation of hESCs into the gastrulation layers (*27*), the number of expressed genes per cell exhibited considerable intraphenotypic variation (Fig. 2A, left). Indeed, when evaluated at a single-cell level, 412 predefined gene sets from our in silico screen outperformed gene counts (fig. S8A and table S2). Because scRNA-seq was designed to capture single-cell gene expression, we reasoned that genes whose expression patterns correlate with gene counts might better capture differentiation states. Indeed, by simply averaging the expression levels of genes that were most highly correlated with gene counts in each dataset (materials and methods), the resulting dataset-specific gene counts signature (GCS) became the top-performing measure in the screen, outranking every predefined gene set and computational tool that we assessed (fig. S8, A to D).

GCS, like gene counts, is inherently insensitive to dropout events, is agnostic to prior knowledge of developmentally regulated genes, and is not solely attributable to multilineage priming (*28*) (fig. S9 and supplementary text) or a known molecular signature (e.g., pluripotency) (fig. S2B and table S5). Despite these characteristics, GCS was still moderately noisy in some datasets (e.g., Fig. 2A, center and fig. S8C). We therefore implemented a two-step procedure to directly smooth GCS on the basis of transcriptional covariance among single cells (Fig. 2A, right, and materials and methods). The resulting method, which we call CytoTRACE [for cellular (Cyto) Trajectory Reconstruction Analysis using gene Counts and Expression; https://cytotrace.stanford.edu], outperformed GCS and the other RNA-based features that we evaluated (fig. S8 and table S2).

Performance evaluation across tissues, species, and platforms

To validate our findings, we assembled an expanded compendium of 33 additional scRNA-seq datasets from 26 studies (fig. S10A, table S1, and materials and methods). These datasets represent diverse developmental and differentiation processes and consist of 141,267 single cells spanning 266 phenotypes, nine biological systems, five species [including two whole organisms (*29*, *30*)], and nine scRNA-seq platforms (three droplet-based and six plate-based protocols, ranging from an average of ~10,000 unique molecular identifiers to ~1 million reads per cell, respectively) (fig. S5A).

When assessed at the single-cell level, CytoTRACE outperformed all evaluated RNA-based features in the validation cohort (Fig. 2B), achieving a substantial gain in performance over the second-highest-ranking approach (median rho = 0.72 versus 0.53 for the second-highest-ranking approach; P = 0.001) (Fig. 2C; fig. S10B; and tables S2 and S4). Similar improvements were observed across many complex systems, including bone marrow differentiation (fig. S10C). In addition, CytoTRACE results were positively correlated with the direction of differentiation in 88% of datasets ($P = 7 \times 10^{-7}$, binomial test). These results were consistent with our findings for the training cohort (Fig. 2B and fig. S10D) and were robust to the use of smoothing (fig. S11). Moreover, no significant biases in performance were observed in relation to tissue type, species, the number of cells analyzed, time series experiments

versus snapshots of developmental states, or plate-based versus droplet-based technologies (fig. S12).

To further evaluate CytoTRACE, we compared it with RNA velocity, a kinetic model that can predict future cell states but is limited to scRNA-seq data with continuous fate transitions (*8*). To analyze RNA velocity output, which consists of an individualized prediction for every cell (fig. S13), we identified all pairs of current and future cell states spanning a known shift in developmental potential (in the order of less to more, or vice versa). We then scored each predicted trajectory against known differentiation states on five datasets with continuous developmental processes (fig. S13B and materials and methods). To permit a fair comparison, CytoTRACE was evaluated on the same cells. Although both methods performed similarly on an embryonic chromaffin dataset from the RNA velocity study (*8*), CytoTRACE achieved higher accuracy overall (median of 74% versus 54%, respectively; fig. S13C). This was likely due to the short mRNA half-lives and developmental time scales assumed for the RNA velocity model (*8*) (supplementary text).

Having assessed performance on individual datasets, we next asked whether CytoTRACE could be applied across independent scRNA-seq datasets unified by batch correction. To address this, we leveraged mutual nearest-neighbor and Gaussian kernel normalization techniques from Scanorama (*31*) (materials and methods). We then merged several datasets with this approach. Regardless of whether we integrated datasets profiled on different scRNA-seq platforms (Fig. 3A) or datasets containing developmentally distinct cell types (fig. S14), single-cell orderings predicted by CytoTRACE were accurate.

Stem cell–related genes and hierarchies

Given the ability of CytoTRACE to recover the direction of differentiation in nearly every evaluated dataset (supplementary text), we next explored its potential to identify markers of immature phenotypes without prior knowledge. By rank-ordering genes on the basis of their correlation with CytoTRACE, markers of immature cells were readily prioritized in 86% of benchmarking datasets (fig. S15A). These included well-established stem and progenitor markers, such as *Kit* and *Stmn1* in mouse bone marrow (*32*) and *Axin2* and *Lgr5* in mouse intestinal crypts (*33*), underscoring the utility of CytoTRACE for the de novo discovery of developmentally regulated genes (fig. S15B and table S6).

Lineage relationships and their associated genes can also be determined by dedicated branch detection tools, such as Monocle 2; however, these approaches do not predict the starting point of the biological process. For example, when applied to 4442 bone marrow cells, Monocle 2 identified 23 possible "roots" from which to calculate pseudotime values (Fig. 3B, left). By contrast, CytoTRACE readily identified the correct root without user input

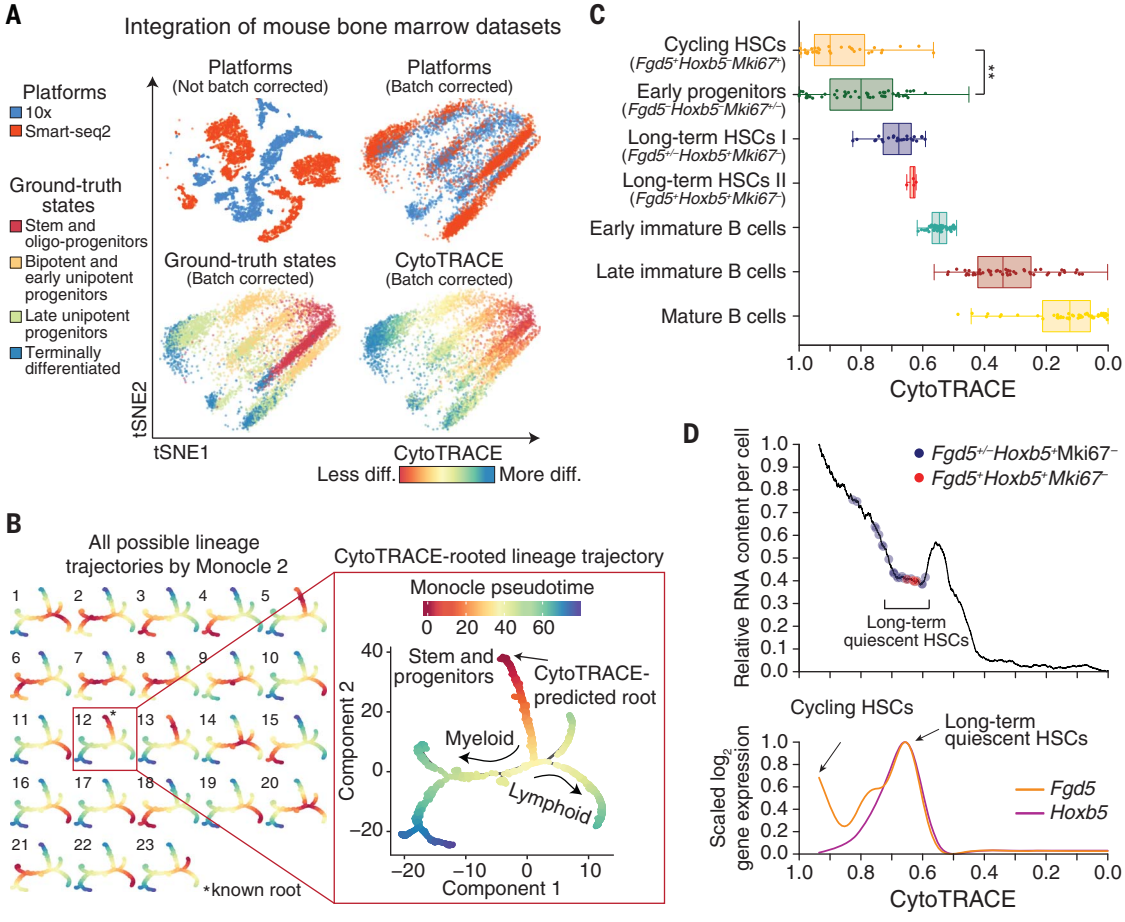


Fig. 3. Characterization of developmental hierarchies and quiescent stem cells using CytoTRACE. (**A**) Impact of batch correction (materials and methods) on two datasets of mouse bone marrow differentiation: Bone Marrow (10x) and Bone Marrow (Smart-seq2) (table S1). diff, differentiated. (**B**) Combined application of CytoTRACE and Monocle 2 to mouse bone marrow differentiation [Bone Marrow (Smart-seq2) dataset] (table S1). (Left) Multilineage tree inferred by Monocle 2 showing all 23 possible pseudotimes when the root is unknown. (Right) Automatic selection of the correct root by CytoTRACE. (**C** and **D**) Prioritization of quiescent and cycling HSCs in index-sorted scRNA-seq data of mouse hematopoiesis [Bone Marrow (Smart-seq2) dataset] (table S1). All plots are identically ordered by CytoTRACE. (C) Boxplots showing CytoTRACE values for candidate cycling HSCs (n = 31), long-term or quiescent HSCs (n = 30), early immature B cells (n = 285), late immature B cells (n = 863), and mature B cells (n = 700). HSCs, long-term or quiescent HSCs, and proliferating cells were defined on the basis of expression of *Fgd5* (*49*), *Hoxb5* (*35*), and *Mki67*, respectively. Although boxplots represent all analyzed cells, a maximum of 50 cells per phenotype are displayed as points for clarity. Statistical significance was assessed by a two-sided Wilcoxon signed-rank test. $^{**}P$ = 0.003. (D) (Top) RNA content per cell, shown as a function of CytoTRACE and displayed as the moving average of 200 cells. (Bottom) Expression of *Fgd5* and *Hoxb5* displayed as a smoothing spline over the moving average of 200 cells. Data from monocytic and granulocytic lineages are consistent with the above results.

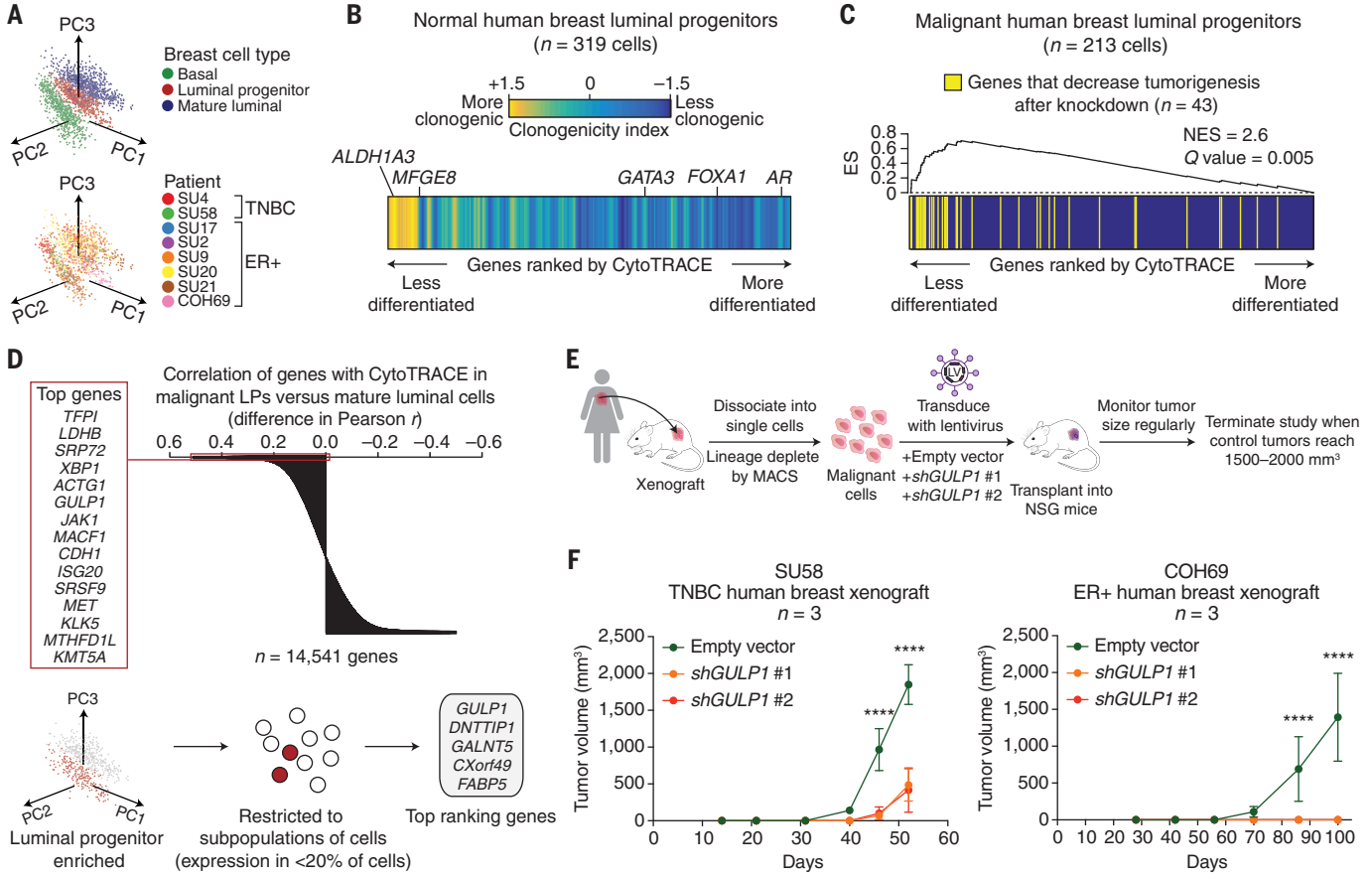


Fig. 4. Identification of immature cell markers in normal and malignant human breast LPs on the basis of CytoTRACE. (**A**) Principal component analysis of scRNA-seq profiles from 1902 human breast epithelial cells, colored according to subpopulation (top) and patient (bottom). (**B**) Heat map showing genes from adjacent normal LPs rank-ordered by their Pearson correlation with CytoTRACE and colored according to a clonogenicity index, defined as the $\log_2$ fold change in expression between highly and lowly clonogenic LPs from normal human breast (*39*) (materials and methods). The clonogenicity index is displayed as a moving average of 200 genes. Key genes associated with less (*ALDH1A3*, *MFGE8*) or more (*GATA3*, *FOXA1*, *AR*) differentiated normal LPs are indicated. (**C**) Enrichment of genes associated with human breast tumorigenesis [RNAi dropout viability screen (*41*)] within a ranked list of genes expressed by malignant LPs, rank-ordered by their Pearson correlation with CytoTRACE. Enrichment was calculated with preranked gene set enrichment analysis. NES, normalized enrichment score; ES, enrichment score. (**D**) Identification of candidate tumorigenic genes associated with immature malignant human LPs. (Top) Genes rank-ordered by the difference in their Pearson correlations with CytoTRACE in malignant LPs versus malignant mature luminal cells. The top 15 genes that are predicted to be specifically associated with less differentiated LPs are indicated on the left. (Bottom) Schema for the identification of genes that are ranked as above but that are also more highly expressed in malignant LPs than MLs ($\log_2$ fold change > 0; Benjamini-Hochberg adjusted $P < 0.05$, unpaired two-sided *t*-test) and that are expressed by a subpopulation of LPs (<20% of cells). The top five filtered genes are shown (right). (**E**) Schema for shRNA knockdown of *GULP1* in a human breast cancer xenograft model. (**F**) Growth of human breast cancer xenografts from two patients, one with TNBC (left) and one with ER+ luminal-type cancer (right), after lentiviral transduction with empty vector or shRNA targeting *GULP1*. Tumor volumes after knockdown with *shGULP1* #1 (orange) and *shGULP1 #2* (red) were indistinguishable in COH69 xenografts (right). Data are expressed as means ± SD ($n = 3$ mice). Statistical significance was assessed by a two-way ANOVA. ****$P < 0.0001$.

(Fig. 3B, right, and fig. S16, A and B). Integration of these methods facilitated automatic identification of lineage-specific regulatory factors and marker genes during granulocyte, monocyte, and B cell differentiation (fig. S16C). Similar results were obtained for mouse intestinal cells (fig. S16, D to F). Notably, other methods also showed strong performance when oriented by CytoTRACE (fig. S16G and table S4).

We next asked whether CytoTRACE could distinguish cycling and long-term or quiescent stem cells from their downstream progenitors (*34*). As these populations have been well-characterized in the bone marrow (*3*), we investigated this question in the mouse hematopoietic system. Although both cycling and quiescent hematopoietic stem cell (HSC) subpopulations (*34*) were correctly predicted to be less differentiated, only proliferative HSCs were significantly ranked above early progenitors (Fig. 3C). This result was not unexpected, however, because quiescent cells have reduced metabolic activity and low RNA content (*1*). By devising a simple approach to visualize inferred RNA content as a function of CytoTRACE (Fig. 3D, top), we observed a distinct valley in RNA abundance that coincided with elevated expression of *Hoxb5*, a marker of long-term or quiescent HSCs (*35*) (Fig. 3D, bottom). Since these cells could not be identified by gene counts or RNA content alone, this analysis confirmed the utility of CytoTRACE and demonstrated an approach for elucidating tissue-specific stem cells from scRNA-seq data.

Application to neoplastic disease

Increasing evidence suggests that human breast tumors contain less-differentiated cells that are resistant to therapy and are associated with the development of relapse and metastasis (*10*, *36*). Subpopulations of tumor cells within the luminal progenitor (LP) epithelium are thought to

give rise to aggressive basal-like breast cancers, such as triple-negative breast cancer (TNBC) (*37*), and possibly also to estrogen receptor–positive (ER+) breast cancers (*38*). However, the differentiation states and tumor-initiating properties of LP subsets remain incompletely understood.

To determine whether CytoTRACE can provide insights into immature LP cells and their associated genes in breast cancer, we performed scRNA-seq profiling of breast tumor epithelial cells and adjacent normal epithelial cells from eight patients with triple-negative (n = 2) or ER+ (n = 6) breast cancer. Using a Smart-seq2 protocol combined with fluorescence-activated cell sorting (FACS), we index-sorted and sequenced cells from three major human epithelial subpopulations: basal (CD49f^{high} EPCAM$^{med\text{-}low}$), luminal progenitor (CD49f^{high} EPCAM^{high}), and mature luminal (ML) subpopulations (CD49f^{low} EPCAM^{high}) (fig. S17A and table S7). After removing low-quality cells and applying principal component analysis to visualize the data, we confirmed three well-separated clusters of basal, LP, and ML cells, each with characteristic expression patterns of previously described lineage markers (Fig. 4A and fig. S17B). No obvious clustering was observed for tumor versus normal cell differences or by patient (Fig. 4A and fig. S18A).

To validate the ability of CytoTRACE to define LP differentiation states, we started by rank-ordering genes expressed in adjacent normal LPs by their Pearson correlation with CytoTRACE. We found that previously described marker genes of less-differentiated normal LPs [*ALDH1A3* and *MFGE8* (*39*)] and more-differentiated normal LPs [*GATA3*, *FOXA1*, and *AR* (*39*, *40*)] were successfully enriched by this approach (Fig. 4B). Moreover, genes that were up-regulated in highly clonogenic normal LPs (*39*) were skewed toward genes predicted to mark less-differentiated cells (Fig. 4B).

We next sought to identify LP genes associated with tumorigenesis. We first ordered genes expressed in malignant LPs by their Pearson correlation with CytoTRACE. In this rank-ordered list, we observed a significant enrichment of genes whose knockdown by RNA interference (RNAi) led to decreased viability of tumor cells in patient-derived xenograft (PDX) models of TNBC (*41*) (Q = 0.002, gene set enrichment analysis) (Fig. 4C; fig. S18, B and C; and table S8). Moreover, when we applied CytoTRACE to prioritize genes in tumor LPs compared to tumor MLs, the latter of which are developmentally downstream of LPs in normal breast (*39*), the top 15 genes included known members of tumorigenic pathways in breast cancer [e.g., *MET* and *JAK1* (*42*, *43*)], as well as unknown candidates (e.g., *GULP1*) (Fig. 4D, top). We focused on genes that were (i) more highly expressed in tumor LPs than MLs and (ii) expressed in a subpopulation of tumor LPs (<20% of cells) (Fig. 4D, bottom). After applying this filter, *GULP1* emerged as the top candidate gene (Fig. 4D, bottom right, and fig. S18C).

GULP1 is an engulfment adaptor protein (*44*) and its murine homolog is a specific marker of mouse HSCs, suggesting a conserved role of this gene in other immature cell states (fig. S19A). We measured the effect of *GULP1* knockdown on the proliferation of the metastatic TNBC cell lines MDA-MB-231 and MDA-MB-468 (fig. S19, B to E). We found that *GULP1* knockdown reduced proliferation of both cell lines as measured by a colorimetric assay for metabolic activity (fig. S19E). In addition, *GULP1* knockdown in PDXs (n = 2) either inhibited tumor growth (TNBC sample) or fully abrogated tumor growth (ER+ sample) (Fig. 4, E and F). These data suggest a possible role for GULP1 in human breast cancer tumorigenesis.

Discussion

Efforts to characterize single-cell transcriptomes in diverse tissues, organs, and whole organisms have underscored the need for RNA-based determinants of developmental potential. In our analysis of RNA-based features across numerous developmental processes, we observed that the number of detectably expressed genes per cell powerfully associates with cellular differentiation status.

Although previous studies demonstrated a global reduction in chromatin accessibility and/or plasticity during lineage commitment in specific developmental settings [e.g., embryonic stem cells, intestinal stem cells, and neural stem cells (*23*–*25*)], this work extends the scope of this finding and quantitatively links it to single-cell gene counts. Moreover, as has been previously shown (*45*), our results indicate that variability in gene counts between phenotypically identical single cells is not exclusively due to dropout events but is also due to differential sampling of the transcriptome (fig. S4). Our results are therefore consistent with a model in which less-mature cells maintain looser chromatin to permit wider sampling of the transcriptome, whereas more-differentiated cells generally restrict chromatin accessibility and transcriptional diversity as they specialize (Fig. 1E and fig. S3) (*46*). Theoretically, this model can be represented by "attractor states" within a genome-wide gene regulatory network (*47*). In this context, differentiating cells descend toward stable regions of the network (attractor states), characterized by restricted gene expression, whereas metastable cells broadly sample the network, maintaining higher differentiation potential (*47*). Future investigations of this phenomenon, and its relationship to single-cell gene counts, may reveal new mechanisms of stem cell regulation and lineage commitment. However, further studies will be needed to confirm the validity of this model across additional tissue compartments, developmental time points, and phenotypic states.

In summary, we have shown that the number of expressed genes per cell is a hallmark of developmental potential. By exploiting this property of scRNA-seq data, we developed a general framework for resolving single-cell differentiation hierarchies. We envision that our approach will complement existing scRNA-seq analysis strategies, with implications for the identification of immature cells and their developmental trajectories in complex tissues throughout multicellular life.

REFERENCES AND NOTES

1. J. E. Visvader, H. Clevers, *Nat. Cell Biol.* **18**, 349–355 (2016).
2. K. Kretzschmar, F. M. Watt, *Cell* **148**, 33–45 (2012).
3. J. Seita, I. L. Weissman, *Wiley Interdiscip. Rev. Syst. Biol. Med.* **2**, 640–653 (2010).
4. J. E. Sulston, E. Schierenberg, J. G. White, J. N. Thomson, *Dev. Biol.* **100**, 64–119 (1983).
5. L. Kester, A. van Oudenaarden, *Cell Stem Cell* **23**, 166–179 (2018).
6. W. Saelens, R. Cannoodt, H. Todorov, Y. Saeys, *Nat. Biotechnol.* **37**, 547–554 (2019).
7. X. Qiu *et al.*, *Nat. Methods* **14**, 979–982 (2017).
8. G. La Manno *et al.*, *Nature* **560**, 494–498 (2018).
9. J. Shin *et al.*, *Cell Stem Cell* **17**, 360–372 (2015).
10. M. F. Clarke, *N. Engl. J. Med.* **380**, 2237–2245 (2019).
11. D. Grün *et al.*, *Cell Stem Cell* **19**, 266–277 (2016).
12. A. E. Teschendorff, T. Enver, *Nat. Commun.* **8**, 15599 (2017).
13. M. Guo, E. L. Bao, M. Wagner, J. A. Whitsett, Y. Xu, *Nucleic Acids Res.* **45**, e54 (2017).
14. F. J. Müller *et al.*, *Nature* **455**, 401–405 (2008).
15. T. M. Malta *et al.*, *Cell* **173**, 338–354.e15 (2018).
16. A. Liberzon *et al.*, *Cell Syst.* **1**, 417–425 (2015).
17. M. B. Gerstein *et al.*, *Nature* **489**, 91–100 (2012).
18. A. Lachmann *et al.*, *Bioinformatics* **26**, 2438–2444 (2010).
19. B. Treutlein *et al.*, *Nature* **509**, 371–375 (2014).
20. I. C. Macaulay *et al.*, *Cell Rep.* **14**, 966–977 (2016).
21. L. Yan *et al.*, *Nat. Struct. Mol. Biol.* **20**, 1131–1139 (2013).
22. R. Jaenisch, R. Young, *Cell* **132**, 567–582 (2008).
23. G. N. Aughey, A. Estacio Gomez, J. Thomson, H. Yin, T. D. Southall, *eLife* **7**, e32341 (2018).
24. N. C. Gomez *et al.*, *Cell Rep.* **17**, 1607–1620 (2016).
25. A. Pękowska *et al.*, *Cell Syst.* **7**, 482–495.e10 (2018).
26. K. M. Loh *et al.*, *Cell* **166**, 451–467 (2016).
27. L. F. Chu *et al.*, *Genome Biol.* **17**, 173 (2016).
28. S. Huang, Y. P. Guo, G. May, T. Enver, *Dev. Biol.* **305**, 695–713 (2007).
29. M. Plass *et al.*, *Science* **360**, eaaq1723 (2018).
30. J. A. Farrell *et al.*, *Science* **360**, eaar3131 (2018).
31. B. Hie, B. Bryson, B. Berger, *Nat. Biotechnol.* **37**, 685–691 (2019).
32. Tabula Muris Consortium. *Nature* **562**, 367–372 (2018).
33. A. L. Haber *et al.*, *Nature* **551**, 333–339 (2017).
34. A. Wilson *et al.*, *Cell* **135**, 1118–1129 (2008).
35. J. Y. Chen *et al.*, *Nature* **530**, 223–227 (2016).
36. D. R. Pattabiraman *et al.*, *Science* **351**, aad3680 (2016).
37. E. Lim *et al.*, *Nat. Med.* **15**, 907–913 (2009).
38. K. Polyak, *J. Clin. Invest.* **117**, 3155–3163 (2007).
39. M. Shehata *et al.*, *Breast Cancer Res.* **14**, R134 (2012).
40. H. Kouros-Mehr, E. M. Slorach, M. D. Sternlicht, Z. Werb, *Cell* **127**, 1041–1055 (2006).
41. R. W. Hsieh *et al.*, *bioRxiv* **317776** [Preprint]. 10 May 2018.
42. C. R. Graveel, D. Tolbert, G. F. Vande Woude, *Cold Spring Harb. Perspect. Biol.* **5**, a009209 (2013).
43. C. E. Barkauskas *et al.*, *J. Clin. Invest.* **123**, 3025–3036 (2013).
44. Q. A. Liu, M. O. Hengartner, *Curr. Biol.* **9**, 1347–1350 (1999).
45. G. K. Marinov *et al.*, *Genome Res.* **24**, 496–510 (2014).
46. S. Melcer, E. Meshorer, *Essays Biochem.* **48**, 245–262 (2010).
47. S. Huang, G. Eichler, Y. Bar-Yam, D. E. Ingber, *Phys. Rev. Lett.* **94**, 128701 (2005).
48. J. S. Packer *et al.*, *Science* **365**, eaax1971 (2019).
49. R. Gazit *et al.*, *J. Exp. Med.* **211**, 1315–1331 (2014).

ACKNOWLEDGMENTS

We thank A. Chaudhuri, M. Vahid, T. Raveh, A. Gentles, and A. Alizadeh for critical feedback on the manuscript. We are grateful to S. Bobo for assistance with patient specimen acquisition, R. Sinha and C. K. F. Chan for provision of data, C. L. Liu for assistance with the website, P. Lovelace and S. Weber for assistance with FACS, and S. Sim for assistance with scRNA-seq libraries. **Funding:** This work was supported by grants from the National Cancer Institute (A.M.N., R00CA187192-03; M.F.C., 5R01CA100225-09; G.S.G., PHS grant no. CA09302), the Stinehart-Reed Foundation (A.M.N.), the Stanford Bio-X Interdisciplinary Initiatives Seed Grants Program (IIP) (A.M.N., M.F.C.), the Virginia and D.K. Ludwig Fund for Cancer Research (A.M.N., M.F.C), the U.S. Department of Defense (W81XWH-11-1-0287 and W81XWH-13-1-0281 to M.F.C.; S.S.S., W81XWH-12-1-0020), a National Science Foundation Graduate Research Fellowship (DGE-1656518 to M.J.B.), Stanford Bio-X Bowes Graduate Student Fellowship (G.S.G.), and the Stanford Medical Science Training Program (G.S.G.). **Author contributions:** G.S.G. and A.M.N. developed the concept for CytoTRACE, designed related experiments, and analyzed the data with assistance from S.S.S., D.J.W., and M.F.C. G.S.G. and A.M.N. wrote the manuscript with assistance from S.S.S. G.S.G. and A.M.N. performed the bioinformatics analyses with assistance from D.J.W., A.B., A.M., M.J.B., and F.I. G.S.G, A.M., and A.B. developed the website with input from A.M.N. S.S.S. generated the human breast cancer scRNA-seq data with assistance from A.H.K., R.W.H., S.C., M.Z., F.A.S., N.A.L., D.Q., and F.B.Y. S.S.S. performed the mouse experiments under the supervision of M.F.C. F.M.D. assisted with the collection of patient specimens. All authors commented on the manuscript at all stages. **Competing interests:** G.S.G., S.S.S, M.F.C., and A.M.N. are inventors on a provisional patent application filed by Stanford University (US 62/852,231) that covers methods described in this work. **Data and materials availability:** Details of publicly available datasets are provided in materials and methods and table S1. The scRNA-seq expression data generated in this study are available at https://cytotrace.stanford.edu and have been deposited with the Gene Expression Omnibus (www.ncbi.nlm.nih.gov/geo) under accession code GSE138536.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/367/6476/405/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S19
Tables S1 to S8
References (*50–128*)

View/request a protocol for this paper from *Bio-protocol*.

21 February 2019; resubmitted 3 August 2019
Accepted 18 December 2019
10.1126/science.aax0249

REPORT

NANOPHOTONICS

Attosecond coherent manipulation of electrons in tunneling microscopy

M. Garg[1]* and K. Kern[1,2]*

Nanoelectronic devices operating in the quantum regime require coherent manipulation and control over electrons at atomic length and time scales. We demonstrate coherent control over electrons in a tunnel junction of a scanning tunneling microscope by means of precise tuning of the carrier-envelope phase of two-cycle long (<6-femtosecond) optical pulses. We explore photon and field-driven tunneling, two different regimes of interaction of optical pulses with the tunnel junction, and demonstrate a transition from one regime to the other. Our results show that it is possible to induce, track, and control electronic current at atomic scales with subfemtosecond resolution, providing a route to develop petahertz coherent nanoelectronics and microscopy.

Efforts to achieve atomic-scale control and resolution concurrently in space and time (*1*) have demonstrated that isolated photon pulses lasting only ~100 attoseconds (as) have the capability to probe electron dynamics at its natural time scales but fare poorly at the spatial resolution (*2*). Electron pulses, on the other hand, have the desired spatial resolution, but their compression to electronic time scales is extremely challenging (*3*). A promising route to attaining this goal is the integration of ultrashort laser pulses with a scanning tunneling microscope (STM). Since the first demonstration of topographic imaging of atomic surfaces with an STM (*4*, *5*), continuous efforts have been made to integrate ultrahigh temporal resolution of laser pulses with extreme spatial resolution of an STM (*6*–*9*). However, thermal effects induced by the high-energy laser pulses have plagued the development in this field (*10*, *11*). Recently, the development of terahertz STM (*12*–*14*) has shown the capability to trace the orbital-scale dynamics in a molecule with picosecond temporal resolution (*15*). Nevertheless, the natural electronic time scale of a few femtoseconds down to hundreds of attoseconds (*2*, *16*, *17*) with atomic spatial resolution has remained out of reach.

Terahertz pulses induce lower thermal effects at the STM junction owing to the reflective nature of most materials to terahertz radiation; ultrashort laser pulses ($\tau < 6$ fs with a few nanojoules of energy) on the other hand, with electronic temporal resolution owing to their extreme temporal confinement, couple extremely weakly to the atomic lattice, hence lowering the thermal effects at the STM junction. In our experiments, carrier-envelope phase (CEP)–stable <6-fs (central wavelength, λ_c ~ 810 nm) laser pulses with ~2.5 nJ of energy coming at a repetition rate of 80 MHz are focused by an off-axis parabolic mirror (OAPM) into the apex of a nanotip in an STM junction (Fig. 1A). The angle of incidence at the junction is ~7° with respect to the sample surface (see supplementary text section S1 in the supplementary materials) (*18*). Linearly polarized laser pulses with their polarization axis along the tip were used, with strong geometric confinement of the laser pulses at the STM junction leading to a field enhancement (α) of the pulses to ~30 times with respect to the incident field strength. A finite-difference time-domain method (see section S2) (*18*) with commercial software (*19*) has been used to calculate the spatial distribution of the laser electric field at the STM junction (Fig. 1D). An interferometric autocorrelation trace of the laser-induced tunneling current demonstrating temporal resolution at the STM junction is shown in Fig. 1E. Laser pulses were mechanically modulated at 5-kHz frequency to lock-in detect the current in the STM. Thermal noise induced by the laser pulses on the tunneling gap has been meticulously characterized by capacitance measurement at high frequencies (see section S3) (*18*); negligible thermal noise (<1 pm) was measured for mechanical modulation frequencies above 1 kHz for the laser pulses.

Laser-induced tunneling current was recorded as a function of increasing intensity of the driver (Fig. 1F). To avoid wavefront distortion and changing of focal conditions at the STM junction, the laser intensity was varied by inserting a combination of ultrathin (~10-μm-thick) pellicle beam splitters into the beam path. A rapid nonlinear increase in the laser-induced tunneling current is observed for intensities below 1.4×10^{13} W/cm^2 (Fig. 1F)—this behavior is a characteristic of multiphoton processes (*2*). The fitting of points on the left side of the dotted green line in Fig. 1F reveals second- and third-order optical nonlinearities ($\propto I^2 + I^3$) playing the dominant role in this regime of interaction (fig. S6) (*18*). With a central photon energy of 1.55 eV and the work function of Pt (E_{Barrier}) at ~5 eV, a two-to three-order photon process is justified. Once an electron absorbs two to three photons and goes to a spatially delocalized virtual state above the Fermi level (below the barrier), where it has finite spatial overlap with the high-lying spatially delocalized vacant states on the other side of the junction, the electron tunnels to the other side of the barrier (Fig. 1B). Assisted by three-photon absorption—photons from the blue tail of the laser spectrum (~2 eV)—the electrons can also make an over-the-barrier transition. The direction of laser-induced electron tunneling is primarily defined by the direct current (DC) bias at the junction; a positive bias and a zero bias imply the direction of laser tunneling from nanotip to the sample. At higher intensities >1.4×10^{13} W/cm^2, to the right of the dotted green line in Fig. 1F, the yield of laser-induced current increases much less nonlinearly. This feature is a hallmark of transition from multiphoton (Fig. 1B) to field-driven tunneling (Fig. 1C) (*2*). The transition from photon to field-driven tunneling regime occurs when the Keldysh parameter γ becomes less than 1 (Fig. 1F) (*2*). The Keldysh parameter is defined as the square root of the ratio between tunnel barrier and the ponderomotive kinetic energy, U_P. $\gamma = \sqrt{\frac{E_{\mathrm{Barrier}}}{2U_P}}$ with $U_P = \frac{e^2\alpha^2E^2}{4m_e\omega^2}$, where e and m_e are the electronic charge and mass, respectively (*2*), and ω is the central frequency of the laser pulses at 1.55 eV. The transition from photon to field-driven tunneling regime is well captured by a time-dependent perturbation theoretical model using a strong-field approximation (see section S4) (*18*, *20*, *21*). The polarization control over electron tunneling is shown in Fig. 1G, the polarization angle of $\pi/2$ corresponds to the laser polarization axis oriented along the direction of the nanotip.

We also implement a feed-forward CEP stabilization scheme (*22*–*24*). This technique produces a train of CEP-stable pulses whose CEP can be very precisely tuned (section S1) (*18*). The field-driven tunneling current is recorded as a function of varying CEP of the laser pulses for various widths of the tunnel gap (Fig. 2A). The tunnel junction is driven by laser pulses of intensity $\alpha^2 I = 2.5 \times 10^{13}$ W/cm^2, and the bias at the tunnel junction is 100 mV. Crosscuts of the CEP map in Fig. 2A are shown in Fig. 2, B and C, for two different tunnel gap widths of $\Delta z = -3$ and -11 Å, respectively. The time-varying electric field of the laser pulses drives electron tunneling to either side of the junction, as shown by temporal evolution of the

[1]Max Planck Institute for Solid State Research, 70569 Stuttgart, Germany. [2]Institut de Physique, Ecole Polytechnique Fédérale de Lausanne, 1015 Lausanne, Switzerland.
*Corresponding author. Email: mgarg@fkf.mpg.de (M.G.); k.kern@fkf.mpg.de (K.K.)

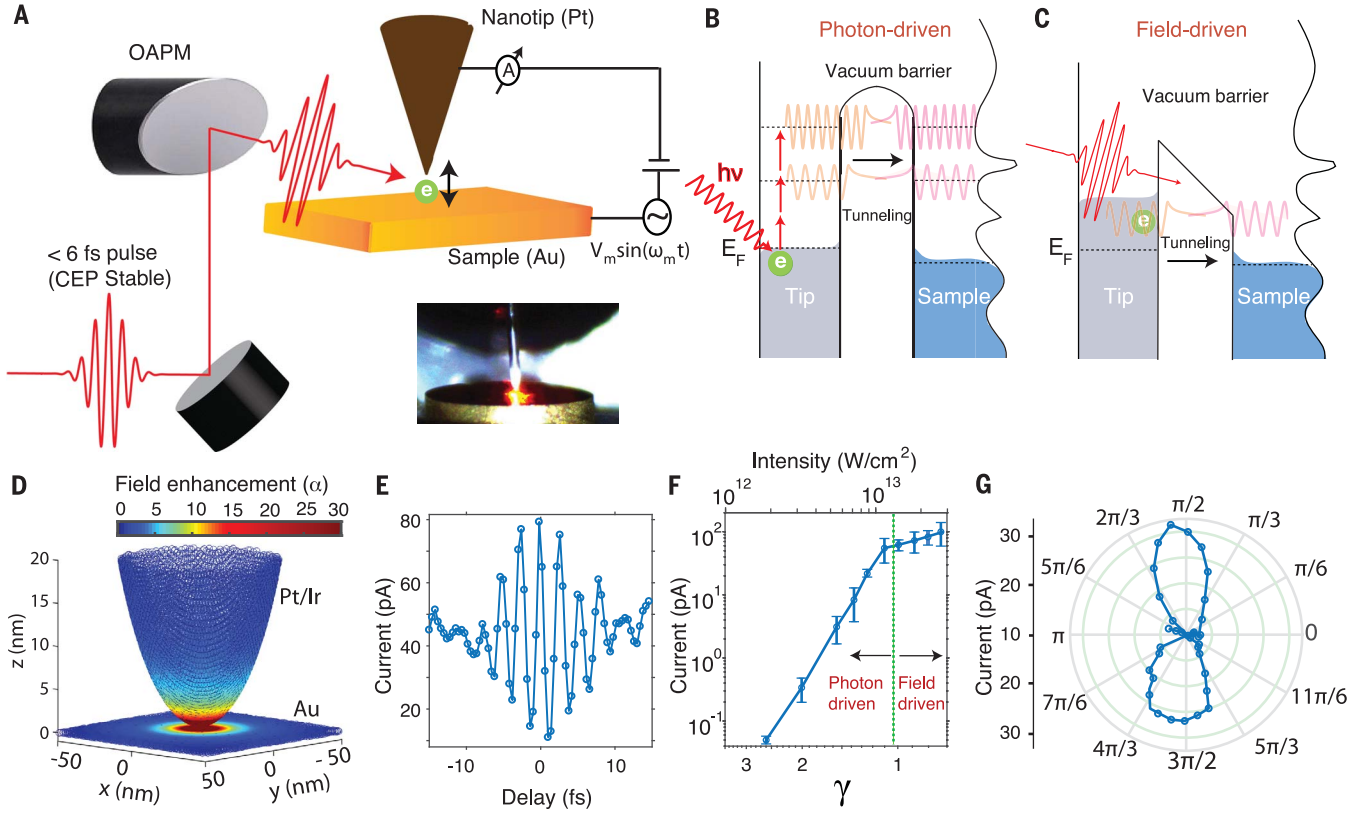


Fig. 1. Transition from photon-driven to optical field–driven electron tunneling. (**A**) Schematic of the experimental setup: <6-fs carrier-envelope phase (CEP)–stable optical pulses are focused by an off-axis parabolic mirror (OAPM) at the apex of a Pt/Ir nanotip in tunneling contact with the Au surface. The inset shows an image of the tunnel junction. Sinusoidal modulation voltage, $V_m\sin(\omega_m t)$, at a frequency ω_m. (**B** and **C**) Illustration of tunneling of electrons driven by photons (B) and by the electric field of the optical pulses (C) at the tunnel junction. $h\nu$, photon energy; E_F, Fermi level; e, electron. (**D**) Evaluated spatial distribution of the electric field across the tunnel junction with a gap of 1 nm between nanotip and the surface. The local field enhancement factor α is denoted by the color code in the color bar. (**E**) Interferometric autocorrelation trace of the tunnel current induced by an ~6-fs pulse at the tunnel junction. (**F**) Laser-induced tunneling current as a function of increasing intensity of the driver (enhanced intensity $\alpha^2 I$). For intensities <1.4 × 10^{13} W/cm^2 ($\gamma > 1$), tunneling is photon-driven; at higher intensities ($\gamma < 1$), optical field–driven tunneling dominates. (**G**) The optical field–driven tunneling current as a function of rotating polarization of the field ($\alpha^2 I$ = 3 × 10^{13} W/cm^2) along the axis of the tip, polarization angles of 0 and π/2 correspond to electric field oscillations perpendicular and parallel to the tip, respectively. Set parameters at the tunnel junction: I_{set} = 1 nA and V_{DC} = 100 mV.

tunnel barrier on exposure to <6-fs laser pulses (Fig. 2D). Strong modulation of the tunnel barrier is shown by crosscuts at two discrete times (Fig. 2E). Tunneling current modulation of nearly 50% was obtained by the variation of CEP of the laser pulses. Depending on the CEP, a maximum or minimal current is produced, and a 2π periodicity is observed. We evaluated current densities using the Simmons tunneling model (*25*) for two different CEPs of the laser pulse (Fig. 2F). A laser pulse of $\varphi_{CEP} = 0$ produces primarily a positive transient current density, whereas a laser pulse of $\varphi_{CEP} = \pi$ produces both positive and negative transients of current density J; nevertheless, an effective (integrated) lower positive current is generated compared with $\varphi_{CEP} = 0$, which is consistent with our experimental observations. A CEP variation of 0.1π corresponds to a temporal shift between carrier wave and envelope of the laser pulse of nearly 200 as and clearly demonstrates our capability to control electron tunneling at such small time scales. These results attest to the fact that no thermal effect is present in the measurements; a CEP change of laser pulses leads to a change only of the temporal profile of the pulse without affecting the average laser power (energy) at the tunnel junction.

The spectral response of atoms to impulsive excitation by strong laser pulses as a function of CEP has previously unraveled the response time of bound electrons in atomic systems (*26*). Relative changes of electron rescattering times have also been measured during high-harmonic generation (*27*). The white dashed line in Fig. 2A shows the shift of the maxima in the CEP curves for decreasing tunnel gaps. A shift of 0.5π in the position of the maxima in the CEP curves in Fig. 2A from high to low gap width corresponds to a shift in the time of nearly 600 as. This time difference represents the relative change in time spent by the electrons in the tunnel barrier, on change of gap width by ~1 nm in the present study. At a gap width of $\Delta z = -10$ Å, the electrons spend an additional 600 as in the tunnel barrier as compared with a tunnel gap width of $\Delta z = -1$ Å. Our measured relative shift of the maxima in the CEP curves with decreasing gap widths is consistent with a one-dimensional time-dependent Schrödinger equation model (see section S5) (*18*). The calibration of CEP in the future will lead us to an absolute experimental measurement of the time an electron needs to tunnel through a vacuum barrier and its dependence on the barrier width.

With control over optical field–driven tunneling with CEP and DC bias at the tunnel junction (Fig. 3A), crosscuts of the CEP map for two different DC biases of 400 and −800 mV are shown in Fig. 3, B and C, respectively. For negative DC bias, the field-driven tunneling current becomes negative. A DC bias at the tunnel junction predefines the slope of the tunneling barrier, and the laser-induced modulation of the tunneling barrier is on top

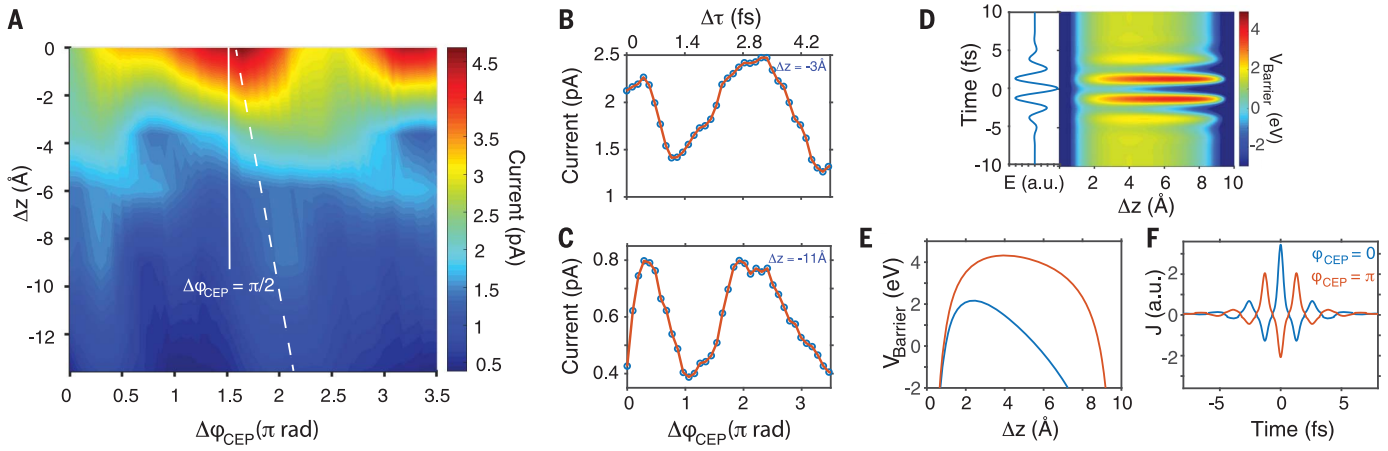


Fig. 2. CEP control over electrons at the tunnel junction. (**A**) Modulation of the field-driven tunneling current as a function of CEP of the laser pulses ($\alpha^2 I = 3 \times 10^{13}$ W/cm^2) for decreasing tunnel gaps. The color bar indicates the amplitude of the field-driven current in picoamperes. The white dashed line shows the shift of the maxima of the tunneling current with decreasing tunnel gap; the solid white line is vertical. (**B** and **C**) Crosscuts of color map in Fig. 2A for two different tunnel gaps, Δz = −3 and −11 Å, respectively. (**D**) Temporal evolution of the tunnel barrier in the presence of an enhanced ~6-fs optical pulse (right panel). Electric field of the pulse is shown in the left panel. Color bar indicates the tunnel barrier in electron volts. (**E**) Tunnel barrier at two different moments in time: −10 fs (red curve) and 0 fs (blue curve). (**F**) Evaluated current density J for two different CEPs of the laser pulse. a.u., arbitrary units.

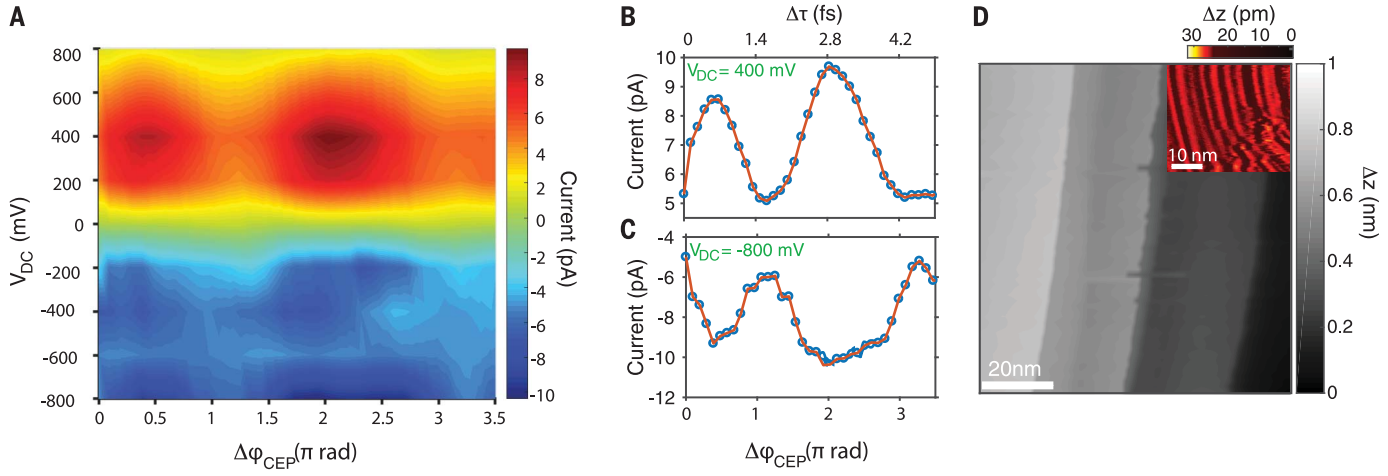


Fig. 3. Taming flow of electrons with optical phase and DC bias at the tunnel junction, and optical field–driven tunneling microscopy. (**A**) Dependence of field-driven tunneling current on CEP of the laser pulses for varying DC voltages at the tunnel junction. The color bar indicates the amplitude of the field-driven current in picoamperes. (**B** and **C**) Crosscuts of the color map in Fig. 3A for two different DC voltages at the tunnel junction, V_{DC} = 400 and −800 mV, respectively. (**D**) Surface topography with atomic terraces (steps) on a Au surface. Inset shows atomic reconstruction on the Au surface. The tunneling current for the topographic imaging is solely generated by the laser pulses. Imaging parameters are V_{DC} = 0 V and I_{set} = 100 pA.

of this: $V_{Barrier}(z,t) = V_{DC}(z) + V_{Laser}(z,t)$ (Fig. 2D). Comparison of crosscuts in Fig. 3, B and C, reveals that maxima in the tunneling current for positive and negative biases are shifted by π rad. This observation is due to a flip in polarity of the DC bias that introduces an extra π rad shift in $V_{Barrier}$, hence causing the maxima in the CEP curves to move by π for negative DC biases.

A spatially localized optical field–driven tunneling current at the atomic scale can be used to map the topography of surfaces with atomic precision (Fig. 3D); the lock-in current of laser pulses is used as the source of constant current, I_{set}, during topographic imaging in this case. Four atomic planes of the Au(111) surface can be seen. The present results rule out the possibility of generation of delocalized photoelectrons from the surface of Au by the laser pulses. Optical field–driven tunneling microscopy also enables the topographic imaging of the atomic reconstruction (herringbone structure) on the Au(111) surface (Fig. 3D, inset). Optical field–driven tunneling microscopy can therefore have the same spatial resolution, both lateral and vertical, as in a conventional STM. Note that atomic imaging with terahertz field–driven electron tunneling has been previously reported on a Si(111) surface (*14*) and on a single molecule (*15*); also, atomically confined optical light has recently been used to spatially map the vibrational modes of a single molecule (*28*).

As an illustrative demonstration of the capability of the powerful union of CEP-stable ultrashort laser pulses with STM to probe electronic dynamics at the nanoscale, we study carrier decay dynamics of collective oscillations of electrons, i.e., localized surface plasmon resonances (LSPRs), in a Au nanorod (Fig. 4B). Au nanorods of dimensions 150 nm by 20 nm by 30 nm (length times width times height) were

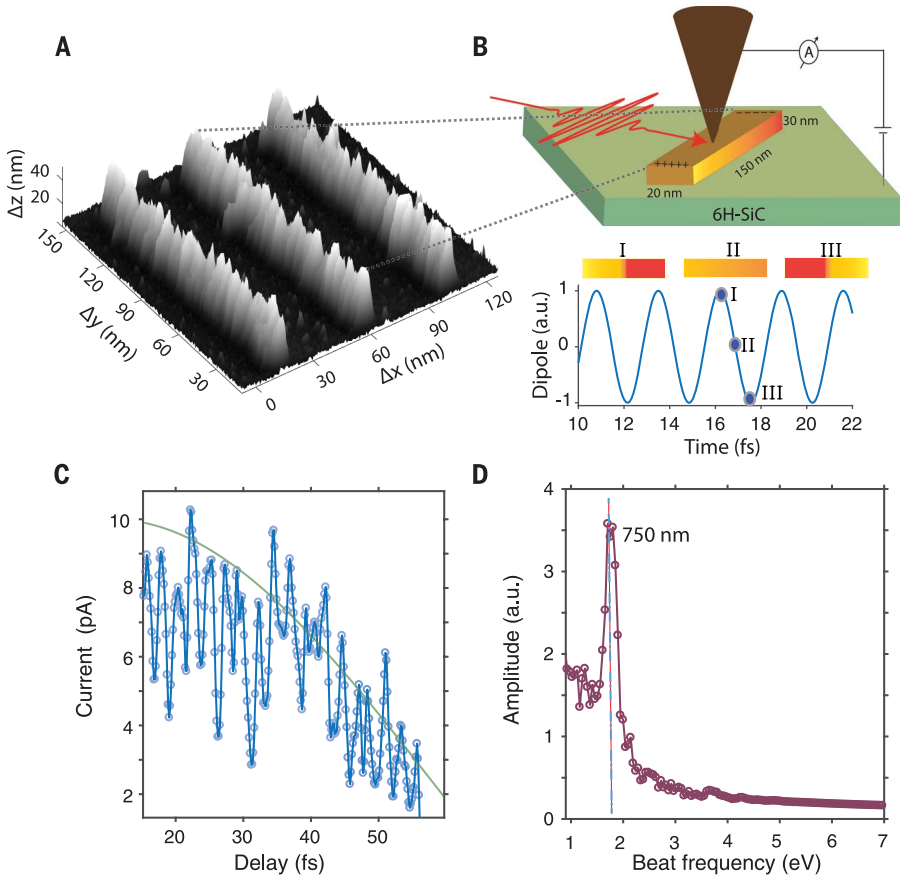


Fig. 4. Nanoscale tracing of the relaxation dynamics of localized surface plasmons in Au nanorods. (**A**) STM image of Au nanorods on an n-doped 6H-SiC, DC bias on the sample was +2 V, and set current during topographic imaging was 100 pA. (**B**) Schematic of a Au nanorod on a 6H-SiC surface (top panel). Schematic of temporal oscillation of the plasmons in the nanorods (bottom panel), along with the spatial distribution of electron density at three discrete times annotated by roman numerals; here, a red color-coded distribution implies higher electron density compared with the yellow color. (**C**) Temporal modulation of the laser-induced tunneling current as a function of delay between pump and probe pulses in the noninterferometric temporal overlap region of the pulses. (**D**) Observed frequency of the temporal oscillation of plasmons in the nanorods.

grown on an atomically flat surface of 6H-SiC (see section S6) (*18*). A constant current STM image of Au nanorods is shown in Fig. 4A. A semiconducting surface of a bandgap of ~3eV was chosen as a substrate for having little interference with the decay dynamics of the LSPR (resonance ~ 1.7 eV) and at the same time allowing DC current to pass through it for the tunneling contact with the nanotip of the STM. A Pt/Ir nanotip was placed in tunnel contact on top of an Au nanorod, and pump-probe measurement with laser pulses passing through a Michelson interferometer was performed. Laser pulses were polarized along the axis of the Au nanorods (p-polarized) to effectively excite the LSPR oscillations—the broad bandwidth of our laser pulses enabled us to coherently excite the longitudinal mode of the surface plasmons (Fig. 4B, top panel). A representation of the coherent space-time oscillation of plasmons in the Au nanorod under the nanotip is shown in Fig. 4B, bottom panel. The delay between pump-probe pulses from −10 to +10 fs is of interferometric nature (Fig. 1E); for decay dynamics of LSPR we look into the temporal window >15 fs. Variation of the laser-induced lock-in current as a function of delay between pump and probe pulses (Fig. 4C) shows a distinct oscillatory behavior of the tunnel current, with a Fourier transform of the temporal oscillation revealing a frequency of ~1.7 eV or 750 nm, which matches quite well with the LSPR of longitudinal modes in Au nanorods of the dimension used in the present experiment (Fig. 4D). For an aspect ratio (length/width) of ~8, the Au nanorod's resonance of the longitudinal mode is expected at ~750 nm (*29*, *30*). The decay of the LSPR dipole oscillations can be clearly seen in the temporal evolution of the tunneling current in Fig. 4C. Coherent oscillation of plasmons excited by the laser pulse decay on time scales of ~40 fs, as shown by the green curve in Fig. 4C. This decay arises because of carrier-carrier scattering (*31*); slower decay of plasmons owing to their coupling to the phonons (carrier-phonon scattering) takes place on time scales of a few picoseconds, which are not investigated in the present study.

The capability, demonstrated here, to probe collective oscillations of quasiparticles concurrently with angstrom-scale spatial and subfemtosecond temporal resolution opens new avenues to study quasiparticle dynamics in superconductor and two-dimensional layered materials with extraordinary resolution. For example, the diffusion dynamics of excitons around atomic defects can be studied and can help boost their lifetime in future excitonic devices (*32*). The combination of CEP-stable two-cycle (<6-fs) pulses with an STM opens a plethora of opportunities in nanoscience and microscopy. The functionalities of light wave–driven nanoelectronics on surfaces can be probed and controlled with extreme spatial (subangstrom) and temporal resolution (subfemtosecond)—understanding of such dynamics holds the key to future developments in this field.

REFERENCES AND NOTES

1. V. A. Lobastov, R. Srinivasan, A. H. Zewail, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 7069–7073 (2005).
2. F. Krausz, M. Ivanov, *Rev. Mod. Phys.* **81**, 163–234 (2009).
3. C. Kealhofer *et al.*, *Science* **352**, 429–433 (2016).
4. G. Binnig, H. Rohrer, C. Gerber, E. Weibel, *Appl. Phys. Lett.* **40**, 178–180 (1982).
5. G. Binnig, H. Rohrer, C. Gerber, E. Weibel, *Phys. Rev. Lett.* **49**, 57–61 (1982).
6. G. Nunes Jr., M. R. Freeman, *Science* **262**, 1029–1032 (1993).
7. Y. Terada, S. Yoshida, O. Takeuchi, H. Shigekawa, *Nat. Photonics* **4**, 869–874 (2010).
8. S. Dey, D. Mirell, A. R. Perez, J. Lee, V. A. Apkarian, *J. Chem. Phys.* **138**, 154202 (2013).
9. J. Lee, S. M. Perdue, D. Whitmore, V. A. Apkarian, *J. Chem. Phys.* **133**, 104706 (2010).
10. V. Gerstner, A. Knoll, W. Pfeiffer, A. Thon, G. Gerber, *J. Appl. Phys.* **88**, 4851–4859 (2000).
11. V. Gerstner, A. Thon, W. Pfeiffer, *J. Appl. Phys.* **87**, 2574–2580 (2000).
12. T. L. Cocker *et al.*, *Nat. Photonics* **7**, 620–625 (2013).
13. K. Yoshioka *et al.*, *Nat. Photonics* **10**, 762–765 (2016).
14. V. Jelic *et al.*, *Nat. Phys.* **13**, 591–598 (2017).
15. T. L. Cocker, D. Peller, P. Yu, J. Repp, R. Huber, *Nature* **539**, 263–267 (2016).
16. P. B. Corkum, F. Krausz, *Nat. Phys.* **3**, 381–387 (2007).
17. E. Goulielmakis *et al.*, *Nature* **466**, 739–743 (2010).
18. See supplementary materials.
19. COMSOL, Multiphysics, Version 4.4; www.comsol.com.
20. H. R. Reiss, *Phys. Rev. A* **22**, 1786–1813 (1980).
21. R. Bormann, M. Gulde, A. Weismann, S. V. Yalunin, C. Ropers, *Phys. Rev. Lett.* **105**, 147601 (2010).
22. F. Lücking, thesis, LMU München, Faculty of Physics (2014).
23. F. Lücking, A. Assion, A. Apolonski, F. Krausz, G. Steinmeyer, *Opt. Lett.* **37**, 2076–2078 (2012).
24. S. Koke *et al.*, *Nat. Photonics* **4**, 462–465 (2010).
25. J. G. Simmons, *J. Appl. Phys.* **34**, 1793–1803 (1963).
26. M. T. Hassan *et al.*, *Nature* **530**, 66–70 (2016).
27. C. Ott *et al.*, *New J. Phys.* **15**, 073031 (2013).
28. J. Lee, K. T. Crampton, N. Tallarida, V. A. Apkarian, *Nature* **568**, 78–82 (2019).
29. Y. J. Xiang *et al.*, *J. Phys. Chem. C* **112**, 3203–3208 (2008).

30. S. Link, M. B. Mohamed, M. A. El-Sayed, *J. Phys. Chem. B* **103**, 3073–3077 (1999).
31. Q. Sun *et al.*, *Light Sci. Appl.* **2**, e118 (2013).
32. D. Unuchek *et al.*, *Nature* **560**, 340–344 (2018).

ACKNOWLEDGMENTS

We thank A. Leitenstorfer, D. Brida, and T. Kurihara for fruitful discussions, and W. Stiepany, L. Krumbein, T. Reindl, and D. Paone for technical support. K.K. thanks A. Kroesen and team for their precision skills, which made his contribution possible. **Author contributions:** M.G. planned the experiments, built the experimental setup, and conducted the experiments and simulations. K.K. planned and supervised the project. Both authors interpreted data and contributed to the preparation of the manuscript. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present here or in the supplementary materials.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/367/6476/411/suppl/DC1
Supplementary Text
Figs. S1 to S8
References (*33*–*36*)

13 August 2019; accepted 30 October 2019
Published online 14 November 2019
10.1126/science.aaz1098

SUPERNOVAE

A type Ia supernova at the heart of superluminous transient SN 2006gy

Anders Jerkstrand[1,2]*, Keiichi Maeda[3,4], Koji S. Kawabata[5]

Superluminous supernovae radiate up to 100 times more energy than normal supernovae. The origin of this energy and the nature of the stellar progenitors of these transients are poorly understood. We identify neutral iron lines in the spectrum of one such supernova, SN 2006gy, and show that they require a large mass of iron (≳0.3 solar masses) expanding at 1500 kilometers per second. By modeling a standard type Ia supernova hitting a shell of circumstellar material, we produce a light curve and late-time iron-dominated spectrum that match the observations of SN 2006gy. In such a scenario, common envelope evolution of a progenitor binary system can synchronize envelope ejection and supernova explosion and may explain these bright transients.

Superluminous supernovae (SNe) are a rare type of astrophysical explosion that emit large amounts of energy, more than can be explained by standard supernova powering mechanisms. One of the first to be observed was SN 2006gy, which showed narrow hydrogen lines (supernova type IIn) indicating interaction with a circumstellar medium (CSM). SN 2006gy radiated about 10^{51} erg in a few months (*1*, *2*). Proposed mechanisms to produce such a transient include large amounts of radioactivity in a pair-instability supernova (PISN) (*1*), a collision between a core-collapse supernova (CCSN) and a Luminous Blue Variable–like eruption (*3*), and a pulsational pair-instability explosion (*4*). However, the nature of SN 2006gy remains unclear and disputed.

A spectrum of the supernova at 394 days after explosion (*5*) revealed a set of emission lines around 8000 Å that could not be identified. Figure 1 shows this spectrum, after removal of light echoes (light from earlier epochs reflected by circumstellar dust) (*6*). By searching atomic line lists, we determined that these lines all coincide with low-excitation, strong transitions in Fe I (*6*).

These lines are predicted by emission line models for slow-expanding supernova ejecta (*7*). They arise from the z^7D multiplet of Fe I at 2.4 eV above the ground state, which is excited by thermal electron collisions at typical supernova temperatures of a few thousand kelvin. Most supernovae have, however, too little neutral iron and expansion velocities that are too high to exhibit these lines in their spectrum. In addition to these Fe I lines, the spectrum of SN 2006gy shows lines from Ca II and Fe II and is thus dominated by heavy elements, likely produced in explosive oxygen and silicon fusion. The FWHM (full-width-at-half-maximum) of these iron and calcium lines is ~1500 km s^{-1}, which corresponds to the characteristic expansion velocity of the gas at +394 days.

To obtain constraints on the iron producing this emission, we calculated a grid of iron line (Fe I and Fe II) emission models with the spectral synthesis code SUMO (*8*), varying the iron mass, temperature, ionization, and clumping (degree of compression compared to a uniform distribution) (*6*). Small masses of iron [$\lesssim$0.1 solar masses ($M_\odot$)] cannot produce the observed luminosity for any physical conditions (Fig. 2). To both fulfill ionization balance and reproduce the observed emission ratio between Fe I and Fe II lines, another constraint $M(\mathrm{Fe}) \gtrsim 0.3\ M_\odot$ can be derived, assuming a filling factor (the inverse of clumping) of 0.1 to 1 (*6*). Lower masses lead to an ionization state that is too high, producing emission mainly from Fe II and Fe III, rather than Fe I. The iron mass

[1]Max Planck Institute for Astrophysics, Karl-Schwarzschild-Straße 1, 85748 Garching, Germany. [2]The Oskar Klein Centre, Department of Astronomy, Stockholm University, Albanova 10691, Stockholm, Sweden. [3]Department of Astronomy, Kyoto University, Kitashirakawa-Oiwake-cho, Sakyo-ku, Kyoto 606-8502, Japan. [4]Kavli Institute for the Physics and Mathematics of the Universe, The University of Tokyo, 5-1-5 Kashiwanoha, Kashiwa, Chiba 277-8583, Japan. [5]Hiroshima Astrophysical Science Center, Hiroshima University, 1-3-1 Kagamiyama, Higashi-Hiroshima, Hiroshima 739-8526, Japan.
*Corresponding author. Email: anders.jerkstrand@astro.su.se

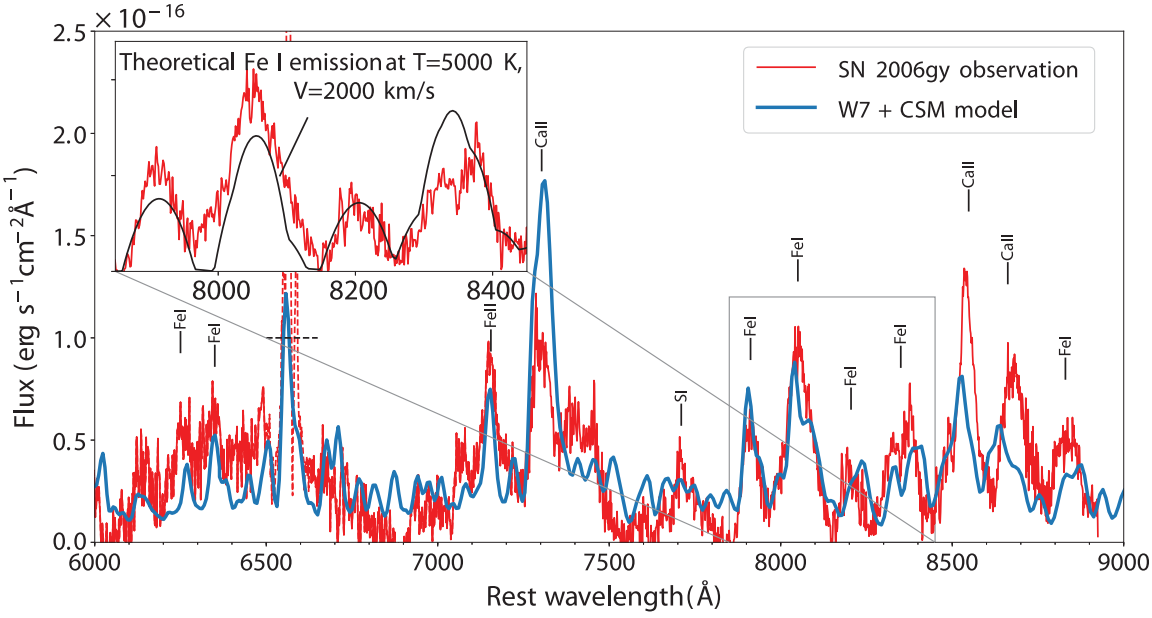


Fig. 1. Observed spectrum of SN 2006gy at +394 days (red) compared to a standard type Ia model (W7) with hydrodynamic velocities reduced by a factor of 7 (blue). The model also has 3 $M_\odot$ of CSM mixed with the Ia ejecta. The black dashed line around 6560 Å shows the upper limit on Hα emission from the supernova. The inset shows an enlarged area on the lines at 7900 to 8500 Å that we identify as due to Fe I. The black line shows a theoretical model of emission from 0.5 $M_\odot$ of Fe at 5000 K, scaled to the same distance as SN 2006gy.

limit holds also under exploration of smaller filling factors (*6*). A large mass of iron is therefore inferred, likely arising from radioactive decay of ^{56}Ni (through the intermediary ^{56}Co), the main product of explosive silicon fusion.

At +394 days after explosion, SN 2006gy was about 100 times fainter compared with previous observations at +200 days. A fundamental property of a localized CSM is that the shock will traverse the CSM on a time scale of 230 days $(R/10^{16}\ \mathrm{cm})/(v_{\mathrm{shock}}/5000\ \mathrm{km\ s^{-1}})$, where R is the radius and v_{shock} is the shock speed. Similar drops in brightness have been seen in other luminous type IIn supernovae (*9*, *10*). In its second and third year after explosion, SN 2006gy became dominated by an echo with slower decay than either interaction or radioactive powering (*11*).

The amount of initial radioactive ^{56}Ni needed to match the estimated luminosity of the supernova at +394 days is 0.5 $M_\odot$ (*6*). Figure 1 (inset) shows a theoretical emission spectrum of 0.5 $M_\odot$ of Fe I at 5000 K, scaled to the same distance as SN 2006gy, which reproduces the observed Fe I lines. These strong iron lines in SN 2006gy are difficult to reconcile with several previously suggested models in which there is no ^{56}Ni production, such as the collisions of pulsational pair instability shells (*4*).

Core-collapse supernovae (arising when the core of a massive star collapses to a neutron star or black hole) produce much less ^{56}Ni, typically $\lesssim$0.1 $M_\odot$ (*12*, *13*), although a small fraction, almost all in the broad-lined type Ic class, has inferred ^{56}Ni masses $\gtrsim$0.3 $M_\odot$ (*14*). Such an engine for SN 2006gy can, however, be excluded on two grounds. First, for a CCSN to produce 0.5 $M_\odot$ of ^{56}Ni, the explosion energy must exceed 10^{52} erg (*15*). Because wind-driven mass loss and pair instability pulsations limit the final mass of the supernova progenitor to about 10 solar masses, these supernovae expand fast (6000 to 12,000 km s^{-1}) , as confirmed by late-time spectroscopic observations (*14*). For such a supernova to reach a velocity of 1500 km s^{-1} after a few hundred days, the ejecta must have been strongly decelerated by a massive CSM, with associated re-radiation of the bulk of the original kinetic energy (~10^{52} erg). The observed radiated energy in SN 2006gy is an order-of-magnitude lower at 10^{51} erg (*6*), preventing any self-consistent CCSN scenario. Second, CCSN ejecta are dominated by oxygen, with strong [O I] lines after a few hundred days, of which SN 2006gy shows none.

Two model scenarios can explain a ^{56}Ni mass of ~0.5 $M_\odot$ expanding with 1500 km s^{-1} at 400 days: a pair-instability explosion of a ~90 $M_\odot$ He core (*16*); or a type Ia supernova [the thermonuclear explosion of a white dwarf (WD)] decelerated by strong circumstellar interaction. The ejecta mass needed to trap gamma rays from radioactive decay (which transfer the decay energy to heat) at 400 days is 1.8 $M_\odot$ (setting the optical depth $\tau_\gamma = \kappa_\gamma \rho R = 1$, where κ_γ is the gamma-ray opacity and ρ is the density), and the gamma rays therefore mainly power the supernova ejecta rather than the CSM in both cases.

We calculated model spectra for the two model scenarios with SUMO and found good agreement for both, as they have similar core structures. Figure 1 shows spectra calculated by using the type Ia explosion model W7 (*17*, *18*), with all velocities in the hydrodynamic model reduced by a factor of 7 to mimic the slowdown due to CSM interaction (leading to higher densities at any given time). We mixed the ejecta with a few solar masses of CSM material; however, the spectrum was not sensitive to this (*6*). This W7+CSM model reproduces the Fe I lines, the [Ca II] 7291, 7323 Å doublet, and the only ionized iron line seen, [Fe II] 7155 Å. The Ca II triplet at 8500 to 8700 Å is underproduced, possibly because the Ca-rich region is not compact enough; higher density favors a stronger calcium triplet.

The degeneracy between type Ia and PISN models in this late (nebular) phase can be broken by considering the earlier phases of the supernova. We calculated the total amount of light emitted by SN 2006gy using all the spectral and photometric data available in the literature (*1*, *19*, *20*). We obtain 9×10^{50} erg, close to that expected in the strong interaction limit of a type Ia supernova where a large fraction of the kinetic energy of (1 to 2) $\times$ 10^{51} erg is converted to radiation (*6*). Some previous estimates of this number were a factor of 2 to 3 higher, but were based either on single-band data with an assumed bolometric correction (*1*), or extrapolated blackbodies with high ultraviolet(UV)/blue flux (*20*). Such UV/blue emission is often blocked by line opacity in supernovae, and the spectra of SN 2006gy show such behavior (*6*). We used the radiation hydrodynamic code SNEC (*21*) to calculate light curves arising when a standard Ia SN ejecta (W7), or PISN ejecta, collide with a dense H-rich CSM. The resulting light curves for the

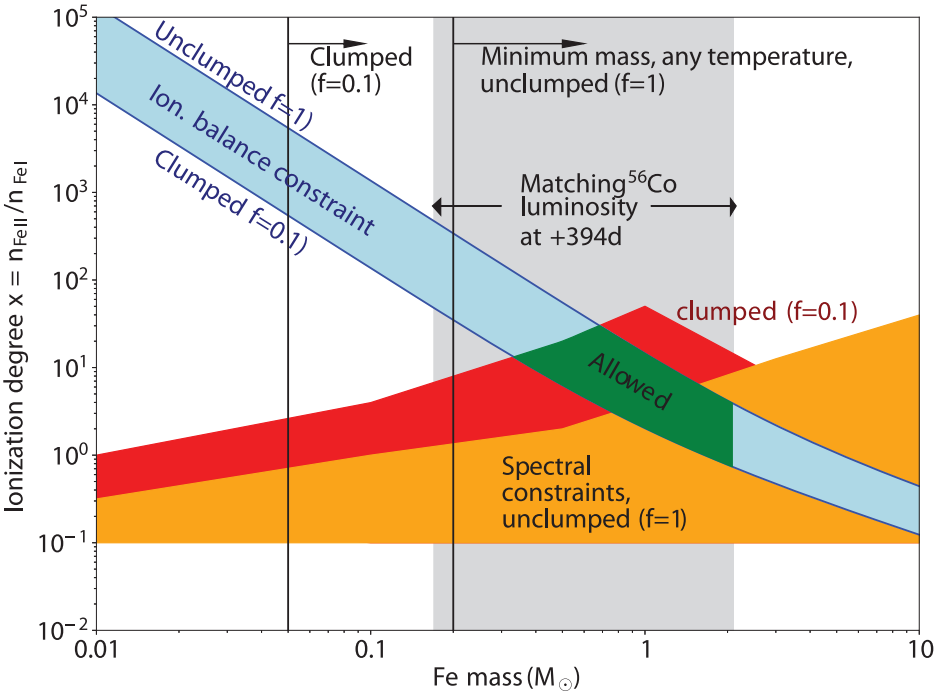


Fig. 2. Allowed domains for iron mass and ionization degree. The vertical lines show the minimum Fe masses needed to reproduce the Fe I lines at +394 days for any temperature and ionization. The blue domain outlines where ionization balance holds. The orange and red domains outline where the observed Fe II/Fe I and Fe I cluster emission ratios are reproduced, for unclumped (orange) and clumped (red) ejecta cases. The gray regime outlines where the luminosity is reproduced by the remaining amount of ^{56}Co, if the iron comes from ^{56}Ni/^{56}Co decay. The regime where all constraints are fulfilled at $f \gtrsim 0.1$ is colored green; it constrains the iron mass to $0.3\ M_\odot \lesssim M(\mathrm{Fe}) \lesssim 2.1\ M_\odot$.

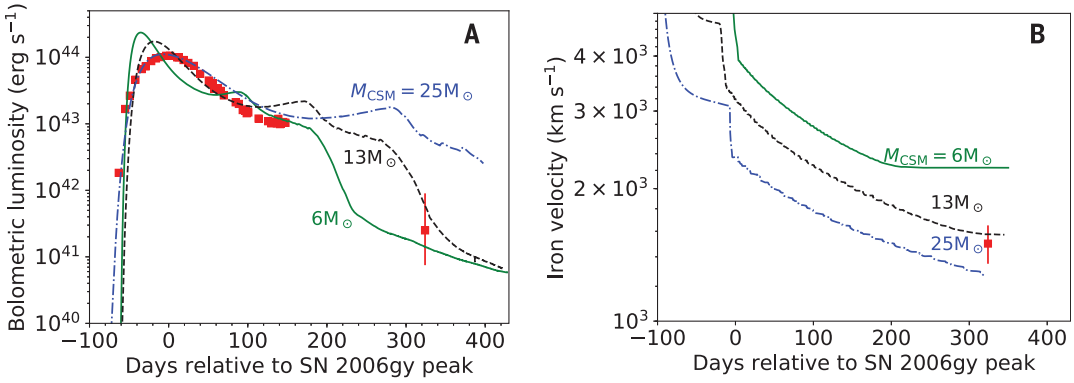


Fig. 3. Type Ia-CSM light-curve models compared to SN 2006gy (red points). (**A**) Bolometric luminosity (emission integrated over all frequencies) and (**B**) velocity of the Ia ejecta (at $M = 0.5\ M_\odot$). This iron only becomes observable in the nebular phase spectrum at +394 days after explosion (= +324 days after peak). All models have outer CSM radius $R_{CSM} = 8 \times 10^{15}$cm. The green curve is representative of models with too small CSM mass ($\lesssim 6 M_\odot$); the light curve peak is too narrow and too bright, and the deceleration is insufficient. The blue curve is representative of models with too much CSM ($\gtrsim 25\ M_\odot$); the CSM interaction powers the light curve for too long and the slowdown is too severe. The black model represents a case of CSM mass ($13\ M_\odot$) where all properties are reproduced.

Ia case match SN 2006gy if a CSM mass of about 10 $M_\odot$ is present (Fig. 3). Pair-instability supernovae, by contrast, produce light curves that are in strong disagreement with observations (fig. S6).

Inspection of the Ia-CSM hydrodynamic models shows that the ejecta are decelerated to 1500 km s^{-1} following interaction with a CSM with properties suitable for reproducing the light curve (Fig. 3). This matches the observed velocities of the Fe I lines at +394 days. The type Ia explosion energy, 1.3×10^{51} erg for the standard scenario (*18*), is accounted for by about 3×10^{50} erg still in kinetic energy at +394 days (~$15 M_\odot$at 1500 km s^{-1}, both ejecta and CSM expand with this asymptotic velocity), with the rest radiated. The "type Ia-CSM" hypothesis thus matches all observables. This scenario has been previously proposed for SN 2006gy (*2*) but was then largely forgotten, as most analyses focused on massive-star progenitors.

From the CSM extension and velocity, the CSM material must have been ejected between 10 and 200 years before the supernova explosion. A candidate scenario to explain this is common envelope evolution of a binary progenitor system, in which a white dwarf spirals into a giant or supergiant companion star. This could causally link the processes of envelope ejection and a merger with the core of the other star, producing the explosion. Such synchronization by common envelope evolution has previously been discussed in other contexts (*22*). The inspiral process has been shown to robustly transfer energy and angular momentum from the orbit to the common envelope, and eject most or all of this, while the orbital separation shrinks by a factor 100 or more (*23*, *24*).

The ejection time scale in SN 2006gy matches the time scales for common envelope ejection obtained in simulations: ~10 years for red giants engulfing WDs (*23*), and 2 to 200 years for more-massive red supergiants (RSGs) (*24*). The released orbital energy for a WD of mass M_{WD} spiralling in toward a companion with core mass M_{core} and radius R_{core} is

$$4 \times 10^{48} \left(\frac{M_{core}}{M_\odot}\right) \left(\frac{M_{WD}}{M_\odot}\right) \left(\frac{R}{R_\odot}\right)^{-1} \text{ erg}$$

where $R_\odot$ is the solar radius. This is sufficient to unbind 10 $M_\odot$ of envelope material in an extended star (binding energy 4×10^{48} erg for a typical $R = 100\ R_\odot$) and also account for the kinetic energy of the ejected envelope (10^{48} erg for 100 km s^{-1}). It is less clear how the two cores merge and explode. These steps are rarely explored in inspiral simulations, because of computational difficulties, although some results have shown that less-evolved giants merge more easily (*24*). Material may also form a disk around the two cores that could drive the final stages of merging (*25*).

A similar scenario may explain type IIa supernovae, a rare class that have spectra of type Ia at early times but later transition to type IIn (but much less luminous than SN 2006gy). One suggestion put forth is the common envelope ejection in a merger of a WD and an Asymptotic Giant Branch (AGB) star (*26*). Such a scenario has been criticized on the grounds that the final merger would have to occur by gravitational waves, which would take much longer than decades or centuries (*27*). However, the last stages of common envelope evolution are not well understood, so that conclusion may be premature.

It is possible that SN 2006gy is an extreme example of the Ia-CSM family, with higher CSM mass located closer to the supernova compared to other cases. This would be more efficient at converting kinetic energy to radiation, over a shorter time scale, leading to the extreme luminosity. It also led to strong ejecta deceleration that trapped gamma rays and produced the distinct narrow Fe lines after a few hundred days. Type IIa supernovae show longer-lasting interactions with a more extended CSM, which would not slow the expanding core sufficiently to produce a distinct signature from the inner ejecta at late times.

Other superluminous type IIn SNe such as SN 2006tf (*28*), SN 2008fz (*29*), and SN 2008am (*30*) share several similarities with SN 2006gy. The total radiated energy in these events is also around 10^{51} erg, so some may also represent a type Ia SN exploding in a massive common envelope–ejected CSM. These other supernovae were, however, much farther away, with no observable signature similar to the +394-day spectrum of SN 2006gy; attempts at late-time observations yielded either no detections or still-ongoing interaction through broad hydrogen lines (*28*–*30*).

REFERENCES AND NOTES

1. N. Smith *et al.*, *Astrophys. J.* **666**, 1116–1128 (2007).
2. E. O. Ofek *et al.*, *Astrophys. J.* **659**, L13–L16 (2007).
3. N. Smith, R. McCray, *Astrophys. J.* **671**, L17–L20 (2007).
4. S. E. Woosley, S. Blinnikov, A. Heger, *Nature* **450**, 390–392 (2007).
5. K. S. Kawabata *et al.*, *Astrophys. J.* **697**, 747–757 (2009).
6. Materials and methods are available as supplementary materials.
7. A. Jerkstrand *et al.*, *Mon. Not. R. Astron. Soc.* **475**, 277–305 (2018).
8. A. Jerkstrand, C. Fransson, C. Kozma, *Astron. Astrophys.* **530**, A45 (2011).
9. C. Fransson *et al.*, *Astrophys. J.* **797**, 118 (2014).
10. C. Inserra *et al.*, *Mon. Not. R. Astron. Soc.* **459**, 2721–2740 (2016).
11. O. D. Fox *et al.*, *Mon. Not. R. Astron. Soc.* **454**, 4366–4378 (2015).
12. T. Müller, J. L. Prieto, O. Pejcha, A. Clocchiatti, *Astrophys. J.* **841**, 127 (2017).
13. S. J. Prentice *et al.*, *Mon. Not. R. Astron. Soc.* **485**, 1559–1578 (2019).
14. P. A. Mazzali, K. Nomoto, F. Patat, K. Maeda, *Astrophys. J.* **559**, 1047–1053 (2001).
15. A. Heger, S. E. Woosley, *Astrophys. J.* **724**, 341–373 (2010).
16. A. Heger, S. E. Woosley, *Astrophys. J.* **567**, 532–543 (2002).
17. K. Nomoto, F.-K. Thielemann, K. Yokoi, *Astrophys. J.* **286**, 644 (1984).
18. K. Iwamoto *et al.*, *Astrophys. J. Suppl. Ser.* **125**, 439–462 (1999).
19. I. Agnoletto *et al.*, *Astrophys. J.* **691**, 1348–1359 (2009).
20. N. Smith, R. Chornock, J. M. Silverman, A. V. Filippenko, R. J. Foley, *Astrophys. J.* **709**, 856–883 (2010).
21. V. Morozova *et al.*, *Astrophys. J.* **814**, 63 (2015).
22. R. A. Chevalier, *Astrophys. J.* **752**, L2 (2012).
23. J. L. Terman, R. E. Taam, L. Hernquist, *Astrophys. J.* **422**, 729 (1994).
24. R. E. Taam, E. L. Sandquist, *Annu. Rev. Astron. Astrophys.* **38**, 113–141 (2000).

25. A. Kashi, N. Soker, . *Mon. Not. R. Astron. Soc.* **417**, 1466–1479 (2011).
26. M. Livio, A. G. Riess, *Astrophys. J.* **594**, L93–L94 (2003).
27. N. N. Chugai, L. R. Yungelson, *Astron. Lett.* **30**, 65–72 (2004).
28. N. Smith *et al*., *Astrophys. J.* **686**, 467–484 (2008).
29. A. J. Drake *et al*., *Astrophys. J.* **718**, L127–L131 (2010).
30. E. Chatzopoulos *et al*., *Astrophys. J.* **729**, 143 (2011).
31. A. Jerkstrand, "SN2006gy", Max Planck Society (2019); https://doi.org/10.17617/3.30.

ACKNOWLEDGMENTS

We thank N. Ivanova, R. Pakmor, R. Iaconi and J. Grumer for discussion. We thank M. Rampp and L. Stanisic at the Max Planck Computing and Data Facility for computing assistance. We also thank the referees and the editor for much useful feedback that improved the manuscript. **Funding:** A.J. acknowledges funding by the European Union's Framework Programme for Research and Innovation Horizon 2020 under Marie Sklodowska-Curie grant No. 702538 and European Research Council (ERC) Starting Grant No. 803189, as well as support from the Swedish Research Council and Swedish National Space Board. K.M. acknowledges support from JSPS Kakenhi grants 18H05223, 18H04585, and 17H02864. K. K. acknowledges support from JSPS grants 18H03720 and 17H06363. **Author contributions:** A. J. identified the Fe I lines in SN 2006gy and performed the spectral modelling. K. M. performed the radiation hydrodynamic modelling. K. K. and A. J. performed the data calibration. All authors contributed to the analysis and paper writing. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** The spectrum shown in Fig. 1 is available in the WiseRep database at https://wiserep.weizmann.ac.il/system/files/uploaded/general/2006gy_2007-09-18_Subaru_FOCAS_None_0.xy. The SNEC code is available at https://stellarcollapse.org/SNEC. The SUMO code was developed by A. Jerkstrand, C. Kozma, and C. Fransson. All input and output files for our SUMO and SNEC modeling, and SUMO executables, are available at the Max-Planck-Society data repository EDMOND in collection "SN2006gy" (*31*).

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/367/6476/415/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S11
Tables S1 and S2
References (*32*–*66*)

23 November 2018; accepted 19 December 2019
10.1126/science.aaw1469

NANOMATERIALS

Rational construction of a scalable heterostructured nanorod megalibrary

Benjamin C. Steimle[1], Julie L. Fenton[1], Raymond E. Schaak[1,2,3]*

Integrating multiple materials in arbitrary arrangements within nanoparticles is a prerequisite for advancing many applications. Strategies to synthesize heterostructured nanoparticles are emerging, but they are limited in complexity, scope, and scalability. We introduce two design guidelines, based on interfacial reactivity and crystal structure relations, that enable the rational synthesis of a heterostructured nanorod megalibrary. We define synthetically feasible pathways to 65,520 distinct multicomponent metal sulfide nanorods having as many as 6 materials, 8 segments, and 11 internal interfaces by applying up to seven sequential cation-exchange reactions to copper sulfide nanorod precursors. We experimentally observe 113 individual heterostructured nanorods and demonstrate the scalable production of three samples. Previously unimaginable complexity in heterostructured nanorods is now routinely achievable with simple benchtop chemistry and standard laboratory glassware.

The controlled placement of multiple materials within a nanoparticle (NP) is important for designing next-generation nanostructures across many fields. The materials and their spatial arrangements define the functions of a NP, whereas the interfaces that connect them control electronic and magnetic coupling. For example, interfacing various semiconductors with appropriately aligned band gaps and band-edge positions leads to directional transport of electrons and holes for photocatalytic water splitting (*1*), photon upconversion (*2*), and light-responsive light-emitting diodes and photodetectors (*3*, *4*). Materials having a large number of interfaces can have increased phonon scattering, which enhances thermoelectric performance (*5*), as well as improved separation of photogenerated charge carriers for improved quantum efficiencies (*6*, *7*), multiwavelength absorption for optimal solar spectrum utilization (*8*), and multicolor emission for ratiometric sensing (*9*). The design and synthesis of heterostructured NPs for these and other applications are limited by the lack of scalable methods that are sufficiently versatile to enable integration of multiple materials while maintaining key morphological features, including size, shape, and uniformity.

Synthetic approaches that produce NPs having the largest number of materials components and interfaces are made using top-down, surface growth, templating, or nanoreactor techniques, which typically require specialized equipment and yield only microgram-scale quantities of NPs (*10*–*14*). Scalable methods, which are often solution based, produce NPs that are limited in complexity and require synthetic pathways that must be developed and optimized for each system, and are not usually generalizable. Scalable and generalizable synthetic platforms are emerging, but only a relatively small number of distinct types of heterostructured NPs have been produced (*15*–*19*).

To develop a simple, scalable, and generalizable platform for producing heterostructured NPs of arbitrary complexity, we focused on cation-exchange reactions, which can postsynthetically modify diverse classes of readily available NPs without requiring specialized equipment (Fig. 1A). In these reactions, cations in metal sulfide, selenide, telluride, phosphide, halide, and oxide NPs are replaced with cations from solution, driven by solvation energies and Lewis acid–base interactions (*20*–*23*). Partial exchange reactions replace only a fraction of the cations to produce phase-segregated heterostructured NPs that contain multiple materials, and provide a convenient pathway for introducing interfaces (*24*–*29*).

We use roxbyite copper sulfide ($Cu_{1.8}S$) nanorods (figs. S1 and S2) as a model system, because Cu^+ cations exchange with several other cations to produce related metal sulfides (*18*–*20*, *30*). In these reactions, a flask containing distilled oleylamine, which serves as a stabilizing ligand, as well as benzyl ether and octadecene as solvents, was heated to 120°C, and $Cu_{1.8}S$ nanorods dispersed in tri-octyl phosphine, which functions as a Lewis base, were injected. Multiple successive exchange solutions, each of which is stoichiometrically limited relative to the number of remaining Cu^+ cations available for exchange, could then be injected to sequentially transform the $Cu_{1.8}S$ nanorods into heterostructured products that contain internal interfaces between multiple discrete material domains (*31*).

Successive injection of five exchange solutions (Zn^{2+}, In^{3+}, Ga^{3+}, Co^{2+}, Cd^{2+}) into the reaction flask containing the $Cu_{1.8}S$ nanorods produced ZnS–$CuInS_2$–$CuGaS_2$–CoS–(CdS–$Cu_{1.8}S$) heterostructured nanorods that contained six distinct materials and six interfaces. In each cation-exchange reaction, Cu^+ in $Cu_{1.8}S$ was selectively replaced owing to the more favorable soft acid, soft base interactions of Cu^+ with tri-octyl phosphine in solution compared to the multivalent cations (*18*–*20*). These reactions were performed in a single flask without having to isolate the nanorods between steps. However, aliquots were removed and analyzed after each step to characterize the intermediate nanorods and to better understand their formation. Element maps generated using scanning transmission electron microscopy with energy-dispersive spectroscopy (STEM–EDS) for a $Cu_{1.8}S$ nanorod and a nanorod analyzed after each injection (Fig. 1B) indicated that each sequential cation-exchange step formed a new segment within the nanorod. First-generation (G-1) $Cu_{1.8}S$ sequentially transformed into G-2 ZnS–$Cu_{1.8}S$, G-3 ZnS–$CuInS_2$–$Cu_{1.8}S$, G-4 ZnS–$CuInS_2$–$CuGaS_2$–$Cu_{1.8}S$, G-5 ZnS–$CuInS_2$–$CuGaS_2$–CoS–$Cu_{1.8}S$, and finally G-6 ZnS–$CuInS_2$–$CuGaS_2$–CoS–(CdS–$Cu_{1.8}S$). High-resolution TEM (HRTEM) imaging (Fig. 1C) showed that each segment in a representative G-6 ZnS–$CuInS_2$–$CuGaS_2$–CoS–(CdS–$Cu_{1.8}S$) nanorod was single crystalline and that the sulfur sublattice structure persisted throughout the entire particle. Additional characterization data, including wider-field STEM images, EDS element maps, corresponding EDS spectra, electron-diffraction patterns, and HRTEM images for each nanorod generation, are shown in figs. S2 to S10.

The data in Fig. 1 and figs. S2 to S7 revealed that each successive injection led to cation exchange adjacent to the previously exchanged region. Thus, the interfacial region between the exchanged region and the remaining $Cu_{1.8}S$ was more reactive than any other section of the $Cu_{1.8}S$. A HRTEM image highlighting the ZnS–$Cu_{1.8}S$ interface of a G-2 ZnS–$Cu_{1.8}S$ nanorod (Fig. 1D) revealed a thin region between $Cu_{1.8}S$ and the most recently exchanged segment (in this case ZnS) composed of Cu and S but structurally distinct from $Cu_{1.8}S$ far from the interface. This Cu–S interfacial region also exhibited a higher density of defects and contained areas that were polycrystalline or poorly crystalline. Regions of copper chalcogenide NPs with high defect densities can increase reactivity in cation-exchange reactions (*32*), possibly providing enhanced diffusion pathways that favor preferential exchange at the preexisting interfaces. Figure S10 shows HRTEM images for all of the intermediates, confirming that similar poorly crystalline interfacial regions occurred for the G-2 through G-5 particles shown in Fig. 1B.

Because the six-component nanorods can be isolated as ~10- to 40-mg powder samples, bulk characterization is also possible. Powder

[1]Department of Chemistry, The Pennsylvania State University, University Park, PA 16802, USA. [2]Department of Chemical Engineering, The Pennsylvania State University, University Park, PA 16802, USA. [3]Materials Research Institute, The Pennsylvania State University, University Park, PA 16802, USA.
*Corresponding author. Email: res20@psu.edu

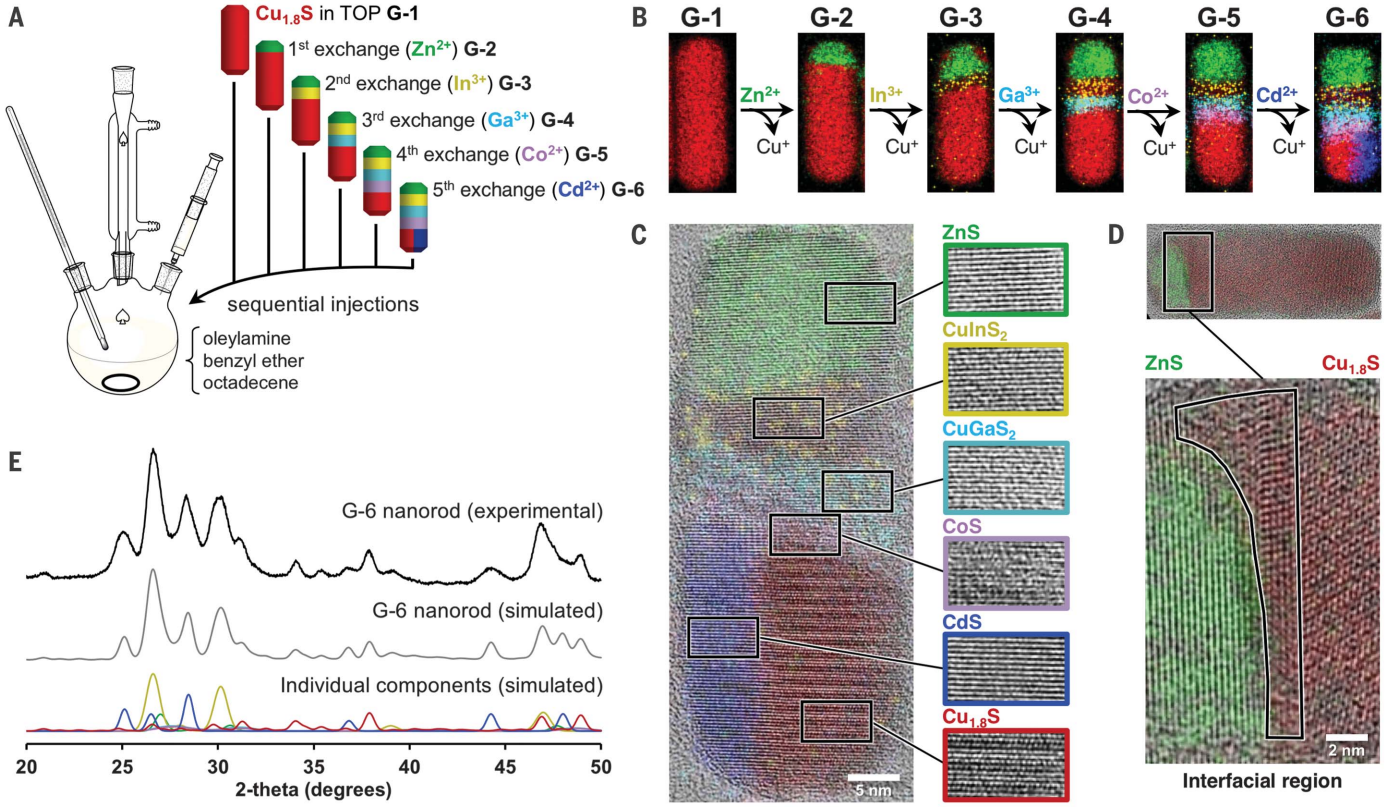


Fig. 1. Synthesis and characterization of G-6 nanorods. (**A**) Schematic showing the reaction setup and injection sequence that incrementally transform G-1 $Cu_{1.8}S$ into G-6 ZnS–$CuInS_2$–$CuGaS_2$–CoS–(CdS–$Cu_{1.8}S$). (**B**) STEM–EDS element maps for each nanorod generation. Cu Kα, Zn Kα, In Kα, Ga Kα, Co Kα, and Cd Lα lines are shown in red, green, yellow, teal, purple, and blue, respectively. (**C**) HRTEM image with overlaid EDS map highlighting the crystallinity of each material within the G-6 nanorod. (**D**) HRTEM image with overlaid EDS map for a G-2 ZnS–$Cu_{1.8}S$ nanorod and an enlarged part of the image showing the 1- to 2-nm structurally distinct interfacial region where the next exchange occurs. (**E**) Experimental powder XRD pattern for the G-6 heterostructured nanorod, along with individual and combined simulated patterns that account for preferred orientation effects and the microscopically observed crystalline domain sizes. Figures S1 to S13 include additional characterization data for each generation and information on how the simulated G-6 XRD pattern was generated.

x-ray diffraction (XRD) data for the G-6 ZnS–$CuInS_2$–$CuGaS_2$–CoS–(CdS–$Cu_{1.8}S$) nanorods (Fig. 1E) were compared with a reference pattern generated by combining simulated diffraction patterns from each of the constituent materials. For each material, the crystalline domain size used in the simulated pattern to define the peak widths was estimated on the basis of the average segment thickness determined by TEM analysis. Because each segment within the heterostructured nanorods is an asymmetric, single crystalline domain and the one-dimensional (1D) morphology favors alignment of the nanorods parallel to the XRD sample holder surface, small preferred orientation effects were also included in the simulated diffraction patterns (supplementary text and figs. S11 to S13). The experimental XRD pattern matches well with the simulated pattern, which helps to confirm that the bulk sample contains each of the six materials in a nanorod morphology with crystallite sizes that match those observed by TEM (*33*).

We developed design guidelines from these results. First, the sequence of materials in the heterostructured nanorod matches the sequence of exchange solution injections, so the order of the materials, as well as the heterointerfaces that form, could be controlled by changing the order of the injections. The extent of exchange can be controlled through the amount of exchange solution injected into the reaction flask. Second, different materials orient differently, such that their interfaces span a range of directions. Some interfaces, such as ZnS–$Cu_{1.8}S$ and $CuInS_2$–$Cu_{1.8}S$, are most commonly observed exactly perpendicular to the long direction of the nanorod, whereas others, such as CoS–$Cu_{1.8}S$ and $CuGaS_2$–$Cu_{1.8}S$, are observed to be either perpendicular to the long direction of the nanorod or ~40° relative to perpendicular. By contrast, CdS–$Cu_{1.8}S$ interfaces are most often observed to be parallel to the length of the nanorod.

To further characterize the interfaces, we subjected four samples of single-tip ZnS–$Cu_{1.8}S$ nanorods to partial cation exchange with In^{3+}, Ga^{3+}, Co^{2+}, and Cd^{2+} to form three-component ZnS–$CuInS_2$–$Cu_{1.8}S$, ZnS–$CuGaS_2$–$Cu_{1.8}S$, ZnS–CoS–$Cu_{1.8}S$, and ZnS–(CdS–$Cu_{1.8}S$) nanorods, respectively. We then obtained high-angle annular dark-field (HAADF)–STEM images and STEM-EDS element maps and compared the results to the crystal structures (Fig. 2, A to L). Many factors contribute to interface formation and stability in these systems (*25*–*29*). All of the exchange products have similar wurtzite-related crystal structures, which allowed for simple comparisons of readily available crystallographic information to qualitatively predict and rationalize which interfaces form (*28*). The three experimentally observed interfaces—perpendicular, ~40° relative to perpendicular, and parallel to the $Cu_{1.8}S$ nanorod length—had the best lattice matching and the least strain (Fig. 2M).

Each of these interface orientations has been observed in NP systems that contain similar materials (*27*, *28*). Wurtzite (WZ), the crystal structure adopted by all of the product phases formed through cation exchange, has a hexagonal close-packed (hcp) sulfur sublattice, whereas roxbyite $Cu_{1.8}S$ has a distorted hcp sulfur sublattice. A pseudohexagonal wurtzite-like subcell of roxbyite can be defined for

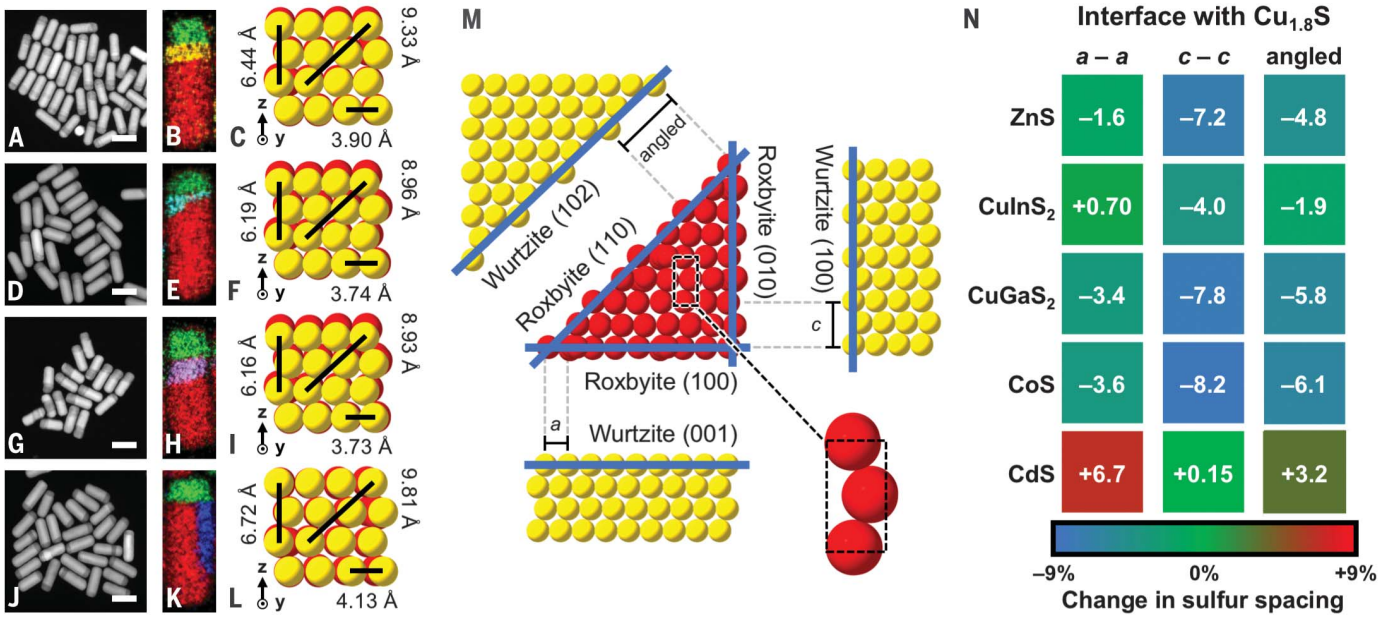


Fig. 2. Preferred interface directions in cation-exchange reactions with In^{3+}, Ga^{3+}, Co^{2+}, and Cd^{2+}. HAADF–STEM images (50-nm scale bars) and STEM–EDS element maps for G-3 (**A** and **B**) ZnS–$CuInS_2$–$Cu_{1.8}S$, (**D** and **E**) ZnS–$CuGaS_2$–$Cu_{1.8}S$, (**G** and **H**) ZnS–CoS–$Cu_{1.8}S$, and (**J** and **K**) ZnS–(CdS–$Cu_{1.8}S$) nanorods. Cu Kα, Zn Kα, In Lα, Ga Kα, Co Kα, and Cd Lα lines are shown in red, green, yellow, teal, purple, and blue, respectively. Crystallographic projections of the sulfur sublattices (yellow spheres) in wurtzite (**C**) $CuInS_2$, (**F**) $CuGaS_2$, (**I**) CoS, and (**L**) CdS overlaid (to scale) on a corresponding region of roxbyite $Cu_{1.8}S$ (red spheres) to visually depict which systems and crystallographic directions have the best lattice matching. Values for the intersulfur spacing corresponding to the *a* and *c* lattice parameters, as well as along the (102) plane, are shown for each WZ phase. (**M**) Idealized projections of the crystal planes that define the observed interfaces between the WZ phases and roxbyite $Cu_{1.8}S$, as well as the simplified pseudohexagonal subcell of roxbyite $Cu_{1.8}S$. (**N**) Heat map representing the percent mismatch in sulfur spacing for each WZ/roxbyite system in each possible interfacial direction [*a*–*a*, *c*–*c*, and angled, as shown in (M)]. The color scale represents sulfur-spacing changes in the WZ phases that would require high compression (blue), very little change (green), or high expansion (red) to match that of roxbyite $Cu_{1.8}S$.

direct comparison with the WZ product phases (Fig. 2M) (*34*). The observed $Cu_{1.8}S$–WZ interfaces were oriented in such a way that there is either direct *a*–*a* or *c*–*c* lattice parameter matching between WZ and the pseudohexagonal subcell of roxbyite, or lattice matching for a sulfur sublattice spacing in a crystallographic direction that is ~40° from the *x*–*y*, or (001), plane in WZ (Fig. 2M).

Figure 2N shows a heat map that corresponds to the percent lattice mismatch, which is a proxy for interfacial strain (*18*, *28*) between roxbyite $Cu_{1.8}S$ and each of the five WZ phases in the observed crystallographic directions. Values in green indicate a small lattice mismatch and a low strain. Values in blue indicate a high compressive strain, and values in red indicate a high expansive strain; both have substantial lattice mismatch. We observed experimentally that interfaces having lattice mismatch values near 0% were preferred. ZnS–$Cu_{1.8}S$ and $CuInS_2$–$Cu_{1.8}S$ interfaces, for example, were almost always observed to be perpendicular to the length of the rod, which created a low-strain interface that aligns the (001) plane of the wurtzite-type products [$WZ_{(001)}$] and the (100) plane in roxbyite $Cu_{1.8}S$ [$Cu_{1.8}S_{(100)}$]. By contrast, the CdS–$Cu_{1.8}S$ interface was parallel to the length of the rod, as alignment of the $CdS_{(100)}$ and $Cu_{1.8}S_{(010)}$ planes had the best lattice match. The CoS–$Cu_{1.8}S$ and $CuGaS_2$–$Cu_{1.8}S$ interfaces were observed, with almost equal frequencies, in two configurations, perpendicular to the rod length and ~40° from perpendicular, which aligned the $WZ_{(102)}$ and $Cu_{1.8}S_{(110)}$ lattice planes. We attributed the ambiguity in observed interface angle to the relatively small difference between the most preferred interfaces, $CoS_{(100)}$–$Cu_{1.8}S_{(001)}$ and $CuGaS_{2(100)}$–$Cu_{1.8}S_{(001)}$, and the second-closest interfaces, $CoS_{(102)}$–$Cu_{1.8}S_{(110)}$ and $CuGaS_{2(102)}$–$Cu_{1.8}S_{(110)}$.

Figure 2M shows crystallographic projections of the idealized interfaces between roxbyite and a WZ product that matched the *a*-parameter sulfur spacing [$WZ_{(001)}$–$Cu_{1.8}S_{(100)}$], the *c*-parameter sulfur spacing [$WZ_{(100)}$–$Cu_{1.8}S_{(010)}$], or a sulfur spacing ~40° from perpendicular [$WZ_{(102)}$–$Cu_{1.8}S_{(110)}$)]. Figure S14 shows measurements for the sulfur spacing for the $Cu_{1.8}S_{(110)}$ and $WZ_{(102)}$, and table S1 shows additional data corresponding to lattice matching across the various interfaces for the materials used in this study. Analysis of the possible interfaces that could form between any two of the six materials that can be incorporated into the heterostructured nanorods revealed those that were favorable and indicates how they were likely to be oriented.

Using the two simple design guidelines that were developed in Figs. 1 and 2 for six different materials, we embarked on the rational synthesis of a large megalibrary of heterostructured nanorods (Fig. 3). Starting with G-1 $Cu_{1.8}S$ nanorods, we synthesized three distinct types of G-2 ZnS/$Cu_{1.8}S$ nanorods in high morphological yield by modulating the reaction conditions (fig. S15). Single-tip ZnS–$Cu_{1.8}S$ nanorods were synthesized by injecting a stoichiometrically limited Zn^{2+} exchange solution into the reaction mixture at 120°C and allowing it to react for 15 min, whereas central-band $Cu_{1.8}S$–ZnS–$Cu_{1.8}S$ nanorods were synthesized by injecting the Zn^{2+} exchange solution at room temperature, increasing the reaction temperature to 100°C, and then continuing the reaction at that temperature for 30 min. Dual-tip ZnS–$Cu_{1.8}S$–ZnS nanorods are known to form at lower temperatures in an excess of Zn^{2+} cations (*18*, *31*).

The formation of multiple distinct G-2 ZnS/$Cu_{1.8}S$ systems can be rationalized by subtle differences in the reactivities of different regions of metal chalcogenide nanorods (*27*, *35*, *36*), which can be preferentially accessed through judicious choice of reaction conditions (fig. S15). By using a Zn^{2+} exchange solution injection temperature (90°C) that closely balances these reactivity contributions, we obtained a sample containing all three ZnS–$Cu_{1.8}S$, $Cu_{1.8}S$–ZnS–$Cu_{1.8}S$, and ZnS–$Cu_{1.8}S$–ZnS nanorods, as well as several other arrangements of ZnS domains that currently could not be synthesized

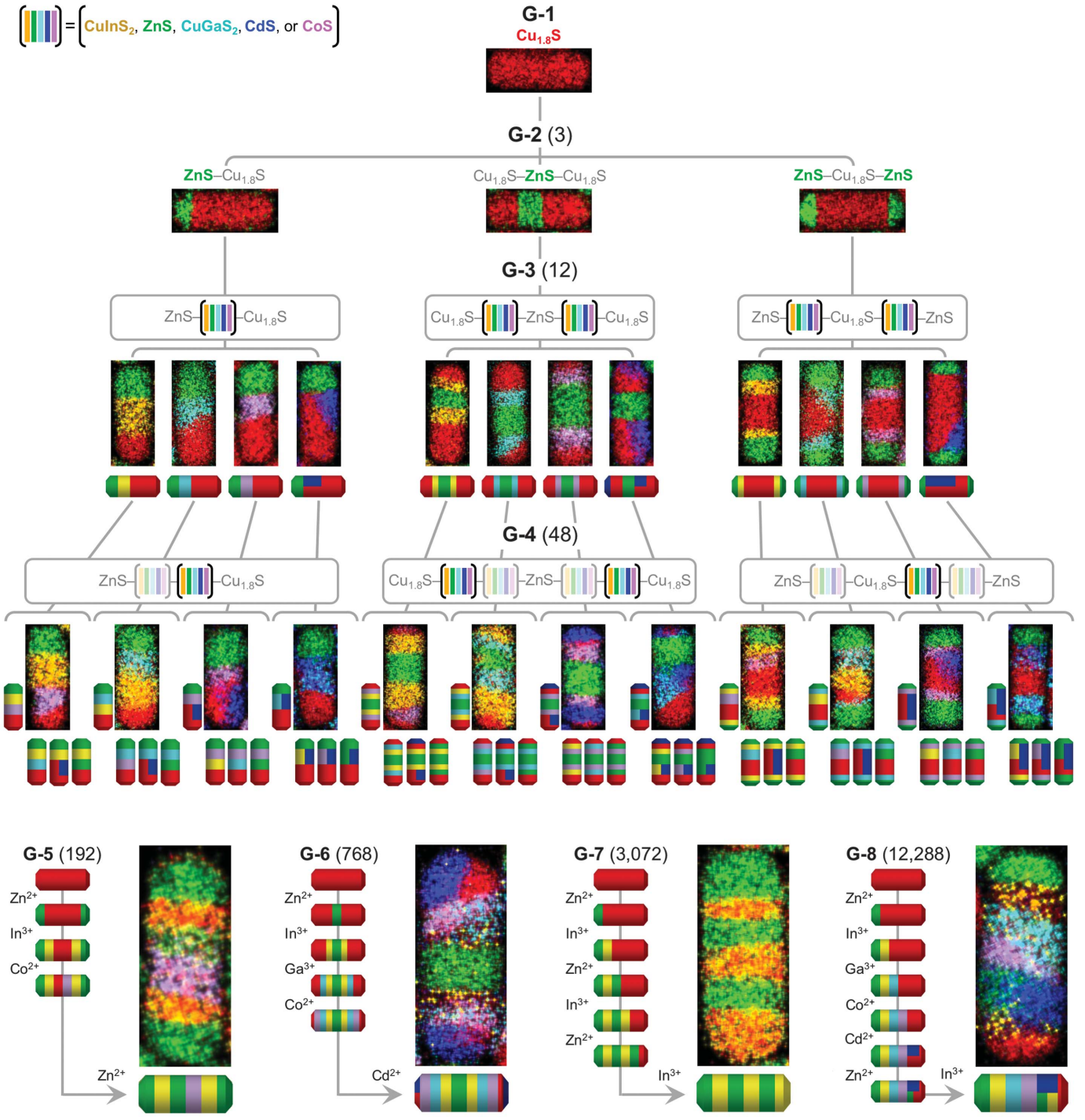


Fig. 3. Synthetic pathways to 16,380 distinct G-3 through G-8 nanorods. Reaction diagram showing the pathways by which G-1 $Cu_{1.8}S$ nanorods can be transformed to 3 distinct G-2 ZnS/$Cu_{1.8}S$ derivatives, which can subsequently be transformed, using pathways that are derived from the design guidelines in Figs. 1 and 2, into 12 G-3, 48 G-4, 192 G-5, 768 G-6, 3072 G-7, and 12,288 G-8 nanorods through various combinations of Zn^{2+}, In^{3+}, Ga^{3+}, Co^{2+}, and/or Cd^{2+} cation-exchange steps. The cation used in the most recent exchange is excluded as a possibility, as it would produce a product having an identical sequence of materials. STEM–EDS element maps, cropped from mixed-population samples, are shown for 28 of the 16,380 possible G-3 through G-8 members of this family. Drawings indicate the additional accessible G-4 nanorods, as well as pathways to selected G-5, G-6, G-7, and G-8 nanorods. STEM-EDS signals from the Cu Kα, Zn Kα, In Lα, Ga Kα, Co Kα, and Cd Lα lines are shown in red, green, yellow, teal, purple, and blue, respectively. For samples that contain both In and Cd, the weaker-intensity In Kα line, also shown in yellow, was mapped instead of the more intense In Lα because the In Lα and Cd Lβ lines overlap. All nanorods are ~55 nm by 20 nm. STEM–EDS data for 75 additional members of this reaction pathway and those that are generated from other G-2 heterostructured nanorods are shown in fig. S18.

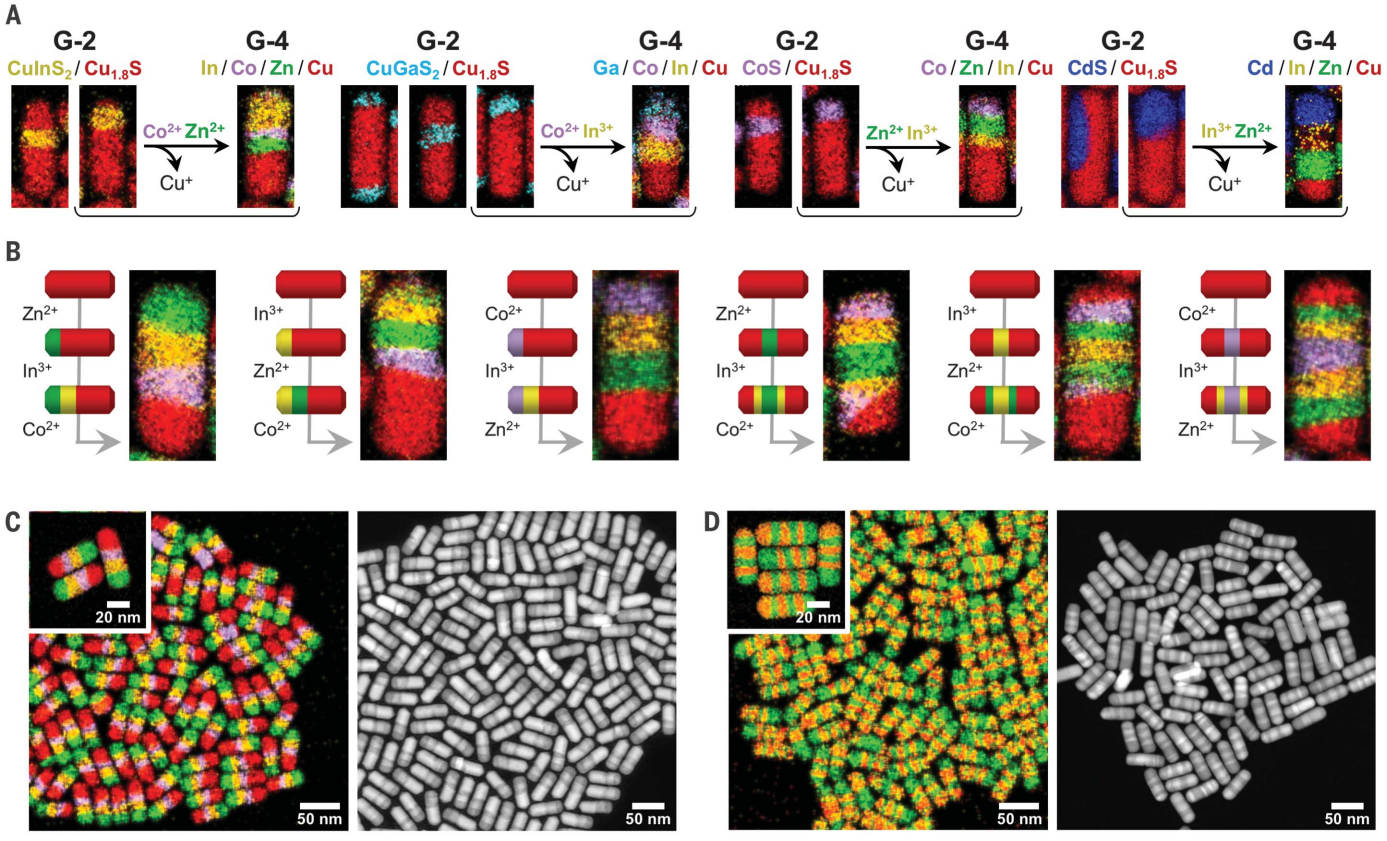


Fig. 4. Additional systems, configurational isomers, and scaled-up reactions. (**A**) STEM–EDS element maps of accessible G-2 nanorods in the $CuInS_2/Cu_{1.8}S$, $CuGaS_2/Cu_{1.8}S$, $CoS/Cu_{1.8}S$, and $CdS/Cu_{1.8}S$ systems, along with single-tip G-4 derivatives made using the indicated injection sequences. (**B**) STEM–EDS element maps of six distinct nanorod isomers containing various spatial arrangements of ZnS, $CuInS_2$, CoS, and $Cu_{1.8}S$. Each isomer is derived from a single-tip or central-band G-2 nanorod, as indicated by the corresponding reaction pathway. (**C** and **D**) HAADF–STEM images and STEM–EDS element maps for two selected reactions with high morphological yield, derived from single-tip G-2 $ZnS–Cu_{1.8}S$, that generate 10 to 40 mg of heterostructured nanorod products; additional characterization data are included in figs. S20 and S21. All nanorods are ~55 nm by 20 nm. STEM-EDS signals from the Cu Kα, Zn Kα, In Lα, Ga Kα, Co Kα, and Cd Lα lines are shown in red, green, yellow, teal, purple, and blue, respectively. For samples that contain both In and Cd, the weaker-intensity In Kα line, also shown in yellow, was mapped instead of the more intense In Lα because the In Lα and Cd Lβ lines overlap.

in high morphological yield (fig. S16). Applying multiple successive partial cation-exchange reactions to this mixed-subpopulation sample allows us to observe a broad range of cation-exchange behavior using a relatively small number of reactions (*31*).

Reacting the sample that contained a heterogeneous mixture of single-tip $ZnS–Cu_{1.8}S$, central-band $Cu_{1.8}S–ZnS–Cu_{1.8}S$, and dual-tip $ZnS–Cu_{1.8}S–ZnS$ nanorods with In^{3+} produced at least three distinct heterostructured nanorod isomers within the same sample. As predicted on the basis of the design guidelines, partial exchange with In^{3+} produced $CuInS_2$ at the existing $ZnS–Cu_{1.8}S$ interfaces, with the $CuInS_2–Cu_{1.8}S$ interface perpendicular to the nanorod length because $CuInS_2$ preferentially formed an *a*–*a* interface with $Cu_{1.8}S$. The three types of G-2 $ZnS/Cu_{1.8}S$ nanorods produced three types of G-3 derivatives: $ZnS–CuInS_2–Cu_{1.8}S$ from the single-tip $ZnS–Cu_{1.8}S$ nanorods, $Cu_{1.8}S–CuInS_2–ZnS–CuInS_2–Cu_{1.8}S$ from the central-band $Cu_{1.8}S–ZnS–Cu_{1.8}S$ nanorods, and $ZnS–CuInS_2–Cu_{1.8}S–CuInS_2–ZnS$ from the dual-tip $ZnS–Cu_{1.8}S–ZnS$ nanorods.

Analogous results were obtained when the G-2 $ZnS/Cu_{1.8}S$ nanorods were instead exchanged with Ga^{3+}, Co^{2+}, and Cd^{2+}, except that the directions of the interfaces were different than for $CuInS_2$, which is consistent with predictions based on the crystal structures, as outlined in Fig. 2. Together, the three G-2 nanorods produced 12 G-3 nanorods, which all contained three distinct materials and between two and four interfaces, depending on their configurations. The STEM-EDS maps in Fig. 3 were cropped from the mixed-population samples, consistent with the goal of observing microscopically the various nanorods that form. As a representative example, fig. S17 includes data for a larger region, showing the distribution of nanorod subpopulations within a sample of G-3 $ZnS/CuInS_2/Cu_{1.8}S$ nanorods.

Two features of the G-3 nanorods were especially noteworthy, as they demonstrated exquisite control over subtle interfacial arrangements and segmentation patterns. First, $CuInS_2$, $CuGaS_2$, and CoS formed at all available $ZnS–Cu_{1.8}S$ interfaces, whereas CdS exchanged at only one $ZnS–Cu_{1.8}S$ interface, even for the dual-tipped $ZnS–Cu_{1.8}S–ZnS$ nanorods that had two $ZnS–Cu_{1.8}S$ interfaces. This difference was rationalized by considering the volume change associated with the formation of CdS from $Cu_{1.8}S$. The pseudo-hexagonal subcell of roxbyite $Cu_{1.8}S$ (87.03 $Å^3$) must expand by ~14% upon formation of wurtzite CdS (99.27 $Å^3$), whereas wurtzite $CuInS_2$, $CuGaS_2$, CoS, and ZnS have a slightly smaller volume than the pseudo-hexagonal subcell of roxbyite and require only a slight compression of the lattice (table S2). Second, when there were two $Cu_{1.8}S$ regions within the nanorods, i.e., for central-band $Cu_{1.8}S–ZnS–Cu_{1.8}S$, a second CdS domain was typically observed at one exposed tip of the $Cu_{1.8}S$ rod. This finding suggests that the reactivity of the $ZnS–Cu_{1.8}S$ interfacial region toward Cd^{2+} exchange was somewhat comparable to that of the tip of the $Cu_{1.8}S$ rods.

Each of the 12 selected G-3 nanorods could be further exchanged with In^{3+}, Ga^{3+}, Co^{2+}, Cd^{2+}, or Zn^{2+} to produce at least 48 distinct G-4 derivatives (Fig. 3). (We exclude the cation that would produce the same material that is adjacent to $Cu_{1.8}S$, i.e., In^{3+} is not included as an exchange option for G-3 ZnS–$CuInS_2$–$Cu_{1.8}S$ because it would produce ZnS–$CuInS_2$–$CuInS_2$–$Cu_{1.8}S$, which is identical in sequence to ZnS–$CuInS_2$–$Cu_{1.8}S$.) Figure 3 shows data for 12 of the 48 G-4 possibilities; drawings are shown for the others. All G-4 nanorods contained at least four distinct materials and between three and five interfaces. We observed that G-4 heterostructures derived from dual-tip ZnS–$Cu_{1.8}S$ typically exhibited exchange at only one of the two $Cu_{1.8}S$ interfacial regions of the G-3 precursor. For example, G-3 ZnS–$CuInS_2$–$Cu_{1.8}S$–$CuInS_2$–ZnS underwent partial Co^{2+} exchange to form G-4 ZnS–$CuInS_2$–CoS–$Cu_{1.8}S$–$CuInS_2$–ZnS rather than ZnS–$CuInS_2$–CoS–$Cu_{1.8}S$–CoS–$CuInS_2$–ZnS. We attribute this asymmetry to increased strain as the $Cu_{1.8}S$ region becomes smaller and more confined, as related strain effects have been observed to result in the formation of periodic superlattices during partial cation exchange (*25*).

The 48 G-4 nanorods could each be exchanged with four of the five cations to produce 192 distinct G-5 derivatives. Similarly, the 192 G-5 nanorods produce 768 G-6 derivatives, the 768 G-6 nanorods produce 3072 G-7 derivatives, and the 3072 G-7 nanorods produce 12,288 G-8 derivatives. Although it was not practical to synthesize and characterize such a large number of samples, Fig. 3 maps out synthetically feasible pathways to a total of 16,380 G-3 through G-8 derivatives based on the G-2 ZnS/$Cu_{1.8}S$ system, along with experimental observation of selected members of the heterostructured nanorod megalibrary. Additional examples of heterostructured nanorods that were observed, including members of the reaction pathways outlined in Fig. 3 as well as similar pathways applied to other G-2 nanorods, are provided in fig. S18.

Figure 3 also provides data for one example each of a G-5, G-6, G-7, and G-8 nanorod, along with the synthetic pathway that was rationally designed and implemented on the basis of the design guidelines outlined earlier. The G-5 ZnS–$CuInS_2$–ZnS–CoS–$CuInS_2$–ZnS nanorod, which contained four distinct materials and five interfaces, was synthesized by applying the injection sequence In^{3+}, Co^{2+}, and Zn^{2+} to a dual-tip ZnS–$Cu_{1.8}S$–ZnS nanorod. The G-6 CdS–$Cu_{1.8}S$–CoS–$CuGaS_2$–$CuInS_2$–ZnS–$CuInS_2$–$CuGaS_2$–CoS–(CdS–$Cu_{1.8}S$) nanorod contains 6 distinct materials and 11 interfaces, with the heterostructured nanorod divided into 11 segments. This G-6 nanorod was synthesized starting with a central-band $Cu_{1.8}S$–ZnS–$Cu_{1.8}S$ nanorod and using the injection sequence In^{3+}, Ga^{3+}, Co^{2+}, and Cd^{2+}. The G-7 $(ZnS–CuInS_2)_3$ nanorod, which had alternating stripes with different materials on each tip, was synthesized by applying alternating In^{3+} and Zn^{2+} injections to a single-tip ZnS–$Cu_{1.8}S$ nanorod. Finally, the G-8 ZnS–$CuInS_2$–$CuGaS_2$–CoS–[CdS–($ZnS–CuInS_2$)]–$Cu_{1.8}S$ nanorod, which contained 6 distinct materials and 10 interfaces, was synthesized by applying the injection sequence In^{3+}, Ga^{3+}, Co^{2+}, Cd^{2+}, Zn^{2+}, and In^{3+} to a single-tip ZnS–$Cu_{1.8}S$ nanorod.

Figure 3 represents synthetically feasible pathways to 16,380 distinct heterostructured G-3 through G-8 nanorods, but many more are possible by starting with other G-2 systems. For example, in addition to the single-tip, central-band, and dual-tip ZnS-based systems, it is currently possible to make single-tip, dual-tip, and central-band $CuGaS_2$ and single-tip and central-band $CuInS_2$ and CoS, as well as single-tip and Janus-band CdS, as mixed-population samples (Fig. 4A and fig. S19). Figure 4A also shows a few examples of G-4 heterostructured nanorods formed from the single-tip members of the non-ZnS G-2 systems. Assuming that analogous reactions to those shown in Fig. 3 can convert each of these 9 G-2 nanorods to 4 G-3, 16 G-4, 64 G-5, 256 G-6, 1024 G-7, and 4096 G-8 derivatives, which corresponds to 5460 G-3 through G-8 nanorods that could be generated from each G-2 example, the total number of synthetically feasible heterostructured nanorods, on the basis of current capabilities, is 65,520 [16,380 + (9 × 5460)]. (A detailed calculation is included in the supplementary text.) The actual number of accessible heterostructured nanorods is likely to be even higher, because additional G-2 nanostructures have been identified, but we do not include systems that have not yet been observed as a large fraction of a sample. Additionally, other cations can exchange for Cu^+ in $Cu_{1.8}S$, including Ag^+, Au^+, Mn^{2+}, Ni^{2+}, Pd^{2+}, Pt^{2+}, Sn^{2+}, and Sn^{4+} (*18*–*20*, *37*, *38*), which could further expand the potential scope of accessible systems to include an even wider range of semiconductors, semimetals, magnets, and catalysts.

As a demonstration of the scope of synthetic capabilities enabled by the design guidelines outlined above, Fig. 4B shows rational pathways to six distinct G-4 heterostructured nanorod isomers of ZnS, $CuInS_2$, CoS, and $Cu_{1.8}S$, which differed only in the configurations of the four materials. ZnS–$CuInS_2$–CoS–$Cu_{1.8}S$ was synthesized by sequentially injecting In^{3+} and Co^{2+} to single-tip ZnS–$Cu_{1.8}S$ nanorods. By instead starting with single-tip $CuInS_2$–$Cu_{1.8}S$, the $CuInS_2$–ZnS–CoS–$Cu_{1.8}S$ isomer was synthesized by injecting Zn^{2+} and then Co^{2+}. Likewise, single-tip CoS–$Cu_{1.8}S$ produced the CoS–$CuInS_2$–ZnS–$Cu_{1.8}S$ isomer upon sequential injection of In^{3+} and Zn^{2+}. Distinct isomers were produced starting with central-band $Cu_{1.8}S$–ZnS–$Cu_{1.8}S$, $Cu_{1.8}S$–$CuInS_2$–$Cu_{1.8}S$, and $Cu_{1.8}S$–CoS–$Cu_{1.8}S$, which added segments to each WZ–$Cu_{1.8}S$ interface upon injecting the exchange solutions. Applying the various exchange sequences to the central-band G-2 nanorods produced the G-4 isomers $Cu_{1.8}S$–CoS–$CuInS_2$–ZnS–$CuInS_2$–CoS–$Cu_{1.8}S$, $Cu_{1.8}S$–CoS–ZnS–$CuInS_2$–ZnS–CoS–$Cu_{1.8}S$, and $Cu_{1.8}S$–ZnS–$CuInS_2$–CoS–$CuInS_2$–ZnS–$Cu_{1.8}S$. The formation of these six distinct isomers demonstrates the precise control over integration and orientation of functional materials using these rational synthetic pathways.

The STEM-EDS maps in Figs. 1 to 4 and fig. S18 show experimental evidence of 113 distinct heterostructured nanorods that were part of the synthetically feasible megalibrary that contains 65,520 possible G-3 through G-8 members. These observations were made by using mixed-population samples that were purposely synthesized to allow rapid screening for a large number of possible products with the minimum number of reactions, and so yield and purity were not primary considerations. However, as shown in Fig. 1, these reactions were scalable, and we could readily generate ~10- to 40-mg batches. We chose two examples, in addition to G-6 ZnS–$CuInS_2$–$CuGaS_2$–CoS–(CdS–$Cu_{1.8}S$) shown Fig. 1, to demonstrate that the syntheses were scalable and that a desired product could be isolated in high morphological yield. Figure 4C shows an additional STEM-EDS map and HAADF-STEM image for the G-4 ZnS–$CuInS_2$–CoS–$Cu_{1.8}S$ nanorods originally shown as a highlighted G-4 system in Fig. 3 and the first of the isomers shown in Fig. 4B.

Analysis of 146 particles indicated that 88% form the expected ZnS–$CuInS_2$–CoS–$Cu_{1.8}S$ configuration, with the remainder of the sample consisting primarily of related nanorod heterostructures derived from central-band G-2 $Cu_{1.8}S$–ZnS–$Cu_{1.8}S$ impurities. Most of the minority subpopulations still followed the established design guidelines. However, a small fraction of the sample (<2%) included products that diverged from these design guidelines, suggesting that additional factors have the potential to be exploited to produce a larger number of more complex nanorods.

Figure 4D shows analogous data for a scaled-up reaction that produced the G-7 $(ZnS–CuInS_2)_3$ nanorods originally shown in Fig. 3. Analysis of 155 particles indicates that 70% form the expected ZnS–$CuInS_2$– ZnS–$CuInS_2$–ZnS–$CuInS_2$ configuration. The remainder of the sample contained nanorods with fewer stripes that resulted from overexchange during any one of the six sequential reaction steps. Such overexchange could come from local concentration or heat gradients in the reaction vessel or reactivity effects based on variances in nanorod morphology, size, and defect densities. Additional characterization for the nanorods in Fig. 4, C and D, can be found in figs. S20 and S21. Considering that the heterostructured nanorods in Fig. 4, C and D, required between four and seven distinct reactions to form, the configurational uniformity within

the sample was high, on par with morphological yields of mainstream shape-controlled NP syntheses (*39*, *40*) and other postsynthetically modified NPs (*18*, *28*). The ability to carry out all injections sequentially without having to isolate and/or purify the nanorods after each step also ensures the highest possible yields, as sample loss from purification of the NP product is minimized.

Engineering previously unimaginable complexity into bulk-scale heterostructured nanorods is now a routine task that can be carried out in a simple benchtop setup with standard laboratory glassware. Simple design guidelines based on readily available crystallographic information provide rational, synthetically feasible pathways to 65,520 heterostructured nanorods having up to 6 distinct materials, 8 segments, and 11 internal interfaces. Tens of thousands of additional heterostructured nanorods may also be accessible by expanding these capabilities to a broader scope of exchangeable ions and materials and incorporating other classes of postsynthetic modification reactions, as well as identifying and implementing additional design guidelines.

REFERENCES AND NOTES

1. S. Chandrasekaran *et al*., *Chem. Soc. Rev.* **48**, 4178–4280 (2019).
2. C. Yan, A. Dadvand, F. Rosei, D. F. Perepichka, *J. Am. Chem. Soc.* **132**, 8868–8869 (2010).
3. N. Oh *et al*., *Science* **355**, 616–619 (2017).
4. Y. E. Panfil, M. Oded, U. Banin, *Angew. Chem. Int. Ed.* **57**, 4274–4295 (2018).
5. Y. Min *et al*., *ACS Nano* **9**, 6843–6853 (2015).
6. R. Xie, U. Kolb, J. Li, T. Basché, A. Mews, *J. Am. Chem. Soc.* **127**, 7480–7488 (2005).
7. I. Hadar *et al*., *Nano Lett.* **17**, 2524–2531 (2017).
8. T.-T. Zhuang *et al*., *Angew. Chem. Int. Ed.* **55**, 6396–6400 (2016).
9. A. Teitelboim, N. Meir, M. Kazes, D. Oron, *Acc. Chem. Res.* **49**, 902–910 (2016).
10. A. G. Mark, J. G. Gibbs, T.-C. Lee, P. Fischer, *Nat. Mater.* **12**, 802–807 (2013).
11. Y. Wang *et al*., *J. Am. Chem. Soc.* **131**, 9926–9927 (2009).
12. S. R. Nicewarner-Peña *et al*., *Science* **294**, 137–141 (2001).
13. P.-C. Chen *et al*., *Science* **352**, 1565–1569 (2016).
14. P.-C. Chen *et al*., *Science* **363**, 959–964 (2019).
15. R. Costi, A. E. Saunders, U. Banin, *Angew. Chem. Int. Ed.* **49**, 4878–4897 (2010).
16. R. G. Weiner, M. R. Kunz, S. E. Skrabalak, *Acc. Chem. Res.* **48**, 2688–2695 (2015).
17. J. M. Hodges, R. E. Schaak, *Acc. Chem. Res.* **50**, 1433–1440 (2017).
18. J. L. Fenton, B. C. Steimle, R. E. Schaak, *Science* **360**, 513–517 (2018).
19. Y. Liu *et al*., *ACS Nano* **12**, 7803–7811 (2018).
20. L. De Trizio, L. Manna, *Chem. Rev.* **116**, 10852–10887 (2016).
21. G. E. Eperon, D. S. Ginger, *ACS Energy Lett.* **2**, 1190–1196 (2017).
22. B. J. Beberwyck, Y. Surendranath, A. P. Alivisatos, *J. Phys. Chem. C* **117**, 19759–19770 (2013).
23. J. Gui *et al*., *Angew. Chem. Int. Ed.* **54**, 3683–3687 (2015).
24. Z. Zhai, J. C. Flanagan, M. Shim, *Chem. Mater.* **29**, 6161–6167 (2017).
25. R. D. Robinson *et al*., *Science* **317**, 355–358 (2007).
26. D.-H. Ha *et al*., *Nano Lett.* **14**, 7090–7099 (2014).
27. B. Sadtler *et al*., *J. Am. Chem. Soc.* **131**, 5285–5293 (2009).
28. J. L. Fenton, B. C. Steimle, R. E. Schaak, *J. Am. Chem. Soc.* **140**, 6771–6775 (2018).
29. G. Gariano *et al*., *J. Am. Chem. Soc.* **139**, 9583–9590 (2017).
30. C. Coughlan *et al*., *Chem. Rev.* **117**, 5865–6109 (2017).
31. Materials and methods are available as supplementary materials
32. V. Lesnyak, R. Brescia, G. C. Messina, L. Manna, *J. Am. Chem. Soc.* **137**, 9315–9323 (2015).
33. C. F. Holder, R. E. Schaak, *ACS Nano* **13**, 7359–7365 (2019).
34. A. E. Powell, J. M. Hodges, R. E. Schaak, *J. Am. Chem. Soc.* **138**, 471–474 (2016).
35. Y. Zhai, M. Shim, *ChemPhysChem* **17**, 741–751 (2016).
36. T. Mokari, E. Rothenberg, I. Popov, R. Costi, U. Banin, *Science* **304**, 1787–1790 (2004).
37. J. Park *et al*., *ACS Nano* **12**, 7996–8005 (2018).
38. J. M. Luther, H. Zheng, B. Sadtler, A. P. Alivisatos, *J. Am. Chem. Soc.* **131**, 16851–16857 (2009).
39. A. R. Tao, S. Habas, P. Yang, *Small* **4**, 310–325 (2008).
40. C. B. Murray, C. R. Kagan, M. G. Bawendi, *Annu. Rev. Mater. Sci.* **30**, 545–610 (2000).

ACKNOWLEDGMENTS

TEM imaging was performed in the Penn State Microscopy and Cytometry facility. STEM imaging, EDS mapping, and HRTEM imaging were performed at the Materials Characterization Laboratory of the Penn State Materials Research Institute. The authors thank A. M. Fagan and R. W. Lord for helpful discussions. **Funding:** This work was supported by the U.S. National Science Foundation under grant DMR-1904122. **Author contributions:** B.C.S., J.L.F., and R.E.S. conceived the concept. B.C.S. and R.E.S. designed the experiments and wrote the paper. B.C.S. synthesized and characterized the nanoparticle samples. **Competing interests:** The authors declare no competing financial interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are presented in the manuscript or in the supplementary materials.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/367/6476/418/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S21
Tables S1 and S2
References (*41*–*44*)

13 August 2019; accepted 6 December 2019
10.1126/science.aaz1172

QUANTUM SENSING

Entanglement-based single-shot detection of a single magnon with a superconducting qubit

Dany Lachance-Quirion[1], Samuel Piotr Wolski[1], Yutaka Tabuchi[1], Shingo Kono[1], Koji Usami[1], Yasunobu Nakamura[1,2]*

The recent development of hybrid systems based on superconducting circuits provides the possibility of engineering quantum sensors that exploit different degrees of freedom. Quantum magnonics, which aims to control and read out quanta of collective spin excitations in magnetically ordered systems, provides opportunities for advances in both the study of magnetism and the development of quantum technologies. Using a superconducting qubit as a quantum sensor, we report the detection of a single magnon in a millimeter-sized ferrimagnetic crystal with a quantum efficiency of up to 0.71. The detection is based on the entanglement between a magnetostatic mode and the qubit, followed by a single-shot measurement of the qubit state. This proof-of-principle experiment establishes the single-photon detector counterpart for magnonics.

Quantum sensing aims to exploit the fragility of quantum states to external perturbations for the development of sensors. Quantum-enhanced sensing has now become one of the leading applications of quantum technologies (*1*, *2*). Entanglement can be harnessed in quantum sensing to indirectly probe a system of interest through a well-controlled auxiliary mode acting as the sensor (*3*–*5*). Such a task requires careful engineering to integrate existing quantum technologies into sensors able to detect various physical quantities.

The recent development of hybrid quantum systems provides a natural platform to engineer such quantum sensors (*6*). The combination of systems that harness complementary features for quantum technologies opens up the possibility of sensing one degree of freedom through another well-controlled system. One of the main challenges for this application lies in achieving high-fidelity control and readout of the quantum sensor in a hybrid device. Hybrid systems based on superconducting circuits (*7*, *8*) offer a versatile platform with which to overcome this challenge. Recent demonstrations include the measurement of the coherence of a bulk acoustic wave resonator (*9*) and the creation and characterization of quantum states of phonons (*10*, *11*).

Quantum magnonics provides another promising architecture for the development of quantum sensors based on hybrid systems (*12*–*15*). In quantum magnonics, magnetostatic modes in magnetically ordered solid-state systems are coherently coupled to superconducting qubits. Here, we demonstrate a quantum sensor able to faithfully detect single magnons, the quanta of excitations in magnetostatic modes, through entanglement with a superconducting qubit. Such single-magnon detection is made possible by combining high-fidelity time-domain control and single-shot readout of a qubit in a hybrid system with the strong dispersive regime of quantum magnonics. Our demonstration brings the equivalent of the single-photon detector to the emerging field of magnon spintronics (*16*) and establishes a new quantum technology for magnetism.

To realize the single-magnon detector, we used a hybrid system composed of a spherical ferrimagnetic crystal of yttrium iron garnet (YIG), a transmon-type superconducting qubit, and a three-dimensional microwave cavity (Fig. 1A) (*12*–*15*). Our system hosts three modes of interest: the uniform magnetostatic mode, or Kittel mode, in the ferrimagnetic crystal of tunable frequency $\omega_m/2\pi$; the qubit of frequency $\omega_q/2\pi \approx 7.92$ GHz; and a microwave cavity mode of frequency $\omega_c/2\pi \approx 8.45$ GHz. The Kittel mode and the superconducting qubit are respectively coupled to the cavity mode through magnetic-dipole (*17*, *18*) and electric-dipole couplings (*7*, *19*, *20*). These interactions lead to an effective beam-splitter interaction between the Kittel mode and the qubit (*12*–*15*). This coherent interaction enters the strong coupling regime with a coupling strength $g_{q-m}/2\pi = 7.13$ MHz, which is much larger than the decay rates of each system (Fig. 1B).

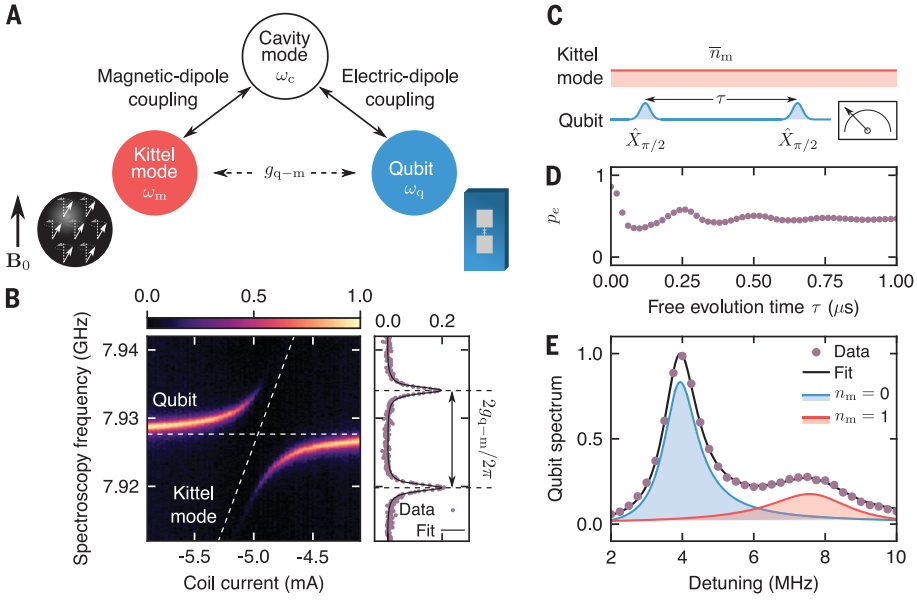


Fig. 1. Strong dispersive regime of quantum magnonics. (**A**) Interaction of strength g_{q-m} between the Kittel mode (with frequency ω_m) of a spherical ferrimagnetic crystal of YIG and a superconducting qubit (ω_q), engineered through magnetic- and electric-dipole couplings to a microwave cavity mode (ω_c). The ferrimagnetic sphere is magnetized with an external magnetic field $\mathbf{B}_0$. (**B**) Normalized qubit spectrum measured as a function of the coil current. Dashed lines are guides for the eye. (Right) Qubit spectrum measured at $\omega_q \approx \omega_m$. The line shows a fit to the data. (**C**) Ramsey interferometry protocol to probe the qubit in the presence of a continuous excitation of $\bar{n}_m$ magnons in the Kittel mode. (**D**) Probability p_e of the qubit being in the excited state $|e\rangle$ as a function of the free evolution time τ in the presence of $\bar{n}_m = 0.53$ magnons in the Kittel mode. (**E**) Normalized qubit spectrum, obtained from the Fourier transform of p_e, indicating a strong dispersive interaction between the Kittel mode and the qubit. The black line shows a fit to the data. The blue line and shaded area show the spectral component corresponding to the magnon vacuum state $|0\rangle$, and the red line and shaded area show the spectral component corresponding to the Fock state $|1\rangle$.

[1]Research Center for Advanced Science and Technology (RCAST), The University of Tokyo, Meguro, Tokyo 153-8904, Japan. [2]Center for Emergent Matter Science (CEMS), RIKEN, Wako, Saitama 351-0198, Japan.
*Corresponding author. Email: yasunobu@ap.t.u-tokyo.ac.jp

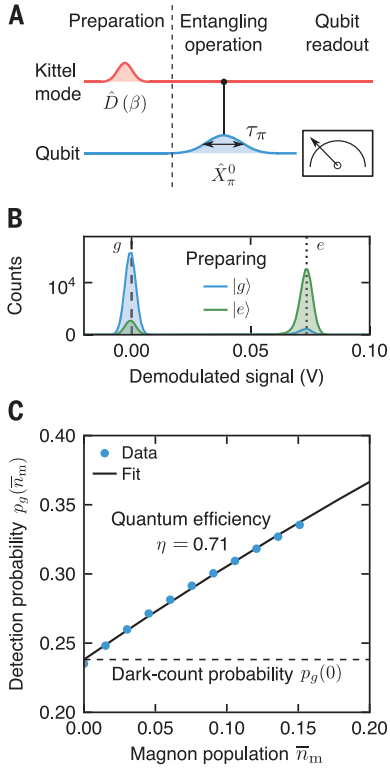


Fig. 2. Single-shot detection of single magnons. (**A**) Protocol to detect the presence of at least a single magnon. The Kittel mode and the qubit are entangled through the qubit excitation $\hat{X}^0_\pi$ conditional on the Kittel mode being in the vacuum state. To characterize the detection protocol, a coherent state of magnons was initially prepared through a displacement operation $\hat{D}(\beta)$. The state of the qubit was read out at the end of the protocol. (**B**) Histograms of the demodulated qubit readout signal for 10^5 single shots when preparing the qubit in the ground state $|g\rangle$ and excited state $|e\rangle$, obtained with the high-power readout technique. The vertical dashed line and vertical dotted line indicate the demodulated signal corresponding to the qubit occupying the ground state $|g\rangle$ and excited state $|e\rangle$, respectively. (**C**) Detection probability $p_g(\bar{n}_m)$ as a function of the magnon population $\bar{n}_m$. The solid black line shows a fit to Eq. 3, indicating a magnon detection efficiency η = 0.71 and a dark-count probability $p_g(0)$ = 0.24 (dashed black line) for a detection time τ_π = 200 ns. Error bars are smaller than the symbols.

Because of the strong coherent coupling, a strong dispersive interaction between the Kittel mode and the qubit can be engineered (*14*, *15*). This dispersive interaction, of strength χ_{q-m}, is described with the Hamiltonian

$$\hat{\mathcal{H}}^{\text{disp}}_{q-m}/\hbar = \frac{1}{2}(2\chi_{q-m}\hat{c}^\dagger\hat{c})\hat{\sigma}_z \qquad (1)$$

where $\hat{c}$ annihilates and $\hat{c}^\dagger$ creates a magnon in the Kittel mode, and $\hat{\sigma}_z = |e\rangle\langle e| - |g\rangle\langle g|$, with $|g\rangle$ the ground state of the qubit and $|e\rangle$ the excited state. The qubit–magnon dispersive interaction leads to a shift of the qubit frequency by $2\chi_{q-m}$ for each magnon in the Kittel mode. To characterize the dispersive interaction, we performed Ramsey interferometry on the qubit while continuously driving the Kittel mode on resonance at $\omega_m/2\pi \approx$ 7.79 GHz, far detuned from the qubit (Fig. 1, C and D). The qubit spectrum (Fig. 1E), obtained from the Fourier transform of the Ramsey oscillations indicates that the qubit frequency is shifted by $2\chi_{q-m}/2\pi$ = –3.82 MHz in the presence of a single magnon, a quantity larger than the linewidths $\gamma_m/2\pi$ = 1.61 MHz of the Kittel mode and $\gamma_q/2\pi$ = 0.33 MHz of the qubit, therefore reaching the strong dispersive regime (*14*, *15*, *21*–*23*).

The single-magnon detection protocol is enabled by the possibility of entangling the Kittel mode and the qubit (*3*). Through the strong dispersive interaction, the qubit can be excited conditionally on the Kittel mode being in the vacuum state $|0\rangle$ (*3*, *24*, *25*). The effect of the conditional excitation $\hat{X}^0_\pi$, with the qubit initially in the ground state $|g\rangle$ and the Kittel mode in an arbitrary magnon state $|\psi\rangle = \sum c_{n_m}|n_m\rangle$, is given by

$$\hat{X}^0_\pi|g\psi\rangle = c_0|e0\rangle + \sum\nolimits_{n_m>0} c_{n_m}|gn_m\rangle \qquad (2)$$

where $|in_m\rangle = |i\rangle \otimes |n_m\rangle$ is the state of the composite system, with $|i = g, e\rangle$ and $|n_m\rangle$ being the qubit states and the magnon Fock states, respectively. More specifically, $\hat{X}^0_\pi$ corresponds to a π-pulse applied to the qubit with a control frequency that corresponds to the qubit frequency with the Kittel mode in the vacuum state $|0\rangle$, ω^0_q. From Eq. 2, measuring the qubit in the ground state indicates the presence of at least a single magnon in the Kittel mode. The detection protocol (Fig. 2A) is composed of the entangling operation $\hat{X}^0_\pi$ and a readout of the qubit state. The fidelity of the entangling gate is mainly determined by the duration τ_π of the excitation, hereafter called the detection time (*3*, *25*). The excitation is conditional only if τ_π is such that the spectral width $\propto 1/\tau_\pi$ is smaller than the amplitude of the shift per excitation $2|\chi_{q-m}|$. The state of the qubit was read out by using the high-power readout technique (*26*), enabling single-shot readout with a fidelity $\mathcal{F}_r \approx 0.9$ without the use of near-quantum-limited amplifiers (Fig. 2B).

To benchmark the detection protocol, a coherent state of magnons $|\beta\rangle$ was initially prepared through a displacement operation $\hat{D}(\beta) = e^{\beta\hat{c}^\dagger - \beta^*\hat{c}}$ (Fig. 2A). The magnon population $\bar{n}_m = |\beta^2|$ at the detection pulse was controlled through the amplitude of the microwave drive resonant with the Kittel mode used for the displacement operation (*27*). The detection probability $p_g(\bar{n}_m)$ was then related to the magnon population $\bar{n}_m$ through the probability $p_{n_m \geq 1} = 1 - e^{-\bar{n}_m}$ of having at least a single magnon in the Kittel mode. More specifically, the detection probability is given by

$$p_g(\bar{n}_m) = \eta(1 - e^{-\bar{n}_m}) + p_g(0) \qquad (3)$$

where η and $p_g(0)$ are respectively the quantum efficiency and the dark-count probability, both critical figures in respect of evaluating the performance of the detector. The detection probability $p_g(\bar{n}_m)$ obtained experimentally for a detection time τ_π = 200 ns is shown in Fig. 2C. Fitting the data to Eq. 3, a quantum efficiency η = 0.71 and a dark-count probability $p_g(0)$ = 0.24 were determined. Considering these values, if the Kittel mode is in the vacuum state $|0\rangle$, the probability that the detector does not click is 1 – $p_g(0)$ = 0.76 (ideally, 1). When the Kittel mode is in the Fock state $|1\rangle$, the detector clicks with a probability η + $p_g(0)$ = 0.95 (ideally, 1). These results demonstrate the single-shot detection of a single magnon, thus bringing the equivalent of the single-photon detector to the field of magnonics.

Signatures of the mechanisms limiting the performance of the single-magnon detector were obtained by measuring the dark-count probability and the quantum efficiency for different detection times τ_π (Fig. 3). As shown in Fig. 3A, the dark-count probability $p_g(0)$ increases with the detection time τ_π because of the finite qubit relaxation time T_1 = 0.80 μs and coherence time $T_2^* = 0.97$ μs. Furthermore, initialization and readout errors set a lower bound on the dark-count probability at

Table 1. Error budget. Contributions from different sources of error determined from numerical simulations for a detection time τ_π = 200 ns. The total error is not equal to the sum of the listed errors owing to additional errors and multiple error processes (*27*). The dash indicates that the entanglement error does not contribute to the dark-count probability.

Source of error	Error	
	Dark-count probability	Inefficiency
	$p_g(0)$	$1 - \eta$
Qubit initialization	0.032	0.023
Qubit decoherence	0.15	0.21
Qubit readout	0.024	0.061
Entanglement	—	0.039
Total	0.22	0.33
Experiment	0.24	0.29

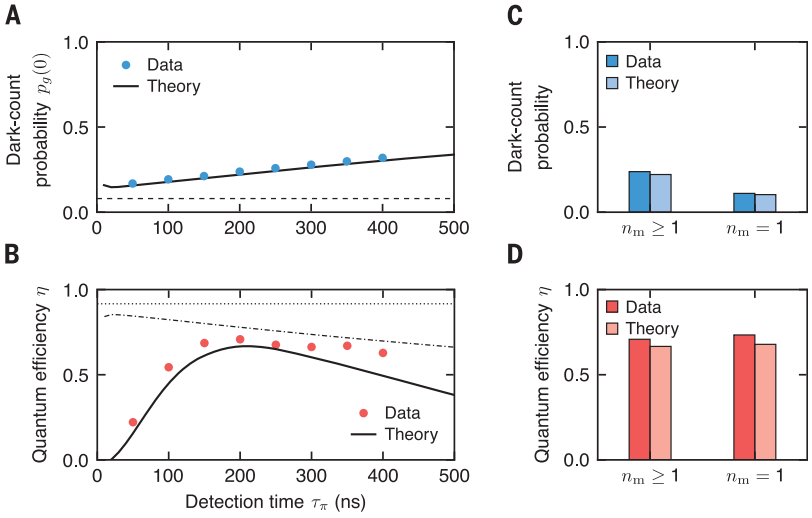


Fig. 3. Characterization of the single-magnon detector. (**A** and **B**) Dark-count probability $p_g(0)$ (A) and quantum efficiency η (B) as a function of the detection time τ_π. Results from numerical simulations are shown as solid lines. In (A), the dashed line shows the dark-count probability due to initialization and readout errors. In (B), the dotted and dot-dashed lines indicate the limits on the quantum efficiency set by readout errors and the dark-count probability, respectively. Error bars are smaller than the symbols. (**C** and **D**) Dark-count probability (C) and quantum efficiency (D) for the detection of at least a single magnon ($n_m \geq 1$) and of exactly a single magnon ($n_m = 1$) for a detection time τ_π = 200 ns.

≈0.08. The quantum efficiency η increases for larger detection times owing to an increase in the selectivity of the entangling operation between the qubit and the Kittel mode (Fig. 3B). For longer detection times, decoherence of the qubit limits the efficiency, leading to an optimal detection time at $\tau_\pi \approx 200$ ns. Two relevant upper bounds on the quantum efficiency are satisfied (Fig. 3B). First, as the magnons were detected by using the qubit as the quantum sensor, the quantum efficiency was bounded by the qubit readout fidelity $\mathcal{F}_r \approx 0.9$. Second, the dark-count probability $p_g(0)$ sets an upper limit on the quantum efficiency at $\eta \leq 1 - p_g(0)$ through the probability $\eta + p_g(0)$ of detecting the single-magnon Fock state $|1\rangle$.

As shown in Fig. 3, numerical simulations of the detection protocol are in good agreement with the experimental results, without any fitting parameters (*27*). Therefore, we used the numerical model to determine the effect of qubit initialization, control, readout, and entangling errors on the dark-count probability $p_g(0)$ and detection inefficiency 1 – η (Table 1). Qubit decoherence constitutes the primary source of error limiting the performance of the detector. A dark-count probability below 0.03 and a quantum efficiency above 0.96 should be within experimental reach with an improved single-magnon detector (*27*).

The performances of the detector can also be improved, without any hardware modifications, by considering an alternative detection scheme. Instead of detecting the presence of at least one magnon (n_m = 1, 2, ...) with the protocol of Fig. 1A, the presence of exactly one magnon (n_m = 1) can be detected by using the conditional operation $\hat{X}_\pi^1$ that excites the qubit only if there is exactly a single magnon in the Kittel mode (*3*, *28*). In the limit in which the probability of having more than one magnon is negligible, both protocols detect the presence of a single magnon. Experimentally, the conditional excitation $\hat{X}_\pi^1$ was realized by attempting to excite the qubit at its frequency with a single magnon in the Kittel mode, ω_q^1. The detection of exactly a single magnon enables us to reduce the dark-count probability by half to 0.12 (Fig. 3C). In this scheme, qubit decoherence does not contribute appreciably to the dark-count probability because the qubit is never actually excited in the absence of magnons. Nevertheless, the quantum efficiency is very similar for both schemes (Fig. 3D) (*27*). A good agreement between the experimental and numerical results is found without any fitting parameters, which highlights a good understanding of the physics at play.

The high-fidelity detection of a single magnon, corresponding to a precession of the magnetization vector of the millimeter-sized ferrimagnetic crystal with an angle of ~10^{-17}°, represents an advance for magnonics and quantum technologies based on magnetism. The magnon detection can be made quantum nondemolition (QND), with a QND readout of the qubit state (*3*, *5*). The relaxation and coherence times of superconducting qubits in quantum magnonics, currently limiting the performance of the single-magnon detector, could be enhanced by reducing internal losses of the microwave cavity modes.

Near-term applications include the heralded probabilistic creation of quantum states of magnons, a critical step toward the development of a magnon-based quantum transducer (*15*, *29*). Furthermore, the single-magnon detector could help to uncover weak magnon excitation processes such as the potential excitation of magnons from galactic axions (*30*). In the longer term, the development of planar devices (*31*–*33*) for the integration of single-magnon detectors could represent the ultimate limit to the conversion between magnons and electrical signals for emerging technologies such as magnon spintronics (*16*).

REFERENCES AND NOTES

1. C. L. Degen, F. Reinhard, P. Cappellaro, *Rev. Mod. Phys.* **89**, 035002 (2017).
2. J. M. Boss, K. S. Cujia, J. Zopes, C. L. Degen, *Science* **356**, 837–840 (2017).
3. B. R. Johnson *et al.*, *Nat. Phys.* **6**, 663–667 (2010).
4. N. Zhao *et al.*, *Nat. Nanotechnol.* **7**, 657–662 (2012).
5. S. Kono, K. Koshino, Y. Tabuchi, A. Noguchi, Y. Nakamura, *Nat. Phys.* **14**, 546–549 (2018).
6. G. Kurizki *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **112**, 3866–3873 (2015).
7. A. Blais, R.-S. Huang, A. Wallraff, S. M. Girvin, R. J. Schoelkopf, *Phys. Rev. A* **69**, 062320 (2004).
8. M. H. Devoret, R. J. Schoelkopf, *Science* **339**, 1169–1174 (2013).
9. Y. Chu *et al.*, *Science* **358**, 199–202 (2017).
10. K. J. Satzinger *et al.*, *Nature* **563**, 661–665 (2018).
11. Y. Chu *et al.*, *Nature* **563**, 666–670 (2018).
12. Y. Tabuchi *et al.*, *Science* **349**, 405–408 (2015).
13. Y. Tabuchi *et al.*, *C. R. Phys.* **17**, 729–739 (2016).
14. D. Lachance-Quirion *et al.*, *Sci. Adv.* **3**, e1603150 (2017).
15. D. Lachance-Quirion, Y. Tabuchi, A. Gloppe, K. Usami, Y. Nakamura, *Appl. Phys. Express* **12**, 070101 (2019).
16. A. V. Chumak, V. I. Vasyuchka, A. A. Serga, B. Hillebrands, *Nat. Phys.* **11**, 453–461 (2015).
17. H. Huebl *et al.*, *Phys. Rev. Lett.* **111**, 127003 (2013).
18. Y. Tabuchi *et al.*, *Phys. Rev. Lett.* **113**, 083603 (2014).
19. J. Koch *et al.*, *Phys. Rev. A* **76**, 042319 (2007).
20. H. Paik *et al.*, *Phys. Rev. Lett.* **107**, 240501 (2011).
21. J. Gambetta *et al.*, *Phys. Rev. A* **74**, 042318 (2006).
22. D. I. Schuster *et al.*, *Nature* **445**, 515–518 (2007).
23. L. R. Sletten, B. A. Moores, J. J. Viennot, K. W. Lehnert, *Phys. Rev. X* **9**, 021056 (2019).
24. E. Solano, *Phys. Rev. A* **71**, 013813 (2005).
25. G. Kirchmair *et al.*, *Nature* **495**, 205–209 (2013).
26. M. D. Reed *et al.*, *Phys. Rev. Lett.* **105**, 173601 (2010).
27. Materials and methods are available as supplementary materials.
28. A. Narla *et al.*, *Phys. Rev. X* **6**, 031036 (2016).
29. R. Hisatomi *et al.*, *Phys. Rev. B* **93**, 174427 (2016).
30. G. Flower, J. Bourhill, M. Goryachev, M. E. Tobar, *Phys. Dark Univ.* **25**, 100306 (2019).
31. L. McKenzie-Sell *et al.*, *Phys. Rev. B* **99**, 140414 (2019).
32. Y. Li *et al.*, *Phys. Rev. Lett.* **123**, 107701 (2019).
33. J. T. Hou, L. Liu, *Phys. Rev. Lett.* **123**, 107702 (2019).

ACKNOWLEDGMENTS

The authors thank A. van Loo for fruitful discussions and J. Koenig for carefully reading the manuscript. **Funding:** This work is partly supported by Japan Society for the Promotion of Science (JSPS) KAKENHI (26220601, 18F18015), JST Exploratory Research for Advanced Technology (ERATO) (JPMJER1601), FRQNT Postdoctoral Fellowships, and Ministry of Education, Culture, Sports, Science and Technology (MEXT) Monbukagakusho Scholarship. D.L.-Q. is an international research fellow of JSPS. **Authors contributions:** D.L.-Q. and S.P.W. performed the experiments and analyzed the data; D.L.-Q. performed the

numerical simulations with assistance from S.K.; Y.T. and Y.N. conceived the hybrid system; and K.U. and Y.N. advised on all efforts. All authors contributed to discussions and production of the manuscript. **Competing interests:** The authors declare no competing financial interests. **Data and materials availability:** All data are available in the manuscript or the supplementary materials. Correspondence and requests for materials should be addressed to corresponding author Y.N.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/367/6476/425/suppl/DC1
Supplementary Text
Figs. S1 to S11
Tables S1 to S5
References (*34*–*40*)

21 October 2019; accepted 19 December 2019
10.1126/science.aaz9236

SURFACE MICROSCOPY

Visualizing H_2O molecules reacting at TiO_2 active sites with transmission electron microscopy

Wentao Yuan[1]*, Beien Zhu[2,3]*, Xiao-Yan Li[2,4]*, Thomas W. Hansen[5], Yang Ou[1], Ke Fang[1], Hangsheng Yang[1], Ze Zhang[1]†, Jakob B. Wagner[5]†, Yi Gao[2,3]†, Yong Wang[1]†

Imaging a reaction taking place at the molecular level could provide direct information for understanding the catalytic reaction mechanism. We used in situ environmental transmission electron microscopy and a nanocrystalline anatase titanium dioxide (001) surface with (1 × 4) reconstruction as a catalyst, which provided highly ordered four-coordinated titanium "active rows" to realize real-time monitoring of water molecules dissociating and reacting on the catalyst surface. The twin-protrusion configuration of adsorbed water was observed. During the water–gas shift reaction, dynamic changes in these structures were visualized on these active rows at the molecular level.

Imaging at the atomic scale with transmission electron microscopy (TEM) has benefited from the developments of aberration correctors and in situ equipment (*1*–*8*). For studies of heterogeneous catalysts, these developments, along with approaches that allow gases and even liquids to contact samples [known as environmental TEM (ETEM)], have enabled imaging of single molecules and atoms adsorbed on a catalyst surface (*9*–*14*). However, the direct visualization of gas molecules reacting at catalytic sites is generally difficult to achieve with TEM. Normally, the molecules that adsorb and react dynamically do not offer sufficient contrast for TEM identification. We now show that this obstacle can be overcome by taking advantage of the highly ordered four-coordinated Ti (Ti_{4c}) rows (termed "active rows," owing to their lower coordination) on the anatase TiO_2 (1×4)-(001) surface [i.e., a TiO_2(001) surface with (1 × 4) reconstruction] to facilitate enhanced contrast of adsorbing molecules along the row direction and allow real-time monitoring of H_2O species dissociating and reacting on the catalyst surface.

The atomic structure of the TiO_2 (1×4)-(001) surface has been characterized by both aberration-corrected ETEM and scanning transmission electron microscopy (STEM) images. The bulk-truncated (1×1)-(001) surface usually reconstructs to a (1×4)-(001) surface (Fig. 1, A to C) by periodically replacing the surface oxygen rows (along the [010] direction) with TiO_3 ridges every four unit cells along the TiO_2[100] direction (*15*–*17*). As a result, protruded Ti_{4c} rows are periodically exposed on the surface and show distinct contrast, so the subtle changes occurring in reactions could be detected by means of ETEM observation without contrast overlap. The ordered Ti_{4c} active rows could provide sufficient contrast for direct ETEM visualization of water if the molecules adsorbed in ordered arrays.

We synthesized TiO_2 nanocrystals with exposed {001} facets by a hydrothermal route (see supplementary materials) (*18*, *19*). The nanocrystals were heated in oxygen in situ (~10^{-3} mbar) at 500° to 700°C to trigger the reconstruction. The reconstructed structures remained stable in this temperature range, in accord with recent ETEM studies (*15*, *16*, *20*). During the ETEM experiments, we used a constant electron beam dose with a small value (<1 A/cm^2), and no appreciable irradiation damage was observed on the TiO_2 surface (*21*). After heating at 700°C for ~10 min, the reconstructed TiO_2 (1×4)-(001) surface of an admolecule (ADM) configuration was obtained, as confirmed by the ETEM image (Fig. 1D), in which the protruding black dots represent the Ti_{4c} rows. The ADM structure did not change appreciably after ~16 min of intermittent TEM observation.

[1]State Key Laboratory of Silicon Materials and Center of Electron Microscopy, School of Materials Science and Engineering, Zhejiang University, Hangzhou, 310027 China. [2]Division of Interfacial Water and Key Laboratory of Interfacial Physics and Technology, Shanghai Institute of Applied Physics, Chinese Academy of Sciences, Shanghai, 201800 China. [3]Shanghai Advanced Research Institute, Chinese Academy of Sciences, Shanghai, 201210 China. [4]University of Chinese Academy of Sciences, Beijing, 100049 China. [5]DTU Nanolab, Technical University of Denmark, DK-2800, Kgs. Lyngby, Denmark.
*These authors contributed equally to this work.
†Corresponding author. Email: yongwang@zju.edu.cn (Y.W.); gaoyi@zjlab.org.cn (Y.G.); jakob.wagner@cen.dtu.dk (J.B.W.); zezhang@zju.edu.cn (Z.Z.)

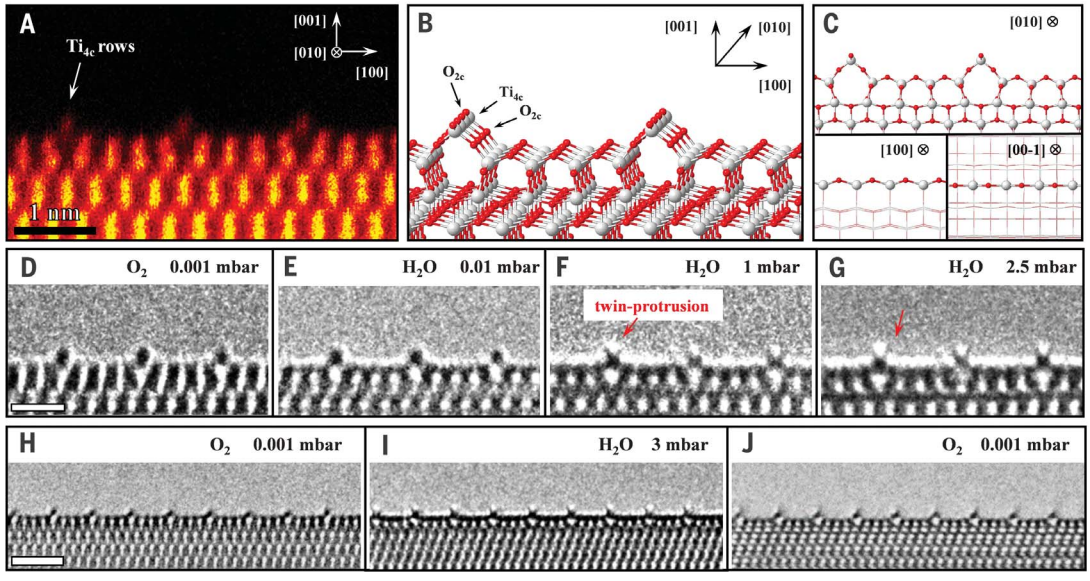


Fig. 1. Dynamic atomic structural evolution of the (1×4) reconstructed TiO_2(001) surface in a water vapor environment. (**A**) High-angle annual dark-field–STEM image of the (1×4)-(001) surface, viewed from the [010] direction. The image was acquired at 700°C in vacuum (TEM column pressure: ~10^{-7} mbar). (**B**) ADM reconstruction models of the (1×4)-(001) surface (Ti, gray; O, red). (**C**) Atomic models of a Ti_{4c} row. (**D** to **G**) Aberration-corrected in situ ETEM images show the same area of TiO_2(001) surface at 700°C under oxygen [(D), 0.001 mbar] and water vapor [(E), 0.01 mbar; (F), 1 mbar; (G), 2.5 mbar] conditions. Scale bar, 1 nm. (**H** to **J**) Another case shows the reversible structural transition induced by a change in the gas environment at 700°C from oxygen [(H), 0.001 mbar] to water vapor [(I), 3 mbar] and then reversion to oxygen [(J), 0.001 mbar]. Scale bar, 2 nm.

The O_2 gas was then evacuated, and H_2O vapor (fig. S1) was introduced at the same temperature. When the H_2O pressure was raised to 1 mbar, two additional small protrusions were observed at the top of the Ti_{4c} rows (Fig. 1F). This twin-protrusion structure became more resolved for a H_2O pressure of 2.5 mbar, owing to a higher water surface coverage (Fig. 1G and movie S1). At both pressures, the twin-protrusion structure remained visible during the TEM observation. When the background environment was changed from H_2O to O_2 or vacuum, the twin-protrusion structure disappeared (Fig. 1, H and J, and fig. S2). The electron beam was switched off after acquisition of the image in Fig. 1H and then H_2O was introduced; a snapshot (Fig. 1I) obtained ~5 min later still shows the twin-protrusion structure, which excludes the effect of the electron beam in its formation. We also ruled out the defocus effect of TEM imaging in different gas environments (figs. S3 to S5). Because the TiO_2 surface did not undergo any other structural changes, we attributed the twin protrusions to an adsorbed water species.

We performed in situ Fourier transform infrared spectroscopy (FTIR) to characterize the surface adsorption species. We heated the TiO_2 crystals to 500°C in vacuum to obtain the (1×4)-(001) surface. Under these conditions, no obvious valley was observed in the hydroxyl region (blue trace in Fig. 2A). Water vapor (5 mbar) was introduced into the in situ FTIR reactor to mimic the in situ TEM experimental condition. About 20 min later, we started to acquire the spectrum and observed two valleys in the hydroxyl region at 3717 and 3663 cm^{-1}. We assigned both features to the adsorbed species on the Ti_{4c} rows (*22*, *23*), because previous studies have shown that the water molecules only chemically adsorb at the Ti_{4c} ridges on the (1×4)-(001) surface (*24*). This indicates that the twin-protrusion structure observed in the ETEM experiments (also at 500°C; see fig. S6) was composed of two different hydroxyl species.

We used density functional theory (DFT) to examine the different adsorbed water structures on the (1×4)-(001) surface (figs. S7 and S8 and appendix S1). At low coverage, one dissociative H_2O adsorbs stably at the Ti_{4c} site by transforming the H atom to the adjacent O_{2c} atom and cleaving the Ti_{4c}–O_{2c} bond. With increasing coverage, the stability of the dissociatively adsorbed H_2O structure decreases because of the increased stress in the reconstructed substrate, in agreement with recently reported results (*25*). Instead, the relative stability of the structure with two symmetric protrusions (each is an OH–H_2O group) (Fig. 2, B to D) increases because it does not induce additional stress at higher coverages (fig. S9). The structure has comparable adsorption energy per H_2O molecule with the dissociatively adsorbed H_2O at ½ coverage. The stability of this twin-protrusion structure becomes compelling when the coverage reaches 1, corresponding to the experimental condition as calculated by combining the adsorption energy with the thermodynamic adsorption isotherm (*26*, *27*). On the basis of this atomic structure, a simulated high-resolution TEM image (fig. S10B) was generated, in agreement with the ETEM image (fig. S10A). In addition, the calculated vibration frequencies of the twin protrusions at 3695 and 3652 cm^{-1}, respectively, were consistent with the in situ FTIR results.

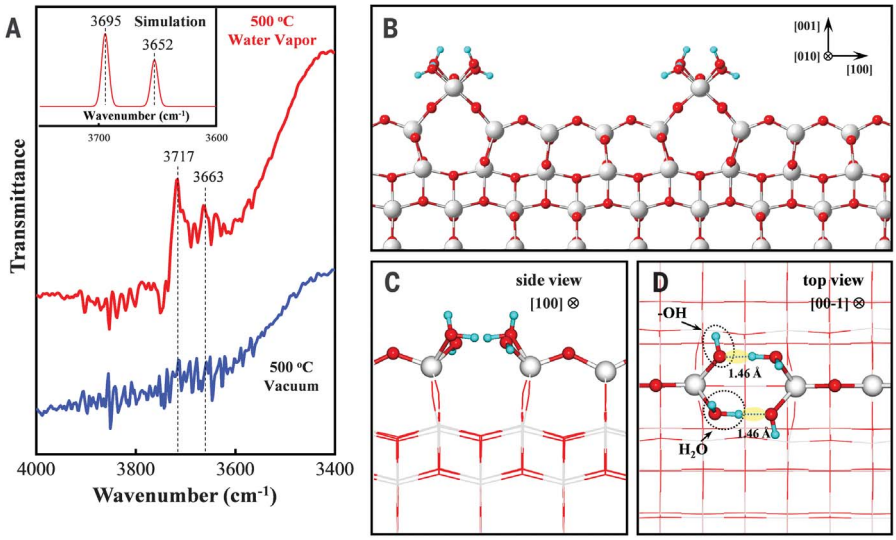


Fig. 2. The twin-protrusion configuration of adsorbed water. (**A**) In situ FTIR spectra of the hydroxyl region for TiO_2 in the presence of water vapor (5 mbar; 500°C) and vacuum (10^{-6} mbar; 500°C). The inset shows results of a theoretical simulation. (**B** to **D**) Atomic structure of the adsorbed H_2O species on the TiO_3 rows, as verified by theoretical calculations, viewed from the [010] direction (B), the [100] direction (C), and the [00-1] direction (D) (gray, Ti; red, O; cyan, H).

Because TiO_2 can catalyze the water–gas shift reaction ($H_2O + CO \rightarrow H_2 + CO_2$) at elevated temperatures (*28*, *29*), we studied this reaction by introducing CO into the ETEM column. The gas environment was changed from pure water vapor (2.5 mbar) to a mixed gas environment (CO and H_2O vapor in a 1:1 ratio; pressure: 5 mbar). Under these conditions, the twin-protrusion structure became unstable (Fig. 3A and movie S2). Its contrast changed dynamically: Most of the time it was blurred, but it would occasionally clear (Fig. 3B), with no substantial contrast change observed in TiO_2 bulk and in other surface areas. For example, in one case the twin protrusion was clearly seen initially [Fig. 3B, (1)], almost disappeared after 2.2 s [Fig. 3B, (2)], and then reappeared at 4 s [Fig. 3B, (3)]. The disappearance and reappearance occurred again at 5.8 s [Fig. 3B, (4)] and 7.8 s [Fig. 3B, (5)], respectively. The contrast change of the twin protrusions was also evidenced by the intensity profiles across the protruding row (Fig. 3C). Similar cases are shown in fig. S11 and movie S3. In a pure water vapor environment, the twin protrusions did not display such contrast changes (fig. S12 and movie S1), hence ruling out electron beam effects for the disappearances and reappearances.

Thus, the dynamic change of twin protrusions in mixed gas environments suggests that the adsorbed hydroxyls were reacting with CO molecules, which indicates that the Ti_{4c} sites are the reaction sites. In addition, because the net free-energy change of this reaction is negative (–3.76 kJ mol^{-1} under the experimental condition) and the known conversion temperatures are generally lower than 700°C (*28*, *29*), it is reasonable to conclude that the observed reaction was not induced by the electron beam. The reaction pathway of the twin-protrusion–adsorbed H_2O species with CO molecules was calculated by DFT (fig. S13). During the reaction, the H_2O species of the twin protrusion are consumed by CO gas and supplemented by H_2O vapor repeatedly, which relates to the dynamic contrast change observed experimentally. In the reaction cycle (fig. S13), the two largest energy barriers come from H_2O dissociation of the twin-protrusion (0.48 eV) and single OH–H_2O (0.57 eV) structures, which indicates that these are two relatively stable structures with comparatively long lifetimes. Thus, a changing mixture of single OH–H_2O and twin-protrusion structures was imaged by TEM. The contrast of the twin protrusions would occasionally clear when they were the majority on one of the active rows [Fig. 3B, (2) and (4)].

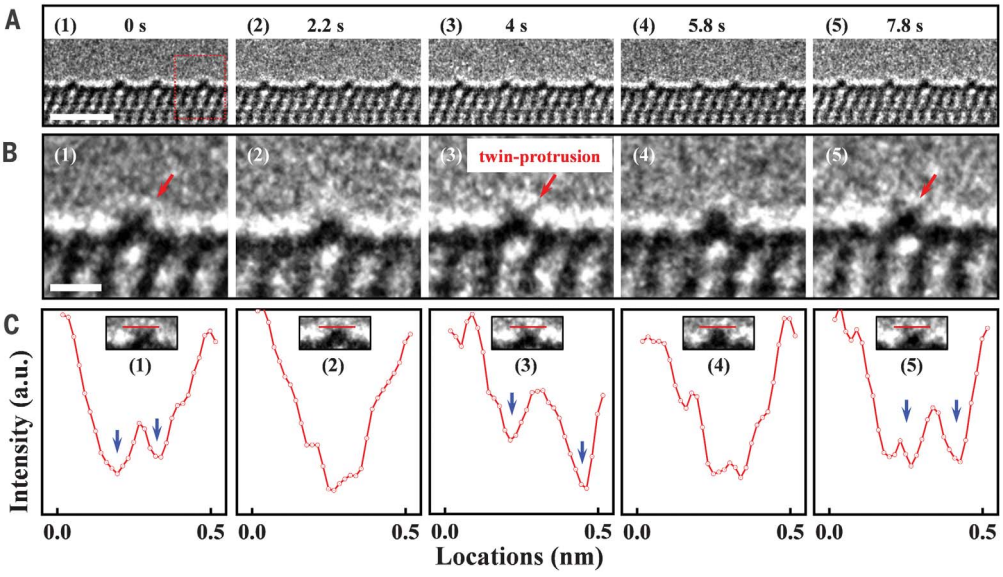


Fig. 3. Dynamic structural evolution of the (1 × 4)-(001) surface in the water–gas shift reaction. (**A**) Sequential ETEM images acquired in the mixed gas environment (1:1 ratio of CO and H_2O vapor; gas pressure: 5 mbar; temperature: 700°C), viewed from the [010] direction. Scale bar, 2 nm. (**B**) Enlarged ETEM images show the dynamic structural evolution of the Ti row outlined by the dotted rectangle in (A). Scale bar, 0.5 nm. (**C**) Intensity profiles along the lines crossed the Ti rows of (B). Blue arrows denote intensity valleys corresponding to the twin protrusions. a.u., arbitrary units.

Most of the time, the contrast is blurred because of the interference between the two structures [Fig. 3B, (1), (3), and (5)]. The single OH–H_2O structure was not obviously visualized via TEM, as shown by the simulated image (fig. S14).

By visualizing and monitoring the adsorbed water species on the ridge of the (1×4)-(001) TiO_2 surface, we confirmed that the Ti_{4c} atoms on the ridge are active sites for H_2O dissociation and reaction. The direct TEM visualization revealed an adsorbed water structure with a twin-protrusion feature on the TiO_2 surface. This work demonstrates that in situ ETEM can be used to monitor a catalytic process taking place at highly ordered active sites.

REFERENCES AND NOTES

1. D. A. Muller, *Nat. Mater.* **8**, 263–270 (2009).
2. D. S. Su, B. Zhang, R. Schlögl, *Chem. Rev.* **115**, 2818–2882 (2015).
3. L. DeRita *et al.*, *Nat. Mater.* **18**, 746–751 (2019).
4. L. Luo *et al.*, *Nat. Mater.* **17**, 514–518 (2018).
5. L. Zou *et al.*, *Nat. Mater.* **17**, 56–63 (2018).
6. L. Zhang, B. K. Miller, P. A. Crozier, *Nano Lett.* **13**, 679–684 (2013).
7. K. Sytwu *et al.*, *Nano Lett.* **18**, 5357–5363 (2018).
8. Y. Lin *et al.*, *Phys. Rev. Lett.* **111**, 156101 (2013).
9. C. L. Jia, M. Lentzen, K. Urban, *Science* **299**, 870–873 (2003).
10. M. Koshino *et al.*, *Science* **316**, 853 (2007).
11. Z. Liu, K. Yanagi, K. Suenaga, H. Kataura, S. Iijima, *Nat. Nanotechnol.* **2**, 422–425 (2007).
12. J. E. Allen *et al.*, *Nat. Nanotechnol.* **3**, 168–173 (2008).
13. Y. Oshima *et al.*, *Phys. Rev. B* **81**, 035317 (2010).
14. H. Yoshida *et al.*, *Science* **335**, 317–319 (2012).
15. W. T. Yuan *et al.*, *Chem. Mater.* **29**, 3189–3194 (2017).
16. W. Yuan *et al.*, *Nano Lett.* **16**, 132–137 (2016).
17. M. Lazzeri, A. Selloni, *Phys. Rev. Lett.* **87**, 266105 (2001).
18. H. G. Yang *et al.*, *Nature* **453**, 638–641 (2008).
19. X. Han, Q. Kuang, M. Jin, Z. Xie, L. Zheng, *J. Am. Chem. Soc.* **131**, 3152–3153 (2009).
20. K. Fang *et al.*, *J. Phys. Chem. C* **123**, 21522–21527 (2019).
21. Y. Kuwauchi, H. Yoshida, T. Akita, M. Haruta, S. Takeda, *Angew. Chem. Int. Ed.* **51**, 7729–7733 (2012).
22. C. Arrouvel, M. Digne, M. Breysse, H. Toulhoat, P. Raybaud, *J. Catal.* **222**, 152–166 (2004).
23. C. Deiana, E. Fois, S. Coluccia, G. Martra, *J. Phys. Chem. C* **114**, 21531–21538 (2010).
24. J. Blomquist, L. E. Walle, P. Uvdal, A. Borg, A. Sandell, *J. Phys. Chem. C* **112**, 16616–16621 (2008).
25. I. Beinik *et al.*, *Phys. Rev. Lett.* **121**, 206003 (2018).
26. M. Duan *et al.*, *Angew. Chem. Int. Ed.* **57**, 6464–6469 (2018).
27. B. Zhu, Z. Xu, C. Wang, Y. Gao, *Nano Lett.* **16**, 2628–2632 (2016).
28. P. Panagiotopoulou, D. I. Kondarides, *J. Catal.* **225**, 327–336 (2004).
29. D. G. Rethwisch, J. A. Dumesic, *Appl. Catal.* **21**, 97–109 (1986).

ACKNOWLEDGMENTS

We gratefully acknowledge J. Fan (Department of Chemistry, Zhejiang University) for support and useful discussions.

Funding: We acknowledge the financial support of the National Natural Science Foundation of China (51390474, 91645103, 11574340, 21773287, 51801182, 11604357 51872260, and 11327901), the Zhejiang Provincial Natural Science Foundation (LD19B030001), the Ministry of Science and Technology of China (2016YFE0105700), and the Fundamental Research Funds for the Central Universities. B.Z. was supported by the Natural Science Foundation of Shanghai (16ZR1443200) and the Youth Innovation Promotion Association CAS. The computations were performed at the National Supercomputing Center in Guangzhou and Shanghai. W.Y. was supported by the China Postdoctoral Science Foundation (2018M642407 and 2019T120502). **Author contributions:** Y.W. initiated the work. Y.W., Y.G., J.B.W., and Z.Z. supervised the work. W.Y., Y.O., and K.F. synthesized the samples. W.Y. and T.W.H. conducted the ETEM experiments. Y.O. and H.Y. carried out the in situ FTIR experiments. B.Z. and X.Y.L. performed the calculations. All authors participated in the analysis and discussion. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper, the supplementary materials, or the Cambridge Crystallographic Data Centre (deposition number: CSD 1970465-1970473).

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/367/6476/428/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S14
References (*30*–*42*)
Movies S1 to S3
Appendix S1

2 June 2019; resubmitted 24 October 2019
Accepted 10 December 2019
10.1126/science.aay2474

PLANT SCIENCE

A two-way molecular dialogue between embryo and endosperm is required for seed development

N. M. Doll[1], S. Royek[2]*, S. Fujita[3]*†, S. Okuda[4]*, S. Chamot[1], A. Stintzi[2], T. Widiez[1], M. Hothorn[4], A. Schaller[2], N. Geldner[3], G. Ingram[1]‡

The plant embryonic cuticle is a hydrophobic barrier deposited de novo by the embryo during seed development. At germination, it protects the seedling from water loss and is, thus, critical for survival. Embryonic cuticle formation is controlled by a signaling pathway involving the ABNORMAL LEAF SHAPE1 subtilase and the two GASSHO receptor-like kinases. We show that a sulfated peptide, TWISTED SEED1 (TWS1), acts as a GASSHO ligand. Cuticle surveillance depends on the action of the subtilase, which, unlike the TWS1 precursor and the GASSHO receptors, is not produced in the embryo but in the neighboring endosperm. Subtilase-mediated processing of the embryo-derived TWS1 precursor releases the active peptide, triggering GASSHO-dependent cuticle reinforcement in the embryo. Thus, a bidirectional molecular dialogue between embryo and endosperm safeguards cuticle integrity before germination.

In angiosperms, seeds comprise three genetically distinct compartments: the zygotic embryo, the endosperm, and the maternal seed coat. Their development must be tightly coordinated for seed viability. In this work, we have elucidated a bidirectional peptide-mediated signaling pathway between the embryo and the endosperm. This pathway regulates the deposition of the embryonic cuticle, which forms an essential hydrophobic barrier separating the apoplasts of the embryo and endosperm. After germination, the cuticle—one of the critical innovations underlying the transition of plants from their original, aqueous environment to dry land—protects the seedling from catastrophic water loss (*1*, *2*).

Formation of the embryonic cuticle has previously been shown to depend on two receptor-like kinases (RLKs)—GASSHO1/SCHENGEN3 (hereafter named GSO1) and GSO2—and on ALE1, a protease of the subtilase family (*2*–*5*). *gso1 gso2* and (to a lesser extent) *ale1* mutants produce a patchy and highly permeable cuticle (*2*). Mutant embryos also adhere to surrounding tissues, causing a seed-twisting phenotype (*6*). Because subtilases have been implicated in the processing of peptide hormone precursors (*7*–*9*), we hypothesized that ALE1 may be required for the biogenesis of the elusive intercompartmental peptide signal required for GSO1/2-dependent cuticle deposition.

[1]Laboratoire Reproduction et Développement des Plantes, University of Lyon, ENS de Lyon, UCB Lyon 1, CNRS, INRAE, F-69342, Lyon, France. [2]Department of Plant Physiology and Biochemistry, University of Hohenheim, 70599 Stuttgart, Germany. [3]Department of Plant Molecular Biology, University of Lausanne, 1015 Lausanne, Switzerland. [4]Structural Plant Biology Laboratory, Department of Botany and Plant Biology, University of Geneva, 1211 Geneva, Switzerland.
*These authors contributed equally to this work.
†Present address: National Institute of Genetics, 1111 Yata, Mishima, Shizuoka 411-8540, Japan.
‡Corresponding author. Email: gwyneth.ingram@ens-lyon.fr

CASPARIAN STRIP INTEGRITY FACTORs (CIFs), a family of small sulfated signaling peptides, are ligands for GSO1 and GSO2 (*10*–*12*). CIF1 and CIF2 are involved in Casparian strip formation in the root endodermis (*10*, *11*). The function of CIF3 and CIF4 is still unknown. To assess the role of CIF peptides in cuticle development, the quadruple mutant (*cif1 cif2 cif3 cif4*) was generated (fig. S1A). Neither cuticle permeability nor seed twisting phenotypes were observed in this quadruple mutant (fig. S1, B to E). However, reduction [in the leaky *sgn2-1* allele (*10*)] or loss [in the *tpst-1* mutant (*13*)] of tyrosyl-protein sulfotransferase (TPST) activity results in seed-twisting and cuticle-permeability phenotypes resembling those observed in *ale1* mutants (Fig. 1, A to D, and fig. S2, A to D). These data suggest that a sulfated peptide may act as the ligand of GSO1/2 during seed development.

Consistent with the hypothesis that TPST acts in the same pathway as GSO1 and GSO2, no difference was observed between the phenotype of *tpst-1 gso1-1 gso2-1* triple and *gso1-1 gso2-1* double mutants (fig. S2E). In contrast, TPST and ALE1 appear to act synergistically, as a phenotype resembling that of *gso1 gso2* double mutants was observed in *tpst-1 ale1-4* double mutants (Fig. 1, E to I, and fig. S2, F to J). This result supports the hypothesis that TPST and ALE1 act in parallel regarding their roles in embryonic cuticle formation, possibly through independent posttranslational modifications that contribute to the maturation of the hypothetical peptide signal.

Identification of the peptide signal was facilitated by a study of TWISTED SEED1 (TWS1) (*14*), which reported a loss-of-function phenotype that was notably similar to that of *gso1 gso2* double mutants. Because existing alleles of *TWS1* are in the Wassilewskija (WS) background, we generated new CRISPR alleles (*tws1-3* to *tws1-10*) in the Col-0 background and confirmed the phenotype of resulting mutants (Fig. 1 and fig. S3). No additivity was observed when loss-of-function alleles of *TWS1* and of other pathway components (*GSO1*, *GSO2*, *TPST*, and *ALE1*) were combined, providing genetic evidence for *TWS1* acting in the GSO signaling pathway (fig. S4). Furthermore, gaps in the cuticle of embryos and cotyledons, similar to those observed in *ale1* and *gso1 gso2* mutants (*2*), were detected in both the *tws1* mutants and *tpst* mutants (Fig. 1, J to N, and fig. S5). Inspection of the TWS1 protein sequence revealed a region with limited similarity to CIF peptides, including a DY motif that marks the N terminus of the CIFs (Fig. 1O) and is the minimal motif required for tyrosine sulfation by TPST (*15*). Corroborating the functional importance of the putative peptide domain, the *tws1-6* allele (deletion of six codons in the putative peptide-encoding region) and the *tws1-5* allele (substitution of eight amino acids, including the DY motif) both showed total loss of function of the TWS1 protein (fig. S3).

We tested whether TWS1 is a substrate of ALE1 by coexpression of ALE1:(His)6 and TWS1:GFP-(His)6 fusion proteins in tobacco (*Nicotiana benthamiana*) leaves. A specific TWS1 cleavage product was observed upon coexpression of ALE1 but not in the empty-vector control, suggesting that TWS1 is processed by ALE1 in planta (Fig. 1P). Likewise, recombinant TWS1 expressed as GST-fusion in *Escherichia coli* was cleaved by purified ALE1 in vitro. (Fig. 1Q). Mass spectroscopy analysis of the TWS1 cleavage product purified from tobacco leaves showed that ALE1 cleaves TWS1 between His^{54} and Gly^{55} (fig. S6). These residues are important for cleavage site selection, as ALE1-dependent processing was not observed when either His^{54} or Gly^{55} was substituted by site-directed mutagenesis (Fig. 1Q). His^{54} corresponds to the C-terminal His or Asn of CIF peptides (Fig. 1O). Thus, the data suggest that ALE1-mediated processing of the TWS1 precursor marks the C terminus of the TWS1 peptide. Because the CIF1 and CIF2 peptides are located at the very end of their respective precursors, C-terminal processing could represent a mechanism of peptide activation operating in the developing seed but not in the root. A summary of TWS1 modifications is provided in Fig. 1R.

To test the biological activity of TWS1, the predicted peptide encompassing the conserved N-terminal DY motif and the C terminus defined by the ALE1 cleavage site was custom-synthesized in tyrosine-sulfated form. As synthetic TWS1 cannot easily be applied to developing embryos, a root bioassay for CIF activity was used. In wild-type roots, TWS1 induced ectopic endodermal lignification, as previously observed for the CIF1 and CIF2 peptides (*12*). TWS1 activity was GSO1-dependent, suggesting that processed TWS1 peptide can replace

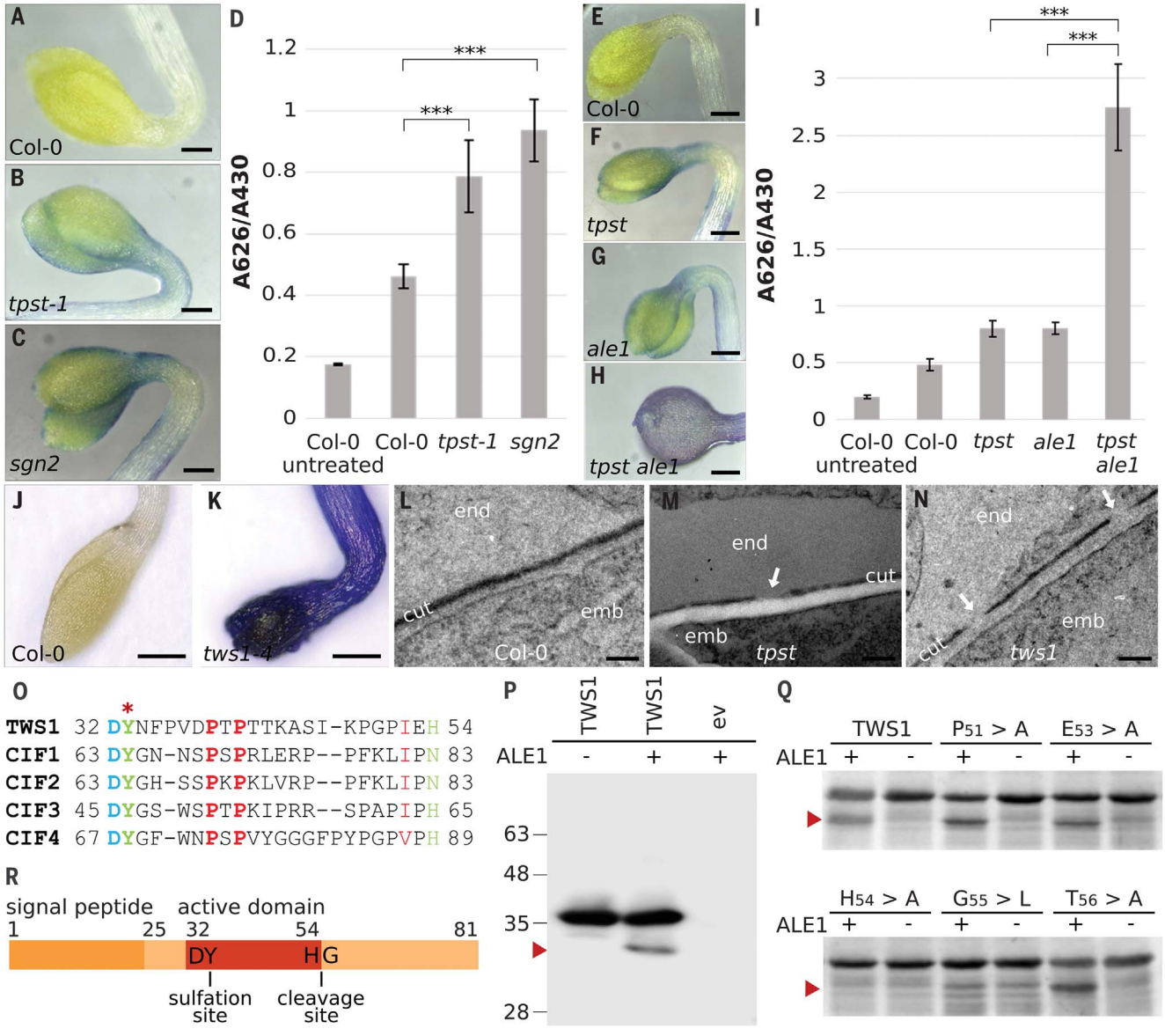


Fig. 1. TPST and ALE1 are required for maturation of the TWS1 peptide. (**A** to **C**, **E** to **H**, and **J**, and **K**) Toluidine blue tests on etiolated cotyledons. Scale bars, 200 μm. (**D** and **I**) Quantification of toluidine blue uptake by the aerial parts of young seedlings, normalized to chlorophyll content. $N = 6$, 10 seedlings per repetition. *** indicates statistical differences with one-way analysis of variance (ANOVA) followed by a post hoc Scheffé multiple comparison test ($P < 0.01$) in (D) and (I). Error bars represent standard deviations. (J and K) Toluidine blue permeability of *tws1-4* compared with Col-0. Scale bars, 400 μm (**L** to **N**) Transmission electron micrographs of the embryo (emb) to endosperm (end) interface at the heart stage. Scale bars, 200 nm. Genotypes are indicated, and gaps in the cuticle (cut) are shown by white arrows. (**O**) The predicted TWS1 active peptide sequence and alignment with four other known GSO ligands (CIF1, CIF2, CIF3, and CIF4). The site of predicted sulfation is indicated with a red asterisk. (**P**) Anti-His Western blot of protein extracts from *N. benthamiana* leaves, agro-infiltrated to express TWS1::GFP (His)6 (TWS1) or the empty vector (−). Coexpression of ALE1::(His)6 or the empty-vector control are indicated by + and −, respectively. (**Q**) Coomassie-stained SDS-PAGE showing recombinant GST-TWS1 and the indicated site-directed mutants digested in vitro with (+) or without (−) ALE1-(His)6 purified from tobacco leaves. Arrows indicate specific cleavage products. (**R**) The full length TWS1 precursor. Sulfation and ALE1 cleavage sites are indicated. Single-letter abbreviations for the amino acid residues are as follows: A, Ala; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; N, Asn; P, Pro; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.

CIF1 and CIF2 as a ligand for GSO1 during Casparian strip formation (Fig. 2A and fig. S7). Supporting this, TWS1 application complemented the *cif1 cif2* mutant, albeit with reduced activity compared with CIF2 (Fig. 2B and fig. S8). TWS1 activity in this assay was reduced when sulfation on the DY motif was missing (Fig. 2B). Versions of TWS1 in which Y33 was mutated to either F or T only partially complemented the mutant phenotype of *tws1-4* (fig S9), consistent with a residual but weak activity for nonsulfated TWS1 in vivo and with the weak loss-of-function phenotype of the *tpst-1* mutant.

To confirm TWS1 as a ligand of GSO1 and GSO2, the interaction of the synthetic peptide with the leucine-rich repeat (LRR) ectodomains of the receptors was analyzed in

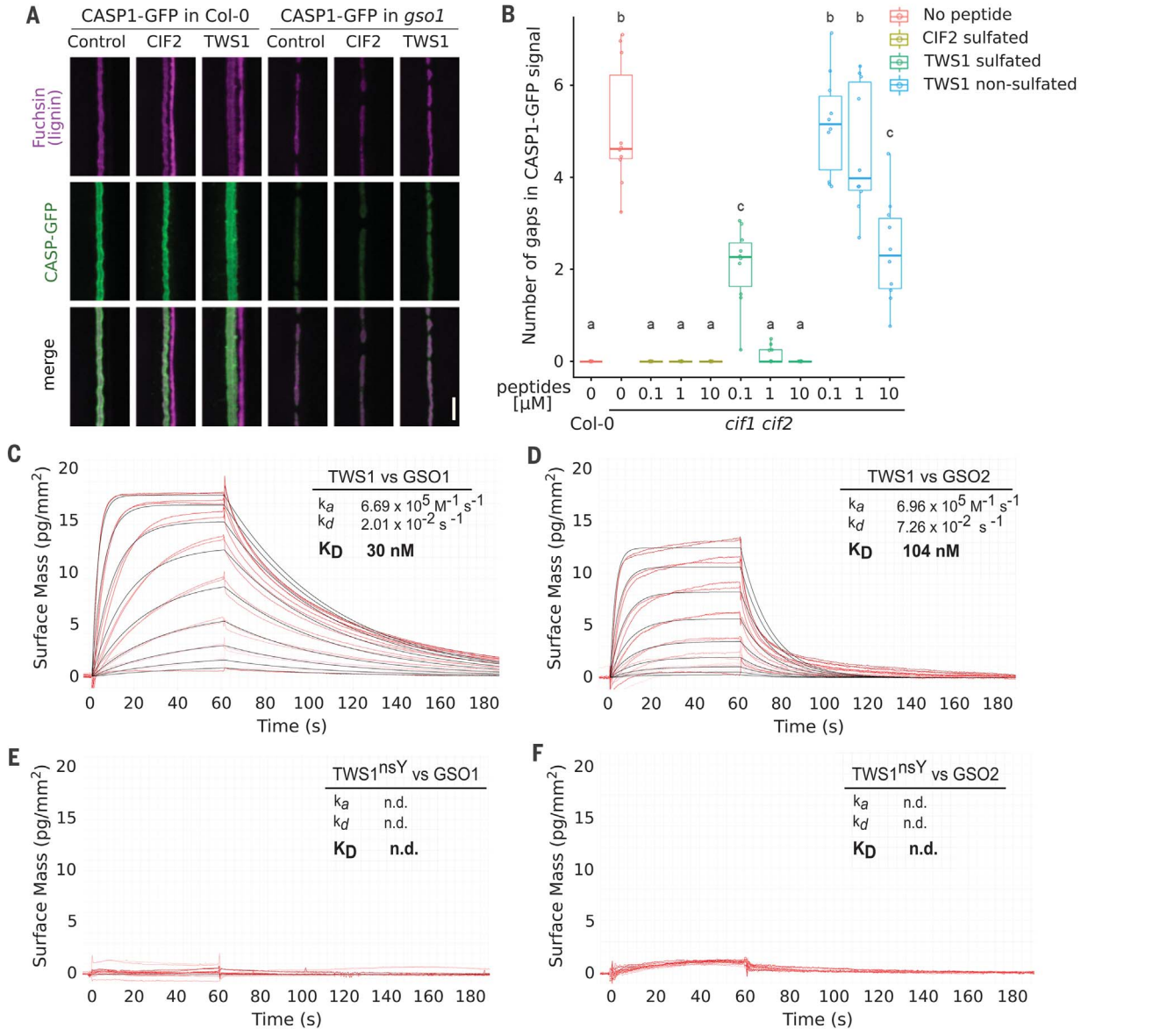


Fig. 2. The TWS1 peptide is a functional GSO1/GSO2 ligand. (**A**) Root overlignification following treatment with the active CIF2 or TWS peptide in Col-0 and in the *gso1* (*sgn3-3*) background. Lignin is stained in purple and CASP-GFP fusion protein, marking the Casparian strip domain, in green. Scale bar, 5 µm. (**B**) Complementation of *cif1-2 cif2-2* Casparian strip integrity phenotype by peptide treatments. Number of gaps in CASP1-GFP signal counted after treatment with CIF2 sulfated peptide, TWS1 sulfated peptide, TWS1 nonsulfated peptide. *N* = 10. a, b, and c correspond to classes statistically supported by one-way ANOVA analysis, followed by Tukey tests ($P < 0.05$). (**C** to **F**) Grating-coupled interferometry (GCI)–derived binding kinetics. Shown are sensorgrams with raw data in red and their respective fits in black. k_a, association rate constant; k_d, dissociation rate constant; K_D, dissociation constant. (C) Data for the GSO1 extracellular domain in the presence of the sulfated TWS1 peptide. (D) Data for the GSO2 extracellular domain in the presence of the sulfated TWS1 peptide. (E) Data for the GSO1 extracellular domain in the presence of the nonsulfated TWS1 peptide. (F) Data for the GSO2 extracellular domain in the presence of the nonsulfated TWS1 peptide. n.d., not determined.

grating-coupled interferometry binding assays. GSO1 bound sulfated TWS1 with a K_D (dissociation constant) of ~30 nM (Fig. 2C). The observed binding affinity is ~1/10 that of the CIF2 peptide (K_D = 2.5 nM) (fig. S10), which is consistent with the reduced ability of TWS1 to complement the root phenotype of the *cif1 cif2* double mutant (Fig. 2B). Sulfated TWS1 also bound to the LRR domain of GSO2, albeit with slightly reduced affinity (K_D ~ 100 nM) (Fig. 2D). As previously shown for other CIF peptides (*11*), tyrosine sulfation was critical for the interaction of TWS1 with GSO1 and GSO2 in vitro (Fig. 2, E and F). Technical issues at high peptide concentrations may explain discrepancies between in vitro binding assays and the in vivo activity of nonsulfated TWS1. In vivo activities for nonsulfated versions of other normally sulfated peptides, including CIF2, have been reported (*11*, *16*–*18*). Adding a 3AA C-terminal extension to the sulfated TWS1 peptide reduced binding affinity to both GSO1 and GSO2 (fig. S10), consistent with the need for ALE1-mediated C-terminal processing for efficient signaling.

Taken together, our results suggest the sulfated TWS peptide as the missing link in the

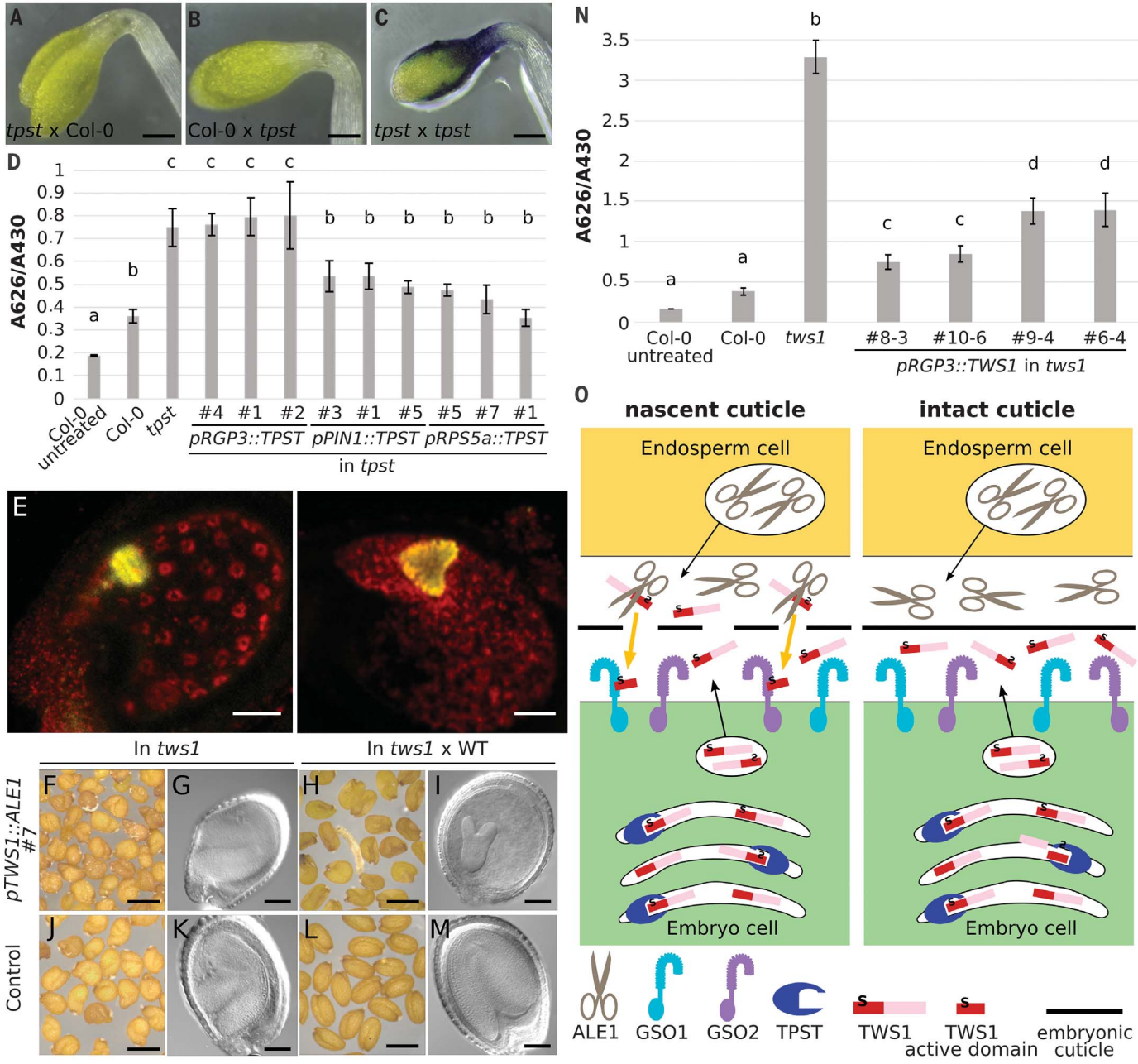


Fig. 3. Spatial separation of *ALE1* and *TWS1* expression is critical for pathway function. (**A** to **C**) F1 seedlings from reciprocal crosses stained with Toluidine blue. (**D** and **N**) Toluidine blue quantification as in Fig. 1. a to d indicate statistical differences with one-way ANOVA followed by a post hoc Scheffé multiple comparison test ($P < 0.01$). (D) Complementation of *tpst-1* mutant with endosperm-specific expression of *TPST* (*pRGP3::TPST*), embryo-specific expression of *TPST* (*pPIN1::TPST*), and ubiquitous expression of *TPST* (*pRPS5a::TPST*) compared with *tpst-1* and Col-0. Three independent lines were analyzed. (**E**) Confocal images of *pTWS1::mCitrine::NLS-mCitrine* reporter lines, signal in yellow, autofluorescence in red. Scale bars, 50 μm. (**F** and **G**) Dry seeds (scale bars, 400 μm) and chloral hydrate cleared seeds (9 DAP) (scale bars, 100 μm), respectively, from a line expressing *ALE1* in the embryo in the *tws1-4* background (*pTWS1::ALE1 line#7*). (**H** and **I**) Seeds from crosses of Col-0 pollen onto *line#7*. (**J** and **K**) Self-fertilized *tws1-4* seeds as a control. (**L** and **M**) Seeds from a cross of Col-0 pollen on a *tws1-4* pistil as a control. Results for three further independent transgenic lines are shown in figs. S18 and S19. (N) Complementation of *tws1-4* mutants by expression of TWS1 in the endosperm. Four independent lines were analyzed. (**O**) Model for embryonic cuticle integrity monitoring. Left shows the wild-type situation before gap-filling (nascent cuticle), illustrating the diffusion and processing of TWS1 across the embryo-endosperm interface. Right shows the wild-type situation when the cuticle is intact, spatially separating signaling components and, thus, attenuating signaling.

intercompartmental signaling pathway for embryonic cuticle formation. The activities of ALE1 and TPST both contribute to the formation of the bioactive peptide (Fig. 1R), which is perceived by GSO1 and GSO2 to ensure appropriate cuticle deposition.

To understand how the elements of the signaling pathway cooperate to ensure the formation of a functional cuticle, we analyzed their spatial organization. In silico data indicate that the *TPST* gene is expressed in all seed tissues (fig. S11) (*19*, *20*). To investigate in

which compartment TPST [which acts cell autonomously (*13*)] is required for TWS1 maturation, reciprocal crosses and complementation assays using tissue-specific promoters were performed. No cuticle permeability defects were observed when homozygous mutants were pollinated with wild-type pollen, confirming their zygotic origin. (Fig. 3, A to C). Expressing TPST under the ubiquitously active *RPS5A* promoter (*21*) or the *PIN1* promoter [which is embryo-specific in seed (fig. S12)] complements *tpst-1* cuticle defects. In contrast, no complementation was observed using the endosperm-specific *RGP3* promoter (*22*), indicating that TPST activity is required for TWS1 sulfation specifically in the embryo to ensure cuticle integrity (Fig. 3D and fig. S13). Consistent with this observation and with a previous report (*14*), the *TWS1* promoter was found to drive expression specifically in the developing embryo from the early globular stage onwards (Fig. 3E and fig S14). The *TPST* promoter (*10*) drove expression throughout the embryo proper at the onset of embryo cuticle establishment (the globular stage) before becoming restricted to the root tip (fig. S11). We conclude that the TWS1 peptide is both sulfated and secreted specifically in the embryo.

However, production of mature TWS1 requires a C-terminal cleavage event that we have shown to be mediated by ALE1. *ALE1* is expressed only in the endosperm (*4*, *23*), on the opposite side of the nascent cuticle to the GSO1 and GSO2 receptors, which are localized on the membranes of the epidermal cells that produce the cuticle (figs. S15 to S17) (*2*). Our data therefore support a model in which activation of the GSO signaling pathway depends on the diffusion of the TWS1 peptide precursor to the endosperm, where it is cleaved and activated by ALE1 before diffusing back to the embryo to trigger GSO1/2-dependent cuticle deposition. An intact cuticle would separate the subtilase from its substrate, terminating signaling.

Expressing *ALE1* in the embryo, under the control of the *TWS1* promoter, provided support for this model. Multiple transformants were obtained in *tws1* mutants, but not in the wild-type background. When *tws1* plants from four independent plants carrying the *pTWS1:ALE1* transgene were pollinated with wild-type pollen—introducing a functional *TWS1* allele into the zygotic compartments and thus inducing colocalization of TWS1 precursors with ALE1, GSO1, and GSO2 in the embryo—premature embryo growth arrest was observed in all seeds. This leads to severe shriveling of all seeds at maturity (Fig. 3, F to M, and figs. S18 and S19). A proportion of seeds could, nonetheless, germinate to give developmentally normal plants (fig. S20), indicating that co-expression of all signaling components in the embryo—although detrimental to embryo development—does not lead to a complete loss of viability. Growth arrest may be due to constitutive embryonic activation of the *GSO1/GSO2* signaling pathway, and stress-responsive genes shown to require GSO1/GSO2 signaling for expression in the seed (*2*) were upregulated in seeds coexpressing GSO1, GSO2, TWS1, and ALE1 in the embryo (fig. S21). We thus postulate that the spatial separation of the TWS1 precursor and the GSO receptors from the activating protease by cuticle is required for signaling attenuation.

We next tested if CIF1, CIF2, and TWS1 could complement *tws1* and *ale1* mutants when expressed in the endosperm (under the *RGP3* promoter). All three peptides complemented *tws1* mutants, confirming that retrograde peptide movement from endosperm to embryo is sufficient to allow integrity monitoring (Fig. 3N and fig. S22). Lack of full complementation could reflect suboptimal N-terminal processing or sulfation in the endosperm. CIF1 and CIF2 (lacking C-terminal extensions) complemented *ale1* mutants much more efficiently than TWS1 (fig. S23). Weak complementation of *ale1* by TWS1 may reflect the presence of redundantly acting subtilases in the endosperm, as suggested by the weak phenotype of *ale1* mutants.

The proposed bidirectional signaling model allows efficient embryo cuticle integrity monitoring. The sulfated TWS1 precursor is produced by the embryo and secreted (probably after N-terminal cleavage of the pro-peptide) to the embryo apoplast. In the absence of an intact cuticular barrier, it can diffuse to the endosperm and undergo activation by ALE1 (and potentially other subtilases). Activated TWS1 peptide then leaks back through cuticle gaps to bind the GSO1 and GSO2 receptors and activate local gap repair (Fig. 3O). When the cuticle is intact, proTWS1 peptides are confined to the embryo where they remain inactive.

Our results demonstrate a role for a subtilase in providing spatial specificity to a bidirectional peptide signaling pathway. In contrast, the related CIF1-, CIF2-, and GSO1-dependent signaling pathway controlling Casparian strip integrity is unidirectional, negating the need for C-terminal cleavage-mediated peptide activation (*10*, *12*). Both pathway components and their spatial organization differ between the two systems, suggesting an independent recruitment of the GSO receptors to different integrity monitoring functions within the plant.

REFERENCES AND NOTES

1. C. Delude, S. Moussu, J. Joubès, G. Ingram, F. Domergue, *Subcell. Biochem.* **86**, 287–313 (2016).
2. A. Creff *et al.*, *PLOS Genet.* **15**, e1007847 (2019).
3. R. Tsuwamoto, H. Fukuoka, Y. Takahata, *Plant J.* **54**, 30–42 (2008).
4. H. Tanaka *et al.*, *Development* **128**, 4681–4689 (2001).
5. Q. Xing *et al.*, *Development* **140**, 770–779 (2013).
6. S. Moussu *et al.*, *Plant Cell* **29**, 1642–1656 (2017).
7. A. Schaller *et al.*, *New Phytol.* **218**, 901–915 (2018).
8. N. Stührwohldt, A. Schaller, *Plant Biol. J.* **21**, 49–63 (2019).
9. K. Schardon *et al.*, *Science* **354**, 1594–1597 (2016).
10. V. G. Doblas *et al.*, *Science* **355**, 280–284 (2017).
11. S. Okuda *et al.*, Molecular mechanism for the recognition of sequence-divergent CIF peptides by the plant receptor kinases GSO1/SGN3 and GSO2. bioRxiv 692228 [Preprint]. 5 July 2019. https://doi.org/10.1101/692228
12. T. Nakayama *et al.*, *Science* **355**, 284–286 (2017).
13. R. Komori, Y. Amano, M. Ogawa-Ohnishi, Y. Matsubayashi, *Proc. Natl. Acad. Sci. U.S.A.* **106**, 15067–15072 (2009).
14. E. Fiume *et al.*, *Plant Physiol.* **172**, 1732–1745 (2016).
15. H. Hanai *et al.*, *FEBS Lett.* **470**, 97–101 (2000).
16. Y. Matsubayashi, Y. Sakagami, *Eur. J. Biochem.* **262**, 666–671 (1999).
17. A. Kutschmar *et al.*, *New Phytol.* **181**, 820–831 (2009).
18. Y. Matsuzaki, M. Ogawa-Ohnishi, A. Mori, Y. Matsubayashi, *Science* **329**, 1065–1067 (2010).
19. B. H. Le *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **107**, 8063–8070 (2010).
20. A. V. Klepikova, A. S. Kasianov, E. S. Gerasimov, M. D. Logacheva, A. A. Penin, *Plant J.* **88**, 1058–1070 (2016).
21. D. Weijers *et al.*, *Development* **128**, 4289–4299 (2001).
22. G. Denay *et al.*, *Development* **141**, 1222–1227 (2014).
23. S. Yang *et al.*, *Development* **135**, 3501–3509 (2008).

ACKNOWLEDGMENTS

We thank L. Lepiniec for providing the *tws1-1* and *tws1-2* seeds; C. Galvan Ampudia, Y. Jaillais, and L. Armengot for materials and helpful discussions; A. Creff, U. Glück-Behrens, A. Lacroix, P. Bolland, J. Berger, I. Desbouchages, and H. Leyral for technical assistance; A. Patole, B. Martin Sempore, C. Vial, and S. Maurin for administrative assistance; and B. Würtz and J. Pfannstiel (Core Facility Hohenheim) for mass spectrometric analyses. Transmission electron microscopy images were acquired at the Centre Technologique des Microstructures, Université Lyon 1. **Funding:** The study was financed by joint funding (project Mind the Gap) from the French Agence National de Recherche (ANR-17-CE20-0027) (to G.I.) and the Swiss National Science Foundation (NSF) (to N.G., supporting S.F.). N.M.D. was funded by a Ph.D. fellowship from the Ministère de l'Enseignement Supérieur et de la Recherche. Funding was also provided by NSR grant no. 31003A_176237 (to M.H.) and an International Research Scholar grant from the Howard Hughes Medical Institute (to M.H.). S.O. was supported by a long-term postdoctoral fellowship by the Human Frontier Science Program (HFSP). S.R. was supported by a Ph.D. fellowship from the Carl-Zeiss Foundation. **Author contributions:** G.I. led the study. G.I. and N.G. obtained funding for the study. G.I., N.G., A.Sc., M.H., T.W., and A.St. supervised the work. N.M.D., S.R., S.F., S.O., and S.C. carried out the experiments. All authors were involved in the analysis of the results. G.I., A.Sc., and N.M.D. wrote the paper with input from all authors. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** All lines used in the study will be provided upon signature of an appropriate material transfer agreement. All data are available in the main text or the supplementary materials.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/367/6476/431/suppl/DC1
Materials and Methods
Figs. S1 to S23
References (*24*–*39*)

View/request a protocol for this paper from *Bio-protocol*.

6 September 2019; accepted 18 December 2019
10.1126/science.aaz4131

LIFE-SPAN EXTENSION

Four glial cells regulate ER stress resistance and longevity via neuropeptide signaling in *C. elegans*

Ashley E. Frakes[1,2], Melissa G. Metcalf[1,2], Sarah U. Tronnes[1,2], Raz Bar-Ziv[1,2], Jenni Durieux[1,2], Holly K. Gildea[1,2], Nazineen Kandahari[1,2], Samira Monshietehadi[1,2], Andrew Dillin[1,2]*

The ability of the nervous system to sense cellular stress and coordinate protein homeostasis is essential for organismal health. Unfortunately, stress responses that mitigate disturbances in proteostasis, such as the unfolded protein response of the endoplasmic reticulum (UPR^{ER}), become defunct with age. In this work, we expressed the constitutively active UPR^{ER} transcription factor, XBP-1s, in a subset of astrocyte-like glia, which extended the life span in *Caenorhabditis elegans*. Glial XBP-1s initiated a robust cell nonautonomous activation of the UPR^{ER} in distal cells and rendered animals more resistant to protein aggregation and chronic ER stress. Mutants deficient in neuropeptide processing and secretion suppressed glial cell nonautonomous induction of the UPR^{ER} and life-span extension. Thus, astrocyte-like glial cells play a role in regulating organismal ER stress resistance and longevity.

During aging, there is an organism-wide loss of protein homeostasis, exacerbated by the inability to mount an effective unfolded protein response of the endoplasmic reticulum (UPR^{ER}), which likely contributes to tissue damage and increased susceptibility to disease (*1*–*3*). The age-dependent decline in the ability to induce the UPR^{ER} can be prevented by the selective overexpression of constitutively active *xbp-1s* in neurons. Neuronal XBP-1s leads to cell nonautonomous activation of the UPR^{ER} in distal intestinal cells, which is sufficient to confer ER stress resistance and prolong life span (*2*). To date, cell nonautonomous stress signaling has been ascribed only to neurons (*2*, *4*–*7*). However, glial cells—the gatekeepers and guardians of the central nervous system—may also play a role in regulating organismal stress resistance and longevity (*8*).

To determine whether glia play a role in regulating protein homeostasis and longevity, we generated *Caenorhabditis elegans* strains overexpressing *xbp-1s* under a glial-specific promoter, *ptr-10*, which is expressed in most glia except amphid sheath glia (fig. S1A) (*9*, *10*). Animals expressing *xbp-1s* in most glia (*ptr-10p::xbp-1s*) exhibited a marked increase in survival compared with control (N2) animals (Fig. 1A). To identify which glial cells were mediating XBP-1s–dependent longevity, we expressed *xbp-1s* within select subtypes of the 56 *C. elegans* glial cells (*11*). Expression of *xbp-1s* specifically in two amphid and two phasmid sheath glia (AMsh and PHsh) using the *fig-1* promoter did not extend life span beyond that of control animals (Fig. 1B and fig. S1B) (*12*). However, expression of *xbp-1s* in the four cephalic astrocyte-like sheath glia (CEPsh) using the *hlh-17* promoter resulted in an extension of life span (Fig. 1C and fig. S1C) (*13*, *14*).

We hypothesized that glial *xbp-1s* was inducing a beneficial UPR^{ER}, leading to life-span extension. To characterize the localization and extent of UPR^{ER} activation, we generated animals expressing *xbp-1s* in glia with the UPR^{ER} reporter strain, *hsp-4p::GFP* (*15*). At day 1 of adulthood, expression of *xbp-1s* in most glia (*ptr-10p::xbp1s)* or all glia (*mir-228p::xbp-1s*) induced *hsp-4::GFP* in glial cells and in the distal intestine (fig. S2, A to C). Animals overexpressing *xbp-1s* in AMsh and PHsh glia (*fig-1p::xbp1s*) exhibited robust *hsp-4::GFP* induction in *fig-1*–expressing glial cells and in the distal intestine. This expression pattern was distinct from that of *fig-1p::tdTomato* reporter animals, in which tdTomato fluorescence was restricted to AMsh and PHsh glia (fig. S1B and fig. S3, A and B). Animals overexpressing *xbp-1s* in the four CEPsh glia (*hlh-17p::xbp-1s*)

[1]Department of Molecular and Cell Biology, University of California, Berkeley, CA 94720, USA. [2]Howard Hughes Medical Institute, University of California, Berkeley, CA 94720, USA.
*Corresponding author. Email: dillin@berkeley.edu

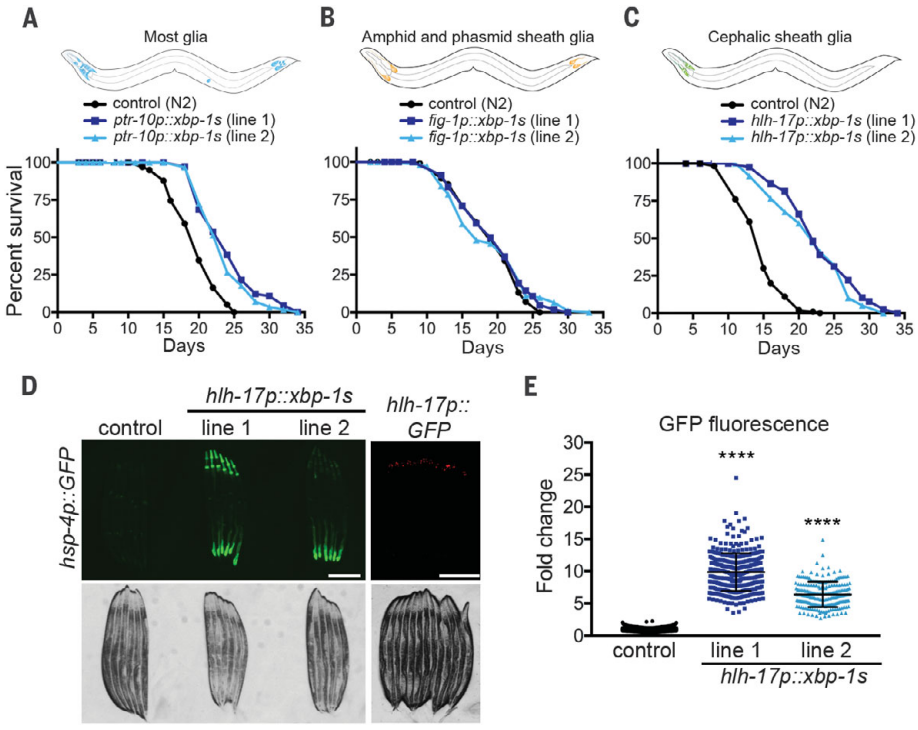


Fig. 1. Glial *xbp-1s* extends life span and induces cell nonautonomous UPR^{ER}. (**A**) Survival of animals expressing *xbp-1s* in most glia [*ptr-10p::xbp-1s*, line 1 (dark blue), line 2 (light blue)] compared with control N2 animals (black). (**B**) Survival of animals expressing *xbp-1s* in four amphid and phasmid sheath glia [*fig-1p::xbp-1s*, line 1 (dark blue), line 2 (light blue)] compared with control N2 animals (black). (**C**) Survival of animals expressing *xbp-1s* in four cephalic sheath glia [*hlh-17p::xbp-1s*, line 1 (dark blue), line 2 (light blue)] compared with control N2 animals (black). (**D** and **E**) Fluorescent micrograph (D) and quantification (E) of UPR^{ER} reporter worms (*hsp-4p::GFP*) expressing *hlh-17p::xbp-1s* (left). *hlh-17p::GFP* reporter worms, pseudo-colored red (right), are shown. Data in (D) are representative of $n > 10$. Scale bars, 250 μm. Quantification of *hsp-4p::GFP* fluorescence using COPAS biosorter was normalized to time of flight (length) and extinction (thickness) of animals. Results are shown relative to *hsp-4p::GFP* alone (control) with error bars representing means ± SD. One-way analysis of variance (ANOVA) Tukey's post hoc test, $n = 2$, ****$P < 0.0001$. Life spans are representative of $n = 3$. See table S1 for life-span statistics.

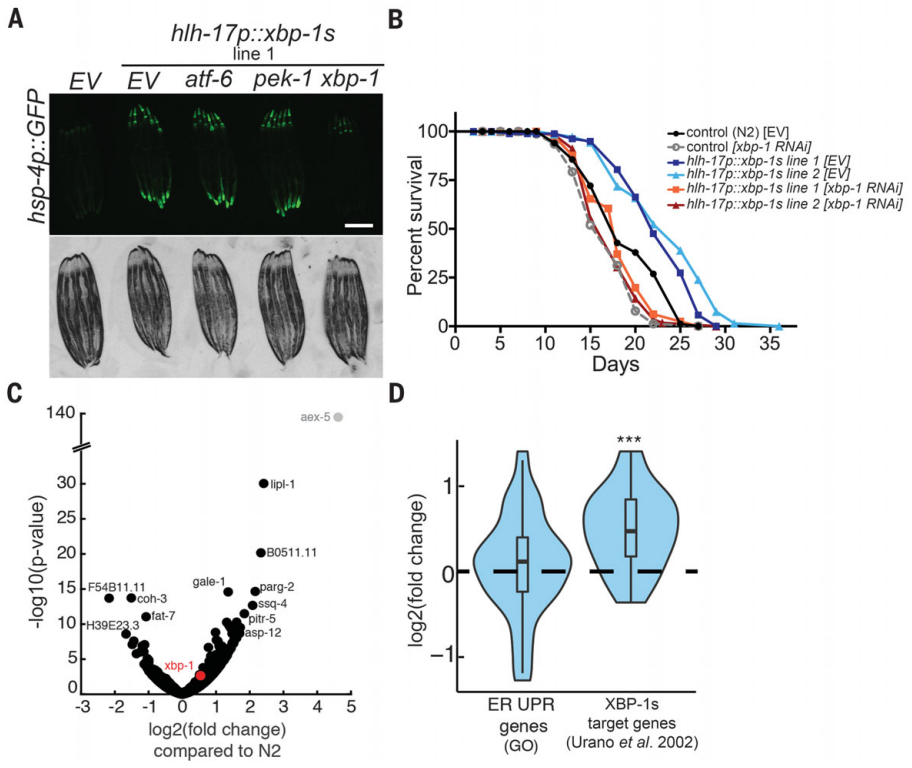


Fig. 2. Cell nonautonomous induction of the UPR^{ER} is dependent on *xbp-1*, but not *atf-6* or *pek-1*. (**A**) Fluorescent micrographs of day 1 *hsp-4p::GFP; hlh-17p::xbp1-s* animals grown on control empty vector (EV), *atf-6*, *pek-1*, or *xbp-1* RNAi from hatch. Scale bar, 250 μm; $n = 3$. (**B**) Survival of control (N2) and *hlh-17p::xbp-1s* animals grown on EV control RNAi or RNAi-targeting *xbp-1*. See table S1 for lifespan statistics; $n = 2$. (**C**) Volcano plot of whole-animal transcriptional profiling from *hlh-17p::xbp-1s* animals compared with wild type (N2). *xbp-1* is highlighted in red. Note that *aex-5* (gray) was detected as highly overexpressed because of a small *aex-5* promoter and exon fragment present in the 3′ untranslated region in the backbone plasmid used for all constructs. All *aex-5* reads aligned to this short fragment. (**D**) The UPR^{ER} is activated in animals expressing *xbp-1s* in CEPsh glia compared with N2, shown by fold change of two gene groups: UPR^{ER} (GO: 0030968 and 1900103) or *xbp-1* targets (*19*). The line inside the box represents the median change of the gene group. ***$P < 0.001$. GO enrichment analysis for genes with a fold change P value <0.05 for terms with a false discovery rate Q value <0.05 can be found in table S3.

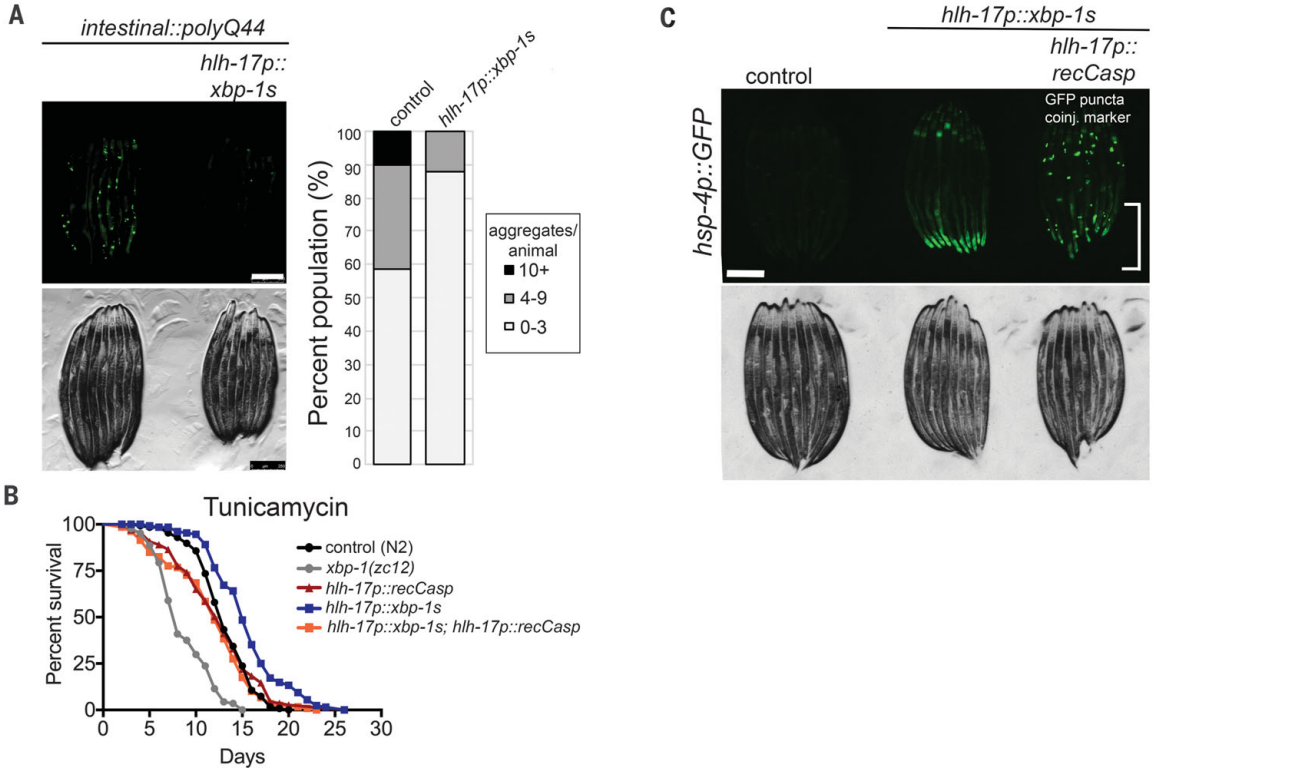


Fig. 3. Expression of *xbp-1s* in glial cells protects animals against protein aggregation and chronic ER stress. (**A**) Fluorescent micrograph and quantification of age-dependent accumulation of polyQ44-YFP aggregates in control animals or animals expressing *hlh-17p::xbp-1s*. Control animals average 3.5 puncta per animal, compared with 1.3 in *hlh-17p::xbp-1s* animals ($P < 0.0001$). Scale bar, 250 μm; $n = 2$. (**B**) Survival of animals transferred to tunicamycin-containing plates at day 1 of adulthood. CEPsh glial ablation via *hlh-17p::recCasp* suppresses *hlh-17p::xbp-1s* ER stress resistance. $n = 2$. (**C**) Fluorescent micrograph of *hsp-4p::GFP* reporter worms expressing *hlh-17p::xbp-1s* and *hlh-17p::xbp-1s; hlh-17p::recCasp*. GFP puncta in *hlh-17p::recCasp* strain represent co-injection (coinj.) marker for *hlh-17p::recCasp* transgene, which is expressed in coelomocytes. White bracket marks distal intestine, where induction of cell nonautonomous UPR^{ER} is reduced in animals expressing *hlh-17p::recCasp*. Scale bar, 250 μm; $n = 3$.

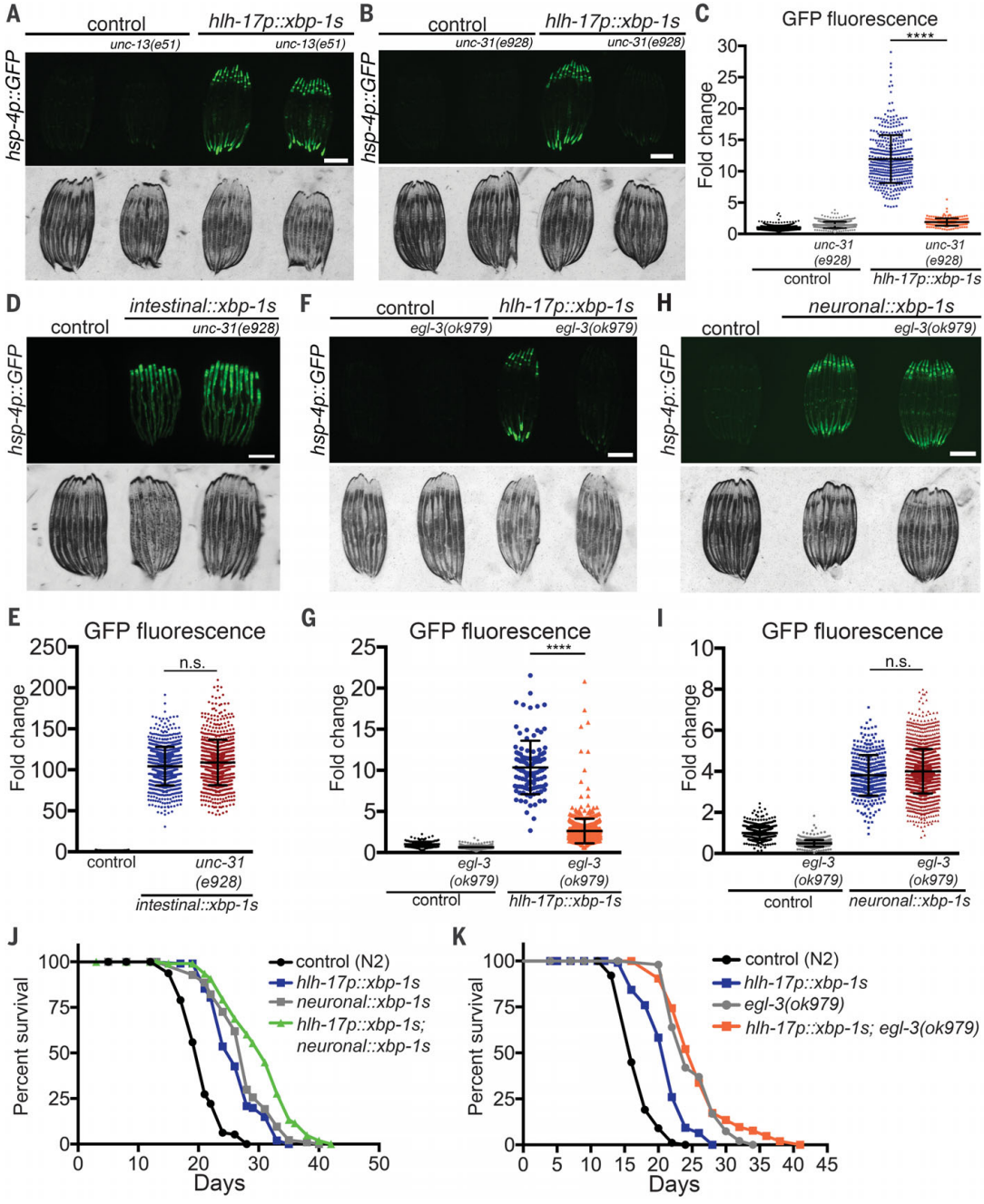


Fig. 4. Neuropeptides are required for glial cell nonautonomous activation of the UPR^{ER} and longevity. (**A** and **B**) Fluorescent micrographs of control (*hsp-4p::GFP*) and *hsp-4p::GFP*; *hlh-17p::xpb-1s* (line 1) animals, with and without the *unc-13(e51)* or *unc-31(e928)* loss-of-function mutations, which render animals deficient in SCV or DCV release, respectively. Scale bars, 250 μm; *n* = 3. (**C**) COPAS quantification of animals in (B), *n* = 3. (**D** and **E**) Fluorescent micrographs (D) and COPAS quantification (E) of control and *hsp-4p::GFP* animals expressing intestinal *xbp-1s* (*vha-6p::xbp-1s*), with and without the *unc-31(e928)* mutation. Scale bar, 250 μm; *n* = 2. n.s., not significant. (**F** and **G**) Fluorescent micrographs (F) and COPAS quantification (G) of control and *hsp-4p::GFP; hlh-17p::xpb-1s* (line 1), with and without the *egl-3(ok979)* mutation, which renders animals unable to cleave pro-neuropeptides. Scale bar, 250 μm; *n* = 3. (**H** and **I**) Fluorescent micrographs (H) and COPAS quantification (I) of control and *hsp-4p::GFP* animals expressing *xbp-1s* in all neurons (*rgef-1p::xbp-1s*), with and without the *egl-3(ok979)* mutation. Scale bar, 250 μm; *n* = 2. (**J**) Survival of control (N2) animals (black), *hlh-17p::xbp-1s* (dark blue), *neuronal(rgef-1p)::xbp-1s* (gray), and *hlh-17p::xbp-1s; neuronal(rgef-1p)::xbp-1s* (green). *n* = 2. (**K**) Survival of control (N2) animals (black), *hlh-17p::xbp-1s* (dark blue), *egl-3(ok979)* (gray), and *hlh-17p::xbp-1s; egl-3(ok979)* (orange). *n* = 3. See table S1 for life-span statistics. COPAS results are shown relative to *hsp-4p::GFP* alone (control), with means ± SD. One-way ANOVA Tukey's post hoc test, ****$P < 0.0001$.

showed induction of *hsp-4p::GFP* in the CEPsh glia and in the distal intestine and pharynx (Fig. 1, D and E). Notably, green fluorescent protein (GFP) expression was limited to CEPsh glial cells in *hlh-17p::GFP* reporter animals (fig. S1C) (*9*, *10*, *13*, *14*, *16*–*18*). These data suggest that *xbp-1s* expression in glial cells can induce cell nonautonomous UPR^{ER} in distal intestinal cells and that CEPsh glia have a unique role in regulating *xbp-1s*–mediated longevity.

To elucidate how CEPsh glia promote longevity via *xbp-1s*, we first tested whether life-span extension and cell nonautonomous activation of the UPR^{ER} from CEPsh glia was dependent on the known signaling components of the UPR^{ER} branches, PERK, ATF6, and XBP1, encoded by *pek-1*, *atf-6*, and *xbp-1*, respectively, in *C. elegans*. No difference was observed in *hsp-4::GFP* induction with *pek-1* or *atf-6* RNA interference (RNAi)–mediated knockdown in *hlh-17p::xbp-1s* animals (Fig. 2A and fig. S4). However, knockdown of *xbp-1* reduced GFP fluorescence of *hlh-17p::xbp-1s; hsp-4p::GFP* animals and abolished the life-span extension of *hlh-17p::xbp-1s* animals (Fig. 2, A and B, and fig. S4). Whole-worm RNA sequencing (RNA-seq) of *hlh-17p::xbp-1s* animals revealed 115 differentially expressed genes (adjusted *P* value <0.05), including a significant increase in *xbp-1s*–dependent transcripts (Fig. 2, C and D, and table S2) (*19*). Gene ontology analysis showed enrichment of genes involved in the immune response, stress response, and, as expected, response to ER stress (table S3).

We hypothesized that the increased activation of the UPR^{ER} in *hlh-17p::xbp-1s* animals

would render these animals more resistant to age-dependent protein aggregation and chronic ER stress. Expression of *xbp-1s* in CEPsh glia notably reduced aggregation of yellow fluorescent protein (YFP)–tagged, Huntington-like polyglutamine protein in the intestine (with age) compared with controls (Fig. 3A). Furthermore, animals expressing *xbp-1s* in CEPsh glia exhibited an increase in survival when chronically exposed to tunicamycin, a chemical inducer of ER stress (Fig. 3B). Perturbing CEPsh glial development, using a partially penetrant reconstituted caspase (recCasp), abrogated the ER stress resistance of *hlh-17p::xbp-1s* animals grown on tunicamycin-containing plates and decreased the median life span of *hlh-17p::xbp-1s* animals grown on control plates (Fig. 3B and fig. S5, A and B). Moreover, distal UPR^{ER} was reduced in *hlh-17p::xbp-1s* animals harboring *hlh-17::recCasp* (Fig. 3C). Consistent with these findings, *hsp-4p::GFP* induction was suppressed in *hlh-17p::xbp-1s* animals harboring a loss-of-function mutation in *vab-3*, a *Pax6/7*-related gene required for CEPsh glial cell development (fig. S5, C and D) (*10*). In contrast to other model organisms, ablation of glial cells does not lead to neuronal cell death in *C. elegans* (*20*).

Next, we assessed whether overexpression of *xbp-1s* in CEPsh glia induces other stress responses known to affect protein homeostasis and longevity, such as the mitochondrial UPR (UPR^{MT}), the heat shock response (HSR), or reduced insulin and insulin-like growth factor 1 (IGF-1) signaling (*5*, *21*, *22*). We did not observe induction of the UPR^{MT} reporter, *hsp-6::GFP*, the HSR reporter, *hsp-16.2::GFP*, or the *sod-3p::GFP* reporter with *hlh-17p::xbp1s* expression. However, *hlh-17p::xbp-1s* animals were still capable of activating these responses (fig. S6). Taken together, these data indicate that expression of *xbp-1s* in CEPsh glia specifically induces the UPR^{ER}, which protects animals from age-dependent protein aggregation and chronic ER stress.

Previously, our laboratory had found that cell nonautonomous activation of the UPR^{ER} by neuronal *xbp-1s* is dependent on the release of small clear synaptic vesicles (SCVs) containing neurotransmitters (*2*). To determine if glial *xbp-1s* signals through a mechanism similar to that of neuronal *xbp-1s*, we generated *hlh-17p::xbp-1s; hsp-4p::GFP* animals containing an *unc-13* mutation, which are deficient in SCV exocytosis (*23*). Notably, cell nonautonomous signaling remained intact in *hlh-17p::xbp-1s* animals harboring either *unc-13(e51)* or *unc-13(s69)* mutations (Fig. 4A and fig. S7, A to D). Therefore, glia do not transmit UPR^{ER} to distal tissues via a SCV-dependent mechanism like neurons.

CEPsh glia reside nearly 300 μm from where we observed robust distal activation of the UPR^{ER}. Therefore, we hypothesized that this transcellular signaling mechanism is dependent on neuropeptides, which are packaged into dense core vesicles (DCVs); can be secreted from neurons, glia, or neuroendocrine cells; and can function as long-range signaling hormones. We crossed *hlh-17p::xbp-1s* animals with an *unc-31* loss-of-function mutant in which DCV exocytosis is disrupted. The *unc-31(e928)* mutation suppressed cell nonautonomous activation of the UPR^{ER}, with GFP fluorescence nearly equal to levels observed in *hsp-4::GFP* controls (Fig. 4, B and C, and fig. S7, E and F) (*24*). The *unc-31(e928)* mutation had no effect on cell autonomous activation of the UPR^{ER} in intestinal cells or neuronal cell nonautonomous activation of the UPR^{ER} (Fig. 4, D and E) (*2*). Furthermore, we tested a loss-of-function mutation in the proprotein convertase, *egl-3*, which is deficient in neuropeptide processing, and found that induction of the cell nonautonomous UPR^{ER} by CEPsh glia was suppressed (Fig. 4, F and G, and fig. S8, A and B) (*25*). Blocking neuropeptide processing had no effect on cell autonomous *hsp-4p::GFP* induction in intestinal cells or cell nonautonomous activation of the UPR^{ER} in animals expressing neuronal *xbp-1s* (fig. S9, A to D, and Fig. 4, H and I). Thus, glial-mediated cell nonautonomous induction of the UPR^{ER} is dependent on neuropeptides, which is an entirely distinct mechanism to that initiated by neurons expressing *xbp-1s*.

As an additional measure of the separation between neuronal and glial induction of peripheral UPR^{ER}, we removed CEPsh glial cells in animals expressing *xbp-1s* solely in neurons, and cell nonautonomous activation of the UPR^{ER} remained intact (fig. S10). Thus, neuronal activation of the peripheral UPR^{ER} via *xbp-1s* is independent of CEPsh glia. Next, we investigated whether combinatorial *xbp-1s* overexpression in both neurons and CEPsh glia would result in an additive increase in activation of the UPR^{ER} and life-span extension. Animals overexpressing *xbp-1s* in both neurons and CEPsh glia induced *hsp-4p::GFP* and extended life span to a greater degree than animals expressing *xbp-1s* only within CEPsh glia or neurons (fig. S11, A and B, and Fig. 4J).

To identify the cell type responsible for secreting the peptides mediating cell nonautonomous UPR^{ER}, we expressed wild-type *unc-31(cDNA)* in either neurons or glia in *hlh-17p::xbp-1s; unc-31(e928)* animals. Neuronal *unc-31(cDNA)* did not restore activation of the UPR^{ER} in the intestine of *hlh-17p::xbp-1s; unc-31(e928)* animals (fig. S12, A and B). In contrast, expression of *unc-31(cDNA)* in CEPsh glia or *egl-3(cDNA)* in CEPsh glia or all glia led to an increase in activation of the UPR^{ER}, albeit a modest increase (fig. S12, C and D). These data suggest that the neuropeptides required for glial-mediated cell nonautonomous activation of the UPR^{ER} do not originate from neurons but are secreted, in part, by glial cells themselves.

Lastly, we sought to determine whether neuropeptide signaling was mediating longevity in *hlh-17p::xbp-1s* animals. Loss-of-function *egl-3* mutants are inherently long-lived because of reduced insulin and IGF-1 signaling (*26*). However, we did not observe an additive increase in survival of *hlh-17p::;xbp-1s* animals harboring the *egl-3(ok979)* mutation, suggesting that life-span extension of *hlh-17p::xbp-1s* animals requires neuropeptides (Fig. 4K).

Previously, cell nonautonomous stress signaling from the brain to the periphery has been ascribed only to neurons. However, our data identify a subtype of astrocyte-like glial cells that coordinate systemic protein homeostasis and aging via neuropeptide signaling—a distinct mechanism from that initiated by neuronal XBP-1s (fig. S13). This suggests there is regional and functional specificity of glial cells to control physiology and aging that evolved as early as the nematode. We speculate that, depending on the physiological cue received by the nervous system, either neurons or glia can signal via XBP-1s to peripheral tissues to coordinate organismal protein homeostasis.

REFERENCES AND NOTES

1. G. Martínez, C. Duran-Aniotz, F. Cabral-Miranda, J. P. Vivar, C. Hetz, *Aging Cell* **16**, 615–623 (2017).
2. R. C. Taylor, A. Dillin, *Cell* **153**, 1435–1447 (2013).
3. A. E. Frakes, A. Dillin, *Mol. Cell* **66**, 761–771 (2017).
4. V. Prahlad, T. Cornelius, R. I. Morimoto, *Science* **320**, 811–814 (2008).
5. J. Durieux, S. Wolff, A. Dillin, *Cell* **144**, 79–91 (2011).
6. P. M. Douglas *et al*., *Cell Reports* **12**, 1196–1204 (2015).
7. Q. Zhang *et al*., *Cell* **174**, 870–883.e17 (2018).
8. B. A. Barres, *Neuron* **60**, 430–440 (2008).
9. G. Rapti, C. Li, A. Shan, Y. Lu, S. Shaham, *Nat. Neurosci.* **20**, 1350–1360 (2017).
10. S. Yoshimura, J. I. Murray, Y. Lu, R. H. Waterston, S. Shaham, *Development* **135**, 2263–2275 (2008).
11. A. Singhvi, S. Shaham, *Annu. Rev. Neurosci.* **42**, 149–168 (2019).
12. T. Bacaj, M. Tevlin, Y. Lu, S. Shaham, *Science* **322**, 744–747 (2008).
13. T. L. McMiller, C. M. Johnson, *Gene* **356**, 1–10 (2005).
14. D. A. Colón-Ramos, M. A. Margeta, K. Shen, *Science* **318**, 103–106 (2007).
15. M. Calfon *et al*., *Nature* **415**, 92–96 (2002).
16. M. Katz *et al*., *Nat. Commun.* **10**, 1882 (2019).
17. M. Katz, F. Corson, S. Iwanir, D. Biron, S. Shaham, *Cell Reports* **22**, 2575–2583 (2018).
18. Z. Shao, S. Watanabe, R. Christensen, E. M. Jorgensen, D. A. Colón-Ramos, *Cell* **154**, 337–350 (2013).
19. F. Urano *et al*., *J. Cell Biol.* **158**, 639–646 (2002).
20. S. Shaham, *Cold Spring Harb. Perspect. Biol.* **7**, a020578 (2015).
21. J. F. Morley, R. I. Morimoto, *Mol. Biol. Cell* **15**, 657–664 (2004).
22. N. Libina, J. R. Berman, C. Kenyon, *Cell* **115**, 489–502 (2003).
23. J. E. Richmond, W. S. Davis, E. M. Jorgensen, *Nat. Neurosci.* **2**, 959–964 (1999).
24. S. Speese *et al*., *J. Neurosci.* **27**, 6150–6162 (2007).
25. S. J. Husson, E. Clynen, G. Baggerman, T. Janssen, L. Schoofs, *J. Neurochem.* **98**, 1999–2012 (2006).
26. B. Hamilton *et al*., *Genes Dev.* **19**, 1544–1555 (2005).

ACKNOWLEDGMENTS

We thank the Shaham laboratory and the Caenorhabditis Genetics Center (CGC) (P40 OD010440) for several strains used in this study and the Garrison laboratory for several plasmids. Thank you to the Dillin laboratory for comments and discussion throughout the project, R. Higuchi-Sanabria for preparing the samples for

RNA-seq, and L. Joe for preparing the RNA-seq libraries. We thank P. Douglas for thoughtful input and discussion at the beginning of this work. **Funding:** A.E.F. is supported by NIH (F32 AG051355); M.G.M. is supported by NIH (F31 AG060660); R.B.-Z. is supported by the EMBO Long-Term Fellowship (462-2017); H.K.G. is supported by NSF (DGE1752814); and A.D. is supported by the Thomas and Stacey Siebel Foundation, the Howard Hughes Medical Institute, and NIH grants AG042679 and AG059566.
Author contributions: A.E.F. conceived the study, generated *C. elegans* strains, performed experiments (life spans, microscopy, COPAS biosorting, and data analysis), and wrote the manuscript. M.G.M. performed life spans, microscopy, tunicamycin ER stress assays, and worm crosses and provided intellectual input. S.U.T. performed life spans, worm crosses, and COPAS biosorting and prepared artwork for Fig. 1 and fig. S13. R.B.-Z. analyzed RNA-seq data, generated figures, and provided intellectual input. J.D. helped generate and analyze strains, prepared artwork in fig. S7, and provided intellectual input. H.K.G. performed backcrosses and promoter characterization and provided intellectual input. N.K. assisted in generating strains and performed crosses and life spans. S.M. performed life spans and assisted with crosses and preparation of samples for RNA-seq. A.D. provided invaluable feedback throughout the project and toward the manuscript. All authors reviewed and edited the manuscript. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper or the supplementary materials. The RNA-seq dataset supporting the conclusions of this article is available in the National Center for Biotechnology Information Sequence Read Archive repository, under accession number BioProject PRJNA589459. Further information and requests for reagents may be directed to dillinlabmaterials@berkeley.edu and will be fulfilled by A.D.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/367/6476/436/suppl/DC1
Materials and Methods
Figs. S1 to S13
Tables S1 to S3
References (*27*–*33*)

View/request a protocol for this paper from *Bio-protocol*.

30 September 2019; accepted 29 November 2019
10.1126/science.aaz6896

NEUROSCIENCE

A common hub for sleep and motor control in the substantia nigra

Danqian Liu[1], Weifu Li[2], Chenyan Ma[1], Weitong Zheng[1], Yuanyuan Yao[1], Chak Foon Tso[1], Peng Zhong[1], Xi Chen[2], Jun Ho Song[3], Woochul Choi[3], Se-Bum Paik[3], Hua Han[2], Yang Dan[1]*

The arousal state of the brain covaries with the motor state of the animal. How these state changes are coordinated remains unclear. We discovered that sleep–wake brain states and motor behaviors are coregulated by shared neurons in the substantia nigra pars reticulata (SNr). Analysis of mouse home-cage behavior identified four states with different levels of brain arousal and motor activity: locomotion, nonlocomotor movement, quiet wakefulness, and sleep; transitions occurred not randomly but primarily between neighboring states. The glutamic acid decarboxylase 2 but not the parvalbumin subset of SNr γ-aminobutyric acid (GABA)–releasing (GABAergic) neurons was preferentially active in states of low motor activity and arousal. Their activation or inactivation biased the direction of natural behavioral transitions and promoted or suppressed sleep, respectively. These GABAergic neurons integrate wide-ranging inputs and innervate multiple arousal-promoting and motor-control circuits through extensive collateral projections.

Many animals are immobile during sleep and in mammals diminished electromyographic (EMG) activity is a major criterion for identifying sleep (*1*–*4*). Although the correlation between brain state and motor activity is widely observed, the underlying mechanism remains poorly understood except for the dedicated circuit for rapid eye movement (REM) sleep atonia (*5*). One way to coordinate the changes is to use shared control circuits. γ-aminobutyric acid (GABA)–releasing (GABAergic) neurons in the substantia nigra pars reticulata (SNr) play a powerful role in movement suppression (*6*–*8*) and innervate multiple wake-promoting neuronal populations (*9*). Here, we investigated whether they participate in brain-state regulation.

First, we characterized the natural brain states and motor behaviors of the mice in their home cages on the basis of electroencephalographic (EEG), EMG, and video recordings (Fig. 1A). A deep learning algorithm was used for image segmentation and automated tracking of the mouse (Fig. 1B and fig. S1, A to E). Two parameters measuring body movement (translation and total movement) fell into three distinct clusters, corresponding to locomotion (LM), nonlocomotor movement (MV, including eating, grooming, and postural adjustments), and immobility. The immobility cluster was further divided into quiet wakefulness (QW) and sleep (SL, including REM and non-REM sleep) on the basis of EEG and EMG recordings (Fig. 1, C to E). These four states exhibit decreasing levels of motor activity and brain arousal, indicated by the decreasing EMG total power and increasing EEG delta power, respectively (Fig. 1F and fig. S1, F and G). Almost all transitions occurred between adjacent states; for example, direct transitions from LM to SL or from MV to SL were never observed (Fig. 1, G and H).

We next examined the activity of SNr neurons during these behavioral states. Recordings from freely moving mice in their home cages showed that some SNr neurons were preferentially active during states of high motor activity, exhibiting positive correlation coefficients (CCs) between their firing rates and EMG power (referred to as "movement-activated neurons"). Others exhibited the opposite profile ("movement-suppressed neurons"). Neurons with high baseline firing rates are more likely to be movement-activated neurons (fig. S2). Such functional diversity is consistent with previous findings (*10*, *11*), suggesting the existence of different cell types. Single-cell gene-expression analysis showed that parvalbumin (PV, encoded by *Pvalb*) and glutamic acid decarboxylase 2 (GAD2, encoded by *Gad2*) are preferentially expressed in separate SNr GABAergic populations (*12*). Indeed, in $Pvalb^{Cre}$ mice, Cre-inducible adeno-associated virus that expresses channelrhodopsin 2 fused with enhanced yellow fluorescent protein (AAV-DIO-ChR2-eYFP) labeled PV-positive neurons almost exclusively (96.1 ± 1.5% SEM), whereas in $Gad2^{Cre}$ mice it primarily labeled PV-negative neurons (92.3 ± 1.1%, referred to as "GAD2 neurons"; Fig. 2A and fig. S3, A and B). The two neuronal populations exhibited distinct spatial distributions, with PV neurons predominantly located in the lateral SNr and GAD2 neurons in the medial SNr (Fig. 2B). This allowed separate tagging of the two populations, which are thought to be involved in sensorimotor versus associative functions (*13*).

We then recorded from ChR2-tagged PV and GAD2 neurons. High-frequency laser pulses (15 and 30 Hz, 10 ms per pulse, 16 pulses per train) were applied intermittently, and single units exhibiting reliable laser-evoked spiking at short latencies were identified as PV or GAD2 neurons in the respective Cre mice (fig. S3, C to F). Many identified PV neurons showed high baseline firing rates and were most active during the LM or MV state (Fig. 2, C, G, and I). Of the 25 identified PV neurons, 22 were movement activated (CC between firing rate and EMG power > 0, $P < 0.05$) and only one was movement suppressed (fig. S3, G to I). Their mean firing rate showed a marked decrease at the MV→QW transition (termination of movement, $P = 0.026$, bootstrap) and an increase at the QW→MV transition (movement initiation, $P = 0.003$, Fig. 2E). By contrast, most of the GAD2 neurons identified in $Gad2^{Cre}$ mice showed low baseline firing rates; they were persistently active during SL and suppressed during periods of motor activity (Fig. 2, D, H, and I, and fig. S3, G to I). Their mean firing rate increased significantly at both movement termination (MV→QW transition, $P = 0.008$) and sleep initiation (QW→non-REM SL transition, $P = 0.018$) and decreased at sleep termination (SL→QW transition, $P = 0.003$, Fig. 2F).

We next tested the functions of GAD2 and PV neurons in regulating motor behaviors and brain states. Bilateral optogenetic activation of GAD2 neurons (2 min per trial applied randomly every 7 to 15 min) caused strong decreases in both LM and MV (Fig. 3, A and B, and movie S1). Compared with laser stimulation in control mice expressing eYFP only (fig. S4A), ChR2-mediated activation of GAD2 neurons significantly decreased movement initiation and increased movement termination ($P < 10^{-7}$, Kolmogorov–Smirnov test; Fig. 3D), consistent with the known function of SNr GABAergic neurons in movement suppression (*6*–*8*). Notably, GAD2 neuron activation also induced a rapid increase in SL ($P < 0.0001$, bootstrap; Fig. 3, A to C, and fig. S4, D and E), primarily by increasing the rate of sleep initiation (Fig. 3D). By contrast, although activation of SNr PV neurons also reduced MV ($P < 0.0001$), it had no effect on SL ($P = 0.81$; Fig. 3, A to C; fig. S4, G to I; and movie S2). Compared with control mice, the main effect of PV neuron activation was to increase movement termination through the MV→QW transition (Fig. 3, D and E).

We also tested the effects of inactivating SNr neurons through a light-activated chloride channel (iC++) (*14*). Inactivation of GAD2 neurons increased both LM and MV and greatly decreased SL (Fig. 3, F to I, and fig. S4, J to L),

[1]Division of Neurobiology, Department of Molecular and Cell Biology, Helen Wills Neuroscience Institute, Howard Hughes Medical Institute, University of California, Berkeley, CA 94720, USA. [2]National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China. [3]Department of Bio and Brain Engineering, Korea Advanced Institute of Science and Technology, Daejeon 34141, Republic of Korea.
*Corresponding author. Email: ydan@berkeley.edu

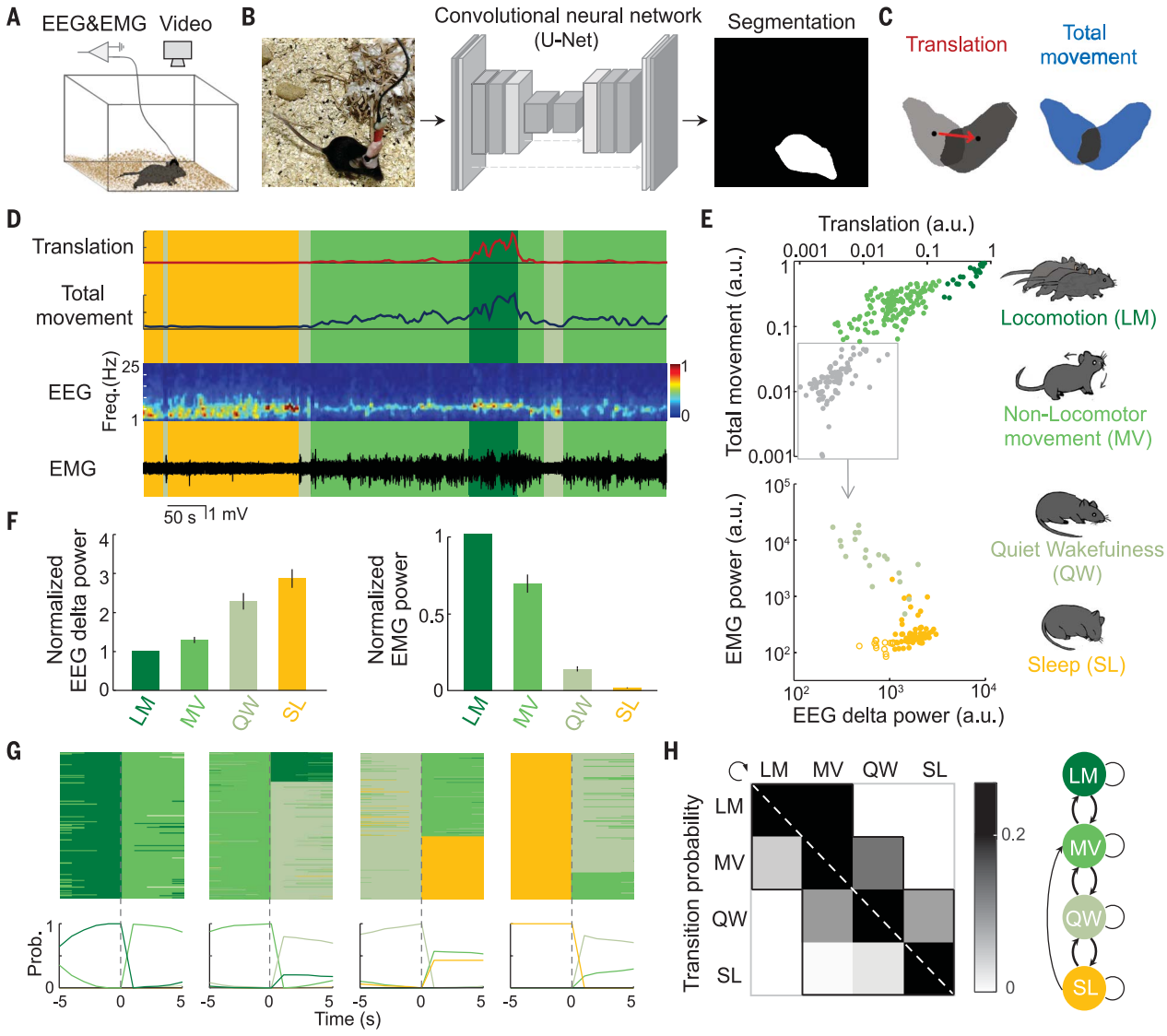


Fig. 1. Automated analysis of mouse home-cage behavior reveals nonrandom state transitions. (**A**) Schematic showing EEG, EMG, and video recordings of freely moving mice in their home cages. (**B**) Automated image segmentation. (**C**) Definition of translation (red arrow) and total movement (total area in blue). Black dots indicate the centroid of the segmented area. (**D**) Example recording showing translation, total movement, EEG spectrogram, and EMG trace. Freq., frequency. (**E**) Top, scatter plot for translation and total movement. Each dot represents data in a 2.5-s bin in (D). The three clusters correspond to LM, MV, and immobile states (gray box). Bottom, scatter plot for EEG delta power (1 to 4 Hz) and EMG total power during the immobile state. The two clusters correspond to QW and SL. Classified states are color-coded. Filled and open yellow, non-REM and REM sleep, respectively. a.u., arbitrary units. (**F**) Mean EEG delta power and EMG power during LM, MV, QW, and SL states (n = 9 mice). EEG delta and EMG power at different states were normalized by EEG delta and EMG power in the LM state, respectively. Error bar indicates ± SEM. (**G**) Top, all transitions from each state. T = 0, time of transition. Bottom, probability (Prob.) of each state before and after transition. (**H**) Left, probability of transition between each pair of states. Right, summary of all natural transitions.

indicating the importance of their endogenous spiking activity in movement suppression and sleep generation. Inactivation of PV neurons also reduced SL, but the effect was much weaker than that of GAD2 neuron inactivation (Fig. 3, F to H, and fig. S4, M to O); a main effect of PV neuron inactivation was to decrease movement termination (Fig. 3, I and J). The different contributions of GAD2 and PV neurons to sleep generation were further confirmed by the effects of chemogenetic activation and inactivation of each population (fig. S5).

Although optogenetic activation of GAD2 neurons caused strong enhancement of movement termination and sleep initiation (Fig. 3D), we observed no direct LM→SL or MV→SL transitions, which were absent in normal home-cage behavior (Fig. 1, G and H). Instead, laser stimulation significantly increased the naturally occurring LM→MV, MV→QW, and QW→SL transitions, all in the direction of decreasing arousal and motor activity; transitions in the opposite direction were strongly suppressed (Fig. 3E). GAD2 neuron activation also increased nest entering, and SL during the activation occurred almost exclusively within the nest, similar to SL without laser stimulation (fig. S4, B and C). Optogenetic inactivation of SNr GAD2 neurons also induced no artificial transition and its effects were caused entirely by enhancing or suppressing the natural transitions in the direction of increasing or decreasing motor activity and EEG activation, respectively (Fig. 3J).

Anterograde tracing of SNr neuron axons expressing ChR2-eYFP showed that PV neurons

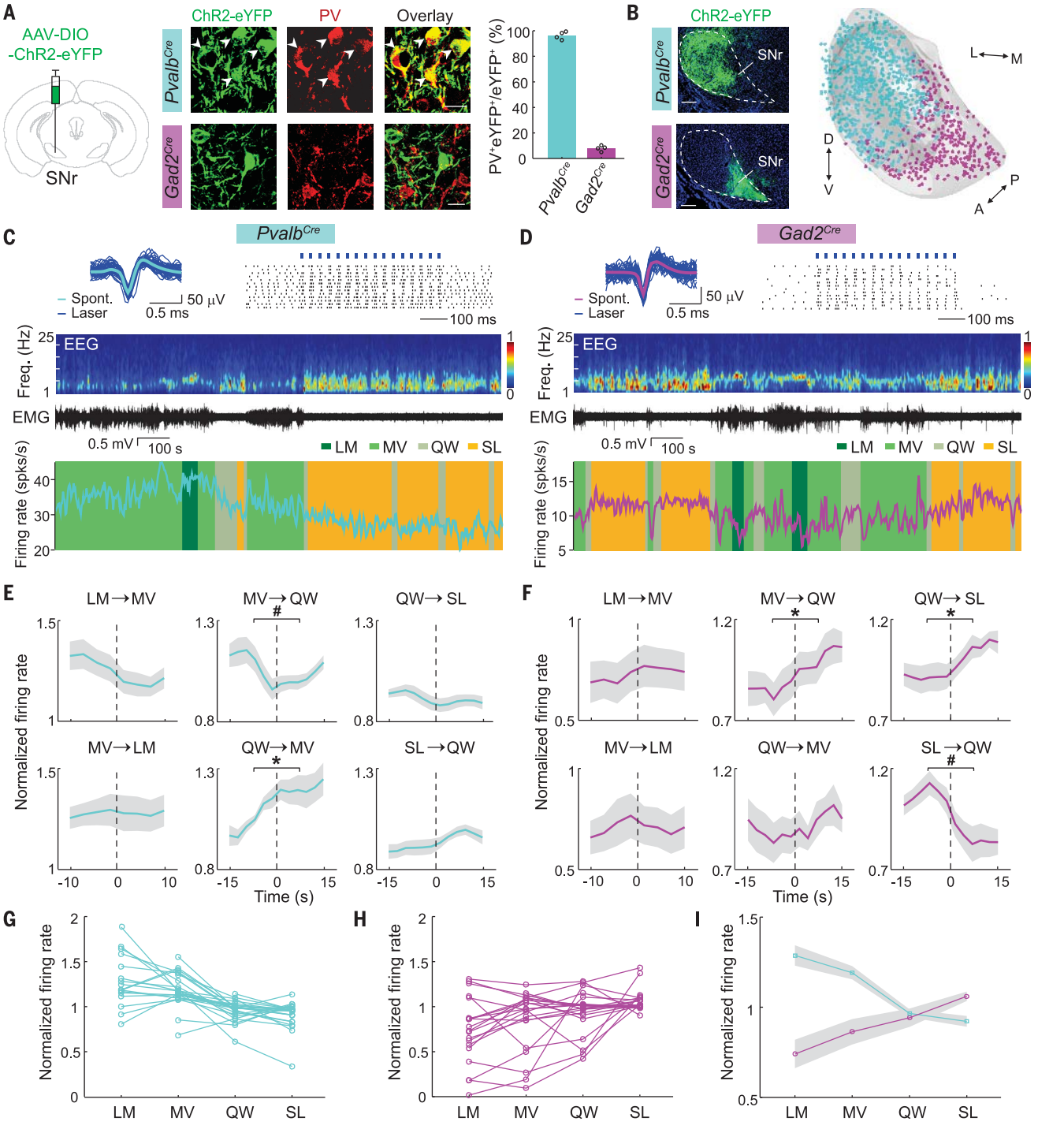


Fig. 2. PV and GAD2 neurons in the SNr show distinct behavioral state–dependent firing rates. (**A**) Left, immunostaining of eYFP and PV in *Pvalb*^{Cre} (top) and *Gad2*^{Cre} (bottom) mice injected with AAV-DIO-ChR2-eYFP. Arrowheads indicate colabeled neurons. Scale bar, 20 μm. Right, percentages of eYFP-labeled neurons that are PV positive (n = 4 mice). (**B**) Fluorescence images (left) and three-dimensional reconstruction (right) of eYFP-labeled PV and GAD2 SNr neurons. Scale bar, 200 μm. (**C**) Example recording from an identified PV neuron. Top left, laser-evoked and spontaneous (Spont.) spike waveforms. Top right, spike raster. Blue ticks, laser pulses (30 Hz). Bottom, firing rate of the PV neuron together with EEG spectrogram, EMG trace, and behavioral states. spks, spikes. (**D**) Similar to (C) but for an identified GAD2 neuron. (**E**) Normalized mean firing rate of all identified PV neurons at behavioral state transitions (n = 25). Vertical line, transition point. Shading indicates ±SEM. *Significant increase at $P < 0.05$, bootstrap. #Significant decrease at $P < 0.05$, bootstrap. (**F**) Similar to (E) but for GAD2 neurons (n = 22). (**G** and **H**) Normalized firing rate for individual neurons during different states. Each line indicates one neuron. For some neurons, no LM was detected during recording. (**I**) Firing rate averaged across all identified PV or GAD2 neurons. Shading indicates ±SEM.

project primarily to brain regions involved in movement control, such as the motor thalamus, motor layers of the superior colliculus (SCm), and the mesencephalic locomotor region (MLR), including the midbrain reticular nucleus (MRN) and the pedunculopontine nucleus (PPN) (*15*) (Fig. 4, A to C; table S1, and movie S3). By contrast, GAD2 neurons also project to several regions involved in brain-state regulation. These include monoaminergic centers such as the dorsal raphe nucleus (DR), locus ceruleus (LC), and ventral tegmental area (VTA; fig. S6, A and B), where SNr neurons directly innervate serotonergic, noradrenergic, and dopaminergic neurons (*9*, *16*, *17*). The

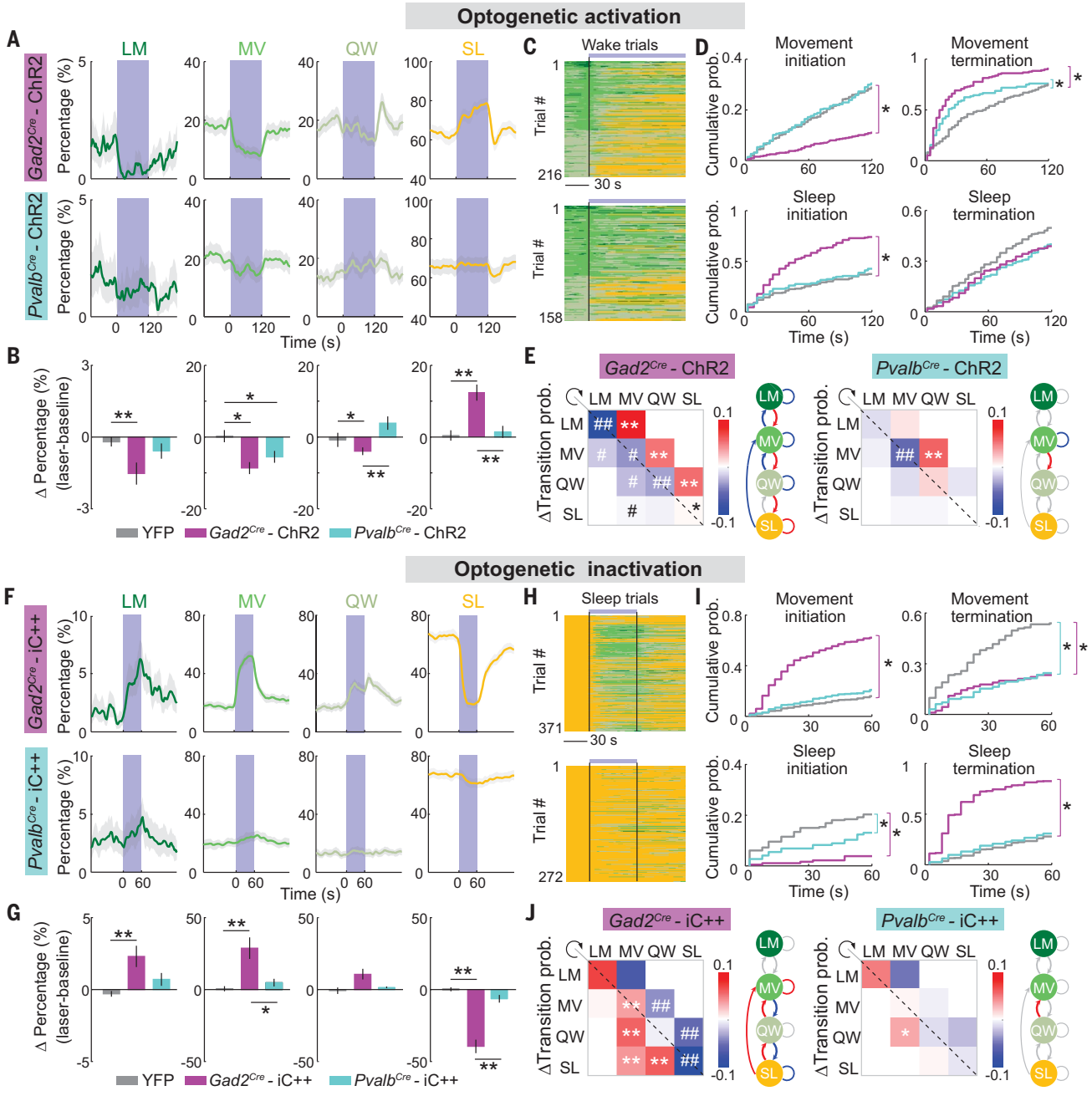


Fig. 3. Causal effects of SNr GAD2 and PV neuron activity on motor behavior and sleep. (**A**) Percentage of time in LM, MV, QW, or SL state before, during, and after laser activation of SNr neurons in *Gad2*^{Cre}-ChR2 mice (top, n = 7 mice, P_{LM}, P_{MV}, P_{QW}, P_{SL} < 0.0001, bootstrap) or *Pvalb*^{Cre}-ChR2 mice (bottom, n = 7, P_{LM} = 0.051, P_{MV} < 0.0001, P_{QW} = 0.0009, P_{SL} = 0.81). Shading indicates 95% confidence intervals. Blue stripe indicates the laser period (constant light, 120 s). (**B**) Laser-induced changes in each state in *Gad2*^{Cre}-ChR2, *Pvalb*^{Cre}-ChR2, or eYFP control mice (n = 6). Error bars indicate SEM. *P < 0.05, **P < 0.01, Mann–Whitney U test. (**C**) Behavioral states in all "wake trials" (mice stayed awake for >30 s before laser onset). (**D**) Cumulative probabilities for movement initiation or termination and sleep initiation or termination during 120-s laser stimulation. *$P < 10^{-7}$ for comparison between ChR2 and eYFP control experiments, Kolmogorov–Smirnov test. (**E**) Laser-induced changes in transition probability. *Increase, P < 0.05, bootstrap. #Decrease, P < 0.05, bootstrap. **Increase, P < 0.0005. ##Decrease, P < 0.0005. (**F**) Similar to (A) but for iC++-mediated inactivation (constant light, 60 s) in *Gad2*^{Cre} mice (top, n = 6, P_{LM}, P_{MV}, P_{QW}, P_{SL} < 0.0001) or *Pvalb*^{Cre} mice (bottom, n = 5, P_{LM} = 0.88, P_{MV} = 0.0034, P_{QW} = 0.88, P_{SL} = 0.0002). (**G**) Similar to (B) but for optogenetic inactivation. (**H**) Similar to (C) but for all sleep trials in inactivation experiments. (**I** and **J**) Similar to (D) and (E) but for optogenetic inactivation.

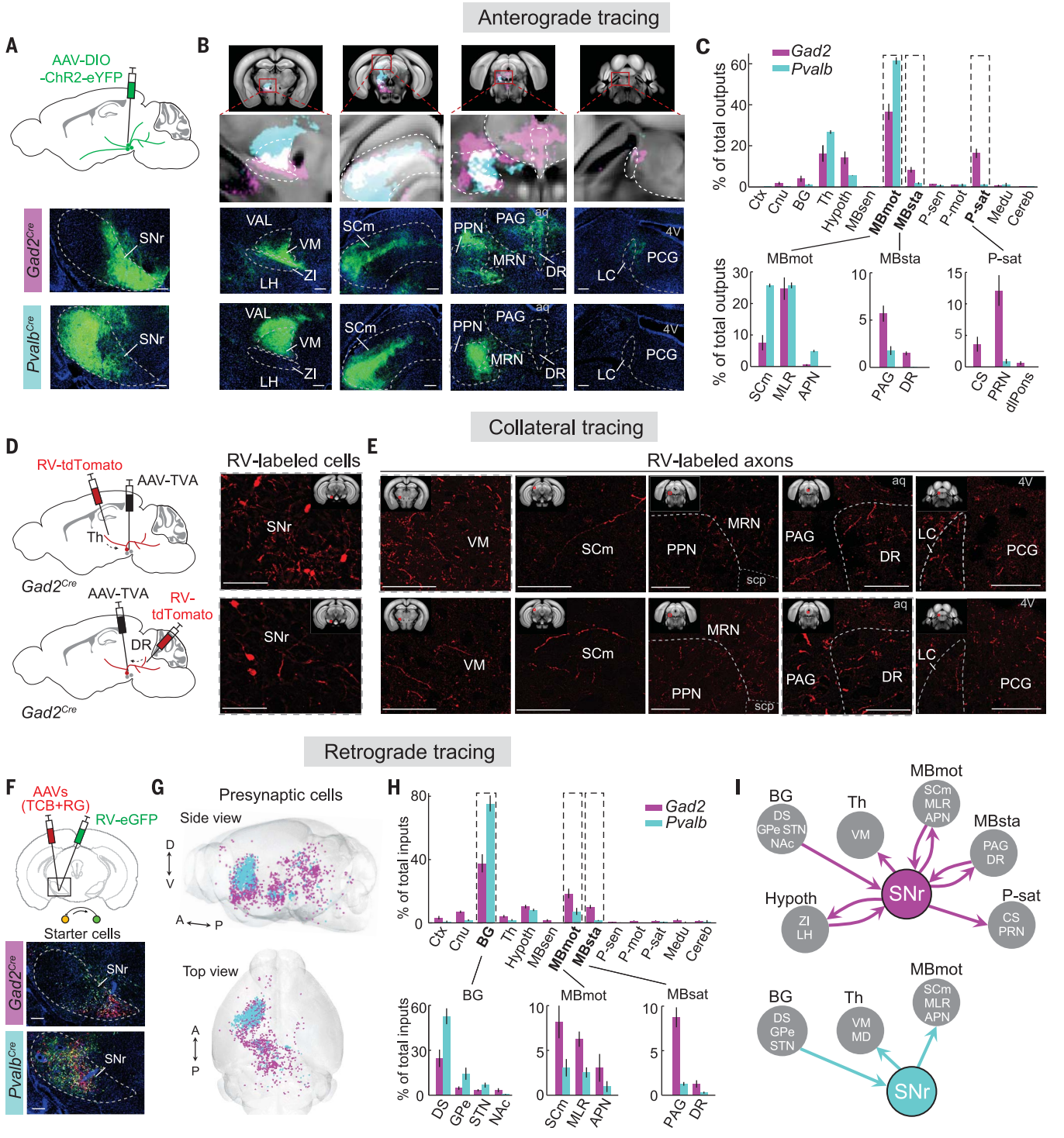


Fig. 4. Whole-brain mapping of outputs and inputs of SNr GAD2 and PV neurons. (**A**) Fluorescence images showing ChR2-eYFP expression in *Gad2*^{Cre} or *Pvalb*^{Cre} mouse SNr. (**B**) Summary (top) and example images showing labeled axons. VM, ventral medial complex of the thalamus; VAL, ventral anterior-lateral complex of the thalamus; ZI, zona incerta; LH, lateral hypothalamic area; PAG, periaqueductal gray; PCG, pontine central gray; aq, cerebral aqueduct; 4V, fourth ventricle. Scale bars, 200 μm for (A) and (B). (**C**) Percentages of labeled axons in 13 main brain regions (top, *n* = 3 mice each) and specific structures (bottom). The MLR includes the MRN and PPN. Abbreviations are defined in table S1. (**D**) Fluorescence images of RV-labeled SNr GAD2 neurons projecting to the thalamus (top) or DR (bottom). (**E**) Axonal collaterals in multiple regions. Scp, superior cerebellar peduncles. Scale bars, 100 μm for (D) and (E). (**F**) Monosynaptic retrograde tracing from GAD2 and PV neurons. Fluorescence images show starter cells (yellow) in the SNr. Scale bars, 200 μm. (**G**) Whole-brain reconstruction of inputs to SNr neurons in a *Gad2*^{Cre} mouse (magenta, *n* = 1200 input cells) or a *Pvalb*^{Cre} mouse (cyan, *n* = 1088). Injection site is excluded. (**H**) Similar to (C) but for input distribution for GAD2 (*n* = 5 mice) and PV neurons (*n* = 5). (**I**) Diagram summarizing major connections for SNr GAD2 and PV neurons.

projections of SNr GAD2 neurons are also different from those of VTA GABAergic neurons, which powerfully promote non-REM sleep but rarely innervate motor thalamus or MLR (*18*, *19*).

The divergent projections of GAD2 neurons could either originate from different SNr subpopulations or represent axon collaterals of the same neurons. To label the axon collaterals of thalamus-projecting GAD2 neurons, we injected a Cre-inducible AAV that expresses avian-specific retroviral receptor, TVA, into the SNr of *Gad2*^{Cre} mice. A modified rabies virus (RV) expressing tdTomato (RV-ΔG-tdTomato+EnvA) was injected 2 weeks later into the thalamus, which allowed RV to enter the TVA-expressing axons, be transported retrogradely to the SNr neurons, and label all of their axon collaterals with tdTomato (Fig. 4D). In addition to the thalamus, we found labeled axons in the SCm, MRN and PPN, DR, LC, and VTA (Fig. 4E and fig. S6, C and D). Similarly, injection of the RV into the DR revealed labeled axons in the thalamus, SCm, MRN and PPN, LC, and VTA. Each SNr GAD2 neuron thus sends axon collaterals to multiple brain regions differentially involved in motor and brain-state control (fig. S7).

Finally, we used RV-mediated transsynaptic tracing to identify monosynaptic inputs to SNr PV and GAD2 neurons. AAVs expressing rabies glycoprotein and TVA fused with mCherry (TCB) were injected into the SNr of each Cre mouse, followed by injection of a modified RV expressing eGFP (RV-ΔG-eGFP+EnvA) 2 weeks later (Fig. 4F). In *Pvalb*^{Cre} mice, the vast majority (75.5 ± 5.0%) of eGFP-labeled input neurons were found within the basal ganglia (Fig. 4, G and H; fig. S8; table S1; and movie S4). By contrast, inputs to GAD2 neurons were much more distributed, with substantial fractions in the hypothalamus and midbrain regions. Thus, whereas PV neurons serve mainly as a basal ganglia output to motor-control regions, GAD2 neurons integrate a much wider range of inputs and project broadly to brain-state as well as motor-control regions (Fig. 4I).

We found that GAD2 but not PV neuron activity promotes sleep generation (primarily non-REM sleep initiation). PV neurons in the lateral SNr fire at higher rates in states of high motor activity and their activation or inactivation increased or decreased movement termination, consistent with a proposed function of the SNr in suppressing unwanted movements during action selection (*20*). By contrast, GAD2 neurons in the medial SNr were preferentially active in states of low motor activity. In addition to motor suppression, their activation powerfully enhanced the transition from QW to SL. This indicates that SNr GAD2 neurons provide general suppression of both motor activity and brain arousal to promote states of quiescence. The involvement of some SNr neurons in sleep regulation is consistent with previous lesion studies in rats (*21*, *22*) and cats (*23*). Within the basal ganglia, activation of adenosine A2A receptor–expressing GABAergic neurons in the striatum or neurotensin-expressing glutamatergic neurons in the subthalamic nucleus also increased sleep (*24*, *25*) (fig. S9). Because these neurons are all part of the basal ganglia indirect pathway (*8*), their sleep-promoting effects are likely mediated, at least in part, by activation of the SNr GAD2 neurons.

Activation of SNr GAD2 neurons suppressed movement and enhanced sleep by biasing the direction of natural state transitions rather than by causing abrupt cessation of all motor activity (behavioral arrest). The LM→MV→QW→SL behavioral sequence promoted by GAD2 neuron activation was characterized by a progressive decrease in motor activity and increase in EEG delta power. Coordination of behavioral and brain-state changes could be mediated by multiple cell types. For example, LC noradrenergic neurons regulate motor activity as well as brain arousal (*26*), and PPN neurons control both locomotion and cortical activation (*15*, *27*–*30*). By innervating these populations through extensive collateral projections while integrating inputs from wide-ranging brain areas, the SNr GAD2 neurons serve as a critical hub in a common circuit for sleep and motor control.

REFERENCES AND NOTES

1. S. S. Campbell, I. Tobler, *Neurosci. Biobehav. Rev.* **8**, 269–300 (1984).
2. J. C. Hendricks *et al*., *Neuron* **25**, 129–138 (2000).
3. P. J. Shaw, C. Cirelli, R. J. Greenspan, G. Tononi, *Science* **287**, 1834–1837 (2000).
4. D. Liu, Y. Dan, *Annu. Rev. Neurosci.* **42**, 27–46 (2019).
5. J. Peever, P. M. Fuller, *Curr. Biol.* **27**, R1237–R1248 (2017).
6. O. Hikosaka, R. H. Wurtz, *J. Neurophysiol.* **53**, 292–308 (1985).
7. A. V. Kravitz *et al*., *Nature* **466**, 622–626 (2010).
8. C. R. Gerfen, D. J. Surmeier, *Annu. Rev. Neurosci.* **34**, 441–466 (2011).
9. C. Ma *et al*., *Neuron* **103**, 323–334.e7 (2019).
10. X. Jin, F. Tecuapetla, R. M. Costa, *Nat. Neurosci.* **17**, 423–430 (2014).
11. G. Rizzi, K. R. Tan, *Cell Rep.* **27**, 2184–2198.e4 (2019).
12. A. Saunders *et al*., *Cell* **174**, 1015–1030.e16 (2018).
13. N. Rajakumar, K. Elisevich, B. A. Flumerfelt, *J. Comp. Neurol.* **350**, 324–336 (1994).
14. A. Berndt *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **113**, 822–829 (2016).
15. T. K. Roseberry *et al*., *Cell* **164**, 526–537 (2016).
16. S. K. Ogawa, J. Y. Cohen, D. Hwang, N. Uchida, M. Watabe-Uchida, *Cell Rep.* **8**, 1105–1118 (2014).
17. B. Weissbourd *et al*., *Neuron* **83**, 645–662 (2014).
18. X. Yu *et al*., *Nat. Neurosci.* **22**, 106–119 (2019).
19. S. Chowdhury *et al*., *eLife* **8**, e44928 (2019).
20. G. Cui *et al*., *Nature* **494**, 238–242 (2013).
21. D. Gerashchenko, C. A. Blanco-Centurion, J. D. Miller, P. J. Shiromani, *Neuroscience* **137**, 29–36 (2006).
22. M. H. Qiu, R. Vetrivelan, P. M. Fuller, J. Lu, *Eur. J. Neurosci.* **31**, 499–507 (2010).
23. Y. Y. Lai *et al*., *Neuroscience* **90**, 469–483 (1999).
24. Y. Oishi *et al*., *Nat. Commun.* **8**, 734 (2017).
25. X. S. Yuan *et al*., *eLife* **6**, e29055 (2017).
26. J. C. Holstege, H. G. Kuypers, *Neuroscience* **23**, 809–821 (1987).
27. A. M. Lee *et al*., *Neuron* **83**, 455–466 (2014).
28. V. Caggiano *et al*., *Nature* **553**, 455–460 (2018).
29. D. Kroeger *et al*., *J. Neurosci.* **37**, 1352–1366 (2017).
30. D. J. Galtieri, C. M. Estep, D. L. Wokosin, S. Traynelis, D. J. Surmeier, *eLife* **6**, e30352 (2017).

ACKNOWLEDGMENTS

We thank F. Weber and M. Xu for data analysis, Q. Xie and H. Ren for deep learning analysis, Y. Zuo and S. Ma for collateral tracing, W. Chang for rabies virus, and J. Ding for helpful comments. **Funding:** This work was supported by the Howard Hughes Medical Institute. **Author contributions:** D.L. and Y.D. designed the experiments and wrote the manuscript; D.L. performed most experiments and data analysis; W.L., X.C., and H.H. performed deep learning analysis; J.H.S., W.C., and S.-B.P. developed the AMaSiNe software; W.Z., C.M., Y.Y., C.F.T., and P.Z. performed histology and FISH experiments; and Y.D. supervised all aspects of the project. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** All data necessary to understand and assess the conclusions of this study are available in the manuscript or the supplementary materials. The viruses used in this work were provided under a materials transfer agreement with Stanford University.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/367/6476/440/suppl/DC1
Materials and Methods
Figs. S1 to S9
Table S1
Captions for Movies S1 to S4
References (*31*–*39*)

View/request a protocol for this paper from *Bio-protocol*.

13 August 2019; accepted 29 November 2019
10.1126/science.aaz0956

CANCER IMMUNOTHERAPY

An RNA vaccine drives expansion and efficacy of claudin-CAR-T cells against solid tumors

Katharina Reinhard[1]*, Benjamin Rengstl[1]*, Petra Oehm[1]*, Kristina Michel[1], Arne Billmeier[1], Nina Hayduk[1], Oliver Klein[1], Kathrin Kuna[1], Yasmina Ouchan[1], Stefan Wöll[1], Elmar Christ[1], David Weber[2], Martin Suchan[2], Thomas Bukur[2], Matthias Birtel[1], Veronika Jahndel[1], Karolina Mroz[1], Kathleen Hobohm[1], Lena Kranz[1], Mustafa Diken[2], Klaus Kühlcke[1], Özlem Türeci[1]†, Ugur Sahin[1,2,3]†‡

Chimeric antigen receptor (CAR)–T cells have shown efficacy in patients with B cell malignancies. Yet, their application for solid tumors has challenges that include limited cancer-specific targets and nonpersistence of adoptively transferred CAR-T cells. Here, we introduce the developmentally regulated tight junction protein claudin 6 (CLDN6) as a CAR target in solid tumors and a strategy to overcome inefficient CAR-T cell stimulation in vivo. We demonstrate that a nanoparticulate RNA vaccine, designed for body-wide delivery of the CAR antigen into lymphoid compartments, stimulates adoptively transferred CAR-T cells. Presentation of the natively folded target on resident antigen-presenting cells promotes cognate and selective expansion of CAR-T cells. Improved engraftment of CAR-T cells and regression of large tumors in difficult-to-treat mouse models was achieved at subtherapeutic CAR-T cell doses.

Adoptive cell therapy (ACT) with genetically engineered T lymphocytes expressing chimeric antigen receptors (CARs) has been clinically successful in patients with B cell malignancies (*1*, *2*). However, in patients with solid tumors, the efficacy of CAR-T cell therapy is challenging and much less effective (*3*). One key hurdle is the limited number of cell-surface targets with high cancer-specific expression to allow for efficient tumor eradication and low risk of off-tumor on-target toxicity (*4*–*6*). We and others have recently reported cancer-associated expression of claudin 6 (CLDN6), a tetraspanin membrane protein that is involved in tight junction formation (*7*). To evaluate the suitability of CLDN6 as a target for CAR-T cell therapy, we profiled its expression in a comprehensive set of human and mouse tissues. In mice, CLDN6 has been reported to be developmentally regulated (*8*). By immunohistochemical (IHC) staining, we found CLDN6 to be broadly expressed in fetal organs but prenatally down-regulated, resulting in lack of expression in most organs of adult mice (fig. S1A). In humans, CLDN6 transcript levels were high in fetal tissues derived from stomach, pancreas, lung, and kidney but undetectable in the corresponding adult tissue samples (fig. S1B). In more than 160 noncancerous healthy human samples from more than 50 adult tissue types analyzed by quantitative real-time–polymerase chain reaction (qRT-PCR), significant CLDN6 transcript expression was ruled out (Fig. 1A and fig. S2B). In addition, CLDN6 protein was not detectable in any of the adult human normal tissue types (>40 tested) assessed by IHC staining (Fig. 1B). In line with previous studies (*9*, *10*), high CLDN6 transcript levels were frequent in various human solid cancers such as testicular, ovarian, uterine, and lung adenocarcinoma (Fig. 1A and fig. S2, A to C). IHC staining showed membrane expression of CLDN6 proteins in these human cancers that was high and homogeneous in many of the tested specimens (fig. S2C). These findings indicate exquisitely tight and complete silencing of CLDN6 in normal human tissues and suggest that CLDN6 is a strictly oncofetal cell-surface antigen with an ideal expression profile for CAR-T cell targeting (*11*).

We designed a second-generation CLDN6-CAR with a 4-1BB costimulatory domain. For the receptor domain, we engineered a single-chain variable fragment (scFv) with exquisite specificity and high binding affinity to CLDN6 in the nanomolar range (Fig. 1C). First, we characterized CLDN6-CAR–engineered human T cells in vitro. CLDN6^{neg} human COLO-699N lung carcinoma cells were transfected with increasing amounts of CLDN6 RNA and assessed for killing by CAR-T cells (Fig. 1D). We observed highly sensitive recognition and lysis of CLDN6-transfected target cells by the CLDN6-CAR, even at the lowest target expression level.

In a similar experimental setting, we evaluated the CLDN6-CAR for cross-recognition of CLDN3, CLDN4, and CLDN9, the most closely related claudin family members that, in contrast to CLDN6, are expressed in toxicity-relevant normal tissues. The homology between the CAR-targeted first extracellular loop of CLDN6 and the corresponding amino acid sequences of these claudins is 81, 85, and 98%, respectively, bearing the risk of cross-reactivity and off-target toxicity of the CAR. We found that the CLDN6-transfected target cells were killed but not those transfected with the related claudins, demonstrating precise targeting by CLDN6-CAR-T cells (Fig. 1E).

To measure cognate immune activation, we cocultured CLDN6-CAR-T cells with human tumor cell lines. We found interferon-γ (IFN-γ) secretion and up-regulation of T cell activation markers upon coculture with CLDN6^{pos} targets but not CLDN6^{neg} cells (Fig. 1F). CLDN6-CAR-T cells were able to efficiently clear CLDN6^{pos} PA-1 ovarian carcinoma spheroids and to kill repetitively upon rechallenge (Fig. 1G). Deletion of CLDN6 by CRISPR-Cas9–mediated genetic knockout (Fig. 1G, top) completely abrogated CAR-T cell recognition of PA-1, further confirming high potency and target-specificity of CLDN6-CAR-T cells.

Next, we studied in vivo antitumor activity of human CLDN6-CAR-T cells in mice xenografted subcutaneously with a human tumor cell line. Of note, the mouse is not a suitable species for studying toxicity of this CAR because the binding affinity of CLDN6-CAR to the mouse CLDN6 ortholog is 15-fold lower than to human CLDN6 and, whereas human CLDN6 is strictly confined to the embryonic stage, murine CLDN6 is expressed in some postembryonic somatic tissues. Immunodeficient NOD-*scid* IL2Rg^{null} (NSG) mice with large ovarian OV90 tumors (mean volume 168 mm^3) underwent ACT with a single dose of human CLDN6-CAR-T cells or control cells. Notably, all CLDN6-CAR-T cell–treated mice experienced complete tumor regression within 2 weeks, compared with control group mice with tumors that progressed rapidly (Fig. 1H). Circulating CLDN6-CAR-T cells were detectable in cured mice for the full observation period of up to 25 days after ACT (fig. S3).

Engraftment and persistence of transferred CAR-T cells are known to be critical for their clinical effect (*12*–*14*). In hematological malignancies, CAR-T cells are directed against lineage antigens of B cells and encounter their targets on the host's normal and malignant B cells. These act as antigen-presenting cells (APCs) that provide strong proliferation signals and promote persistence of CAR-T cells (*13*, *14*).

However, in the solid-tumor setting, the frequency of CAR-T cells typically declines rapidly (*15*–*17*) owing to the impaired accessibility of CAR-T cells to tumor cells within solid lesions and the absence of proliferation signals when CAR-T cells encounter the target in an immunosuppressive tumor microenvironment. We hypothesized that expression of the CAR target in its native conformation on the surface of professional APCs in lymphoid tissues would render it accessible for cognate CAR-T cell

[1]Biopharmaceutical New Technologies (BioNTech) Corporation, BioNTech Cell & Gene Therapies GmbH, BioNTech Innovative Manufacturing Services GmbH, An der Goldgrube 12, 55131 Mainz, Germany. [2]TRON–Translational Oncology at the University Medical Center of Johannes Gutenberg University gGmbH, Freiligrathstr. 12, 55131 Mainz, Germany. [3]Helmholtz Institute for Translational Oncology Mainz, HI-TRON Mainz, Obere Zahlbacher Str. 63, 55131 Mainz, Germany.
*These authors contributed equally to this work. †These authors contributed equally to this work.
‡Corresponding author. Email: ugur.sahin@biontech.de

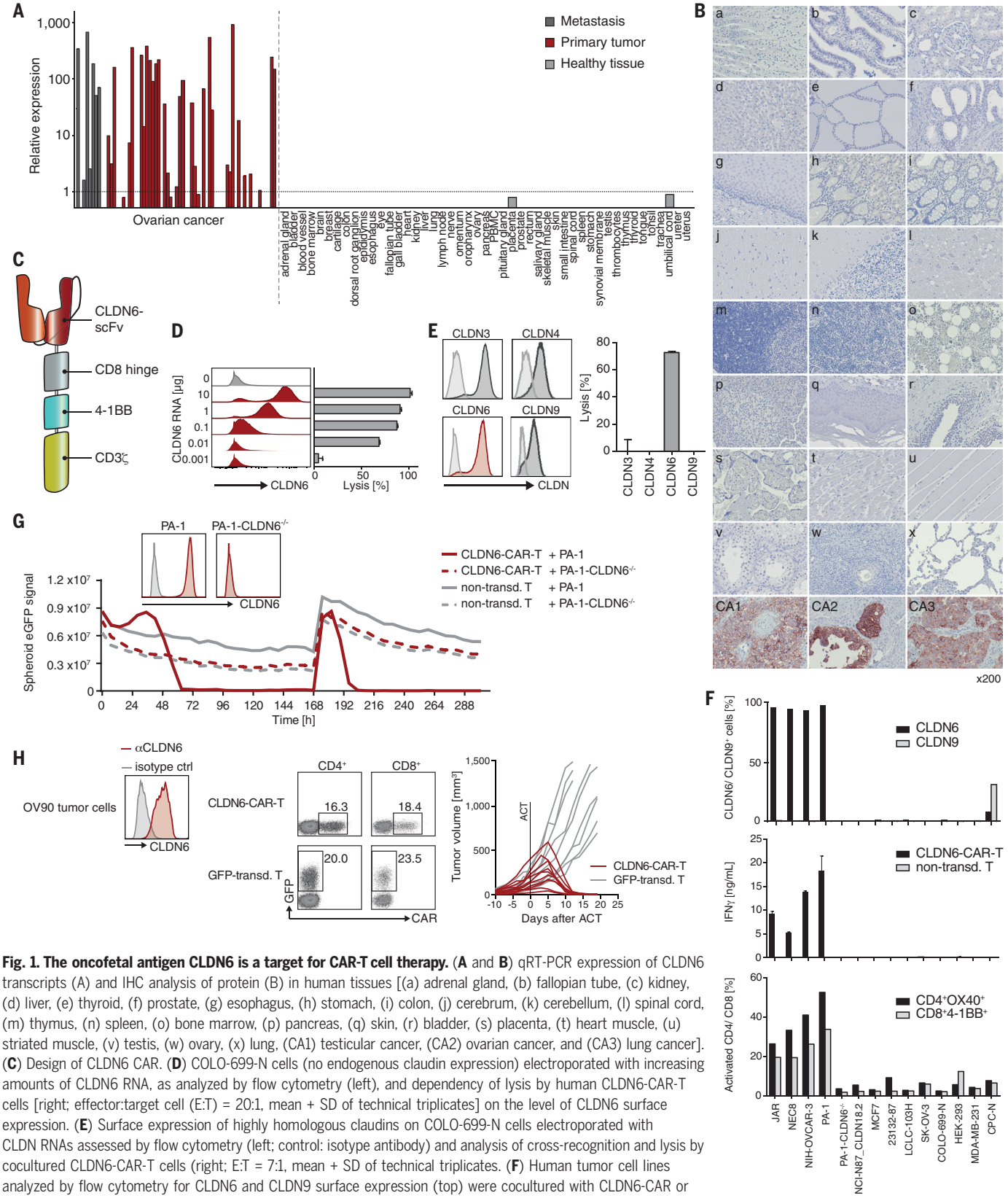


Fig. 1. The oncofetal antigen CLDN6 is a target for CAR-T cell therapy. (**A** and **B**) qRT-PCR expression of CLDN6 transcripts (A) and IHC analysis of protein (B) in human tissues [(a) adrenal gland, (b) fallopian tube, (c) kidney, (d) liver, (e) thyroid, (f) prostate, (g) esophagus, (h) stomach, (i) colon, (j) cerebrum, (k) cerebellum, (l) spinal cord, (m) thymus, (n) spleen, (o) bone marrow, (p) pancreas, (q) skin, (r) bladder, (s) placenta, (t) heart muscle, (u) striated muscle, (v) testis, (w) ovary, (x) lung, (CA1) testicular cancer, (CA2) ovarian cancer, and (CA3) lung cancer]. (**C**) Design of CLDN6 CAR. (**D**) COLO-699-N cells (no endogenous claudin expression) electroporated with increasing amounts of CLDN6 RNA, as analyzed by flow cytometry (left), and dependency of lysis by human CLDN6-CAR-T cells [right; effector:target cell (E:T) = 20:1, mean + SD of technical triplicates] on the level of CLDN6 surface expression. (**E**) Surface expression of highly homologous claudins on COLO-699-N cells electroporated with CLDN RNAs assessed by flow cytometry (left; control: isotype antibody) and analysis of cross-recognition and lysis by cocultured CLDN6-CAR-T cells (right; E:T = 7:1, mean + SD of technical triplicates. (**F**) Human tumor cell lines analyzed by flow cytometry for CLDN6 and CLDN9 surface expression (top) were cocultured with CLDN6-CAR or nontransduced T cells (E:T = 10:1). IFN-γ secretion (middle; mean + SD of technical duplicates) and expression of activation markers OX40 on $CD4^+$ and 4-1BB on $CD8^+$ T cells after coculture (bottom), as assessed by flow cytometry, are shown. (**G**) Serial killing of $CLDN6^{pos}$ and $CLDN6^{-/-}$ PA-1 tumor spheroids cocultured with either CLDN6-CAR or nontransduced (non-transd.) T cells (E:T = 10:1), as measured by enhanced GFP (eGFP) real-time imaging (mean of technical triplicates). (**H**) NSG mice bearing subcutaneous $CLDN6^{pos}$ OV90 xenografts were treated with human T cells transduced with CLDN6-CAR or eGFP. Tumor and T cell characteristics (left and middle) and tumor growth kinetics in individual mice (right) were analyzed. ctrl., control.

stimulation in an optimal immune-activating environment.

Recently, we introduced intravenously administered liposomal antigen-encoding RNA (RNA-LPX) to stimulate tumor-associated T cells in the natural repertoire of cancer patients (*18*). This nanoparticulate vaccine delivers antigen to APCs in the spleen, lymph nodes, and bone marrow and concomitantly initiates a Toll-like receptor–dependent type I IFN–driven immune-stimulatory program, promoting priming and strong expansion of antigen-specific T cells.

To test whether this approach could be adapted to act as a CAR-T cell–amplifying RNA vaccine (referred to hereafter as CARVac), we conducted a series of experiments. First, we tested if CLDN6 can be natively displayed on dendritic cells (DCs) to stimulate CLDN6-CAR-T cells in vitro. We measured concentration-dependent surface expression of CLDN6 on DCs treated with different amounts of CLDN6-encoding RNA-LPX (herein CLDN6-LPX) (Fig. 2A, top). The resulting expression of CLDN6 on DCs induced stimulation, cytokine secretion, and proliferation of co-cultured CLDN6-CAR-T cells in a dose-dependent manner (Fig. 2B, top). When BALB/c mice were injected intravenously with CLDN6-LPX, CLDN6 surface expression was detected on splenic DCs and macrophages but not on lymphocytes (Fig. 2C and fig. S4A), confirming in vivo delivery of the CAR antigen exclusively to APCs. APCs were activated and underwent maturation (fig. S4B), and strong activation of natural killer (NK), B, and T cells was detected in the spleen and lymph nodes of RNA-LPX–injected mice (fig. S4C).

Next, naïve C57BL/6 mice were engrafted with CLDN6-CAR-T cells labeled with a cell proliferation dye and vaccinated with CLDN6- or control-LPX. Spleen and lymph nodes from all major body regions resected from CLDN6-LPX–vaccinated, but not control-treated mice, displayed significantly increased proportions

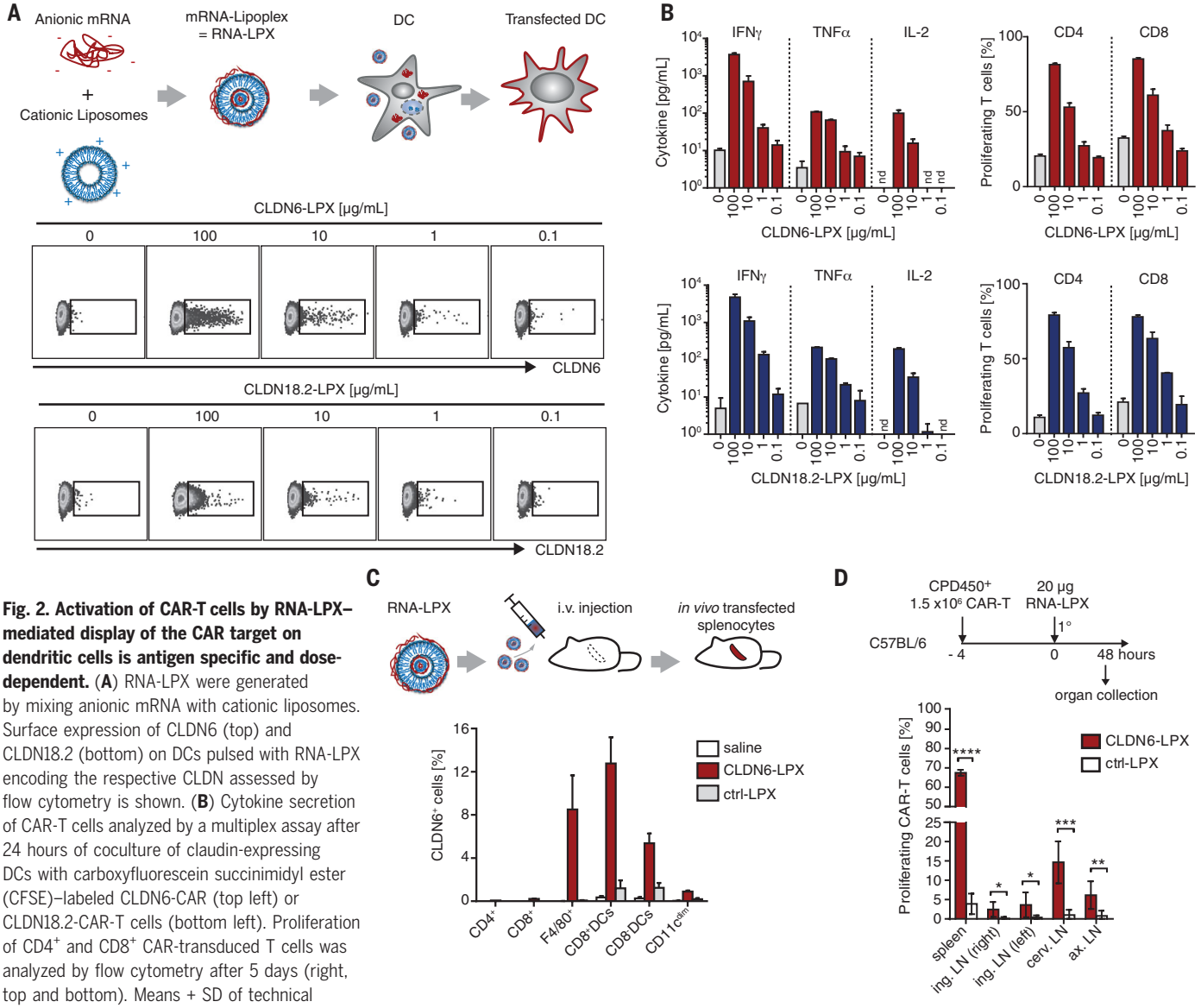


Fig. 2. Activation of CAR-T cells by RNA-LPX–mediated display of the CAR target on dendritic cells is antigen specific and dose-dependent. (**A**) RNA-LPX were generated by mixing anionic mRNA with cationic liposomes. Surface expression of CLDN6 (top) and CLDN18.2 (bottom) on DCs pulsed with RNA-LPX encoding the respective CLDN assessed by flow cytometry is shown. (**B**) Cytokine secretion of CAR-T cells analyzed by a multiplex assay after 24 hours of coculture of claudin-expressing DCs with carboxyfluorescein succinimidyl ester (CFSE)–labeled CLDN6-CAR (top left) or CLDN18.2-CAR-T cells (bottom left). Proliferation of $CD4^+$ and $CD8^+$ CAR-transduced T cells was analyzed by flow cytometry after 5 days (right, top and bottom). Means + SD of technical triplicates are indicated; nd indicates not detected. (**C**) Surface expression of CLDN6 on splenic immune cell populations of BALB/c mice analyzed by flow cytometry 24 hours after a single intravenous (i.v.) injection of 25 µg RNA-LPX encoding either CLDN6 or an irrelevant control (mean + SEM of biological duplicates). See fig. S4A for histograms. (**D**) CAR-T cell proliferation in secondary lymphoid tissues resected 48 hours after intravenous administration of RNA-LPX (CLDN18.2 as control). Data indicate mean ± SEM of biological replicates (n = 5 mice per group). LN, lymph node; ing. LN, inguinal LN; cerv. LN, cervical LN; ax. LN, axillary LN. P values were determined by unpaired Student's t test. *P < 0.05; **P < 0.01; ***P < 0.001; ****P < 0.0001.

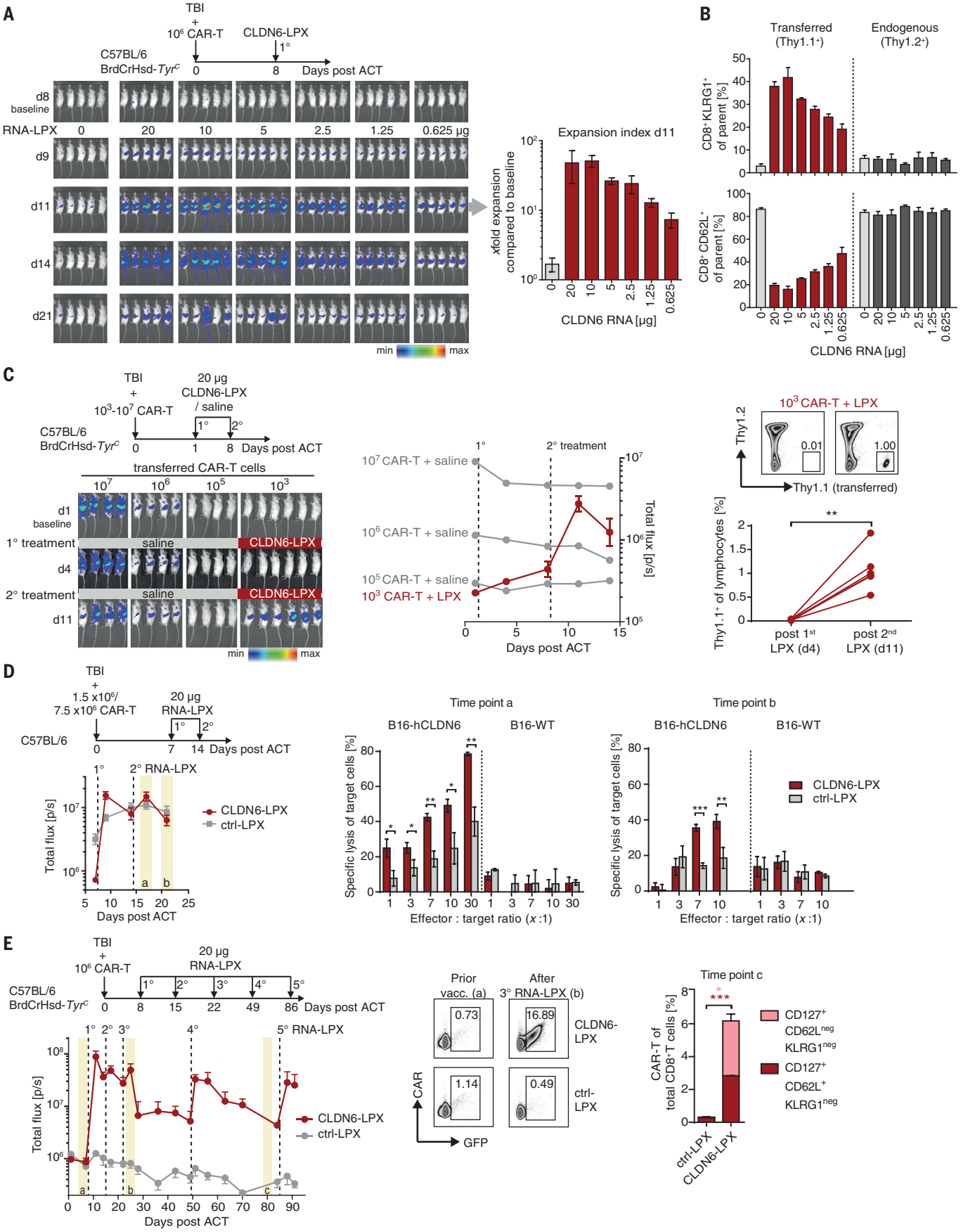


Fig. 3. CARVac promotes efficient in vivo expansion, superior functionality, and memory formation of CAR-T cells. (continued on next page)

(A and **B**) Impact of dose level of intravenously administered target-antigen encoding RNA-LPX on expansion of CAR-T cells in vivo. Luc-expressing Thy1.1$^+$ CLDN6-CAR-T cells (10^6 per animal) were transferred into lymphodepleted Thy1.2$^+$ C57BL/6-albino mice (n = 5 mice per group). Eight days later, mice were injected intravenously with 40 μg RNA-LPX in total containing the indicated titrated doses of CLDN6-LPX. Kinetics of CAR-T cell expansion [(A), left] by bioluminescence imaging (BLI) and the expansion index of CAR-T cells [(A), right] and the frequencies of KLRG1- and CD62L-expressing endogenous (Thy1.2$^+$) and transferred (Thy1.1$^+$) CD8$^+$ T cells in peripheral blood (B) 11 days after ACT by flow cytometry (mean + SEM) are shown. d, day. (**C**) Impact of repetitive intravenous dosing of target-antigen encoding RNA-LPX on expansion of CAR-T cells in vivo. BLI kinetics of different dose levels of Thy1.1$^+$ Luc-expressing CLDN6-CAR-T cells transferred into lymphodepleted Thy1.2$^+$ C57BL/6-albino mice are shown. Mice in the lowest CAR-T cell dose group (10^3 cells) were vaccinated twice with 20 μg CLDN6-LPX (n = 6 mice), whereas all other groups received saline (n = 4 mice per group). Representative imaging (left) and mean ± SEM of treatment groups (middle) are shown. Thy1.1$^+$ subsets in the peripheral blood of individual mice determined by flow cytometry (right) are shown. (**D**) Ex vivo cytotoxic activity of low-dose CAR-T cells from CLDN6-LPX–vaccinated mice (1.5 × 10^6 CAR-T + CLDN6-LPX) compared with high-dose CAR-T cells sorted from control-vaccinated mice (7.5 × 10^6 CAR-T + CLDN18.2-LPX) 3 days (time point a) and 7 days (time point b) after second vaccination (n = 5 mice per treatment group per time point). Sorted, pooled CAR-T cells per treatment group were cocultured for 20 hours in the presence of human CLDN6-transduced B16 mouse melanoma cells or wild-type (WT) control at indicated E:T ratios (mean ± SD of technical triplicates). (**E**) Luc-eGFP–expressing Thy1.1$^+$ CLDN6-CAR-T cells transferred into lymphodepleted Thy1.2$^+$ C57BL/6-albino mice (n = 2 or 3 mice per group) followed by repetitive vaccination with RNA-LPX (Oval as control). CAR-T cell kinetics by BLI (left; mean ± SEM of treatment groups) are shown. Frequency of eGFP$^+$ CAR-T cell subsets in the peripheral blood in pretreatment samples (time point a, day 7 after ACT) and after third RNA-LPX treatment (time point b, day 26 after ACT) (middle) are shown. Frequency of memory CAR-T cells in the CD8$^+$ T cell population 31 days after fourth treatment (right; time point c, day 80 after ACT) is shown. P values were determined by paired (C) and unpaired [(D) and (E)] two-tailed Student's t test. *P < 0.05; **P < 0.01; ***P < 0.001.

of proliferating CLDN6-CAR-T cells, suggesting body-wide functional expression of the CAR antigen within lymphoid compartments (Fig. 2D).

To assess the broader applicability of this approach, we selected CLDN18.2, which is a distantly related cancer-associated member of the claudin family. CLDN18.2 is expressed in various high–medical need tumors, such as gastroesophageal and pancreatic cancers (*19*–*21*). Both in human and mice, its expression in normal tissues is restricted to tight junctions of differentiated cells of the gastric mucosa, in which it is shielded. Only upon cancer-associated perturbation of the tight junction architecture does the CLDN18.2 antibody-binding epitope become exposed (*21*). Monoclonal antibodies (*22*, *23*) and CAR-T cells (*24*) against CLDN18.2 are being evaluated now in clinical studies. We engineered a CLDN18.2-CAR by substituting the CLDN6-specific scFv with an anti-CLDN18.2 scFv that exhibits specific binding with similar affinity to both human and mouse CLDN18.2 (*22*). CLDN18.2-CAR-T cells were shown to exert similar functional features as observed for the CLDN6-CAR, including strictly antigen-specific activation and killing of tumor cells in vitro (fig. S5A) and complete rejection of advanced CLDN18.2^{pos} tumors in vivo (fig. S5B). CLDN18.2-CAR-T cells cocultured with CLDN18.2-LPX–treated DCs showed cognate activation and proliferation (Fig. 2, A and B, bottom).

Next, we studied the in vivo performance of the CARVac strategy in a series of mouse experiments. Thy1.2$^+$ C57BL/6 mice underwent total body irradiation (TBI) for lymphodepletion and were then engrafted with congenic Thy1.1$^+$ CLDN6-CAR-T cells coexpressing luciferase (Luc) and green fluorescent protein (GFP) and subsequently vaccinated with CLDN6-LPX. In vivo bioluminescence imaging revealed that a single intravenous dose of CLDN6-LPX induced a profound expansion of circulating CLDN6-CAR-T cells (Fig. 3A and fig. S6). The expansion correlated with the CLDN6-LPX dose level and was substantial at even the lowest dose of 0.625 μg CLDN6 RNA. Quantitative and phenotypic analysis of peripheral blood T cells in treated animals confirmed increased frequencies of Thy1.1$^+$ CAR-T cells exhibiting an activated phenotype (KLRG1^{hi}, CD62L^{low}), whereas endogenous T cells were not affected at any dose after RNA-LPX treatment (Fig. 3B). The CAR-T cell numbers peaked 3 to 4 days after RNA-LPX vaccination followed by a decline, mimicking the dynamics of a physiological response of antigen-specific T cells to stimulation, with an initial expansion and subsequent retraction phase (Fig. 3A and fig. S6A).

In another experiment, groups of mice received different dose levels of CLDN6-CAR-T cells, starting as low as 10^3 cells per mouse, and either were left untreated or received a CLDN6-LPX regimen shortly after ACT. In mice that did not receive CLDN6-LPX, primary CAR-T cell engraftment (as quantified by bioluminescence) correlated linearly with the number of adoptively transferred cells and remained stable or slowly declined over time (Fig. 3C, left and middle, and fig. S7, A and B). Notably, in mice treated using the CARVac strategy, CAR-T cells were expanded irrespective of the starting dose. CLDN6-LPX mediated expansion of only 10^3 CAR-T cells resulted in detectable frequencies in peripheral blood (Fig. 3C, right). Almost the entire adoptively transferred CAR-T cell population underwent activation and proliferation by RNA-LPX, as indicated by transient up-regulation of Ki67 on the majority of transferred T cells (fig. S8A). The RNA-LPX expanded CLDN6-CAR-T cells were fully functional. As compared with CAR-T cells isolated from unvaccinated mice, they produced higher levels of IFN-γ (fig. S8B) and exerted significantly higher and strictly antigen-dependent cytolytic activity upon ex vivo coculture with CLDN6^{pos} tumor cells (Fig. 3D).

The low-dose CAR-T cell groups benefited more from repetitive RNA-LPX treatment, as indicated by increased expansion. In vivo expansion in the high-dose CAR-T cell groups stagnated after reaching high levels, suggesting a saturation threshold that was presumably due to T cells competing for homeostatic γc-cytokines and niches (Fig. 3C, middle, and fig. S7).

To assess the impact of repetitive RNA-LPX vaccination on long-term persistence of CAR-T cells, CLDN6-CAR-T cell–engrafted mice received three weekly doses of RNA-LPX followed by two further RNA-LPX administrations with longer treatment-free intervals (4 and 4.5 weeks). The first CLDN6-LPX exposure rapidly amplified CAR-T cells by more than two orders of magnitude, and subsequent weekly treatments maintained CAR-T cells at a high level, resulting in a frequency of more than 15% of total peripheral blood lymphocytes (Fig. 3E, left and middle). For the treatment group in which CLDN6-LPX treatment–free intervals were extended to up to 35 days, the blood CAR-T cell frequency declined. CAR-T cell numbers did not drop to the baseline level of engraftment but rather stabilized at a 10-fold higher frequency. After each treatment-free interval, CLDN6-CAR-T cells could be robustly reexpanded by CLDN6-LPX, indicating memory formation of CAR-T cells. Enrichment of CAR-T cells with an effector memory (CD127$^+$, CD62L^{neg}, KLRG1^{neg}) and a central memory (CD127$^+$, CD62L$^+$, KLRG1^{neg}) phenotype was confirmed by flow cytometry (Fig. 3E, right).

Cytokine release syndrome as a clinical manifestation of excessive and prolonged secretion

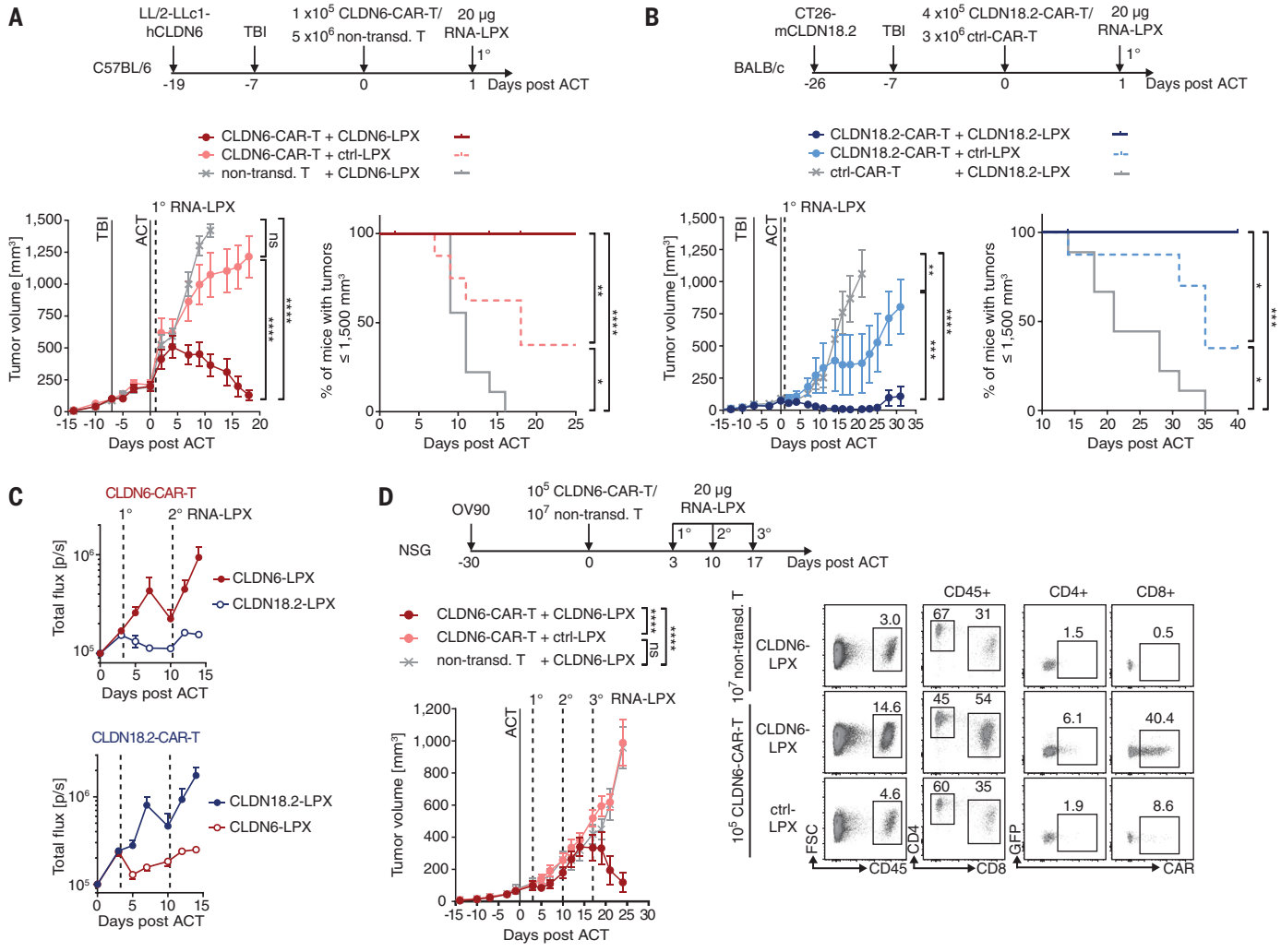


Fig. 4. Subtherapeutic CAR-T cell doses are efficacious against solid tumors after RNA-LPX vaccination. (**A** and **B**) Mice with large established tumors were lymphodepleted and treated with syngeneic nontransduced or CLDN-CAR–redirected mouse T cells followed by single intravenous administration of CLDN or control RNA-LPX. Tumor growth (left; mean ± SEM) and survival (right) were determined. CLDN6-CAR was tested in C57BL/6 mice bearing LL/2-LLc1 tumors transduced with human CLDN6 (n = 9 or 10 mice per group; tumor size at start of treatment was 209 mm^3) (A), and CLDN18.2-CAR was tested in BALB/c mice bearing mouse CLDN18.2-transduced CT26 (n = 9 mice per group; tumor size at start of treatment was 78 mm^3) (B). (**C**) Human Luc-expressing CLDN-specific CAR-T cells in naïve NSG mice vaccinated twice with CLDN-LPX. CAR-T cell expansion was analyzed by measuring the splenic BLI signal (mean ± SEM of 2 or 3 mice per group). (**D**) NSG mice with OV90 xenograft tumors (tumor size at start of treatment was 60 mm^3) were treated with a subtherapeutic dose of human CLDN6-CAR (10^5 cells per animal) or nontransduced T cells followed by three weekly repetitions of RNA-LPX coding for CLDN6 or a control. Tumor growth curves (left; mean ± SEM of 9 or 10 mice per group) and representative CAR-T cell frequencies after third RNA-LPX treatment in peripheral blood, as assessed by flow cytometry (right), are shown. (**E**) Maintaining frequency of circulating CAR-T cells within a therapeutic window by CARVac. P values were determined by two-way analysis of variance with Tukey's multiple-comparisons test [(A), left; (B), left; and (D), left]. The time period from ACT until >50% of mice in the control group were euthanized was used for calculation. Survival benefit was determined with the log-rank test [(A), right; and (B), right]. ns indicates not significant; *P < 0.05; **P < 0.01; ***P < 0.001; ****P < 0.0001.

of proinflammatory cytokines in the expansion phase is the most prominent severe adverse event of CAR-T cells against B cell markers (*25*). To explore the possibility of increased systemic cytokine release in conjunction with the CARVac strategy, we analyzed IFN-γ, interleukin-6 (IL-6), and tumor necrosis factor–α (TNFα) serum concentrations in gently preconditioned CLDN6-CAR-T cell–engrafted mice after exposure to CLDN6-LPX. Except for an early mild and transient increase of IFN-γ, no relevant increases of the tested proinflammatory cytokines were observed (fig. S9A), and treated mice were of normal appearance, displaying regular weight gain over time (fig. S9B).

Because repeated application of RNA-LPX and strong expansion of cytotoxic T cell effectors might bear the risk of depletion of APCs in the lymphoid tissues, we analyzed the spleens of treated mice because this is the organ with the highest RNA-LPX exposure. Spleens exposed to single or repetitive doses of RNA-LPX did not display any overt pathological alterations in spleen architecture or in appearance of red and white pulp (fig. S10, A and B). Flow cytometry of the cellular composition of spleen at different time points after repetitive RNA-LPX treatment showed mild and transient reductions of $CD11c^+$ DC and $F4/80^+$ macrophage populations and no quantitative changes in T, B, and NK cell populations (fig. S10C). No changes were noted in the cellular distribution of APC subsets in spleen tissue sections from corresponding time points (fig. S10D).

Finally, we studied the impact of RNA-LPX on the therapeutic efficacy of CAR-T cells in tumor-bearing mice. Lymphodepleted C57BL/6 mice with large $CLDN6^{pos}$ LL/2-LLc1 Lewis lung tumors (mean tumor volume 209 mm^3) underwent ACT with a subtherapeutic dose of mouse CLDN6-CAR-T cells followed by a single injection of CLDN6-LPX or control. Tumor control by CLDN6-CAR-T cells alone was incomplete, and tumor growth was only delayed. By contrast, 6 of 10 mice receiving CAR-T cells together with CLDN6-LPX vaccination showed complete rejection of large tumors, with a significantly higher median survival (Fig. 4A). We reproduced these findings in BALB/c mice with $CLDN18.2^{pos}$ CT26 colon carcinomas (mean tumor volume 78 mm^3) for CLDN18.2 CAR-T cells in conjunction with a single administration of CLDN18.2-LPX, further supporting the applicability of improving the antitumor effect of CAR-T cells with the CARVac (Fig. 4B).

To explore the possibility of CARVac for human CAR-T cells, we used the $CLDN6^{pos}$ OV90 xenograft tumor model in NSG mice. In pilot experiments, we confirmed that NSG mice are capable of splenic uptake of RNA (Fig. 4C and fig. S11A) and of promoting specific expansion of human CAR-T cells upon repetitive RNA-LPX administration (fig. S11B). NSG mice bearing advanced $CLDN6^{pos}$ OV90 tumors received a subtherapeutic dose of 1×10^5 $CLDN6\text{-}CAR^+$ T cells (fig. S12) followed by repetitive CLDN6-LPX or control treatment (Fig. 4D). The advanced tumors were completely rejected in CLDN6-LPX–treated mice, whereas they rapidly progressed in the control group engrafted with the same CAR-T cell dose (Fig. 4D, left). Effective tumor control correlated with a high frequency of CAR-T cells in the peripheral blood, proving their efficient in vivo expansion and improved persistence upon CLDN6-LPX vaccination (Fig. 4D, right). As with CLDN6-CAR-T cells, these findings were reproduced for human CLDN18.2-CAR-T cells in conjunction with CLDN18.2-LPX in the NSG mice xenograft model (fig. S13).

Our study has established two key findings. First, our data support CLDN6 as an oncofetal cell-surface antigen that appears suitable for CAR-T cell targeting. In humans, the CLDN6 gene is strictly silenced in healthy adult tissues but aberrantly activated in various solid tumors of high medical need, resulting in expression of high protein levels. This, together with the feasibility of engineering a CLDN6-directed CAR of high sensitivity, precise specificity, and strong potency against this surface molecule, proposes it as an ideal target for CAR-T cell therapy of solid cancers. Tumors without homogeneous CLDN6 expression bear the risk of outgrowth of antigen-loss variants. However, activated CLDN6-CAR-T cells are strongly IFN-γ–secreting effectors, and hence their antitumor activity is thought to drive inflammatory remodeling of the suppressive tumor microenvironment and release of endogenous tumor antigens, which together promote antigen-spread and counteract the rapid outgrowth of antigen-loss variants (*26*).

We also present the CARVac strategy as an approach to improve the antitumor efficacy of CAR-T cells. The CAR antigen is displayed in its native conformation on the surface of APCs residing in lymphoid compartments, which is the ideal setting for costimulation and potent expansion of T cells. Of note, it is likely that the same APCs concurrently process and present CLDN6 on major histocompatibility complex molecules, which may result in priming and activation of endogenous CLDN6-specific T cells (*18*). Recently, different approaches have been explored for antigen-specific expansion of CAR-T cells (*27*–*31*). The CARVac approach presented here combines various distinctive features. One of the advantages of CARVac is that single-stranded RNA as a natural Toll-like receptor ligand combines delivery of the antigen and adjuvanticity in one molecule. Importantly, the approach does not require a reengineering of the CAR scaffold or adaptation of T cell transduction protocols nor does it depend on the cumbersome identification and characterization of peptide ligands as vaccine mimotopes. Nanoparticulate RNA-LPX is fast and inexpensive to produce for any protein-based antigen. In ongoing clinical trials, the RNA-LPX vaccine platform is being used for inducing $CD4^+$ and $CD8^+$ T cell responses against a spectrum of different tumor antigens (NCT02410733, NCT02316457, and NCT03815058), with early clinical data supporting the lymphoid targeting and execution of the intended mode of action in humans (*18*, *32*). As exemplified for CLDN6, CLDN18.2, and also CD19 (fig. S14), a matched RNA vaccine can be immediately generated and manufactured in good manufacturing practice (GMP) grade for essentially any existing CAR, including those directed against conformational epitopes, indicating that CARVac is a universally applicable approach.

Our data establish the feasibility and safety of single as well as repetitive administration of CARVac for tunable expansion of engineered T cells. RNA-LPX–stimulated CAR-T cells appear superior to nonstimulated versions with regard to cytokine response and cytolytic activity upon antigen recognition. They form memory T cells and persist at higher frequencies. CARVac not only improves the engraftment of transferred CAR-T cells but also enables therapeutic tumor control at lower CAR-T cell doses (Fig. 4E).

The expansion, retraction, and restimulation kinetics of CAR-T cells mediated by RNA-LPX mimics the physiological process T cells undergo upon antigen-specific priming and boosting. Given that the magnitude of CAR-T cell expansion depends on RNA-LPX dose, control of the levels of circulating CAR-T cells and titration of CAR-T cell frequencies can be achieved within an optimal therapeutic window.

In addition to lack of suitable targets and fast decline of CAR-T cells in the circulation, other barriers for efficacy of CAR-T cells in human solid cancer exist, including tumor antigen heterogeneity, impaired T cell trafficking and extravasation to tumor sites, exhaustion, and an immunosuppressive microenvironment. Maintaining optimally stimulated CAR-T cells within a therapeutic window may provide a good foundation for overcoming those constraints as well.

REFERENCES AND NOTES

1. S. S. Neelapu *et al.*, *N. Engl. J. Med.* **377**, 2531–2544 (2017).
2. S. L. Maude *et al.*, *N. Engl. J. Med.* **378**, 439–448 (2018).
3. I. Scarfò, M. V. Maus, *J. Immunother. Cancer* **5**, 28 (2017).
4. R. A. Morgan *et al.*, *Mol. Ther.* **18**, 843–851 (2010).
5. C. H. Lamers *et al.*, *Mol. Ther.* **21**, 904–912 (2013).
6. S. A. Richman *et al.*, *Cancer Immunol. Res.* **6**, 36–46 (2018).
7. K. Turksen, T. C. Troy, *Dev. Dyn.* **222**, 292–300 (2001).
8. K. Turksen, T. C. Troy, *J. Cell Sci.* **117**, 2435–2447 (2004).
9. T. Ushiku, A. Shinozaki-Ushiku, D. Maeda, S. Morita, M. Fukayama, *Histopathology* **61**, 1043–1056 (2012).
10. P. Micke *et al.*, *Int. J. Cancer* **135**, 2206–2214 (2014).
11. C. H. June, R. S. O'Connor, O. U. Kawalekar, S. Ghassemi, M. C. Milone, *Science* **359**, 1361–1365 (2018).
12. S. L. Maude *et al.*, *N. Engl. J. Med.* **371**, 1507–1517 (2014).
13. M. Kalos *et al.*, *Sci. Transl. Med.* **3**, 95ra73 (2011).
14. D. L. Porter *et al.*, *Sci. Transl. Med.* **7**, 303ra139 (2015).
15. T. Gargett *et al.*, *Mol. Ther.* **24**, 1135–1149 (2016).
16. K. Feng *et al.*, *Sci. China Life Sci.* **59**, 468–479 (2016).
17. D. M. O'Rourke *et al.*, *Sci. Transl. Med.* **9**, eaaa0984 (2017).
18. L. M. Kranz *et al.*, *Nature* **534**, 396–401 (2016).
19. S. Wöll *et al.*, *Int. J. Cancer* **134**, 731–739 (2014).
20. C. Rohde *et al.*, *Jpn. J. Clin. Oncol.* **49**, 870–876 (2019).
21. U. Sahin *et al.*, *Clin. Cancer Res.* **14**, 7624–7634 (2008).
22. U. Sahin *et al.*, *Eur. J. Cancer* **100**, 17–26 (2018).
23. O. Türeci *et al.*, *Ann. Oncol.* **30**, 1487–1495 (2019).
24. H. Jiang *et al.*, *J. Natl. Cancer Inst.* **111**, 409–418 (2019).
25. J. N. Brudno, J. N. Kochenderfer, *Blood Rev.* **34**, 45–55 (2019).
26. J. H. Sampson *et al.*, *Clin. Cancer Res.* **20**, 972–984 (2014).
27. C. Berger *et al.*, *Cancer Immunol. Res.* **3**, 206–216 (2015).
28. C. Y. Slaney *et al.*, *Clin. Cancer Res.* **23**, 2478–2490 (2017).
29. M. Tanaka *et al.*, *Clin. Cancer Res.* **23**, 3499–3509 (2017).
30. L. Ma *et al.*, *Science* **365**, 162–168 (2019).
31. X. Wang *et al.*, *Clin. Cancer Res.* **21**, 2993–3002 (2015).
32. S. Pektor *et al.*, *EJNMMI Res.* **8**, 80 (2018).

ACKNOWLEDGMENTS

We thank B. Jesionek, N. Brüne, C. Stofft, and S. Krapp for technical support and A. Goß for project management. **Funding:** This work was supported by the CI3 Cutting Edge Cluster for Individualized Immune Intervention and funded by the German Federal Ministry of Education and Research (BMBF). **Author contributions:** U.S. was responsible for the conception and experimental strategy of the study. Design and analysis of the experiments were done by K.R., B.R., P.O., S.W., and E.C. K.Mi., A.B., N.H., O.K., K.K., Y.O., K.H., and M.S. performed experiments and acquired the data. D.W. and T.B. performed analysis of RNA sequencing datasets. M.B. and K.Mr. established in vitro assays. L.K. and M.D. established the RNA-LPX technology. V.J. and K.K. coordinated the GMP manufacturing. K.R., B.R., P.O., Ö.T., and U.S. interpreted the data and drafted the manuscript. **Competing interests:** K.R., B.R., P.O., K.Mi., A.B., N.H., O.K., K.K., Y.O., S.W., E.C., M.B., K.Mr., K.H., L.K., K.K., Ö.T., and U.S. are employees at BioNTech SE (Mainz, Germany). M.D. works as a consultant for BioNTech SE (Mainz, Germany). U.S., Ö.T., K.R., B.R., P.O., K.Mi., and K.Mr. are inventors on patents and patent applications, which cover parts of this article. K.R., P.O., O.K., L.K., V.J., M.D., K.K., Ö.T., and U.S. are stockowners. Ö.T. and U.S. are management board members of BioNTech SE (Mainz, Germany). All other authors declare no competing interests. **Data and materials availability:** The results shown here are based in part on data generated by the TCGA Research Network (www.cancer.gov/tcga; dbGap accession: phs000178) and the Genotype-Tissue Expression (GTEx) project (https://gtexportal.org; dbGap accession: phs000424.v4.p1). Claudin-specific antibodies IMAB027 and IMAB362 and their respective anti-idiotype antibodies are courtesy of Astellas Pharma GmbH. Material requests should be directed to Astellas Pharma GmbH.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/367/6476/446/suppl/DC1
Materials and Methods
Figs. S1 to S14
References (*33*–*37*)

View/request a protocol for this paper from *Bio-protocol*.

2 July 2019; accepted 18 December 2019
Published online 2 January 2020
10.1126/science.aay5967

DEVELOPMENTAL BIOLOGY

A tensile ring drives tissue flows to shape the gastrulating amniote embryo

Mehdi Saadaoui[1,2], Didier Rocancourt[1,2], Julian Roussel[1,2]*, Francis Corson[3]†, Jerome Gros[1,2]†

Tissue morphogenesis is driven by local cellular deformations that are powered by contractile actomyosin networks. How localized forces are transmitted across tissues to shape them at a mesoscopic scale is still unclear. Analyzing gastrulation in entire avian embryos, we show that it is driven by the graded contraction of a large-scale supracellular actomyosin ring at the margin between the embryonic and extraembryonic territories. The propagation of these forces is enabled by a fluid-like response of the epithelial embryonic disk, which depends on cell division. A simple model of fluid motion entrained by a tensile ring quantitatively captures the vortex-like "polonaise" movements that accompany the formation of the primitive streak. The geometry of the early embryo thus arises from the transmission of active forces generated along its boundary.

During amniote gastrulation, endodermal and mesodermal derivatives internalize through the primitive streak, a transient structure at the midline of the early embryo. In avians, the primitive streak forms from an initially crescent-shaped region at the margin between the embryo proper (EP) and extra-embryonic tissue (EE) (Fig. 1A), which converges toward and extends along the midline. Although myosin-II–driven oriented cell intercalation is known to underlie convergent extension of the prospective primitive streak (*1*, *2*), how the concomitant vortex-like tissue flows arise (*3*, *4*) and how they relate to the formation of the primitive streak has remained elusive.

To analyze gastrulation movements, transgenic quail embryos expressing a membrane-bound green fluorescent protein (memGFP) (*5*) were cultured ex vivo (*6*) and imaged in their entirety for 12 hours. The resulting movies were processed using particle image velocimetry (PIV) to reconstruct cell trajectories and tissue deformation maps (Fig. 1, B to D, and movie S1). Embryonic territories, originally characterized using anatomical or molecular criteria, could be recognized in these maps (compare Fig. 1, A and C). We designed automated fate-mapping methods that identify and track these territories on the sole basis of tissue movement (Fig. 1, E and F; fig. S1; and supplementary text). To validate this approach, we checked that the inferred location of the embryo margin aligned with the boundary of expression of the ectodermal/EP marker Sox3 in embryos that were fixed after live imaging (fig. S1, K and L). The angular motion of points along the margin, which wind around the EP as they converge to the posterior, captured the progress of gastrulation (Fig. 1, E to G). On the basis of these landmarks, we registered movies of six embryos in space and time to construct an average embryo (movie S2, fig. S1, and supplementary text), which was then used as a reference. The development of cultured embryos was virtually indistinguishable from an embryo imaged directly in the egg (fig. S2 and movie S3).

Noting that the EP maintains an approximately constant area, whereas the EE tissue steadily expands (Fig. 1H), we sought to distinguish area changes from other contributions to tissue movement. A decomposition into divergent (area changes) and rotational (incompressible) components indicated that gastrulation movements can be understood as the sum of three simpler flows: (i) a radial, outward movement of the expanding EE tissue; (ii) an area-preserving flow with two vortices within the EP; and (iii) at later stages, inward movement driven by areal contraction along the streak (Fig. 1, I and J, and movie S4). Although large-scale flows in the epiblast have been proposed to passively ensue from the deformation of the mesendoderm (*1*, *2*, *7*), we found that rotational movement persists after the mesendodermal crescent has converged onto the midline (Fig. 1, I and J) and that areal contraction makes a limited contribution to continued movement toward the streak (Fig. 1, K and L), suggesting that other forces must be at play.

For a viscous fluid that is described by the Stokes equations, these forces could be derived from the Laplacian of the velocity field (see the supplementary text). When applied to tissue flows in the epiblast, this suggested a pattern of tangential forces along the embryo margin extending well into its anterior half (Fig. 2A and movie S4). When flow is driven by active internal stresses, here by cell contractility, forces inferred in this way should be understood as apparent external forces arising from the spatial variations of the active stress. The force pattern of Fig. 2A thus pointed to a tensile margin around the EP, with a tissue-scale tension that decays from posterior to anterior: Tissue along the sides of the EP is drawn toward the posterior, where tension is higher; tissue in the posterior is thrust forward because the margin is curved (Fig. 2B). To test this hypothesis, we formulated a fluid-mechanical model that is based on the Stokes equations, with source terms for nonuniform area changes and for active tensile stresses along the margin (Fig. 2, C and D, and materials and methods). Area changes are taken from the experiment, whereas tensions along the margin, which moves with the tissue, are fit to the observed motion at each time step (Fig. 2C). Although strongly constrained—aside from the tensions, the initial position of the margin and its width are the only free parameters—the model recapitulates the full course of tissue movements over 8 hours (movie S5), with >90% accuracy for the average reference embryo (Fig. 2, F and G, and materials and methods). Considering the deformations that would result from each source term taken separately (Fig. 2E), active tensions largely account for the shaping of the embryo, whereas area changes are mostly responsible for EE expansion. As further abstractions, a "synthetic embryo" in which the source terms are replaced by simple mathematical functions of space and time (fig. S3 and movie S6; see also supplementary text, table S1, and fig. S4) is sufficient to quantitatively capture the movement of the tissue, and its essential features can be recovered analytically in the limit of a thin margin (fig. S5 and supplementary text).

Our results suggest that gastrulation is best understood as a tissue-wide process, arising from a tensile margin and a fluid-like response of the epithelial epiblast. We designed additional experiments to challenge this description. First, to detect a tensile margin, we performed circular ultraviolet laser cuts (*8*) at different locations in the epiblast (Fig. 3, A to C, and movie S7). Cuts along the margin revealed anisotropic tissue strains. Cuts inside the EP showed significantly lower strains, which can be ascribed to passive tissue deformation (see the supplementary text). As a control, linear cuts at the same locations showed a strong correlation between the final opening width (i.e., tissue strain) and the initial opening velocity (a more direct correlate of tissue stress; supplementary text and fig. S7). These experiments further revealed that tension runs all the way to the anterior margin. Although this could not be inferred from motion alone (motion in the model depends essentially on differences in tension), allowing for tension in the anterior recovers patterns of shear stress that agree with the observed tissue strains (supplementary text

[1]Department of Developmental and Stem Cell Biology Institut Pasteur, 75724 Paris, Cedex 15, France. [2]CNRS UMR3738, 75015 Paris, France. [3]Laboratoire de Physique de l'Ecole Normale Supérieure, CNRS, ENS, Université PSL, Sorbonne Université, Université de Paris, 75005 Paris, France.
*Present address: Institut du Cerveau et de la Moelle Épinière, Paris, France.
†Corresponding author. Email: corson@lps.ens.fr (F.C.); jgros@pasteur.fr (J.G.)

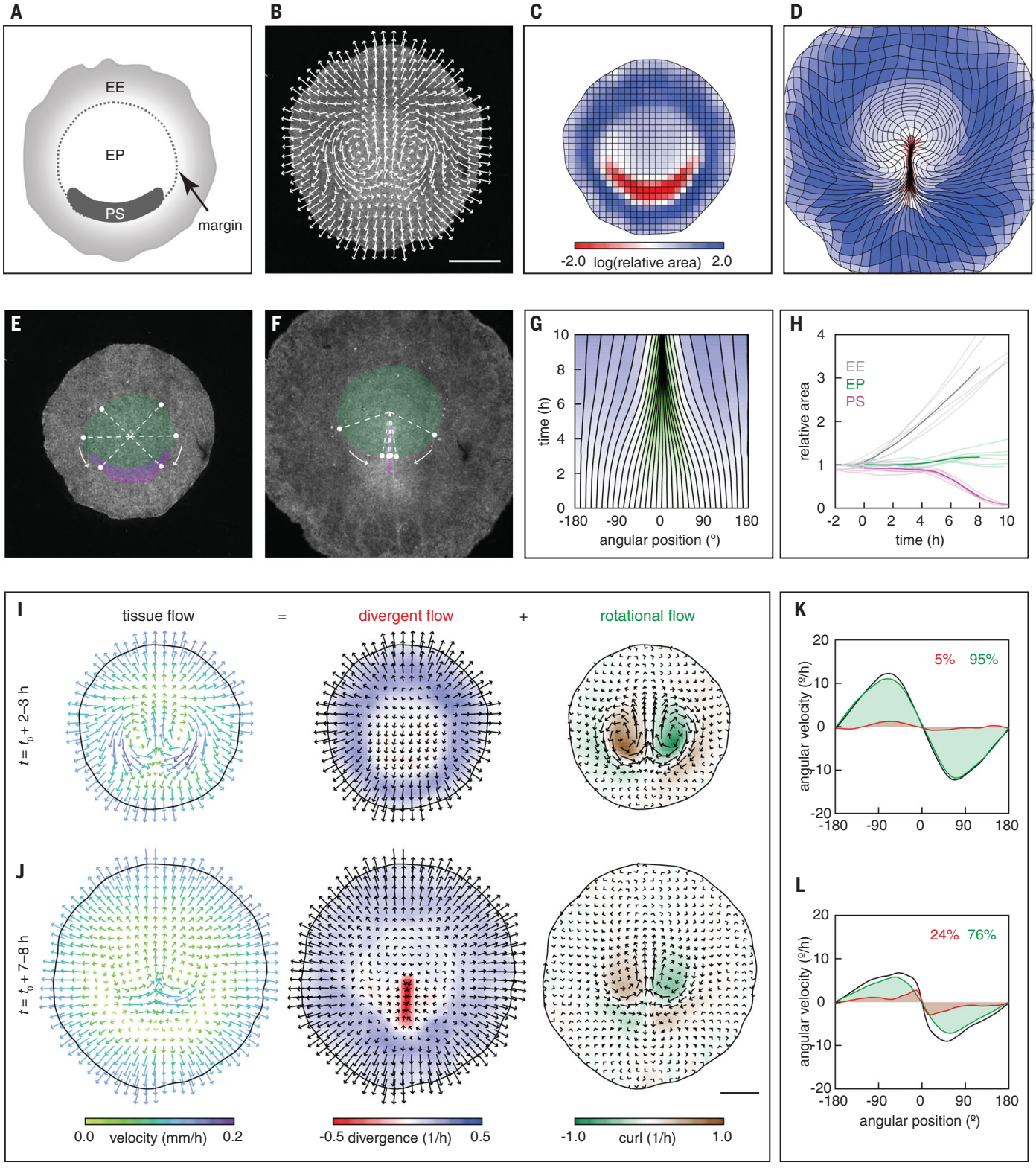


Fig. 1. Quantitative description of gastrulation movements. (**A**) Early epiblast. (**B** to **D**) Trajectories [(B), t = 4 to 6 hours] and deformation of an initially square grid [(C) and (D)] from the PIV analysis of a memGFP embryo movie. Colors in (C) and (D) show area changes between the initial [(C), t = 0] and final [(D), t = 10 hours] configurations. (**E** to **H**) Automated fate mapping (green, EP; magenta, primitive streak); dots show winding motion (arrows) along the margin, quantified in (G) by the time evolution of angular positions [dotted lines in (E) and (F); 0° is posterior]. (H) shows the area of tissue regions versus time (n = 6 embryos; bold lines are averages). (**I** to **L**) Decomposition of the tissue-velocity field into divergent and rotational components [(I) and (J)] and contributions to motion along the margin [(K) and (L); colors are as in (I)]. Percentages quantify shaded areas (averages over n = 6 embryos and the indicated time intervals). t_0, time of motion onset. Scale bars, 1 mm.

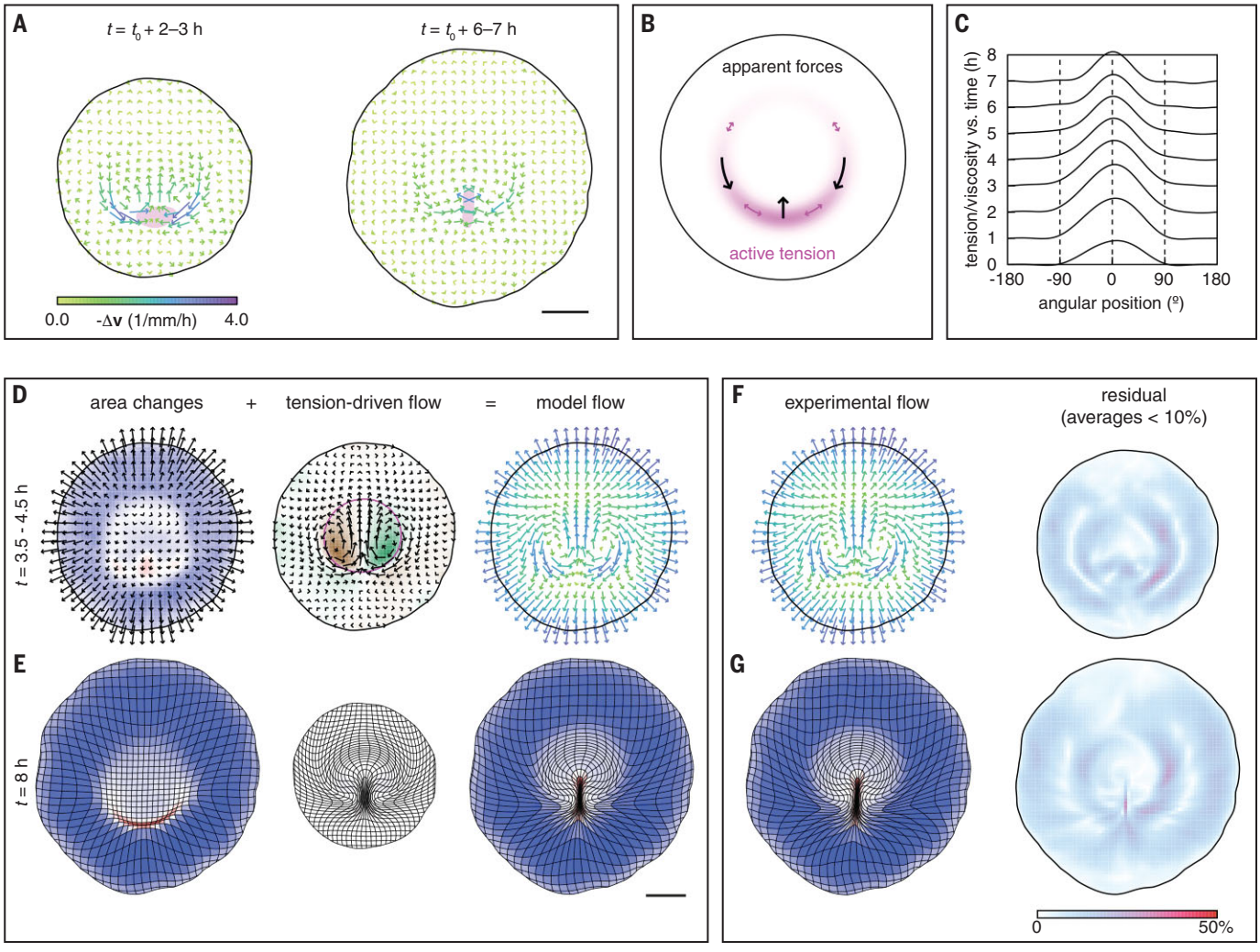


Fig. 2. Quantitative fluid-mechanical model for gastrulation. (**A**) Apparent forces (negative of the Laplacian of the velocity field; averages over n = 6 embryos; magenta, presumptive primitive streak). (**B**) Diagram illustrating how apparent forces (black) arise from graded tensions along the margin (magenta). (**C** to **E**) Quantitative model for gastrulation movements. (C) Tension/viscosity profiles (millimeters per hour; averages over 1-hour intervals) from a fit to the reference average embryo (n = 6 embryos). (D and E) Tissue flows in the model as the resultant of area changes (taken from experiment) plus an incompressible flow driven by tension along the margin [magenta line in (D)]. (D) Velocity fields (colors as in Fig. 1I). (E) Deformation maps from each source term taken separately and together. (**F** and **G**) Velocity field (F) and deformation map (G) for the average embryo (right panels show deviation between model and experiment). Scale bars, 1 mm.

and fig. S20B). To connect tissue-scale motion and cellular-scale behaviors, we analyzed embryos that were fixed after live imaging for different time intervals (Fig. 3, D to O, and figs. S8 to S13). Cell segmentation of entire embryos showed a gradual increase in cell areas in the EE tissue, likely contributing to its expansion (Fig. 3, E, I, K, and O). Cell shapes, which initially had an isotropic distribution, became elongated along the margin, consistent with a state of tension (Fig. 3, F, I, L, and O). Quantification of junctional phosphorylated myosin II revealed localized myosin anisotropy, a correlate of active force generation (*9*), at the margin (Fig. 3, G to I and M to O). The location of this large-scale supracellular ring aligned with the location of the embryo margin determined from the motion of the tissue before fixation, and its width agreed with that inferred from the model (fig. S13, supplementary text, and fig. S6). High-resolution live imaging of transgenic embryos expressing a tdTomato-myosin II reporter revealed the progressive formation of dynamic, tangential actomyosin supracellular cables spanning five to 20 cells at the margin (fig. S14A and movie S8), coincident with the site of apparent forces inferred from tissue motion (fig. S15). In the posterior margin, these cables contracted, driving oriented intercalations (*2*). In the anterior, supracellular cables were also visible but extended tangentially, concomitant with cell elongation and oriented divisions (fig. S14, B and C, and movie S9)—indicative of stress dissipation (*10*). Thus, the margin exerts active tension in the posterior and passive tension in the anterior.

Second, we sought to identify the cellular basis of tissue fluidity. Because cell rearrangements contribute to stress relaxation in epithelial tissues (*11*), and most cell rearrangements in the early avian embryo are associated with cell division (*12*), we reasoned that cell division may be required for a fluid-like behavior. Treatment with hydroxyurea (HU) efficiently suppressed cell division (fig. S16A) but induced apoptosis in the long term (fig. S16B). When HU was combined with the apoptosis inhibitor

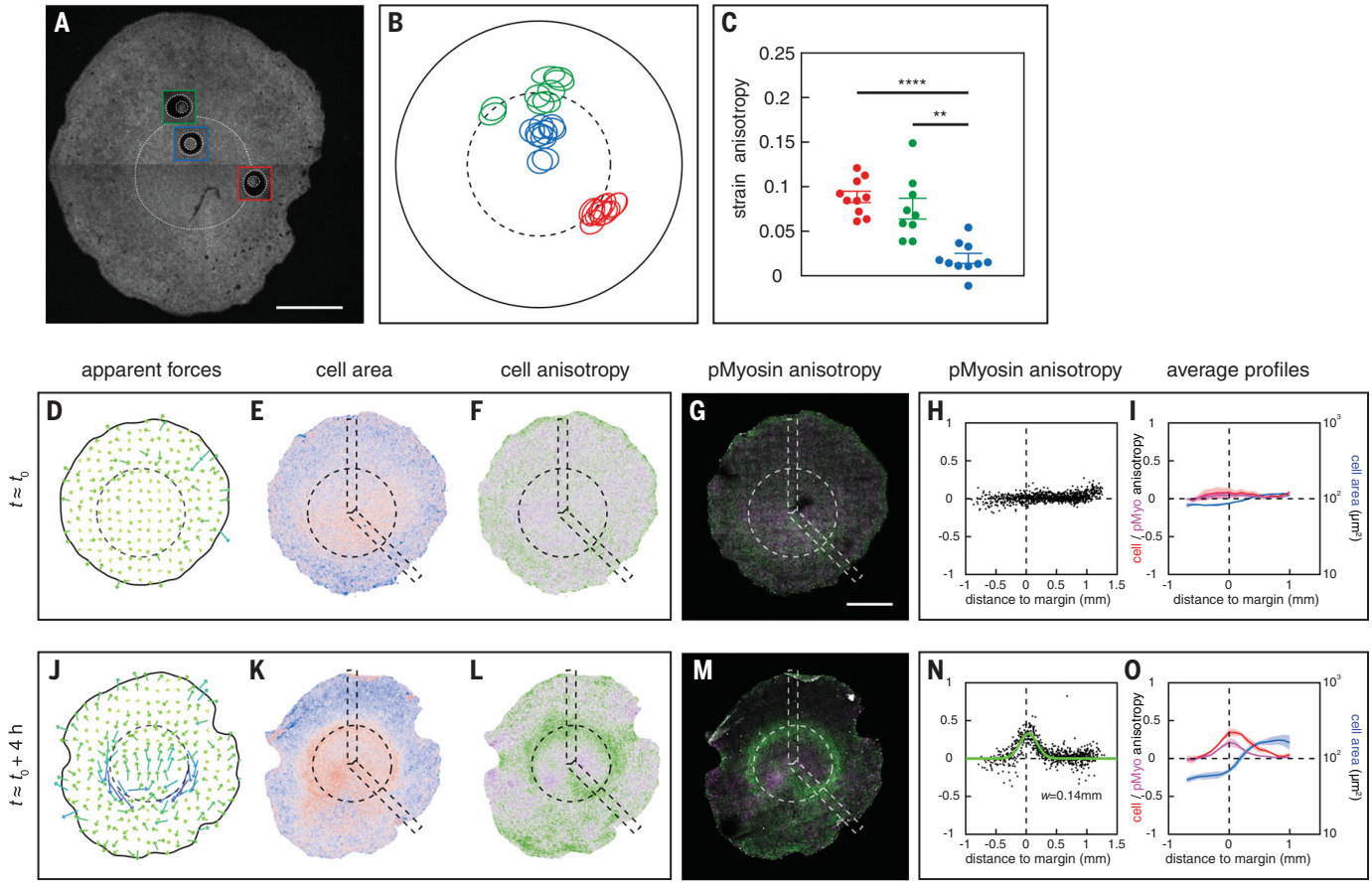


Fig. 3. Mechanical, cellular, and molecular characterization of the embryo margin. (**A** to **C**) 250-μm circular laser cuts in a single memGFP embryo (A) and representation of all laser cut experiments (B); ellipses show anisotropic strain amplified fourfold for visibility. (C) Tangential versus radial strain anisotropy (bars, mean±SE; $^{**}P < 0.01$, $^{****}P < 0.0001$, paired *t* test). Red indicates the posterior margin, green the anterior margin, and blue the EP. (**D** to **O**) Apparent forces [(D) and (J)] in embryos imaged until the indicated stages (dashed line, margin) and cell areas [(E) and (K)], cell shape anisotropy [(F) and (L)], and junctional phosphorylated myosin anisotropy [tangential (green) versus radial (magenta) in (G) and (M) and averaged in 100 × 100 μm boxes in (H) and (N)] mapped in these embryos and averaged across embryos [(I) and (O); mean ± SE of radial profiles in $n = 2$ and $n = 3$ embryos]. Boxed regions are shown in figs. S8 to S11 and individual embryos from (I) and (O) in figs. S12 and S13. Green curve in (N) is the Gaussian fit with standard deviation *w*, as indicated. Scale bars, 1 mm.

Q-VD-OPh, both cell division and apoptosis were suppressed (fig. S16, A, C, and D) and the topology of the epithelium was greatly stabilized compared with controls (fig. S16, E and F, and movie S10). Whereas Q-VD-OPh alone only slightly delayed the progress of gastrulation (figs. S1J and S17, A to E), embryos incubated in both HU and Q-VD-OPh showed a marked slowdown by 6 to 8 hours of treatment and failed to form a primitive streak ($n = 6$; Fig. 4, A and B; fig. S1J; fig. S17, D and F; and movie S11). Tissue expansion persisted, but rotational movements were abolished, consistent with a suppression of fluidity (fig. S17G and movie S12). As a control, treatment with aphidicolin and Q-VD-OPh produced quantitatively similar results (fig. S17, H to J). The model suggested that the amplitude of the tension/viscosity ratio dropped over time (Fig. 4C). Laser cuts in embryos treated with HU and Q-VD-OPh revealed that tensions were still present, if not increased, along the margin (fig. S17, K to M, and movie S13), and a supracellular ring was still observed in fixed embryos (Fig. 4D and fig. S18), implying that the slowdown resulted from an increase in viscosity and not from a decrease in tension. Thus, fluidity of the embryonic epithelium is required for primitive streak formation and emerges from cell division.

Third, we challenged model predictions for hydrodynamic effects in gastrulation. The embryo, which draws the surrounding tissue to the posterior, is akin to a swimmer and should move forward over time. Indeed, embryos exhibited a slow anterior-ward movement, in quantitative agreement with the model (fig. S19). At odds with the view that vortex-like flows are shaped by a confining boundary (*7*, *13*, *14*), our model suggests that they are governed by the distribution of active forces, with boundary conditions playing a limited role in the intact epiblast. The progress of gastrulation is predicted to be weakly sensitive to the distance to epiblast border, reaching >80% of its maximum rate when it is just 50% larger in radius than the EP (fig. S5D and supplementary text). Indeed, circular cuts centered on the margin, which removed most of the EE tissue and brought the epiblast border closer to the EP, had almost no effect on tissue flow and streak formation (Fig. 4, E and F, and movie S14). By contrast, in the case of off-centered cuts, which bring the border even closer to one side of the EP, the model predicted, and experiments confirmed, that the interaction between the EP and the border

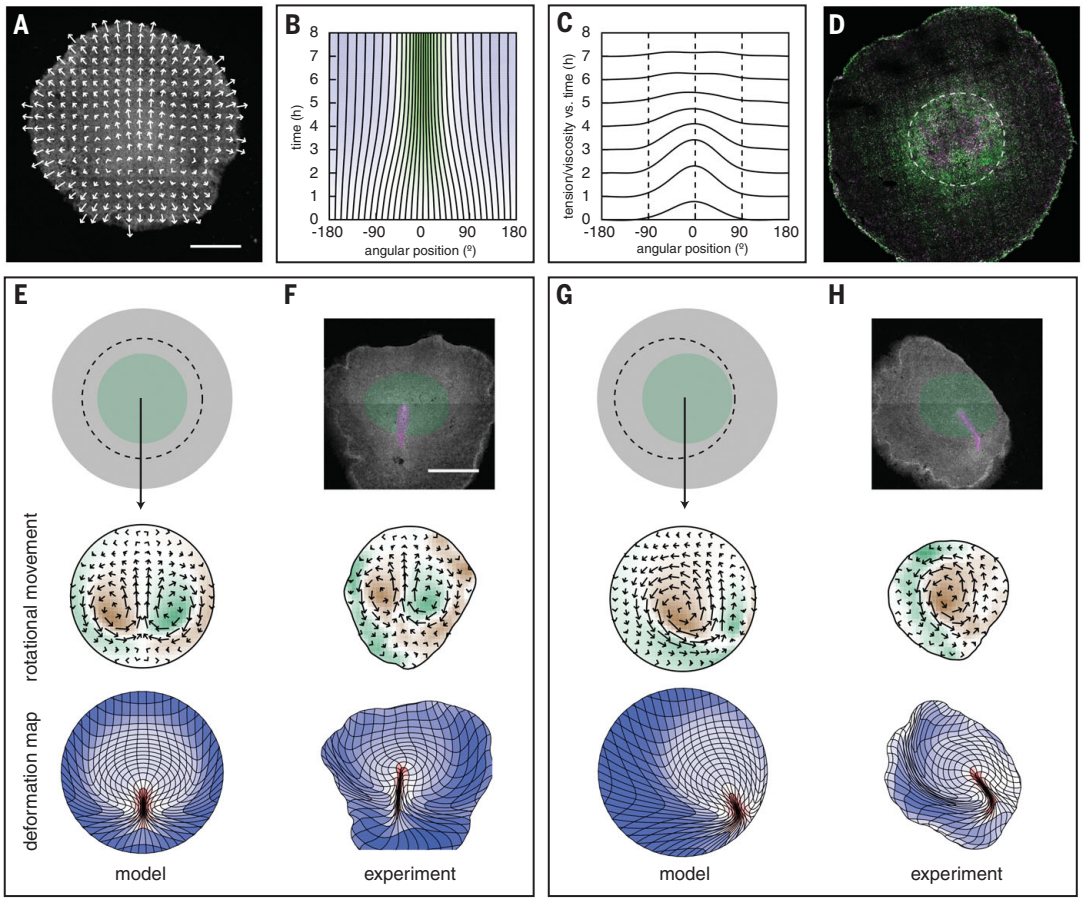


Fig. 4. Manipulation of tissue viscosity and hydrodynamic effects in gastrulation. (**A** to **D**) Effect of HU+Q-VD-OPh on gastrulation movements (n = 5 embryos). (A) Trajectories (individual embryo, t = 4 to 6 hours). (B) Time evolution of angular positions along the margin (average embryo). (C) Tension/viscosity profiles from model fit to average embryo. (D) Junctional phosphorylated myosin anisotropy (compare Fig. 3M). (**E** to **H**) Predictions from the synthetic model [(E) and (G)] and experimental response to centered (F) and off-centered (H) cuts generating a new tissue border. Scale bars, 1 mm.

induces a rotation of the axis, leaving only one apparent vortex and resulting in a bent streak (Fig. 4, G and H, and movie S14).

Our study demonstrates the power of fluid-mechanical approaches (*9*, *15*, *16*) to capture large-scale morphogenetic movements and identifies a simple mechanical basis for gastrulation. Although tissue-wide flows in the embryonic disk were previously interpreted as being a passive consequence of primitive streak formation (*1*, *2*, *7*), we find instead that both are part of a broader process that is driven by tensile forces all along the margin and shapes the embryo as a whole (see the supplementary text and fig. S20 for further discussion of alternative models). Supracellular actomyosin cables, which have been shown to drive local cell rearrangements, stabilize compartment boundaries, and act as purse strings in wound healing and embryonic tissue closure (*17*, *18*), effect large-scale remodeling of the surrounding tissue through nonuniform contraction. Our finding that the embryo margin, which was previously identified as a molecular organizer of early development (*19*–*21*), is also defined by a specific mechanical state and cellular behaviors hints that mechanical and molecular cues may combine in the establishment of the amniote body plan.

REFERENCES AND NOTES

1. O. Voiculescu, F. Bertocchini, L. Wolpert, R. E. Keller, C. D. Stern, *Nature* **449**, 1049–1052 (2007).
2. E. Rozbicki *et al*., *Nat. Cell Biol.* **17**, 397–408 (2015).
3. R. Wetzel, *Wilhelm Roux Arch. Entwickl. Mech. Org.* **119**, 188–321 (1929).
4. L. Gräper, *Dev. Genes Evol.* **116**, 382–429 (1929).
5. C. Moreau *et al*., *Curr. Biol.* **29**, 35–50.e4 (2019).
6. S. C. Chapman, J. Collignon, G. C. Schoenwolf, A. Lumsden, *Dev. Dyn.* **220**, 284–289 (2001).
7. O. Voiculescu, L. Bodenstein, I.-J. Lau, C. D. Stern, *eLife* **3**, e01817 (2014).
8. I. Bonnet *et al*., *J. R. Soc. Interface* **9**, 2614–2623 (2012).
9. S. J. Streichan, M. F. Lefebvre, N. Noll, E. F. Wieschaus, B. I. Shraiman, *eLife* **7**, e27454 (2018).
10. P. Campinho *et al*., *Nat. Cell Biol.* **15**, 1405–1414 (2013).
11. C. Guillot, T. Lecuit, *Science* **340**, 1185–1189 (2013).
12. J. Firmino, D. Rocancourt, M. Saadaoui, C. Moreau, J. Gros, *Dev. Cell* **36**, 249–261 (2016).
13. M. Chuai, C. J. Weijer, in *Multiscale Modeling of Developmental Systems*, S. Schnell, P. K. Maini, S. A. Newman, T. J. Newman, Eds., vol. 81 of *Current Topics in Developmental Biology* (2008), pp. 135–156.
14. S. A. Sandersius, M. Chuai, C. J. Weijer, T. J. Newman, *Phys. Biol.* **8**, 045008 (2011).
15. B. He, K. Doubrovinski, O. Polyakov, E. Wieschaus, *Nature* **508**, 392–396 (2014).
16. M. Smutny *et al*., *Nat. Cell Biol.* **19**, 306–317 (2017).
17. K. Röper, *Bioarchitecture* **3**, 45–49 (2013).
18. C. Schwayer, M. Sikora, J. Slováková, R. Kardos, C.-P. Heisenberg, *Dev. Cell* **37**, 493–506 (2016).
19. K. Joubin, C. D. Stern, *Cell* **98**, 559–571 (1999).
20. R. F. Bachvarova, I. Skromne, C. D. Stern, *Development* **125**, 3521–3534 (1998).
21. S. B. Shah *et al*., *Development* **124**, 5127–5138 (1997).

ACKNOWLEDGMENTS

We thank V. Hakim, P.-F. Lenne, A. Martinez Arias, and F. Schweisguth for critical reading of the manuscript and P. Caldarelli for the in situ hybridization shown in fig. S1, K and L. **Funding:** The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013)/ERC Grant Agreement no. 337635, Institut Pasteur, CNRS, Cercle FSER, Fondation pour la Recherche Medicale, and the Vallee Foundation. **Author contributions:** M.S., F.C., and J.G. conceived the study. D.R. and J.R. generated transgenic lines. M.S. performed experiments. F.C. developed quantitative analysis methods and theoretical models. F.C. and J.G. wrote the paper with input from M.S. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** All experimental data are available in the main text or the supplementary materials.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/367/6476/453/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S20
Table S1
References (*22*–*38*)
Movies S1 to S14

View/request a protocol for this paper from *Bio-protocol*.

28 November 2018; accepted 18 December 2019
10.1126/science.aaw1965

ORGANIC CHEMISTRY

Total synthesis reveals atypical atropisomerism in a small-molecule natural product, tryptorubin A

Solomon H. Reisberg[1], Yang Gao[1], Allison S. Walker[2], Eric J. N. Helfrich[2], Jon Clardy[2]*, Phil S. Baran[1]*

Molecular shape defines function in both biological and material settings, and chemists have developed an ever-increasing vernacular to describe these shapes. Noncanonical atropisomers—shape-defined molecules that are formally topologically trivial but are interconvertible only by complex, nonphysical multibond torsions—form a unique subset of atropisomers that differ from both canonical atropisomers (e.g., binaphthyls) and topoisomers (i.e., molecules that have identical connectivity but nonidentical molecular graphs). Small molecules, in contrast to biomacromolecules, are not expected to exhibit such ambiguous shapes. Using total synthesis, we found that the peptidic alkaloid tryptorubin A can be one of two noncanonical atropisomers. We then devised a synthetic strategy that drives the atropospecific synthesis of a noncanonical atrop-defined small molecule.

In 1894, Emil Fischer proposed a lock-and-key analogy for how biological molecules interact to carry out biological functions, and the three-dimensional (3D) shapes of molecules have been a major focus of biological chemistry ever since (*1*). Accordingly, the structure of small molecules has been assumed to be defined solely by atomic connectivity and point or axial chirality. For example, the steroid hormones all have the same basic carbon skeleton—a rigid assembly of four rings fused one to another—and their different biological roles depend on the modifications to the periphery of this basic skeleton. In contrast, large molecules such as proteins can reversibly self-organize into well-defined 3D structures, and the rules governing this ability are increasingly well understood (*2*). This structural feature of biological macromolecules encodes many of the functions that form the basis of life (*1*). For example, hydrogen-bonding, hydrophobic, arene-π, and solvation interactions drive proteins to fold into specific tertiary structures that render them operable (*2*). Molecular shapes (i.e., tertiary structures) for most macromolecules are derived from atomic connectivity but are fundamentally separate from it; that is, many proteins can be folded and unfolded without breaking or forming covalent bonds (*3*).

For certain macromolecules, however, shape is directly tied to atomic connectivity rather than to conformational changes (Fig. 1A, left). In the case of cyclic DNA, for example, the wound and unwound topologies are interconvertible only by the scission and reformation of phosphate linkages (*4*). Likewise, molecular catenanes have been synthesized with defined topology (*5*). Such nonsuperimposable and noninterconvertible topologies are called topoisomers. Two molecules are topoisomers of each other if they have identical connectivity but nonidentical molecular graphs—that is, molecular pairs that are noninterconvertible without the breaking and reformation of chemical bonds (*6*).

This type of defined topoisomerism is conspicuously absent from small-molecule natural products. A distinct, if seemingly analogous, isomerism in a small-molecule context is atropisomerism (i.e., shape isomerism through hindered bond rotation). Canonically, atropisomerism involves a single torsionally hindered bond that bestows axial chirality; hindered biaryls (Fig. 1A, right) represent a prototypical example.

In contrast to both canonical (singly axially chiral) atropisomerism and topoisomerism, there exist a variety of shape-defined molecules that are theoretically interconvertible by bond rotation but are categorically distinct from canonical atropisomers because of the multiple and nonphysical bond torsions required for their interconversion. Many mechanically interlocked molecules fit into this middle ground; for example, both rotaxanes (*7*) and lasso peptides (*8*) (Fig. 1A, center) are topologically trivial and should formally be considered atropisomers with their unthreaded counterparts, but are clearly categorically distinct from simple prototypical examples of atropisomerism. [For another compelling case of noncanonical atropisomerism, see (*9*).] In a physical (rather than theoretical) sense, most members of the lasso peptide class of natural products can be interconverted from unthreaded to threaded shapes only by breakage and repair of the peptide backbone. It is worth noting that the terms "topology" and "topoisomer" are often erroneously used in the literature to refer to these shape-defined (but topologically trivial) structures; noncanonical atropisomerism is a more accurate vernacular.

Shapes defined by features beyond atomic connectivity and point chirality—both topoisomerism and noncanonical atropisomerism—are not mere structural curiosities: The shape of DNA completely drives its function (*1*), and an entire highly conserved class of enzymes (topoisomerases) have evolved to control DNA's tertiary structure (*10*). Similarly, the superior proteolytic stability of lasso peptides is attributed to their threaded shape (*11*).

Whereas noncanonical atropisomerism is prevalent in macromolecular examples, it is generally not even considered in the realm of small-molecule natural product chemistry: Point chirality, *E/Z* alkene geometry, and more typical biaryl atropisomerism are all prevalent definers of structure in secondary metabolites, but structures with ambiguous molecular shape are conspicuously absent. Indeed, upon isolation of novel small-molecule natural products, considerations of shape may never even enter into the structural assignment process.

Tryptorubin A (**1**, Fig. 1B) is a small-molecule peptidic indole alkaloid containing a highly strained bridging bis(macrocycle) (*12*). Upon isolation, potential atropisomerism was not considered. Therefore, no explicit shape was initially assigned; the 2D rendering in the isolation disclosure [see figure 1 in (*12*), as well as Fig. 1B, left] was drawn in a shape-ambiguous manner, and the 3D rendering [see figure 2A in (*12*)] was arbitrarily drawn with the "bridge below" shape (i.e., equivalent to compound **1b** in Fig. 1B, right).

We have found that tryptorubin A (**1**), as a result of chirality and connectivity alone, could theoretically present as two possible noncanonical atropisomers. We describe an atroposelective synthesis of *atrop*-tryptorubin A (**1b**), the discovery of its atypical atropisomerism, and a hypothesis-driven atropospecific strategy that led to the synthesis of the natural product (**1a**) and its unambiguous atropisomeric assignment. Additionally, we report genomic data that help to clarify the biogenesis of **1a**; these data suggest a biosynthetic pathway involving ribosomal peptide synthesis followed by atroposelective posttranslational modification.

We were interested in pursuing the total synthesis of tryptorubin A for its sheer structural complexity. At the outset of synthetic work, atropisomerism was not considered, and the synthesis of **1** was pursued according to Fig. 2A: Protected dipeptide **2** was prepared at multidecagram scale (*13*). In an effort to cyclize the first macrocycle, we subjected **2** to Ullmann-Goldberg-Buchwald-Ma conditions (*14*) (i.e., CuI, diamine ligand, and carbonate base). This reaction resulted in substantial epimerization adjacent to the methyl ester; **2** was thus first saponified to lower its α-acidity. The Ullmann cyclization proceeded smoothly on the free acid without α-epimerization. Cyclization was

[1]Department of Chemistry, The Scripps Research Institute, La Jolla, CA 92037, USA. [2]Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, MA 02115, USA.
*Corresponding author. Email: jon_clardy@hms.harvard.edu (J.C.); pbaran@scripps.edu (P.S.B.)

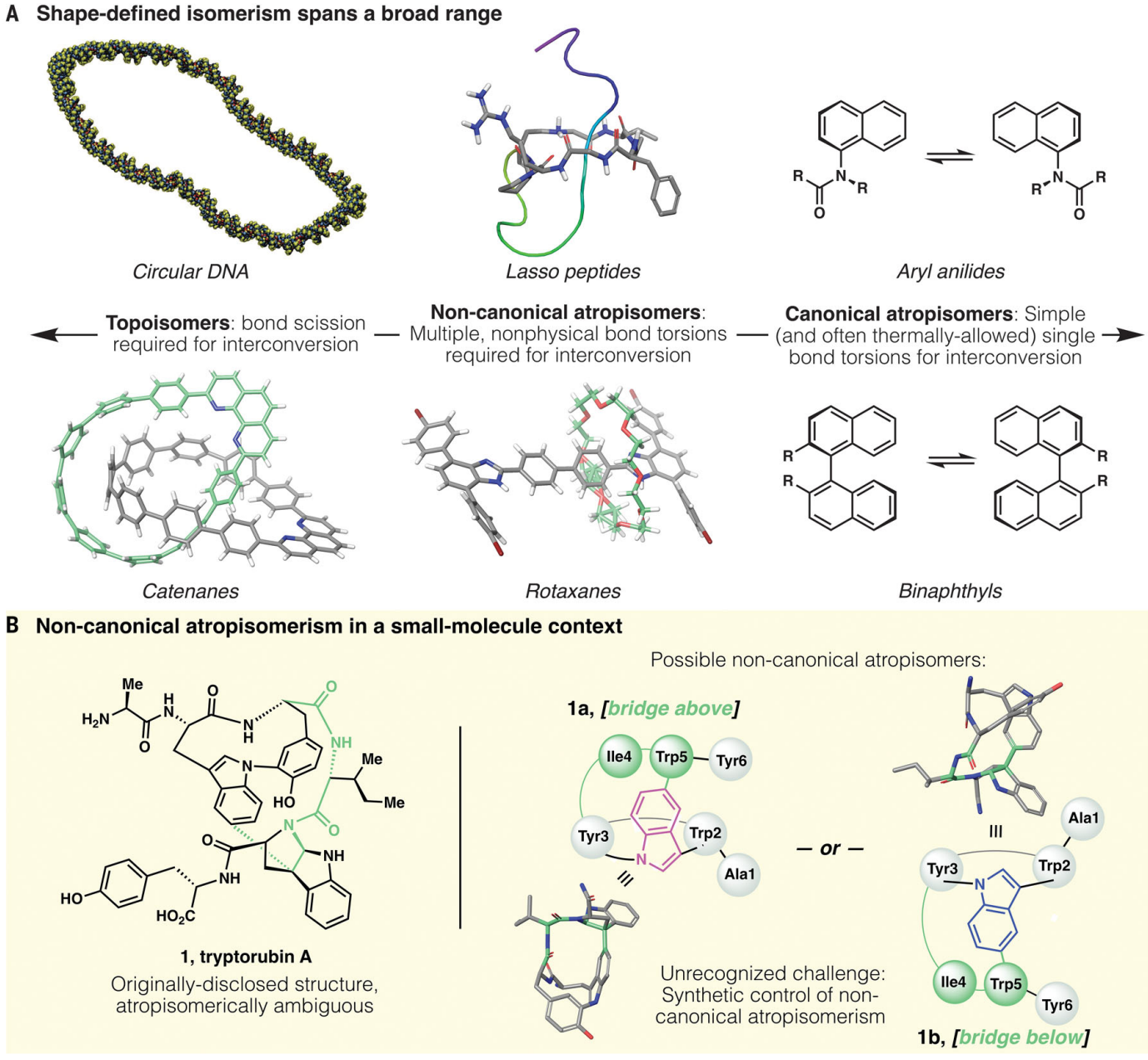


Fig. 1. Shape isomerism in macro- and small molecules. (**A**) Shape-based isomerism in synthetic and natural products spans a broad range. At one end (left), defined topology encodes topoisomers. At the other end (right), canonical atropisomerism is defined by simple axial differences (i.e., torsion of a single bond). Under the broad umbrella of atropisomerism, but distinct from more canonical examples, are noncanonical atropisomers (center) that are formally topologically trivial, but whose interconversion requires complex multibond rotations and unphysical torsions. Historically, this area has been occupied only by macromolecules; in this work, we disclose a small-molecule natural product that presents this type of noncanonical atropisomerism. Structures obtained from PDB and/or CCDC database: circular DNA, reproduced from (*30*); lasso peptide, PDB 5TJ1 (*8*); catenane, CCDC #1835146 (*5*); rotaxane, CCDC #1576710 (*7*). (**B**) Left: Originally proposed structure of tryptorubin A. Right: Two noncanonical atropisomers are possible within the limits of the originally proposed 2D structure. Note that 3D structures of **1a** and **1b** are computed, not crystallographic, and their terminal residues are truncated for clarity.

possible at concentrations as high as 40 mM (>20 g scale) without substantial loss of material to dimeric or oligomeric species.

Subsequent protecting-group manipulation yielded protected indole scaffold **3**, which was subjected to Movassaghi's silver-mediated alkylation (*15*, *16*) to establish the key nonproteogenic C-C bond. Reaction of **3** with alkyl bromide **4** resulted in low yields and complex mixtures of regioisomers at C5 and C6. We hypothesized that reducing **3** to its indoline congener (**5**) would both increase the overall nucleophilicity of the arene and inhibit C6-based reactivity. This tactic was realized by reduction of **3** with trifluoroacetic acid (TFA) and triethylsilane, followed by an alkylation/oxidation sequence, affording the desired product (**6**) as a single regioisomer in moderate yield. In parallel to this work, homologation of indoline **5** with an additional alanine unit (to form **7**) allowed for unambiguous stereochemical assignment by x-ray crystallography.

Subsequent elongation of the peptide chain by appending Ala and Ile residues, followed by

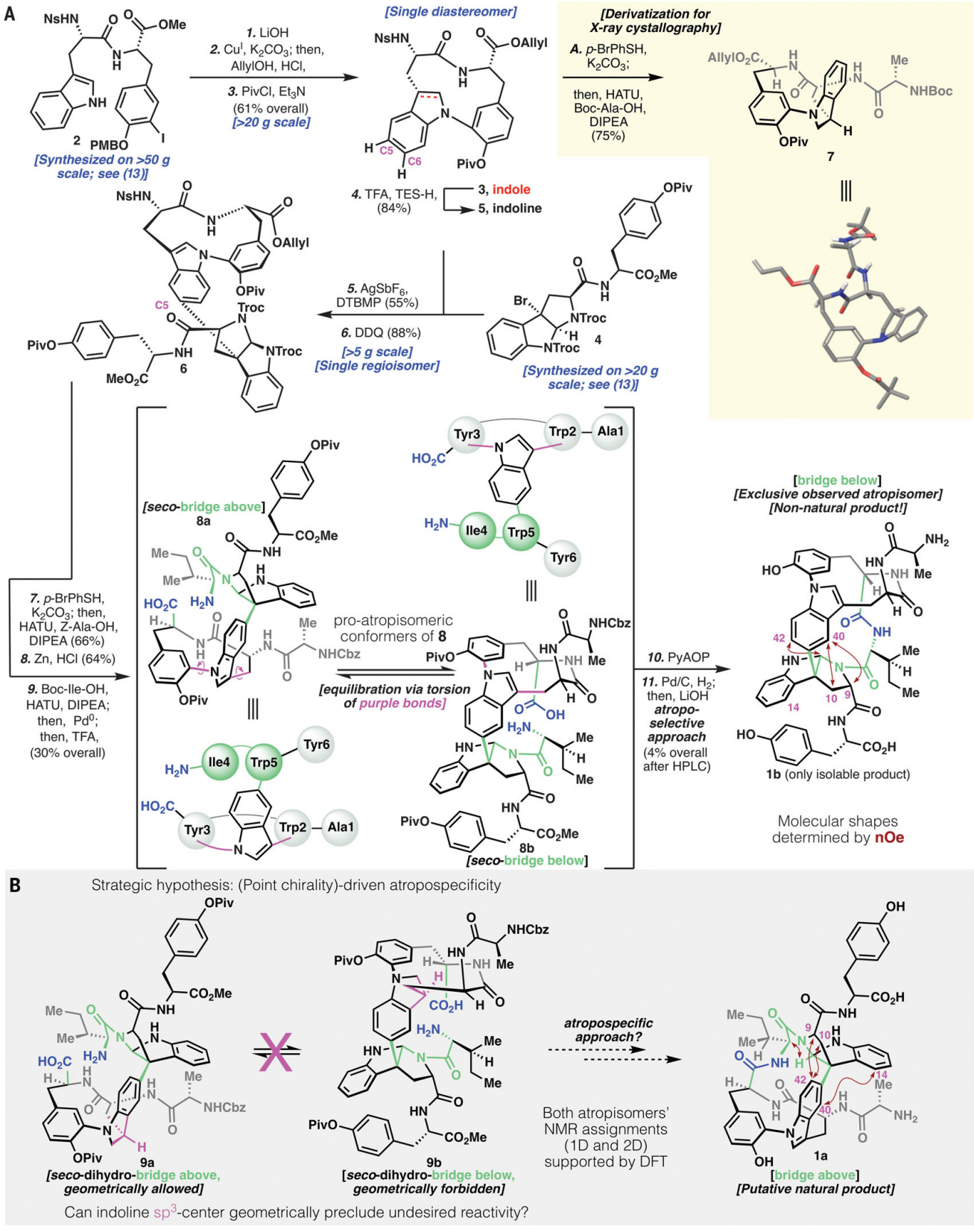


Fig. 2. Tryptorubin A's noncanonical atropisomerism: Discovery and synthesis of the unnatural atropisomer. (**A**) Synthetic route to *atrop*-tryptorubin A (**1b**). (**B**) Strategic hypothesis to use point chirality to drive an atropospecific synthesis of tryptorubin A. Piv, pivalate; PMB, *para*-methoxybenzyl; Ns, nosyl; DTBMP, 2,6-di-*tert*-butyl-4-methylpyridine; HATU, hexafluorophosphate azabenzotriazole tetramethyl uronium; PyAOP, (7-azabenzotriazol-1-yloxy)tripyrrolidino-phosphonium hexafluorophosphate; nOe, nuclear Overhauser effect.

deprotection, gave *seco*–amino acid **8**. At this juncture, characterization by nuclear magnetic resonance (NMR) spectroscopy became challenging (even at high temperature), presumably because of cis/trans amide isomerization of the tertiary pyrroloindolinyl amide, various rotameric populations, and conformational equilibrium between **8a** and **8b**. Nonetheless, **8** appeared as a single sharp peak in high-performance liquid chromatography (HPLC) and exhibited a high-resolution mass spectrum (HRMS) consistent with the postulated structure. After extensive experimentation (*13*), this structure could be cyclized in low yield to a bis (macrocycle). Global deprotection yielded **1b**, with HRMS data indicating the same molecular formula as the natural isolate (**1**). Unfortunately, the NMR data [^{1}H, ^{13}C, heteronuclear multiple bond correlation (HMBC), heteronuclear single quantum coherence (HSQC), rotating-frame nuclear Overhauser effect correlation spectroscopy (ROESY)] and LC retention of **1b** were distinct from the natural product (**1**) [see below and (*13*)].

With these contrasts in spectral data in mind, we began to consider possible explanations for the structural discrepancy between **1** and **1b**. We considered the possibilities of stereochemical misassignment (e.g., a D–amino acid) or regiochemical misassignment (e.g., alternate regiochemistry in the indole-pyrroloindoline C-C bond) in the natural and/or synthetic products. After exhaustive review of natural **1** and synthetic **1b**'s respective spectral data as well as a separate total synthesis of C26-epimeric species ***epi*-8** [see (*13*) for this additional synthesis], we confirmed that natural **1** and synthetic **1b** had the same connectivity and point-stereochemistry (*13*). It was only upon careful analysis of the two compounds' ROESY spectra that a key insight was discovered: Although the natural product (**1a**) showed strong nuclear Overhauser effect correlations from H9 and H10 to H42 (Fig. 2B), the analogous H9 and H10 protons in the synthetic (**1b**) compound's ROESY spectrum showed correlations to H40 (Fig. 2A). This key geometric constraint, combined with additional spectral evidence [**1b** and **1a** in Fig. 2, A and B; see (*13*) for additional details and full skeletal numbering system], illuminated our understanding that even within the limits of identical connectivity and stereochemistry, **1** could potentially exist as two noncanonical atropisomers ("bridge above," **1a**; "bridge below," **1b**). Subsequent density functional theory (DFT)–based NMR prediction validated our assignment of the synthesized product as **1b** and supported our hypothesis that the natural product was the other atropisomer (**1a**). [For a compelling example of use of these predictions to validate structural assignments, see (*17*).] In addition to computational NMR prediction, computation of the respective atropisomers' polar and hydrophobic surface area showed that the unnatural atropisomer (**1b**) had substantially less polar surface area, consistent with its increased retention relative to **1a** on reverse-phase LC (*13*).

Retrospectively, we reasoned that *seco*–compound **8** likely exists as two interconverting pro-atropisomeric conformers (**8a** and **8b**, Fig. 2A) and that the "bridge below" conformer (**8b**) has superior kinetic facility for cyclization, leading exclusively to **1b** as the observed cyclization product. We hypothesized that by geometrically locking the cyclization precursor into the "bridge above" conformation, we could achieve inversion of atroposelectivity. Combining this hypothesis with crystallographic evidence of the geometry of indoline **7**, we recognized that in a substrate such as indoline **9**,

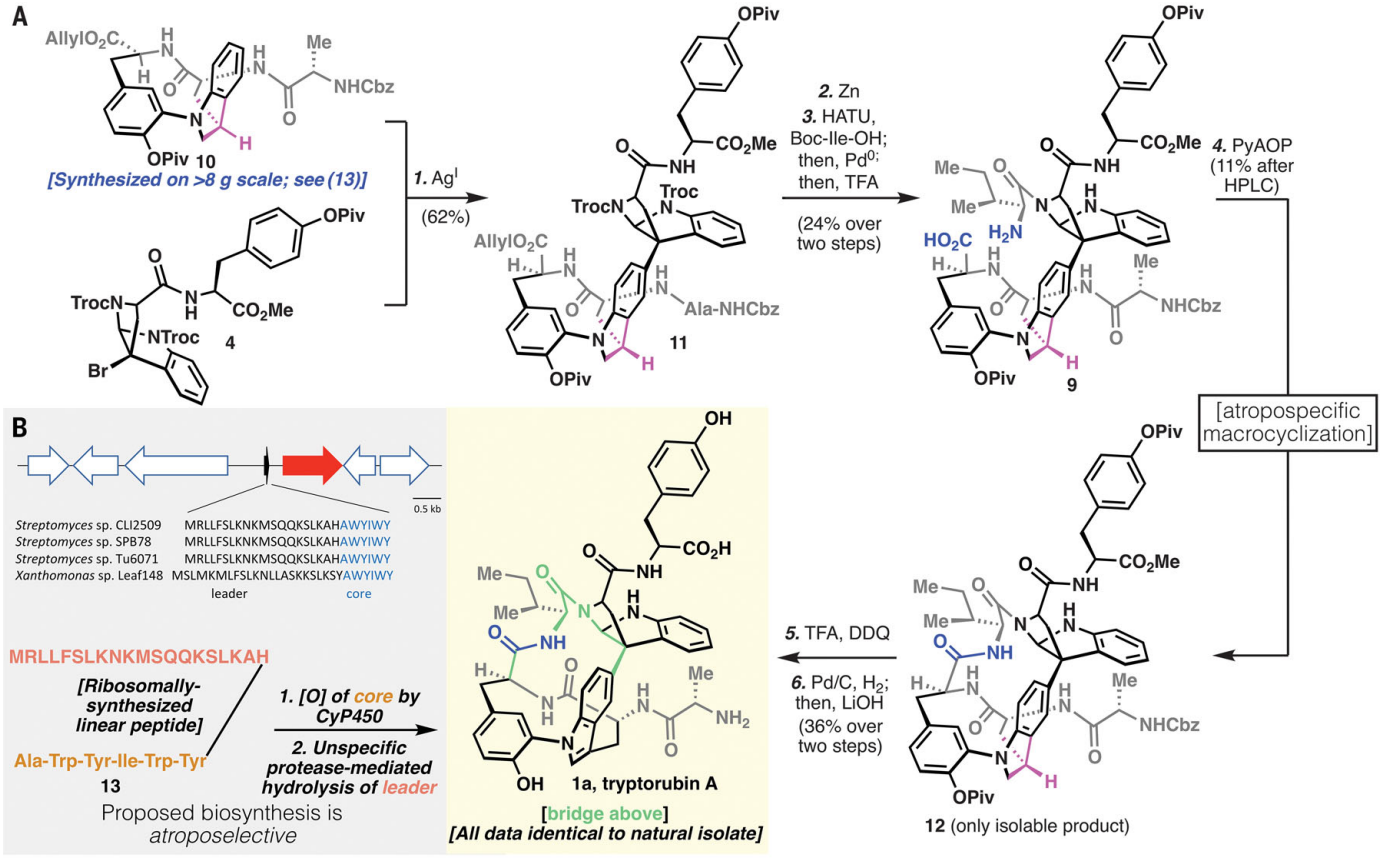


Fig. 3. Total synthesis of tryptorubin A. (**A**) Atropospecific synthesis of tryptorubin A (**1a**). (**B**) Top: A RiPP sequence that encodes tryptorubin A's linear peptide sequence. Bottom: Proposed biosynthetic pathway to **1a**. Amino acid abbreviations: A, Ala; F, Phe; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; Q, Gln; R, Arg; S, Ser; W, Trp; Y, Tyr.

the point chirality at indoline (Fig. 2B, purple methine) would geometrically preclude the "bridge below" conformer (**9b**); indeed, geometric limitations of **9** would render the cyclization atropospecific for the "bridge above" atropisomer **1a** (resulting from cyclization of **9a**). Such a strategy is reminiscent of methods to control more canonical atroposelectivity by point-to-axial chirality transfer (*18*).

Figure 3A describes our successful execution of the atropospecific strategy laid out in Fig. 2B and the subsequent total synthesis of the natural isomer of tryptorubin A (**1a**). We prepared indoline analog **10** in multigram quantities, and a similar Friedel-Crafts sequence as previously described yielded pentamer **11**. We subjected this pentamer to a peptide homologation sequence to give the key *seco*–amino acid (**9**), with the indoline oxidation state geometrically enforcing a "bridge above" conformation. Subsequent cyclization gave **12** as a single isomer. However, simple dehydrogenation of indoline **12** proved challenging. Attempts at metal-catalyzed dehydrogenation (i.e., Pd/C) gave only recovered starting material, and use of cerium ammonium nitrate (CAN) resulted in complete decomposition of the substrate. Although 2,3-dichloro-5,6-dicyanobenzoquinone (DDQ) had given the desired reactivity in the acyclic system (see above), in the present case we observed unselective oxidation of the pyrroloindoline aminal to an amidine, as well as the desired indoline-to-indole oxidation. We reasoned that this overoxidation pathway was due to the excessively electron-rich, unprotected anilinic nitrogen. Transient protection of this basic nitrogen with a proton from solvent-quantity TFA allowed for the desired monodehydrogenation.

Global deprotection of the oxidized material gave **1a**, whose spectral data matched closely with the natural product (*13*). In light of the complex shape-driven assignment of **1b** and **1a**, we tried to obtain solid-state structures in several ways, including classical small-molecule x-ray crystallographic methods, high-throughput protein crystallography approaches (*19*), and small-molecule microcrystal electron diffraction (MicroED) (*20*, *21*). Unfortunately, no solid-state structure could be determined, and our structural assignment remains reliant on solution-phase NMR. Although we observed slight deviations in ^{1}H NMR resonances, these types of deviations are common for natural products rich in potential hydrogen-bonding groups, and the ^{13}C NMR and all 2D (HSQC, HMBC, ROESY) correlations were near-identical. Co-injection of authentic natural product and synthetic material to LC resulted in identical retention as well as a single species in ^{1}H NMR. We thus assign tryptorubin A as having the "bridge above" geometry of **1a**.

Isomers **1a** and **1b** do not interconvert at elevated temperature (*13*), further showcasing the (likely infinitely) high barrier to interconversion in this atypical example of atropisomerism. Geometrically, interconversion of the atropisomers would require one macrocycle to pass through the other (Fig. 4); it is unsurprising that this is thermally unfeasible, as it would appear from first principles to be physically impossible without the breakage and reformation of covalent bonds. One might naïvely assign these compounds as topoisomers because of this physical limitation; however, noncanonical atropisomerism is a more accurate descriptor, as the word topoisomerism is defined without respect to such limitations. In other words, if one could hypothetically stretch bonds beyond the limits of covalency, **1a** and **1b** would be interconvertible, and thus they are topologically trivial. This contrasts with a topologically defined catenane, for example, where even with infinite (nonphysical) bond stretching, the rings could not be disentangled.

The discovery of tryptorubin A's geometric isomerism in the total synthesis effort prompted a reexamination of its biosynthesis. The original bioinformatic analysis identified 18 biosynthetic gene clusters (BGCs), none of which could be confidently predicted to encode the biosynthesis of tryptorubin A (*12*). The most plausible candidate was a modular nonribosomal peptide synthetase by which the hexapeptide chain would be assembled sequentially by dedicated enzymes. However, the selectivity of the module-encoded adenylation domains did not convincingly match the tryptorubin A peptide sequence, and additional genes involved in the biosynthesis of amino acids that are not incorporated into tryptorubin A were present in the direct vicinity (*22*, *23*). We decided to evaluate other possible biosynthetic origins and thus considered the possibility that tryptorubin A is a ribosomally synthesized and posttranslationally modified peptide (RiPP) that is missed by conventional bioinformatic analysis tools because of its small size, its lack of homology to characterized ribosomal peptides, and the presence of noncanonical tailoring genes involved in carbon-carbon bond formation. Lack of robust genetic tools to modify genes in the native producer limits the degree to which such predictions can be checked against experimental outcomes, but we tested our hypothesis with bioinformatics and analytical tools.

Screening the translated *Streptomyces* sp. CLI2509 genome sequence for the tryptorubin core peptide sequence (Ala-Trp-Tyr-Ile-Trp-Tyr) resulted in a single hit. Close inspection of the unannotated region revealed a ribosomal binding site followed by a transcriptional start site, a putative RiPP precursor gene encoding

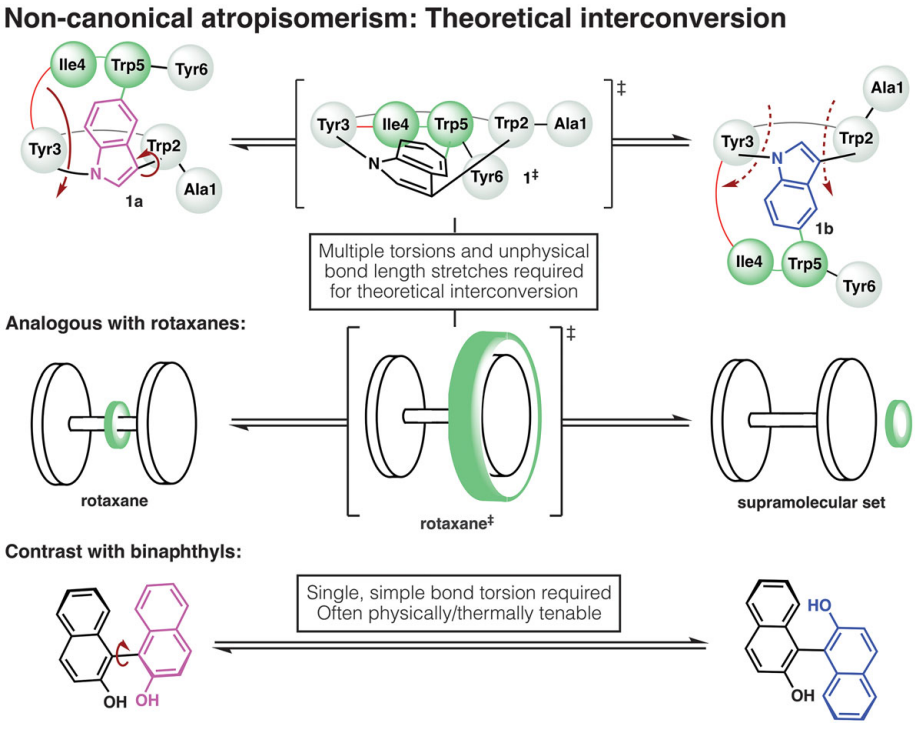


Fig. 4. Graphical thought experiment considering putative interconversion of tryptorubin (1a) and its noncanonical atropisomer (1b). Top: Theoretically, interconversion would require an unphysical inside-out flipping of the molecule, in which one macrocycle passed through the other. Center: This is analogous to atropisomeric inversion of a rotaxane, which would require unphysical stretching of the ring (green) over the dumbbell. Bottom: Such noncanonical atropisomers are contrasted with prototypical atropisomers such as binaphthol, which can interconvert through simple bond torsion.

a 20–amino acid leader, a core peptide, and a stop codon downstream of the core sequence (Fig. 3B and fig. S17). This sequence is followed by a gene encoding a cytochrome P450 enzyme that is likely involved in the formation of the nonproteogenic carbon-carbon and carbon-nitrogen bridges. Although cytochrome P450 enzymes that catalyze carbon-carbon bond formation in ribosomal peptides have not been reported (*24*), analogous carbon-carbon linkages between the aromatic residues in the nonribosomal peptide vancomycin have been shown to be installed by cytochrome P450 enzymes (*25*–*28*). A specific protease in the same vicinity is missing, so an unspecific protease may cleave the cyclized tryptorubin. The lack of a dedicated protease would also explain the presence of tryptorubin B, a congener of tryptorubin A that differs by a terminal alanine residue (*13*). The same gene tandem was identified in the tryptorubin producer *Streptomyces* sp. SPB78 as well as a variety of other *Streptomyces* and more distantly related bacteria (*13*) (figs. S18 and S19). Comparison of this likely tryptorubin locus with other putative producers of tryptorubin-like peptides suggests that the precursor gene and the cytochrome P450 are the only genes conserved in all putative tryptorubin-like BGCs and hence may be sufficient for the biosynthesis of tryptorubin-like metabolites (*13*) (fig. S19). Because we were not able to knock out the putative *trp* BGC in the native producer, we subjected to metabolic analysis *Xanthomonas* sp. Leaf 148, which does not share any other peptide BGC with the reported tryptorubin producers, other than the candidate *trp* locus. Using LC–HRMS spectrometry analysis, we detected a metabolite with the same mass (mass/charge ratio *m/z* of 826.356 $[M+H]^+$) and fragmentation patterns, as revealed by molecular network analysis, similar to those of tryptorubin B (*13*, *29*) (fig. S21). This preliminary work suggests that the tryptorubins are natural products produced by a RiPP biosynthetic pathway, and that a new family of cytochrome P450s is likely responsible for the unusual carbon-carbon and carbon-nitrogen bridges embedded in their structures.

Despite the extensive vernacular to describe regio-, stereo-, and atropisomers, the nuances of molecular shape can be lost within the realm of small-molecule natural product chemistry. Although most practicing synthetic chemists are intimately familiar with the canonical examples of biaryl atropisomerism, the much more complex examples of atropisomerism in polycyclic and mechanically interlocked molecules often remain underexamined. Indeed, the possibility of noncanonical atropisomerism was initially missed during both the isolation and synthesis of tryptorubin A. We present this case as a cautionary tale in structural definition, a demonstration of the power of transferring point chirality to molecular shape, and a reminder that small-molecule organic chemists can greatly benefit from the deep understanding of 3D structure known in the biomacromolecular and supramolecular literature.

REFERENCES AND NOTES

1. D. Remus, E. L. Beall, M. R. Botchan, *EMBO J.* **23**, 897–907 (2004).
2. T. E. Creighton, *Proteins: Structures and Molecular Properties* (Freeman, ed. 2, 1992).
3. J. N. Onuchic, P. G. Wolynes, *Curr. Opin. Struct. Biol.* **14**, 70–75 (2004).
4. J. Roca, *Trends Biochem. Sci.* **20**, 156–160 (1995).
5. Y. Y. Fan *et al.*, *Nat. Commun.* **9**, 3037 (2018).
6. H. L. Frisch, E. Wasserman, *J. Am. Chem. Soc.* **83**, 3789–3795 (1961).
7. K. Zhu, G. Baggi, S. J. Loeb, *Nat. Chem.* **10**, 625–630 (2018).
8. C. Zong, M. J. Wu, J. Z. Qin, A. J. Link, *J. Am. Chem. Soc.* **139**, 10403–10409 (2017).
9. P. J. Canfield *et al.*, *Nat. Chem.* **10**, 615–624 (2018).
10. S. Withoff, S. De Jong, E. G. De Vries, N. H. Mulder, *Anticancer Res.* **16**, 1867–1880 (1996).
11. J. D. Hegemann, M. Zimmermann, X. Xie, M. A. Marahiel, *J. Am. Chem. Soc.* **135**, 210–222 (2013).
12. T. P. Wyche, A. C. Ruzzini, L. Schwab, C. R. Currie, J. Clardy, *J. Am. Chem. Soc.* **139**, 12899–12902 (2017).
13. See supplementary materials.
14. K. Kunz, U. Scholz, D. Ganzer, *Synlett* **15**, 2428–2439 (2003).
15. J. Kim, M. Movassaghi, *J. Am. Chem. Soc.* **133**, 14940–14943 (2011).
16. N. Boyer, M. Movassaghi, *Chem. Sci.* **3**, 1798–1803 (2012).
17. S. D. Rychnovsky, *Org. Lett.* **8**, 2895–2898 (2006).
18. N. Z. Burns, I. N. Krylova, R. N. Hannoush, P. S. Baran, *J. Am. Chem. Soc.* **131**, 9172–9173 (2009).
19. R. K. Spencer, J. S. Nowick, *Isr. J. Chem.* **55**, 698–710 (2015).
20. C. G. Jones *et al.*, *ACS Cent. Sci.* **4**, 1587–1592 (2018).
21. T. Gruene *et al.*, *Angew. Chem. Int. Ed.* **57**, 16313–16317 (2018).
22. K. Blin *et al.*, *Nucleic Acids Res.* **47**, W81–W87 (2019).
23. M. A. Skinnider, N. J. Merwin, C. W. Johnston, N. A. Magarvey, *Nucleic Acids Res.* **45**, W49–W54 (2017).
24. A. Greule, J. E. Stok, J. J. De Voss, M. J. Cryle, *Nat. Prod. Rep.* **35**, 757–791 (2018).
25. B. Hadatsch *et al.*, *Chem. Biol.* **14**, 1078–1089 (2007).
26. D. Bischoff *et al.*, *ChemBioChem* **6**, 267–272 (2005).
27. D. Bischoff *et al.*, *Angew. Chem. Int. Ed.* **40**, 1693–1696 (2001).
28. D. Bischoff *et al.*, *Angew. Chem. Int. Ed.* **40**, 4688–4691 (2001).
29. M. Wang *et al.*, *Nat. Biotechnol.* **34**, 828–837 (2016).
30. E. Pennisi, *Science* **356**, 996 (2017).

ACKNOWLEDGMENTS

We thank L. Pasternack and D.-H. Huang (Scripps Research Institute) for assistance with NMR spectroscopy; A. Rheingold, M. Gembicky, C. Moore, and J. B. Bailey (University of California, San Diego) for assistance with x-ray crystallography; J. Nowick and M. Wierzbicki (University of California, Irvine) for attempts at high-throughput protein crystallography screening; H. M. Nelson, L. J. Kim, and C. Jones (University of California, Los Angeles) for attempts at elucidating structures by MicroED; A. L. Vagstad (ETH Zurich) for assistance in identifying the *trp* BGC; J. Vorholt and M. Bortfeld-Miller (ETH Zurich) for providing *Xanthomonas* sp. Leaf148; J. S. Chen and B. B. Sanchez (Scripps Research Institute) for assistance with HPLC purification and HRMS; and T. P. Wyche and E. Mevers (Harvard University) for useful discussions. **Funding:** Supported by NIH grants GM-118176 (P.S.B.) and R01AT009874 (J.C.); NSF graduate research fellowship 2017237151 and a Donald and Delia Baxter fellowship (S.H.R.); Fujian Juhong Trade & Business Co. postdoctoral fellowship (Y.G.); NIH postdoctoral fellowship F32GM128267 (A.S.W.); and a Swiss National Science Foundation postdoctoral mobility fellowship (E.J.N.H.). **Author contributions:** S.H.R., P.S.B., and Y.G. designed the synthetic routes; S.H.R. and Y.G. performed the chemical synthesis; S.H.R. and P.S.B. identified the putative noncanonical atropisomerism; A.S.W. performed all DFT and molecular modeling computations; E.J.N.H. performed all biosynthetic studies; all co-authors wrote and edited the manuscript. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** Experimental and analytical procedures and full spectral data are available in the supplementary materials. X-ray data and models are available at the Cambridge Crystallographic Data Centre under accession numbers CCDC-1972904 (**7**) and CCDC-1972902 (***epi*-SI-21**).

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/367/6476/458/suppl/DC1
Materials and Methods
Spectral Data
Figs. S1 to S21
Tables S1 to S14
Jmol Structure File
References (*31*–*38*)

View/request a protocol for this paper from *Bio-protocol*.

5 August 2019; accepted 19 December 2019
Published online 2 January 2020
10.1126/science.aay9981

PHASE SEPARATION

Phase separation provides a mechanism to reduce noise in cells

A. Klosin[1]*, F. Oltsch[1,2]*, T. Harmon[1,3], A. Honigmann[1,4], F. Jülicher[2,3,4]†, A. A. Hyman[1,2,4]†, C. Zechner[1,2,4]†

Expression of proteins inside cells is noisy, causing variability in protein concentration among identical cells. A central problem in cellular control is how cells cope with this inherent noise. Compartmentalization of proteins through phase separation has been suggested as a potential mechanism to reduce noise, but systematic studies to support this idea have been missing. In this study, we used a physical model that links noise in protein concentration to theory of phase separation to show that liquid droplets can effectively reduce noise. We provide experimental support for noise reduction by phase separation using engineered proteins that form liquid-like compartments in mammalian cells. Thus, phase separation can play an important role in biological signal processing and control.

Stochasticity in gene expression causes substantial noise in protein concentration, even in genetically identical cell populations grown in the same environmental conditions (*1*–*4*). Despite this noise, living organisms display an extraordinary degree of robustness and exhibit precise spatial and temporal organization.

Liquid-liquid phase separation provides a potential mechanism to reduce noise in protein concentration (*5*–*7*). This is because in a phase-separating system, the concentrations inside and outside the droplets are constrained by thermodynamic laws. When the total concentration of protein changes, the droplets will change in number and size, but the concentration outside of the droplets may be insensitive to these changes (Fig. 1, A and B).

The thermodynamic constraints on coexisting concentrations are well established for macroscopic phase-separating systems at equilibrium. However, phase-separating systems inside cells exhibit mesoscopic noise and are, in general, out of equilibrium. Therefore, whether cells can effectively use phase separation to control protein concentration levels and potentially reduce noise is unclear.

To study under which conditions noise reduction can be effective, we developed a mesoscopic theory that links protein concentration fluctuations to the physics of liquid-liquid phase separation (*8*–*11*). The theory is based on the thermodynamics of a binary mixture segregating into a dilute and droplet phase when the total protein concentration exceeds a threshold value (Fig. 1C and supplementary text 1.1). In our model, we account for nonequilibrium fluctuations due to stochastic synthesis and turnover of protein (Fig. 1D). This system can be described by a master equation, which captures the statistics of the dilute-phase and total protein concentrations, ϕ_+ and $\bar{\phi}$, which can be characterized by their mean $\langle\phi\rangle$ and noise strength $CV^2[\phi] = \sigma^2[\phi]/\langle\phi^2\rangle$, where $\sigma^2[\phi]$ is the variance.

We first considered a situation in which the exchange of protein between phases is much faster than protein synthesis and degradation, which we refer to as quasi-equilibrium. As the mean of the total protein concentration $\langle\bar{\phi}\rangle$ increases and approaches the threshold value ϕ^*, droplets begin to form while the noise strength $CV^2[\phi_+]$ of the dilute-phase concentration ϕ_+ starts to decline (Fig. 1E). For larger mean total concentration, the noise strength settles at a minimum with approximately Poissonian noise (supplementary text 1.1.3). We further show that around the minimal noise strength, concentration fluctuations due to protein expression are suppressed, and the remaining fluctuations are predominantly thermal fluctuations of the phase-separating system (supplementary text 1.2.2). Therefore, as long as phase separation is much faster than protein synthesis and degradation, droplets can reduce noise in protein concentration down to the Poisson limit.

We next considered situations in which the time scales of protein expression and phase separation approach each other. Figure 1F shows the noise strength for three different time-scale ratios of protein diffusion and turnover (τ_D/τ_p) versus mean total concentration. As before, the noise strength in the dilute phase first approaches a minimum as the mean total concentration increases. However, beyond this minimum, the noise strength starts to increase. This is because for highly expressing cells, the rate of protein synthesis is fast compared with the time it takes for a protein to diffuse into a droplet. As a consequence, proteins can accumulate in the dilute phase, which hampers the system's ability to reduce noise at high protein-synthesis rates. The minimal noise strength depends on the time scales of protein diffusion and turnover and is generally above the Poisson limit (Fig. 1G). Thus, noise reduction by phase separation is predicted to be most effective for long-lived and fast-diffusing proteins.

To test the concept of noise reduction inside cells, we used a recombinant 2NT-DDX4^{YFP} protein that phase-separates in vitro (fig. S21). We expressed 2NT-DDX4^{YFP} inside HeLa cells and examined protein concentration and spatial distribution using live-cell microscopy (Fig. 2A). Transient transfection generated a broad range of protein expression levels because of large variability in plasmid transfection efficiency. Similar to the previously described (*12*) DDX4^{YFP}, the 2NT-DDX4^{YFP} variant formed heterologous compartments inside nuclei of transfected HeLa cells (fig. S22). The 2NT-DDX4^{YFP} droplets fused with each other, coarsened over time, and showed high internal recovery after photobleaching, which together confirmed their liquid-like behavior (*13*) (fig. S23). Pixel fluctuation analysis (*14*) showed no evidence for small clusters below the diffraction limit, which could lead to an overestimation of dilute-phase concentration (materials and methods and fig. S24). We used statistical methods (*15*) to estimate the parameters of our nonequilibrium model from time-lapse 2NT-DDX4^{YFP} expression data (Fig. 2B, materials and methods, and supplementary text 2.2).

We next determined the mean concentration and noise strength of 2NT-DDX4^{YFP} in more than 10,000 cells 24 hours after transfection (Fig. 2C) and performed a comparison with our theory. To achieve this, we randomly selected subpopulations from the total pool of cells with prescribed mean total protein concentration and noise strength. For each subpopulation, we then quantified the mean and noise strength of dilute-phase 2NT-DDX4^{YFP} concentration. The results show that protein concentration noise is indeed reduced in the dilute phase as soon as the mean 2NT-DDX4^{YFP} concentration approaches the threshold concentration (Fig. 2D). To compare noise reduction with theory, we used subpopulations of cells with mean total concentration and noise strength that correspond to the statistics of total protein concentration in our model (supplementary text 2.3). The data reveal the features of noise reduction as predicted by our nonequilibrium theory—in particular, a minimum of noise strength at a particular mean total concentration (Fig. 2D). In the experimental data, noise reduction is less pronounced than in the theory (compare green dotted and solid lines), possibly because of additional cell-to-cell differences within

[1]Max Planck Institute of Molecular Cell Biology and Genetics, 01307 Dresden, Germany. [2]Center for Systems Biology Dresden, 01307 Dresden, Germany. [3]Max Planck Institute for the Physics of Complex Systems, 01187 Dresden Germany. [4]Cluster of Excellence Physics of Life, TU Dresden, 01062 Dresden, Germany.
*The authors contributed equally to this work.
†Corresponding author. Email: julicher@pks.mpg.de (F.J.); hyman@mpi-cbg.de (A.A.H.); zechner@mpi-cbg.de (C.Z.)

subpopulations that are not captured by our model (supplementary text 2.4).

The condensates formed by 2NT-DDX4^{YFP} dissolve during mitosis (Fig. 3A), similar to other membraneless organelles (*16*). To test whether dissolving the condensates leads to an increase in noise strength, we followed 141 individual droplet-containing cells through mitosis (fig. S25) and quantified the fluorescence intensity in the dilute phase before, during, and after mitosis. We found that droplet dissolution in mitosis is associated with a more than twofold increase of noise strength in the dilute phase (Fig. 3, B to D, and movie S1). In most postmitotic cells, droplets reform, which again leads to a reduction of noise strength in the dilute nucleoplasmic phase. These data strongly suggest that the lower noise levels in the dilute phase observed during interphase are indeed governed by phase separation and, furthermore, that concentration

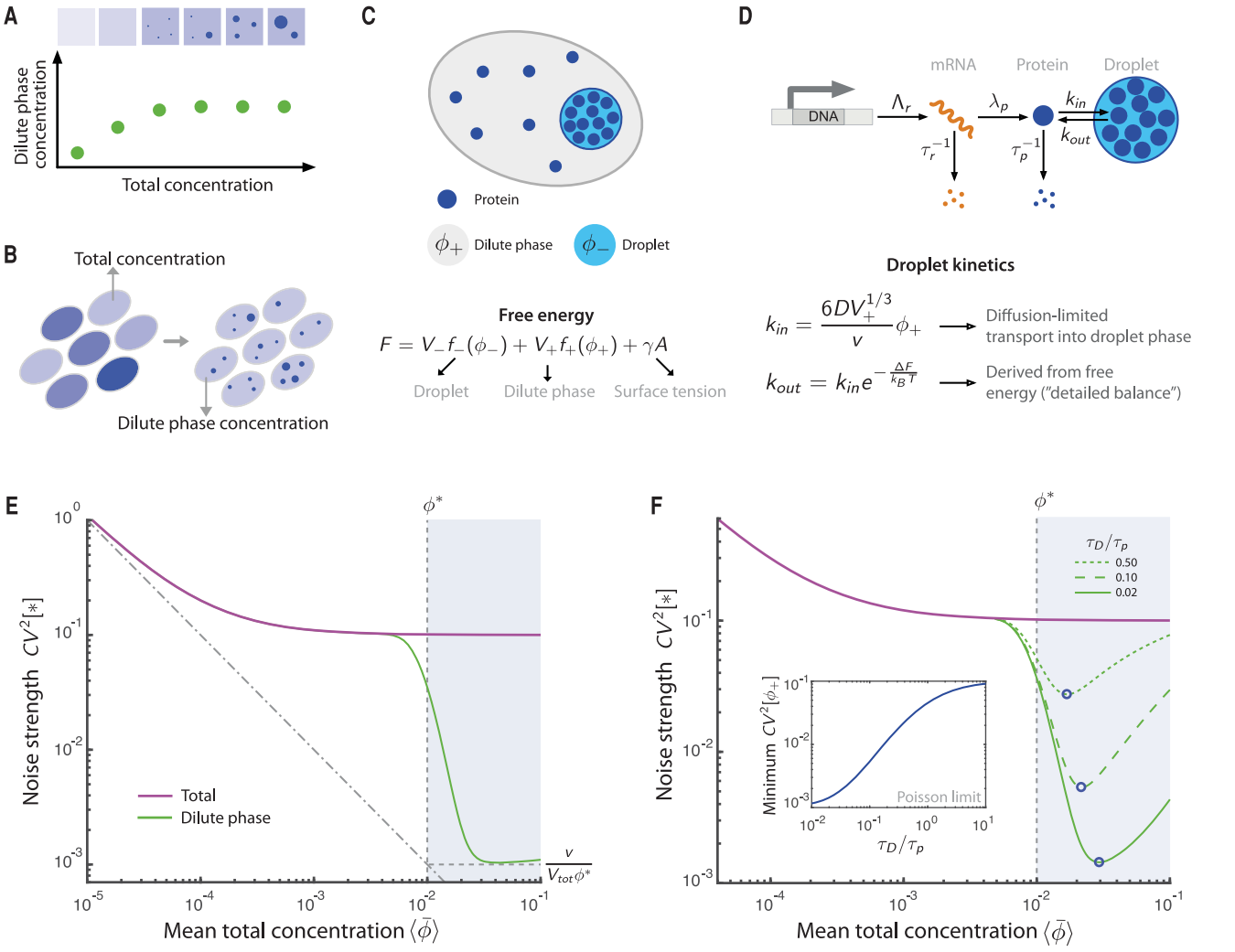


Fig. 1. Concentration fluctuations in a mesoscopic model of phase separation. (**A**) In a simple phase-separating system, the dilute-phase concentration saturates when the total concentration exceeds a threshold. In this regime, droplets can vary in size and number, but the dilute-phase concentration is insensitive to changes in total concentration. (**B**) Phase separation could reduce variability in protein concentration (represented by color intensity) by forming liquid compartments of variable number and size. (**C**) Mesoscopic droplet model of a binary mixture at equilibrium. The mixture consists of a solvent and a protein, which can phase-separate into a dilute and a dense droplet phase. We normalize concentrations c by molecular volume v of the protein and use normalized concentrations (volume fractions) $\phi = vc$. We denote the normalized concentrations in the dilute and droplet phase and the total system as ϕ_+, ϕ_-, and $\bar{\phi}$, respectively. Phase separation is governed by a free energy, which has contributions from both coexisting phases as well as the separating interface, which has a surface tension (supplementary text 1.1.1). (**D**) Droplet model in the presence of protein production and turnover. Fluctuations in total protein concentration are described by a gene expression model accounting for stochastic production and degradation of mRNA and protein (*27*). Additionally, cell-to-cell variability in the transcription rate is taken into account. We capture the dynamics of protein phase separation by stochastic exchange of protein between the dilute and droplet phases. We consider the transport of molecules into the droplet phase (k_{in}) to be diffusion limited. The rate of the corresponding reverse transition (k_{out}) follows from detailed balance (supplementary text 1.1.5). We consider the average protein lifetime τ_p to be the same in both phases. (**E**) Noise strength of total and dilute-phase concentration as a function of mean total concentration in the quasi-equilibrium situation. The minimum of the noise strength in the dilute phase is approximately given by $CV^2[\phi_+] \approx v/(\phi^* V_{tot})$, corresponding to Poissonian noise (dashed horizontal line). The dashed vertical line indicates the threshold concentration ϕ^*. Parameter values are given in table S1. (**F**) Influence of time scales on noise reduction. Noise strength is shown as a function of mean total concentration for three different ratios of the protein diffusion time τ_D and protein lifetime τ_p. The blue circles indicate the minima of the noise strength. (Inset) Minimal noise strength shown as a function of τ_D/τ_p. Parameter values are given in table S2.

noise in cells can be controlled by regulating phase separation.

To test whether noise reduction occurs in an endogenous system, we examined the nucleolus. The nucleolus consists of three coexisting phases that vary in their material properties (*17*). The outermost phase, known as the granular component (GC), exhibits liquid-like properties and is enriched in a protein called nucleophosmin (NPM1), which forms liquid-like droplets in vitro (*17*). Using CRISPR-Cas9, we tagged native NPM1 with mNeonGreen inside HCT116 cells and measured fluorescence intensities in the dilute nucleoplasmic phase of individual cells, which coexists with the GC phase of the nucleolus (Fig. 3, E to H). Similar to the 2NT-DDX4^{YFP} experiments, we found that the dissolution of the GC phase in mitosis is associated with an approximate twofold increase in noise strength in the dilute phase (Fig. 3H and movie S2). Our data suggest that for NPM1 at native expression levels, concentration noise is reduced in the presence of phase-separated condensates.

We used experiment and theory to show that the formation of liquid condensates can reduce concentration noise of proteins participating in phase separation. Many liquid compartments have been identified recently, but their biological function is often

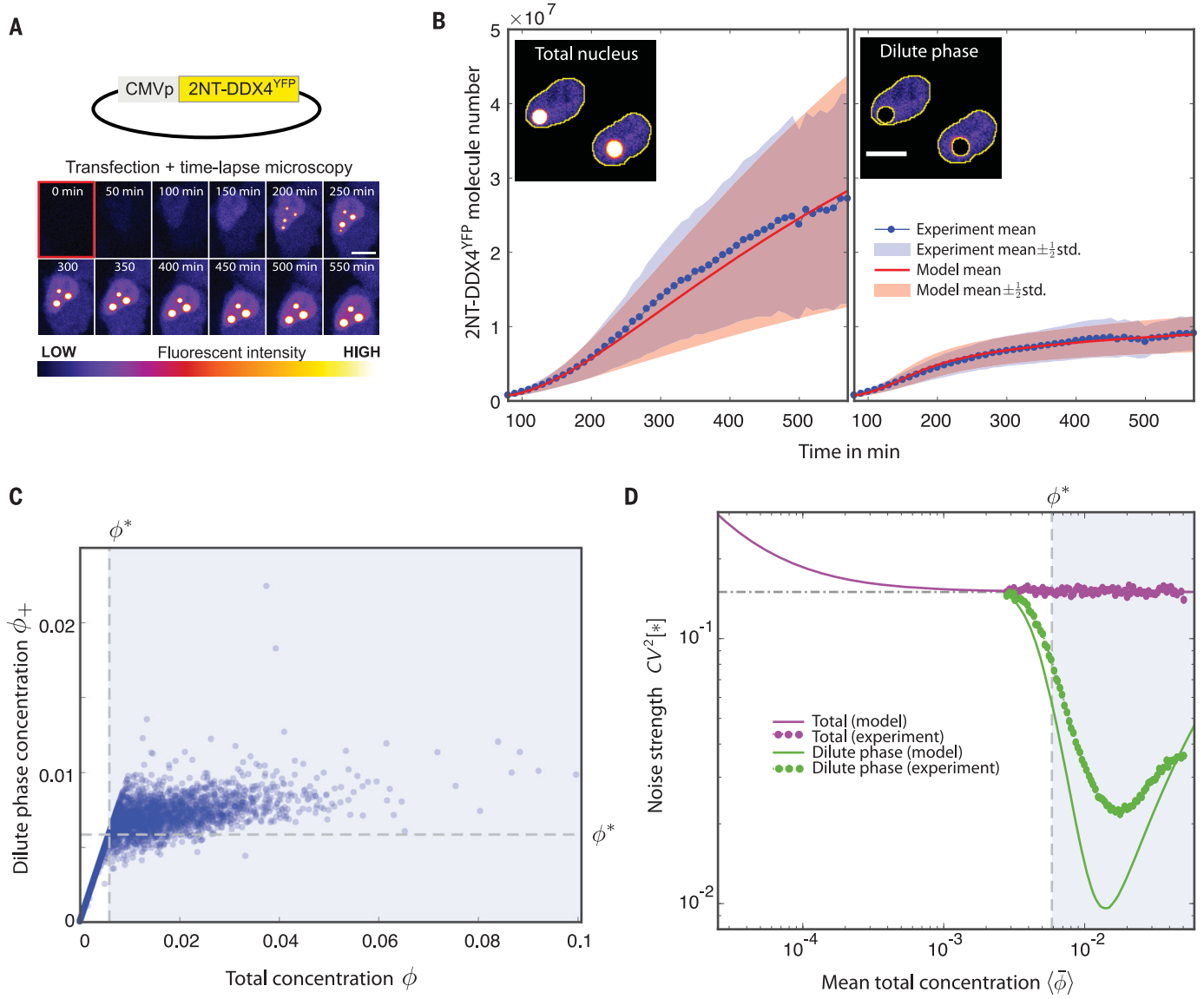


Fig. 2. Reduction of 2NT-DDX4^{YFP} concentration noise in HeLa cells. (**A**) Time-lapse microscopy of a typical cell initiating expression of 2NT-DDX4^{YFP} and phase separation after transfection and plasmid incorporation. CMVp, cytomegalovirus promoter. (**B**) Mean and variance of 2NT-DDX4^{YFP} total (left) and dilute-phase (right) molecule numbers as a function of time. Model parameters were estimated from the time-lapse data (blue) by using statistical methods (supplementary text 2.2). Corresponding model behaviors are shown in red. Shaded areas correspond to half a standard deviation (std.) above and below the mean. The estimated model parameters are given in tables S4 and S5. (**C**) Scatterplot of total versus dilute-phase 2NT-DDX4^{YFP} concentration quantified in more than 10^4 cells 24 hours after transfection. The threshold concentration ϕ^* was estimated as the average of the dilute-phase concentration of the 5% cells with smallest droplets (dashed lines) (supplementary text 2.3). (**D**) Noise strength in dilute-phase (green) 2NT-DDX4^{YFP} concentration as a function of the mean total 2NT-DDX4^{YFP} concentration determined by generating subpopulations with imposed statistics of total concentration (purple). The experimental data are compared with the noise strength predicted by the model by using the parameters estimated from the time-lapse data shown in (B) (green solid line) (supplementary text 2.3).

unclear (*5*, *18*, *19*). Our results suggest that some of them could serve to maintain protein levels within narrow ranges of concentration. For instance, splicing speckles could stabilize splicing activity by controlling the dilute-phase concentration of splicing factors (*20*).

Our work discusses the effects of phase separation in the context of active molecular turnover and provides insights into the interplay between nonequilibrium fluctuations and phase separation. Noise reduction predicted by our simple model was similar to the experimentally measured values for the 2NT-DDX4^{YFP} system. In our theory, we have not considered systems with multiple phase-separating components, which provide more complex thermodynamic constraints on concentrations (*21*, *22*). Extending this theory to multicomponent situations will, therefore, be an interesting subject for future research. Moreover, active

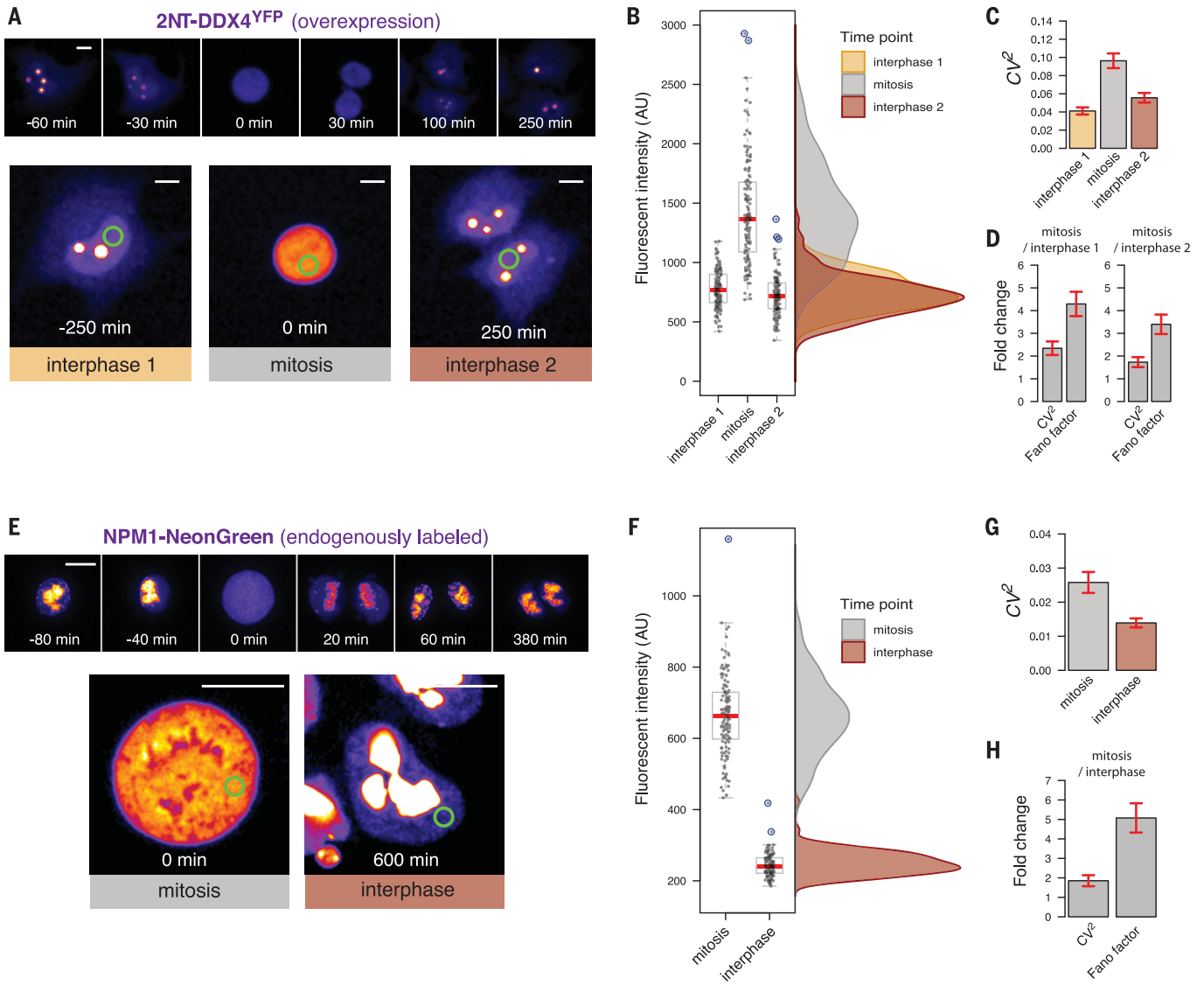


Fig. 3. Droplet dissolution during mitosis results in increased noise in protein concentration. (**A**) HeLa cell expressing 2NT-DDX4^{YFP} from a transiently transfected plasmid imaged during mitosis. (**B**) Quantification of mean 2NT-DDX4^{YFP} fluorescence intensity in the dilute phase [green circle in (A)] of the same 141 cells before, during, and after mitosis. Distribution of the measured values is displayed as a density histogram. AU, arbitrary units. (**C**) Noise in the protein concentration of the dilute phase increased from 0.038 (±0.004) in interphase to 0.081 (±0.008) in mitosis and reduced down to 0.046 (±0.005) in subsequent interphase. (**D**) Fold increase in the noise of 2NT-DDX4^{YFP} dilute-phase concentration ϕ_+ during mitosis compared with preceding and subsequent interphase. Bar plots represent ratios of the noise strength $CV^2[\phi_+]$ for each group. For comparison, the increase in noise has been quantified also in terms of the Fano factor ($FF[\phi_+] = \sigma^2[\phi_+]/\langle\phi_+\rangle$), another common measure to quantify fluctuations. (**E**) HCT116 cell expressing NPM1 labeled endogenously with NeonGreen imaged during mitosis. (**F**) Quantification of mean NPM1-NeonGreen fluorescence intensity in the dilute phase of 127 cells imaged during and after mitosis [green circle in (E)]. Distribution of the measured values is displayed as a density histogram. (**G**) Noise in the protein concentration of the dilute phase decreased from 0.026 (±0.003) in mitosis to 0.014 (±0.001) in subsequent interphase. (**H**) Noise of NPM1-NeonGreen dilute-phase concentration during mitosis and subsequent interphase. Bar plots represent ratios of $CV^2[\phi_+]$ and $FF[\phi_+]$. Scale bar is 10 μm. Bottom panels in (A) and (E) were set to high contrast to demonstrate the change of intensity in the dilute, nucleoplasmic phase. Boxplots indicate median (red bar) and first and third quartiles. Lower and upper whiskers extend to 1.5 times the interquartile range from the first and third quartile, respectively. Outliers that fell 2.5 standard deviations away from the mean (blue circles in boxplots) were excluded from the analysis. Error bars in bar plots represent standard errors of the measurements calculated with bootstrapping.

chemical processes such as posttranslational modifications influence endogenous phase-separated compartments (*23*) and could contribute to noise in notable ways.

Previous studies have proposed that spatial compartmentalization of molecules can be a potential mechanism to enhance the robustness of biological systems (*24*–*26*). For instance, delayed nuclear export of transcripts can lead to reduced variability in cytoplasmic RNA (*24*), and protein clustering can enhance the robustness of biological switches (*25*) and subcellular gradient formation (*26*). Understanding how membrane- and non–membrane-bound compartments affect noise in living systems remains a substantial challenge.

REFERENCES AND NOTES

1. M. B. Elowitz, A. J. Levine, E. D. Siggia, P. S. Swain, *Science* **297**, 1183–1186 (2002).
2. Y. Taniguchi *et al*., *Science* **329**, 533–538 (2010).
3. J. R. S. Newman *et al*., *Nature* **441**, 840–846 (2006).
4. E. M. Ozbudak, M. Thattai, I. Kurtser, A. D. Grossman, A. van Oudenaarden, *Nat. Genet.* **31**, 69–73 (2002).
5. S. F. Banani, H. O. Lee, A. A. Hyman, M. K. Rosen, *Nat. Rev. Mol. Cell Biol.* **18**, 285–298 (2017).
6. T. Stoeger, N. Battich, L. Pelkmans, *Cell* **164**, 1151–1161 (2016).
7. A. S. Holehouse, R. V. Pappu, *Biochemistry* **57**, 2415–2423 (2018).
8. A. A. Hyman, C. A. Weber, F. Jülicher, *Annu. Rev. Cell Dev. Biol.* **30**, 39–58 (2014).
9. M. L. Huggins, *J. Phys. Chem.* **46**, 151–158 (1942).
10. P. J. Flory, *J. Chem. Phys.* **10**, 51–61 (1942).
11. C. A. Weber, D. Zwicker, F. Jülicher, C. F. Lee, *Rep. Prog. Phys.* **82**, 064601 (2019).
12. T. J. Nott *et al*., *Mol. Cell* **57**, 936–947 (2015).
13. C. P. Brangwynne *et al*., *Science* **324**, 1729–1732 (2009).
14. M. A. Digman, R. Dalal, A. F. Horwitz, E. Gratton, *Biophys. J.* **94**, 2320–2332 (2008).
15. C. Zechner *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **109**, 8340–8345 (2012).
16. A. K. Rai, J. X. Chen, M. Selbach, L. Pelkmans, *Nature* **559**, 211–216 (2018).
17. M. Feric *et al*., *Cell* **165**, 1686–1697 (2016).
18. A. H. Fox, S. Nakagawa, T. Hirose, C. S. Bond, *Trends Biochem. Sci.* **43**, 124–135 (2018).
19. A. A. Hyman, C. P. Brangwynne, *Dev. Cell* **21**, 14–16 (2011).
20. D. Berchtold, N. Battich, L. Pelkmans, *Mol. Cell* **72**, 1035–1049.e5 (2018).
21. J. A. Riback *et al*., bioRxiv 809210 [Preprint]. 22 October 2019. https://doi.org/10.1101/809210.
22. J. M. Choi, F. Dar, R. V. Pappu, *PLOS Comput. Biol.* **15**, e1007028 (2019).
23. M. Hofweber, D. Dormann, *J. Biol. Chem.* **294**, 7137–7150 (2019).
24. N. Battich, T. Stoeger, L. Pelkmans, *Cell* **163**, 1596–1610 (2015).
25. B. M. Slepchenko, M. Terasaki, *Mol. Biol. Cell* **14**, 4695–4706 (2003).
26. T. E. Saunders *et al*., *Dev. Cell* **22**, 558–572 (2012).
27. J. Paulsson, *Nature* **427**, 415–418 (2004).

ACKNOWLEDGMENTS

We acknowledge support from the Genome Engineering Facility and Light Microscopy Facility at the MPI-CBG. We thank T. Nott for sharing the recombinant DDX4^{YFP} plasmid. We thank C. Modes, A. Nadler, S. Grill, P. Tomancak, J. Schmiedel, and J. Brugues for critical comments on the manuscript.
Funding: This project was supported by funding from the Max Planck Society. A.A.H. and C.Z. were funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) (ZE1216/1-1; HY3/6-1). A.K. was supported by an EMBO Long-Term Fellowship (ALTF 1069-2017). F.J., A.A.H., and C.Z. were supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany´s Excellence Strategy (EXC-2068; 390729961) Cluster of Excellence Physics of Life of TU Dresden. **Author contributions:** All authors contributed to data analysis and interpretation. A.K. conceived and performed experiments. A.K. performed image analysis with input from A.H. F.O., F.J., and C.Z. developed the theoretical approach. F.O. performed the numerical and theoretical analyses. A.H. performed the pixel fluctuation analysis. T.H. performed the lattice Monte Carlo simulations. A.K., F.O., F.J., A.A.H., and C.Z. wrote the manuscript with input from T.H. and A.H. F.J., A.A.H., and C.Z. conceived the study. **Competing interests:** A.A.H. is cofounder and member of the scientific advisory board of Dewpoint Therapeutics. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper or the supplementary materials.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/367/6476/464/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S26
Tables S1 to S5
References (*28*–*35*)
Movies S1 and S2

View/request a protocol for this paper from *Bio-protocol*.

19 October 2018; resubmitted 18 September 2019
Accepted 6 December 2019
10.1126/science.aav6691

Mona Minkara, assistant professor at Northeastern University and visually impaired, examines a molecular model.

INCLUSIVITY FOR ALL: HOW TO MAKE YOUR RESEARCH GROUP ACCESSIBLE

Creating an inclusive culture in your research group requires installing mechanisms for your whole team to be successful. And when it comes to accessibility, by ensuring everyone has what they need to be their most creative and productive, you ensure a win for science AND your scientific workforce. **By Alaina G. Levine**

Mona Minkara, an assistant professor at Northeastern University, Boston, Massachusetts, is a bioengineer, a world traveler and adventurer, writer, and speaker. She also happens to be blind. Her disability provides what she refers to as her "unseen advantage" in science.

"I cannot see the data, so I have had to come up with other ways to conduct research," Minkara says. As a graduate student, she ran a molecular dynamics simulation to observe and analyze the movement of a protein over 500 nanoseconds. The first step in this process is typically to watch a video of the protein's actions, "but I couldn't do it. Someone tried to explain it to me, but it meant nothing, and I didn't know what to ask," she recalls. But Minkara wasn't deterred. She devised a way to plot the data, so she could interpret the protein's movement mathematically, as opposed to visually. This approach turned out to be extremely meaningful to her as a scientist and to the science she was pursuing: "I observed patterns in the data that were missed by sighted people who were just looking at the trajectory."

As Minkara's example demonstrates, it is critical for science to collect and examine data in numerous ways and to adopt multiple perspectives when endeavoring to solve a problem. To produce the best science, to tackle the toughest problems, and to create and innovate interventions that advance our understanding of the universe, everyone's voice, view, and brain needs to be at the table. Institutions, companies, and governments—as well as principal investigators (PIs)–need to ensure that everyone can be successful in science.

Indeed, accessibility should be a part of your lab or group's core values and actions. "Accessibility is not a one-off," says **Sarah Lewthwaite**, United Kingdom Research and Innovation Future Leaders Fellow, Centre for Research in Inclusion, at the University of Southampton, UK. "It should be a thread that will run through all your research planning." The stakes couldn't be higher, she notes, given the data: According to the UK's Office for National Statistics, 19% of working-age adults in the UK have a disability (*1*). A recent *Harvard Business Review* article states that in the United States, "30% of the professional workforce fits the current federal definition of having a disability … and only 39% of employees with disabilities have disclosed to their manager. Even fewer have disclosed to their teams (24%) and [human resources] (21%)" (*2*).

Asking the right questions

Inclusivity as it relates to accessibility starts with adopting certain tenets. "100% perfect accessibility is extraordinarily difficult to achieve," says **Jesse Shanahan**, an astronomer, data scientist, and freelance accessibility consultant, with clients ranging from individual faculty to university departments, associations, and conferences. "We can't conceive of every single disability or need someone has, so one of the best approaches to take with that in mind is to share as much information with your lab employees as possible," she adds. **cont.>**

Upcoming features

Postdocs—August 21 ■ Faculty—September 18 ■ Top Employers—October 30

Young Investigator Grant and Postdoctoral opportunities in São Paulo, Brazil

The **São Paulo Research Foundation (FAPESP)**, one of the leading Brazilian agencies dedicated to the support of research, offers opportunities for young researchers of any nationality to work in excellence centers in São Paulo, Brazil:

The **Young Investigator Grant** support scientists who have had a few years of postdoctoral experience and demonstrated research leadership capabilities to lead the establishment of new research groups in São Paulo, Brazil. The grants, with a duration of 5 years, include a fellowship for the principal investigator plus funds for research equipment, consumables, specialized services, travel and fellowships for undergraduate and graduate students. Candidates of any nationality are welcome and proposals can be submitted in English. FAPESP offers assistance to interested parties in finding a host institution. The selected candidates will start and lead their own research groups working in internationally competitive themes.

FAPESP Postdoctoral Fellowships support researchers with a recent doctorate degree and a promising research track record. Candidates from any nationality are welcome. The duration of the fellowship is of up to 3-years (fellowships associated to FAPESP's 5-year grants and Center grants can have a 4-year duration). FAPESP postdoctoral fellowship holders can request an additional fellowship (travel plus stipend) to work for up to one year in a qualified research laboratory outside Brazil in a project that will enhance the performance of the research they are performing in Brazil. The fellowships are offered through calls for applications issued internationally and announced at http://www.fapesp.br/oportunidades/ and in relevant research journals. Postdoctoral fellowships can also be requested at any time by Principal Investigators in São Paulo who identify promising candidates.

MORE INFORMATION
www.fapesp.br/en/6251
www.fapesp.br/en/postdoc

FAPESP
SÃO PAULO RESEARCH FOUNDATION

RUA PIO XI, 1500 • ALTO DA LAPA 05468-901
SÃO PAULO, SP, BRAZIL
PHONE: +55-11-3838-4000 • WWW.FAPESP.BR

For example, you can offer information about the location of accessible entrances and bathrooms and ensure that food is labeled to correspond with common allergies. If you are conducting a social event offsite, such as at a restaurant, you can call ahead to make sure it has accessible entrances, menus in braille or large-print, and quieter tables—and inform your staff that you have already researched this. You can also write your research group's manual to include resources on campus, such as the disability resource center.

Furthermore, when crafting a lab policy towards accessibility for all, take a multi-level approach, advises Shanahan. Your priority should be to remove as many barriers to success as possible, which could include something as strategic as policy changes—such as enacting and enforcing anti-harassment policies—or something as simple as supporting your team to have time off for doctor appointments or a dedicated quiet space. According to Shanahan, accessibility is a human right under the UN Declaration of Human Rights, and in the United States, it is a legal requirement. That said, the ability to ask for accommodation is often a privilege. She suggests including the following wording in a lab manual: "I understand that getting a diagnosis or access to accomodation is a privilege, and if this is something you are having challenges with, please come talk to me. We can potentially work out a way for our lab to be accomodating."

One of Shanahan's most meaningful professional experiences occurred when she was hired as a data scientist at consulting firm Booz Allen Hamilton. Her supervisor didn't request her to disclose her disability; rather, he inquired "What do you need to do your best work?" This question showed that the organization wanted to help her be her most successful, and shifted the burden of requesting accommodations from employee to employer, thereby empowering the employee. Shanahan was "taken aback" when this was asked of her, because it had never happened before. But "it instantaneously made me feel comfortable asking for things." She urges researchers to utilize this language as the framework for their inclusivity practices. "This is an incredible move that a PI or advisor can make in their lab," she says.

Providing a nurturing culture

Ultimately, individual actions that are taken in a lab, while critical, are not going to lead to a more inclusive community if there is no foundational culture that nurtures that inclusivity. "It is a lot easier to put a ramp in front of a building than to get people to change the culture of a workplace. Yet it is the culture, not the [unbuilt] ramp, that will drive people away," says Shanahan. Most importantly, and in service to the culture, she says, make all of these opportunities available to everyone, even for people who don't have an official accommodation. "Whether it's a lab or a fancy tech company, people need to be empowered to make decisions that support their health and work–life balance, and this could include offers of telecommuting and flexible work schedules," she says.

Another crucial element of the workplace culture is that of flexibility. "Yes, experiments can be rigid processes, but when creating an environment that is more inclusive, it's not the 'what' that should be different, it's the 'how'," says **Meg O'Connell**, CEO and founder of Global Disability Inclusion, a consulting firm that helps organizations understand how to be disability inclusive, based in St. Augustine, Florida. "How someone gets there and does the experiment should be built on flexibility. ... The most important aspect is how to accommodate everyone who comes into your classroom or lab, so they can get the most out of their experience in the learning environment." This could involve team-members performing the experiment with someone who needs accommodation, and confirming that there is space for people who use mobility aids to maneuver throughout the lab.

"People might not believe they have a disability, but they do have a need," says Shanahan. Respecting "that needs are fluid and change ... is a way to provide those needs without someone identifying themselves as disabled." For example, when Shanahan was in graduate school, her grandfather passed away. She notes that while this wasn't a disability, it's the kind of situation that has distinct parallels with the idea of creating an inclusive atmosphere for people with disabilities. Because she was in a research group that supported success of the team in all ways, she was able to take a month off from work to grieve. "Having that paradigm shift respects that these people coming into your labs are not robots or automatons. We have life events, families, and illnesses. A flexible, accessible work environment can shift to accommodate people in all phases of their life."

cont.>

PHOTO: TIM COBURN

TIP BOX

- "Nothing about us, without us."
- Don't assume.
- Know you won't be able to have a catch-all solution.
- Ask your team: "What do you need to be successful?"
- Share information from resources, such as your institution's disability resource center.
- Recognize that not everyone wants to or has the privilege to disclose their disability.
- Remove as many obstacles as possible without requiring disclosure.
- Read articles on accessibility written by people with disabilities.

Featured participants

The Bendy Biologist
bendybiologist.com

Booz Allen Hamilton
www.boozallen.com

Global Disability Inclusion
www.globaldisabilityinclusion.com

Jesse Shanahan
enceladosaur.us

Northeastern University
www.northeastern.edu

University of Michigan
umich.edu

University of Southampton
www.southampton.ac.uk

Amy-Charlotte Devitz, a Master's student at the University of Michigan, Ann Arbor, conducts field work and views her disability as a "superpower."

PHOTO: ALICIA DEVITZ

Recognizing that disability is a different ability

Since Minkara began her faculty job at Northeastern in August 2019, she has been mindful of building a culture that allows everyone to be successful—in any scholarly endeavor, whether it involves research, management, team building, or communication. For example, when her protégés are going to give a talk, she requires them to provide their slides to her in advance. "When giving a presentation they must be very verbally explicit and describe everything," she explains. "I want to make sure I understand what they are saying—but this [also] helps them be better presenters and communicators." This task extends beyond just improving one's ability to efficiently articulate ideas. "It's a skill," Minkara says. "We are propagating a change in the field—they are learning how to give a presentation that is inherently more inclusive."

Amy-Charlotte Devitz, a Master's student in the University of Michigan Department of Ecology and Evolutionary Biology, Ann Arbor, has learned to see the advantages in her disability, which requires the use of a wheelchair. Her blog, "The Bendy Biologist," profiles all of her research in animal behavior in urban ecosystems, and has covered a broad range of topics at the intersection of disability, science, chronic illness, and academia, such as accessibility tips for those planning scientific conferences and how to approach writing an accessibility statement for a course syllabus. When she started using a wheelchair six years ago due to a previously undiagnosed genetic connective tissue disorder, she was told that science would not be for her. "It's become a personal mission, especially for young and upcoming scientists, to make them realize that disability doesn't have to be a barrier in a STEM field, and in a lot of ways, it can be really beneficial," she says. "You have to problem solve and adapt and figure out new ways to do things. The world is inherently inaccessible, so adapting to the challenges this presents is a skill you need in science—because nothing goes according to plan. In research, this skill translates well."

In fact, Devitz views her disability as a "superpower" that has made her a stronger scientist. "Experiencing the world through the eyes of someone with a disability, you are forced to look at things in [greater] detail and with different perspectives that an able-bodied person wouldn't see," she says. "[My disability] has sharpened my ability to pick out these finer details, and when issues arise in doing fieldwork, to work around challenges with greater ease than when I first started." Her decision to pursue research in urban environments is a direct result of her mobility concerns. "I don't think I would have ever discovered my passion for mixing animal behavior and urban ecology had I not been forced to question what kinds of field work would be compatible with my wheelchair," she explains.

For PIs who have not provided accommodations before, launching a research group with inclusive accessibility in mind is new territory. "Be open to those conversations," advises O'Connell. "Get comfortable with that uncomfortable feeling that you're not an expert and that you have to find something you may not be familiar with to make this work for the student."

She adds, "It is so important to ask the person with the disability what works best for them and not make assumptions. They have encountered things. They know what the workarounds are." And always consider the guiding principle in the disability community, "nothing about us, without us," says O'Connell. "This means you should not make decisions for people with disabilities without including them in the process. No one should assume what accommodations will work for a person with a disability without asking them first what would be most useful."

Inclusive accessibility is about giving everyone not just a voice in science, but the tools and methods to be successful in using that voice. "Make sure you don't have a very narrow concept of what success looks like," says Shanahan. "Some students might take longer or need a break, and this doesn't mean they are not successful. Support them and reassure them that they can be successful." Adds Minkara, "We get so caught up in the whole competition and ambition, and might think of someone who is different as slowing us down … This is such a negative value that is hindering us in all areas of science. My advice is to work with students of all abilities because it will pay off on the end, even if it takes a little more time."

References

1. Office for National Statistics, "Measuring disability: comparing approaches," August 6, 2019, available at www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/disability/articles/measuringdisabilitycomparingapproaches/2019-08-06#disability-prevalence-by-age"www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/disability/articles/measuringdisabilitycomparingapproaches/2019-08-06#disability-prevalence-by-age.
2. P. Jain-Link, J. T. Kennedy, "Why People Hide their Disabilities at Work," *Harv. Bus. Rev.*, June 3, 2019, available at https://hbr.org/2019/06/why-people-hide-their-disabilities-at-work" hbr.org/2019/06/why-people-hide-their-disabilities-at-work.

Alaina G. Levine is a science writer, science careers consultant, professional speaker, and author of Networking for Nerds *(Wiley, 2015).*

UCONN

INSTITUTE FOR SYSTEMS GENOMICS

Faculty Position at Any Rank

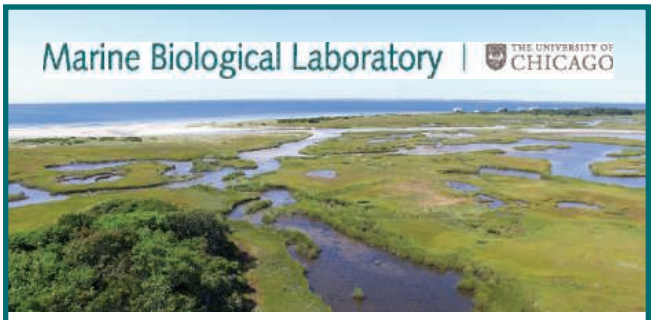
The Institute for Systems Genomics (ISG) at the University of Connecticut (UConn) invites applications for an open-rank (Assistant, Associate or Full Professor) tenure-track faculty position at the Storrs campus. The successful candidate is expected to bring a transformative, innovative, cross-disciplinary, and high-impact research program in genomics to UConn while complementing existing strengths across the departments within the College of Liberal Arts and Sciences participating in this search, Ecology and Evolutionary Biology, Marine Sciences, Molecular and Cell Biology, and Physiology and Neurobiology. The ISG, https://isg.uconn.edu/, draws upon the research strength of 10 schools and colleges at UConn and the nearby Jackson Laboratory for Genomic Medicine (https://jax.org/about-us/locations/farmington) and currently is comprised of 148 members with a collective grant portfolio of over $235 million. Located at one of the nation's premier public research universities, UConn's ISG offers world class core facilities, including several centers established in support of genomics research such as the Center for Genome Innovation (https://cgi.uconn.edu/) and the Computational Biology Core (https://bioinformatics.uconn.edu/).

The successful candidate is expected to develop and sustain an externally-funded research program through collaboration and engagement with the ISG research community, broaden participation among under-represented groups, and contribute to an inclusive culture on campus and in the laboratory. The candidate will advise and mentor students and postdoctoral fellows in research, outreach, and professional development and offer innovative course content in genomics. Minimum qualifications include a Ph.D. in a relevant subject, postdoctoral experience, a track record of research publications in their field, and background that provides preparation for teaching excellence at the undergraduate or graduate levels. Candidates that will establish or bring a research program incorporating innovative and interdisciplinary approaches in genomics complementing the existing strengths of the ISG and contribute to the diversity and excellence of the learning experience and academic community through research, teaching, and service, are preferred. At higher ranks, a successful candidate should demonstrate successes in conducting and leading collaborative research, obtaining extramural support to maintain and grow an independent, dynamic research program, and an outstanding publication record.

This is a full-time, 9-month, tenure-track position with an anticipated earliest start date of August 23, 2020. Salary, rank and start-up packages are highly competitive and will be commensurate with qualifications and experience. To apply please see minimum and preferred qualifications for Applicants seeking Assistant Professor, Associate Professor or Full Professor rank: **https://academicjobsonline.org/ajo/jobs/15839**

Applicants are required to upload the following material: (1) Cover Letter describing your interest the position; (2) Full Curriculum Vitae; (3) Research Statement (2-3 pages) describing your scientific contributions and future research program; (4) Teaching Statement (1 page) including a description of previous teaching and mentoring experience; (5) Diversity, Equity and Inclusion Statement (1 page) including your perspective on barriers to success for underrepresented groups in STEM, your past efforts to address these issues, and your future plans to foster a diverse, equitable, and inclusive research community at UConn; (6) PDFs of your 1-2 most significant publications; and (7) Name and contact information for three professional references. Review of applications will begin immediately and continue until the position is filled. For inquiries about the position, please contact Jessica Williamson, ISG Program Assistant, with any questions (Jessica.Williamson@uconn.edu).

The University of Connecticut is an EEO/AA Employer.



The Ecosystems Center of the Marine Biological Laboratory (MBL) is hiring faculty at all levels to expand our program in coastal ecosystems ecology.

Scientists with an interest in collaborative, interdisciplinary studies on coastal estuaries, bays, marshes, and/or coastal watersheds across the globe will be considered. Applicants from communities underrepresented in science, or with a strong history of service to these communities, are particularly encouraged. Candidates applying at the Associate or Senior level should demonstrate the potential to take a leadership role in the Plum Island Ecosystems Long-Term Ecological Research program (pie-lter.ecosystems.mbl.edu) or the Semester in Environmental Science (mbl.edu/ses). We seek candidates with diverse areas of research expertise, including, but not limited to, biogeochemistry and its controls, trophic interactions, ecological modeling, and community and ecosystem ecology. Top priority will be given to candidates demonstrating interest in conducting research within the broad context of global climate change and other anthropogenic influences on the coastal zone.

The Ecosystems Center (mbl.edu/ecosystems) was founded four decades ago to investigate the structure and functioning of ecological systems and predict their responses to changing environmental conditions. The current faculty is highly collaborative, with strength in biogeochemistry, ecological modeling, microbial ecology, microbial dynamics, plant-soil interactions, coastal processes, and adaption to life on land (mbl.edu/ecosystems/faculty/). Ecosystems faculty also collaborate with other groups at MBL with expertise in molecular evolution, functional genomics, microbial diversity, developmental and regenerative biology, bioinformatics, and advanced imaging techniques. MBL's initiative in coastal ecosystems ecology complements other strategic initiatives at MBL involving microbiome research, the development of aquatic organisms as new research tools, and advanced imaging and image analysis.

Qualifications: Applicants must hold a Ph.D. (or equivalent advanced degree) in a relevant field. The successful candidate will demonstrate an interest in collaborative, interdisciplinary work, as well as a strong potential for establishing a vigorous extramurally supported research program that can complement existing areas of strength.

The MBL is an affiliate of the University of Chicago and an Equal Opportunity/Affirmative Action employer committed to diversity. All qualified applicants will receive consideration for employment without regard to race, color, religion, sex, national origin, disability, gender identity, sexual orientation or protected veteran status.

Applications should be submitted at go.mbl.edu/Eco-jobs by March 15.

Applications received by March 15 will receive full consideration; however, applications will be accepted until the position is filled. Inquiries about the position should be directed to Dr. Anne Giblin, Chair of the Search Committee (agiblin@mbl.edu).

UAMS | Winthrop P. Rockefeller Cancer Institute

Join the Winthrop P. Rockefeller team and help us fight the battle against cancer

The Winthrop P. Rockefeller Cancer Institute is leading the recruitment of multiple investigators (up to 25) for a major strategic increase in its portfolio across basic, translational, population science and clinical research efforts. Recruitment is for tenured, tenure-track, and non-tenure-earning faculty at all ranks and includes academic appointment in one or more academic departments across the University. All applications will be considered, but priority will be given to those applicants with research focused in the following areas: cancer cell biology, population science, tumor immunology and inflammation, cancer cell metabolism, epigenetics, oncolytic viruses, cancer outcomes and survivorship, role of the microbiome in cancer, cancer imaging, chemoprevention and informatics. Likewise, we are interested in research involving the following organ systems: ovary, breast, lung, and hematologic malignancies including multiple myeloma. This large-scale recruiting effort is made possible through significant institutional support and a $10M annual commitment from the state of Arkansas to support the Cancer Institute's efforts to attain NCI Designation.

Successful applicants will join one of the most prestigious cancer centers in the country. The Winthrop P. Rockefeller Cancer Institute has been one of the nation's premier cancer centers for more than three decades, and it is the only cancer center in the state of Arkansas with a robust cancer research portfolio and a mission to improve cancer outcomes for all Arkansans. The Cancer Institute treats over 2,500 new cancer patients annually and has an extensive effort focused on delivering cancer care and conducting research in underserved populations. Its 135 members conduct outstanding cancer research in multiple scientific programs. Cancer Institute members receive approximately $10 million in extramural cancer research funds annually, including multiple "team science" grants. The Cancer Institute has 5 institutional shared resource facilities. A robust clinical trial infrastructure currently supports nearly 270 cancer clinical trials. The University of Arkansas for Medical Sciences is one of 57 institutions with an NIH Clinical and Translational Science Award, which supports translational research and creates a supportive environment that synergizes with the Cancer Institute to promote junior investigators and transdisciplinary research.

Applicants should send their Curriculum Vitae, a one page letter of interest, and 3 professional references directly to Cancer Institute director, Michael J. Birrer, MD, PhD at **mjbirrer@uams.edu**. Application deadline is **June 1, 2020**. Review of applications will continue until all positions are filled. Applicants must have an MD, PhD, MD/PhD or equivalent. Selected applicants will join a diverse and vibrant academic community that values its researchers and is committed to diversity, equity and inclusion. Applicants must be able to work in a team environment.

UAMS is an inclusive Affirmative Action and Equal Opportunity Employer of individuals with disabilities and protected veterans and is committed to excellence.

University of Nebraska Medical Center
COLLEGE OF DENTISTRY

The University of Nebraska Medical Center (UNMC) College of Dentistry invites applications for **full-time, tenure-leading faculty position(s)** available immediately to complement and/or expand the College's current research and academic programs, which can be found at https://www.unmc.edu/dentistry/. The successful candidate is expected to have developed or have potential to develop a strong, externally-fundable research program in a facet of oral health research. Superb opportunities for research collaborations are available with faculty at the College as it expands its research programs and with investigators at UNMC, University of Nebraska -Lincoln (UNL), and throughout the University system. Preference will be given to candidates with a documented record of or immediate potential for external funding. The primary responsibility of the successful candidate(s) is to develop and maintain a funded research program and to support the educational programs of the College through didactic and/or clinical teaching and student mentoring. Qualified applicants will have a terminal doctoral degree. Academic rank and salary are commensurate with qualifications and experience. The College is searching for that (those) unique individual(s) with an interest in building and growing the research enterprise at the College as part of a vibrant University system. Screening of applications will begin immediately. Inquiries regarding the position may be sent to the search committee chair, Dr. Thomas Petro (tpetro@unmc.edu). Please note that to be considered for this position, applicants must submit an application and supporting documentation via UNMC's online employment website, **http://unmc.peopleadmin.com/postings/49030**.

Drexel UNIVERSITY

Professor and Department Head

Biology is made up of 16 tenured/tenure-track faculty and 8 teaching faculty, whose research and teaching activities include a vibrant population of over 800 Biology undergraduate majors, and over 40 PhD/MS students and postdoctoral trainees. Faculty with active research programs and a strong record of extramural funding pursue research in the areas of molecular, cellular and organismal biology, and STEM education. Several Biology faculties hold appointments/collaborations in the College of Medicine; the School of Biomedical Engineering, Science and Health Systems; and Department of Biodiversity, Earth, and Environmental Science, which is integrated with the Academy of Natural Sciences. Biology is housed in the Papadakis Integrated Sciences Building (PISB), a state-of-the-art five-story building with modern classrooms, conference rooms; a comprehensive light microscopy core facility with advanced imaging capabilities, including confocal, super resolution, and 2 photon excitation microscopy. Biology also collaborates with the Center for the Advancement of STEM Teaching and Learning Excellence (CASTLE), and the renowned, HHMI-supported SEA-PHAGES program, both are funded by the Howard Hughes Medical Institute (HHMI).

Job Summary: Drexel University, a top-ranked Carnegie R1-classified research institution located in Philadelphia, seeks an innovative leader and research scientist for the position of Professor and Head, Department of Biology (https://drexel.edu/coas/academics/departments-centers/biology), in the College of Arts and Sciences starting in September 2020. Successful candidate will teach graduate and/or undergraduate courses. Maintain an active research portfolio. Provide service to the University, College, and Department as well as to the general discipline. Department Head Leadership Responsibilities: - Strategic planning - Educational programming and assessment - Management and stewardship - Faculty recruitment, retention, and advancement - Scholarly work, research, and creative activity - Shared governance, collaboration, and departmental culture - Staff management.

Candidates should also exhibit: A proven record of teaching and service activities; Research achievements consistent with qualification for the rank of full professor at an R-1 institution; Academic administrative experience; Exceptional oral, written, and interpersonal communication skills; A strategic vision and ability to guide and nurture the research, teaching, and service missions undertaken by a diverse faculty; Commitment to building a culture of transparency and shared governance; Experience in mentoring students and faculty; An ability to bolster the Department's extramurally funded research programs and strengthen undergraduate and graduate student training; Understanding of the enrollment and retention challenges of an undergraduate-serving Department; An ability to advocate for the Department's personnel and priorities

Required Qualifications: PhD, Biological Sciences

Submit application, a cover letter, curriculum vitae, two-page statement of leadership philosophy and the names of five references, via https://careers.drexel.edu/en-us/job/494272/professor-and-department-head-biology. **See Drexel' Career link to view ad in its entirety.** Preferences will be given to applications received before February 15, 2020. Inquiries should be addressed to the Chair of the Search Committee, Dr. Jacob Russell (jar337@drexel.edu). The Committee aims to select candidates for a first round of interviews to take place in February and March 2020.

Who's the Top Employer for 2019?

Science Careers' annual survey reveals the top companies in biotech & pharma voted on by *Science* readers.

Read the article and employer profiles and listen to podcasts at sciencecareers.org/topemployers

Science 2019
TOP EMPLOYERS

By **Anurag Srivastava**

A lesson from Bollywood

My Ph.D. adviser called me into his office, saying I needn't bring my notebook. Puzzled, I followed him and sat down. We'd met for 2 hours the day before to finalize our project plan for the coming months, and it wasn't clear what more we had to discuss. He started by saying, "Anurag, this conversation isn't going to be easy," instantly sending my mind into a flurry of thoughts about what was to follow. After 15 minutes of listing positive things about my academic capabilities, he looked me in the eye and said, "You are fired from the lab." I stared back, blinking in disbelief. "Is he joking?" I wondered. "How is this possible?"

I had moved to Israel from my native India the year before, excited to experience a new culture and pursue a Ph.D. I'd already completed a master's degree in the Netherlands, and at first things went well in my new lab: I got along with my Ph.D. adviser, and my experiments progressed as planned. Then, 3 months before I was fired, I ran into some problems. I made a few mistakes in the lab that slowed my research, but I wasn't aware that my adviser noticed them, and he never spoke to me about any concerns.

That's why I was caught off guard in his office that day. I'm still not sure why he fired me, but I suspect it was because of those mistakes. He wasn't confident that I could complete my research in the time frame we'd planned.

"When life gives you lemons, make lemonade."

The first few days after my dismissal were especially difficult. I spent hours staring at my computer screen, unable to get anything done. One day all I could do was sit on a beach, crying as I looked out across the Mediterranean Sea and wondered what I should do.

My adviser gave me 2 months to wrap up my work. I tried to change his mind with promising results, but he remained resolute. I could not break the news to my family in India, as the fear of disappointing them overwhelmed me. I soon spiraled into a state of depression and anxiety. Meanwhile, the date for me to leave the country was drawing near, as my visa required me to be enrolled as a student. I was lonely and without hope.

I started to wonder whether my experience was unique. Poking around on the internet, I was relieved to discover that many Ph.D. students never finish their studies for various reasons, one of which is a broken relationship with their adviser. At least I wasn't alone.

Around that time, I watched *Dasvidaniya*, a Bollywood movie that's about a man who is told that he has 3 months to live. He responds by reframing his perspective on life and setting out to make the most of his remaining months. Even though it is a common saying, one line from the movie stuck out to me: "When life gives you lemons, make lemonade." What kind of "lemonade" could I make out of my current situation?

My desire to complete a Ph.D. was never in doubt; it was my confidence that had taken a hit after my dismissal. After much reflection, I told myself that one failed attempt was not the end of the world, and that I needed to give it another try. I reminded myself that even if I am not the most skilled researcher in the lab, I am a good teacher and I care passionately about mentoring students. My goal is to go back to India to work as a professor, a job I think I would excel at.

With renewed confidence, I emailed prospective advisers and applied to other programs. My previous adviser had not yet secured tenure; this time, I sent my applications to more senior, tenured professors. I thought they would have more experience working with international students and would be more patient as I developed my research abilities. Within 2 months of that fateful conversation in my adviser's office, I landed an offer from a Ph.D. program in Italy. I accepted it and relocated to Europe, happy that my goal of completing a Ph.D. was alive once again.

I've faced other challenges during my current Ph.D. program, but my adviser has been supportive, and I've felt comfortable going to him for help and guidance. I'm thankful that I didn't give up on my dream and that I found another professor willing to take me on. So, if you find yourself in a similar situation and life gives you lemons, ask yourself: "How can I make lemonade?" ■

Anurag Srivastava is a Ph.D. student at the University of Turin in Italy.

ILLUSTRATION: ROBERT NEUBECKER