What AI can learn from the minds of babies *pp. 845*

Leader-follower dynamics in white storks *pp. 852 & 911*

Engineering bacteria to monitor disease *pp. 856 & 915*

Science

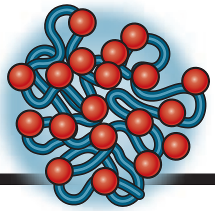
$15
25 MAY 2018
sciencemag.org
AAAS

SINGLE-MOLECULE MAKER

Optical tweezers assemble NaCs atom by atom *pp. 855 & 900*

CONTENTS

25 MAY 2018 • VOLUME 360 • ISSUE 6391



859, 918, & 922
Phase separations in biology

NEWS

IN BRIEF

836 News at a glance

IN DEPTH

839 RIVAL GIANT TELESCOPES JOIN TO SEEK U.S. FUNDING
Projects will offer "public access" in bid for support from astronomy community *By E. Hand*

840 CHINA'S AMBITIOUS BRAIN SCIENCE PROJECT INCHES FORWARD
Beijing and Shanghai open new research centers ahead of the launch of a 15-year national effort *By D. Normile*

841 GERMAN LAW ALLOWS USE OF DNA TO PREDICT SUSPECTS' LOOKS
Bavaria opens the door to DNA phenotyping, provoking debate about the technique's ethics and accuracy *By G. Vogel*

842 B612 PLANS ASTEROID HUNT WITH FLEET OF SMALL SATELLITES
After abandoning plans for a large space telescope, private foundation pins hopes on new technologies *By A. Mann*

844 A CALL TO ARMS AGAINST THE OTHER RETROVIRUS
HTLV-1, discovered just before HIV but almost forgotten, infects millions and causes cancer *By K. Kupferschmidt*

FEATURES

845 BASIC INSTINCTS
Some say artificial intelligence needs to learn like a child *By M. Hutson*
▸ VIDEO

848 THE WAR ON GLUTEN
Wheat sensitivity isn't imaginary, most researchers now agree. But what's really behind it? *By K. Servick*
▸ VIDEO

INSIGHTS

PERSPECTIVES

852 FOLLOWING THE LEADER, FOR BETTER OR WORSE
Juvenile white storks use different behavioral strategies for thermal soaring *By G. A. Nevitt*
▸ REPORT P. 911

854 ENHANCING ENERGY TRANSPORT IN CONJUGATED POLYMERS
Scalable synthesis enables long-distance exciton transport in crystalline polymer nanofibers *By R. J. Holmes*
▸ REPORT P. 897

855 COLD CHEMISTRY WITH TWO ATOMS
Two atoms react to form a molecule in an optical "beaker" *By E. Narevicius*
▸ REPORT P. 900

856 ILLUMINATING DARK DEPTHS
Microelectronic processing and engineered bacteria provide real-time insights into the gut
By P. R. Gibson and R. E. Burgell
▸ REPORT P. 915

858 CANCER IMMUNITY THWARTED BY THE MICROBIOME
Microbial bile acid metabolites promote liver metastasis
By N. Hartmann and M. Kronenberg
▸ RESEARCH ARTICLE P. 876

859 THE RNA FACE OF PHASE SEPARATION
RNA regulates the formation, identity, and localization of phase-separated granules *By M. Polymenidou*
▸ REPORTS PP. 918 & 922

POLICY FORUM

861 DISPARITIES IN SCIENCE LITERACY
Cognitive and socioeconomic factors don't fully explain gaps *By N. Allum* et al.

BOOKS *ET AL.*

863 OUR INHERITANCE
A fascinating history of heredity research reveals the field's highs and lows *By A. Meyer*

864 TRUTH WITH A VENGEANCE
A filmmaker with an ax to grind takes aim at Thomas Kuhn's legacy
By J. D. Martin

LETTERS

865 BRAZIL'S GOVERNMENT ATTACKS BIODIVERSITY
By F. A. Bockmann et al.

866 THE ROAD TO WILD YAK PROTECTION IN CHINA
By M. Chen et al.

866 TROUT IN HOT WATER: A CALL FOR GLOBAL ACTION
By C. C. Muhlfeld et al.

867 TECHNICAL COMMENT ABSTRACTS

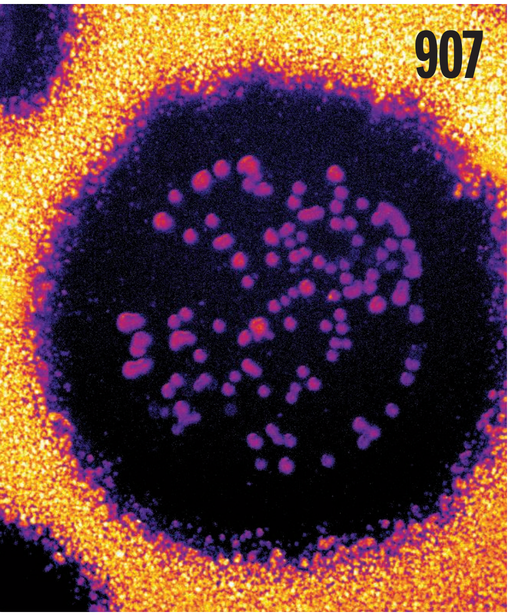
GRAPHIC: KELLIE HOLOSKI/*SCIENCE*; PHOTO: PAUL TAYLOR/GETTY IMAGES

CONTENTS

874 & 875

Flatworm transcriptome atlas

25 MAY 2018 • VOLUME 360 • ISSUE 6391



RESEARCH

IN BRIEF

870 From *Science* and other journals

REVIEW

873 INORGANIC CHEMISTRY
Beyond fossil fuel–driven nitrogen transformations *J. G. Chen* et al.
REVIEW SUMMARY; FOR FULL TEXT: dx.doi.org/10.1126/science.aar6611

RESEARCH ARTICLES

SINGLE-CELL ANALYSIS

874 Cell type transcriptome atlas for the planarian *Schmidtea mediterranea* *C. T. Fincher* et al.
RESEARCH ARTICLE SUMMARY; FOR FULL TEXT: dx.doi.org/10.1126/science.aaq1736
875 Cell type atlas and lineage tree of a whole complex animal by single-cell transcriptomics *M. Plass* et al.
RESEARCH ARTICLE SUMMARY; FOR FULL TEXT: dx.doi.org/10.1126/science.aaq1723

876 MICROBIOME
Gut microbiome–mediated bile acid metabolism regulates liver cancer via NKT cells *C. Ma* et al.
RESEARCH ARTICLE SUMMARY; FOR FULL TEXT: dx.doi.org/10.1126/science.aan5931
▸ PERSPECTIVE P. 858

877 PALEOCLIMATE
A 550,000-year record of East Asian monsoon rainfall from ^{10}Be in loess
J. W. Beck et al.

881 NEURODEVELOPMENT
Evolution of pallium, hippocampus, and cortical cell types revealed by single-cell transcriptomics in reptiles
M. A. Tosches et al.

REPORTS

888 CATALYSIS
Cobalt-catalyzed asymmetric hydrogenation of enamides enabled by single-electron reduction
M. R. Friedfeld et al.

893 VALLEYTRONICS
Imaging of pure spin-valley diffusion current in WS$_2$-WSe$_2$ heterostructures
C. Jin et al.

897 ORGANIC ELECTRONICS
Long-range exciton transport in conjugated polymer nanofibers prepared by seeded growth
X.-H. Jin et al.
▸ PERSPECTIVE P. 854

900 COLD MOLECULE PHYSICS
Building one molecule from a reservoir of two atoms *L. R. Liu* et al.
▸ PERSPECTIVE P. 855

904 PHOTOVOLTAICS
Flexo-photovoltaic effect
M.-M. Yang et al.

907 EVOLUTION
High parasite diversity accelerates host adaptation and diversification
A. Betts et al.

911 MIGRATORY BEHAVIOR
From local collective behavior to global migratory patterns in white storks
A. Flack et al.
▸ PERSPECTIVE P. 852

915 SYNTHETIC BIOLOGY
An ingestible bacterial-electronic system to monitor gastrointestinal health
M. Mimee et al.
▸ PERSPECTIVE P. 856

MOLECULAR BIOLOGY
918 RNA buffers the phase separation behavior of prion-like RNA binding proteins *S. Maharana* et al.
922 mRNA structure determines specificity of a polyQ-driven phase separation *E. M. Langdon* et al.
▸ PERSPECTIVE P. 859

DEPARTMENTS

835 EDITORIAL
Canada's call *By Mona Nemer*

934 WORKING LIFE
What I learned from teaching
By Moamen Elmassry

ON THE COVER

Apparatus for building a single NaCs molecule from two individual sodium and cesium atoms. The single chemical reaction takes place in an 80-millimeter-long glass cell vacuum chamber, where the atoms are captured and manipulated by laser beams focused through a microscope objective (shown at right). The reaction is triggered by a pulse of light after the atoms are brought together. See pages 855 and 900. *Photo: Ken Richardson*

Science Staff 834
AAAS News & Notes 868
New Products 928
Science Careers 929

CREDITS (FROM TOP): C. T. FINCHER *ET AL.*; ISABEL FROST

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. Copyright © 2018 by the American Association for the Advancement of Science. 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): $1808; Foreign postage extra: Mexico, Caribbean (surface mail) $55; other countries (air assist delivery): $89. First class, airmail, student, and emeritus rates on request. Canadian rates with GST available upon request, GST #125488122. Publications Mail Agreement Number 1069624. **Printed in the U.S.A.**
Change of address: Allow 4 weeks, giving old and new addresses and 8-digit account number. **Postmaster:** Send change of address to AAAS, P.O. Box 96178, Washington, DC 20090–6178. **Single-copy sales:** $15 each plus shipping and handling; bulk rate on request. **Authorization to reproduce** material for internal or personal use under circumstances not falling within the fair use provisions of the Copyright Act is granted by AAAS to libraries and others who use Copyright Clearance Center (CCC) Pay-Per-Use services provided that $35.00 per article is paid directly to CCC, 222 Rosewood Drive, Danvers, MA 01923. The identification code for *Science* is 0036-8075. *Science* is indexed in the *Reader's Guide to Periodical Literature* and in several specialized indexes.

Science

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

AAAS

Editor-in-Chief Jeremy Berg

Executive Editor Monica M. Bradford **News Editor** Tim Appenzeller

Deputy Editors Lisa D. Chong, Andrew M. Sugden(UK), Valda J. Vinson, Jake S. Yeston

Research and Insights

DEPUTY EDITOR, EMERITUS Barbara R. Jasny **SR. EDITORS** Gemma Alderton(UK), Caroline Ash(UK), Julia Fahrenkamp-Uppenbrink(UK), Pamela J. Hines, Stella M. Hurtley(UK), Paula A. Kiberstis, Marc S. Lavine(Canada), Steve Mao, Ian S. Osborne(UK), Beverly A. Purnell, L. Bryan Ray, H. Jesse Smith, Jelena Stajic, Peter Stern(UK), Phillip D. Szuromi, Sacha Vignieri, Brad Wible, Laura M. Zahn **ASSOCIATE EDITORS** Michael A. Funk, Brent Grocholski, Priscilla N. Kelly, Seth Thomas Scanlon(UK), Keith T. Smith(UK) **ASSOCIATE BOOK REVIEW EDITOR** Valerie B. Thompson **LETTERS EDITOR** Jennifer Sills **LEAD CONTENT PRODUCTION EDITORS** Harry Jach, Lauren Kmec **CONTENT PRODUCTION EDITORS** Amelia Beyna, Jeffrey E. Cook, Amber Esplin, Chris Filiatreau, Cynthia Howe, Catherine Wolner **SR. EDITORIAL COORDINATORS** Carolyn Kyle, Beverly Shields **EDITORIAL COORDINATORS** Aneera Dobbins, Joi S. Granger, Jeffrey Hearn, Lisa Johnson, Maryrose Madrid, Scott Miller, Jerry Richardson, Anita Wynn **PUBLICATIONS ASSISTANTS** Ope Martins, Nida Masiulis, Dona Mathieu, Hilary Stewart(UK), Alana Warnke, Alice Whaley(UK), Brian White **EXECUTIVE ASSISTANT** Jessica Slater **ADMINISTRATIVE SUPPORT** Janet Clements(UK)

News

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

Executive Publisher Rush D. Holt

Publisher Bill Moran **Chief Digital Media Officer** Josh Freeman

DIRECTOR, BUSINESS STRATEGY AND PORTFOLIO MANAGEMENT Sarah Whalen **DIRECTOR, PRODUCT AND CUSTOM PUBLISHING** Will Schweitzer **MANAGER, PRODUCT DEVELOPMENT** Hannah Heckner **BUSINESS SYSTEMS AND FINANCIAL ANALYSIS DIRECTOR** Randy Yi **DIRECTOR, BUSINESS OPERATIONS & ANALYST** Eric Knott **ASSOCIATE DIRECTOR, PRODUCT MANAGMENT** Kris Bishop **ASSOCIATE DIRECTOR, INSTITUTIONAL LICENSING SALE** Geoffrey Worton **SENIOR SYSTEMS ANALYST** Nicole Mehmedovich **SENIOR BUSINESS ANALYST** Cory Lipman **MANAGER, BUSINESS OPERATIONS** Jessica Tierney **BUSINESS ANALYSTS** Meron Kebede, Sandy Kim, Jourdan Stewart **FINANCIAL ANALYST** Julian Iriarte **ADVERTISING SYSTEM ADMINISTRATOR** Tina Burks **SALES COORDINATOR** Shirley Young **DIRECTOR, COPYRIGHT, LICENSING, SPECIAL PROJECTS** Emilie David **DIGITAL PRODUCT ASSOCIATE** Michael Hardesty **RIGHTS AND PERMISSIONS ASSOCIATE** Elizabeth Sandler **RIGHTS, CONTRACTS, AND LICENSING ASSOCIATE** Lili Catlett **RIGHTS & PERMISSIONS ASSISTANT** Alexander Lee

MARKETING MANAGER, PUBLISHING Shawana Arnold **SENIOR ART ASSOCIATES** Paula Fry **ART ASSOCIATE** Kim Huynh

DIRECTOR, INSTITUTIONAL LICENSING Iquo Edim **ASSOCIATE DIRECTOR, RESEARCH & DEVELOPMENT** Elisabeth Leonard **SENIOR INSTITUTIONAL LICENSING MANAGER** Ryan Rexroth **INSTITUTIONAL LICENSING MANAGERS** Marco Castellan, Chris Murawski **SENIOR OPERATIONS ANALYST** Lana Guz **MANAGER, AGENT RELATIONS & CUSTOMER SUCCESS** Judy Lillibridge

WEB TECHNOLOGIES TECHINICAL DIRECTOR David Levy **TECHNICAL MANAGER** Chris Coleman **PORTFOLIO MANAGER** Trista Smith **PROJECT MANAGER** Tara Kelly, Dean Robbins **DEVELOPERS** Elissa Heller, Ryan Jensen, Brandon Morrison

DIGITAL MEDIA DIRECTOR OF ANALYTICS Enrique Gonzales **SR. MULTIMEDIA PRODUCER** Sarah Crespi **MANAGING DIGITAL PRODUCER** Kara Estelle-Powers **PRODUCER** Liana Birke **VIDEO PRODUCERS** Chris Burns, Nguyên Khôi Nguyễn **DIGITAL SOCIAL MEDIA PRODUCER** Brice Russ

DIGITAL/PRINT STRATEGY MANAGER Jason Hillman **QUALITY TECHNICAL MANAGER** Marcus Spiegler **DIGITAL PRODUCTION MANAGER** Lisa Stanford **ASSISTANT MANAGER DIGITAL/PRINT** Rebecca Doshi **SENIOR CONTENT SPECIALISTS** Steve Forrester, Antoinette Hodal, Lori Murphy, Anthony Rosen **CONTENT SPECIALISTS** Jacob Hedrick, Kimberley Oster

DESIGN DIRECTOR Beth Rakouskas **DESIGN MANAGING EDITOR** Marcy Atarod **SENIOR DESIGNER** Chrystal Smith **DESIGNER** Christina Aycock **GRAPHICS MANAGING EDITOR** Alberto Cuadra **GRAPHICS EDITOR** Nirja Desai **SENIOR SCIENTIFIC ILLUSTRATORS** Valerie Altounian, Chris Bickel, Katharine Sutliff **SCIENTIFIC ILLUSTRATOR** Alice Kitterman **INTERACTIVE GRAPHICS EDITOR** Jia You **SENIOR GRAPHICS SPECIALISTS** Holly Bishop, Nathalie Cary **PHOTOGRAPHY MANAGING EDITOR** William Douthitt **PHOTO EDITOR** Emily Petersen **IMAGE RIGHTS AND FINANCIAL MANAGER** Jessica Adams **INTERN** Mike Shanahan

SENIOR EDITOR, CUSTOM PUBLISHING Sean Sanders: 202-326-6430 **ASSISTANT EDITOR, CUSTOM PUBLISHING** Jackie Oberst: 202-326-6463 **ASSOCIATE DIRECTOR, BUSINESS DEVELOPMENT** Justin Sawyers: 202-326-7061 science_advertising@aaas.org **ADVERTISING PRODUCTION OPERATIONS MANAGER** Deborah Tompkins **SR. PRODUCTION SPECIALIST/GRAPHIC DESIGNER** Amy Hardcastle **SR. TRAFFIC ASSOCIATE** Christine Hall **DIRECTOR OF BUSINESS DEVELOPMENT AND ACADEMIC PUBLISHING RELATIONS, ASIA** Xiaoying Chu: +86-131 6136 3212, xchu@aaas.org **COLLABORATION/CUSTOM PUBLICATIONS/JAPAN** Adarsh Sandhu + 81532-81-5142 asandhu@aaas.org **EAST COAST/E. CANADA** Laurie Faraday: 508-747-9395, FAX 617-507-8189 **WEST COAST/W. CANADA** Lynne Stickrod: 415-931-9782, FAX 415-520-6940 **MIDWEST** Jeffrey Dembksi: 847-498-4520 x3005, Steven Loerch: 847-498-4520 x3006 **UK EUROPE/ASIA** Roger Goncalves: TEL/FAX +41 43 243 1358 **JAPAN** Kaoru Sasaki (Tokyo): + 81 (3) 6459 4174 ksasaki@aaas.org

GLOBAL SALES DIRECTOR ADVERTISING AND CUSTOM PUBLISHING Tracy Holmes: +44 (0) 1223 326525 **CLASSIFIED** advertise@sciencecareers.org **SALES MANAGER, US, CANADA AND LATIN AMERICA SCIENCE CAREERS** Claudia Paulsen-Young: 202-326-6577, **EUROPE/ROW SALES** Sarah Lelarge **SALES ADMIN ASSISTANT** Kelly Grace +44 (0)1223 326528 **JAPAN** Miyuki Tani(Osaka): +81 (6) 6202 6272 mtani@aaas.org **CHINA/TAIWAN** Xiaoying Chu: +86-131 6136 3212, xchu@aaas.org **GLOBAL MARKETING MANAGER** Allison Pritchard **DIGITAL MARKETING ASSOCIATE** Aimee Aponte

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

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

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.

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

EDITORIAL

Canada's call

Mona Nemer is the chief science adviser of Canada.
science@canada.ca

Next month, Canada will host the Group of 7 (G7) summit in picturesque Charlevoix, Québec. As leaders from Canada, France, Germany, Italy, Japan, the United Kingdom, and the United States come together, along with European Union representatives, to discuss the progressive agenda, science will be on everyone's mind. With science and technology playing a prominent role in everyday life, access to science education and to science-based careers is ever more essential for inclusive growth and for women's empowerment.

In addition to the contribution that science makes to economic prosperity, the public is demanding that it also be used to guide decision-making. Scientists and knowledge communities have risen to strengthen their case to policy-makers: Witness the global March for Science last month and the renewed interest in science advice for policy at national and international levels. Nowhere has this trend been more evident than in Canada, where the government has signaled its respect for science-informed policy. This was made clear in some remarkable developments over the past 10 months, including my appointment as the first government chief science adviser in over a decade and a historic science budget for 2018. The Canadian scientific community is generally upbeat and reenergized. Now is the time to ensure that the infrastructure and programs are in place to meet rising expectations and deliver on science promises and potential.

"...scientists will be part of the solutions to global challenges..."

The G7 summit themes range from climate change and ocean sustainability to preparing for the jobs of the future while ensuring a peaceful and secure world. All of them require quality data, special infrastructure, and coherent multi-institutional and multinational approaches. Although these issues will occupy the work of Canada's minister of science and the Office of the Chief Science Advisor, they are not unique to Canada or to any other country. All the more reason, then, that the international science community should encourage sharing of best practices and the development of universally accepted principles for research governance. Using and benefiting from major science initiatives and facilities should not be confined within national borders. Increased international access to highly specialized facilities, such as telescopes, cyclotrons, and Arctic research stations, should be supported through proper resources and mobility programs. Such efforts will promote international collaboration and accelerate discovery and innovation. This can be a cornerstone for international relations.

Developing policies to meet national and international aspirations and commitments requires a well-organized science advice ecosystem inside and outside of government. This includes, in addition to national science advisers, a network of departmental and subnational science advisers, science academies, and scientific associations. Each plays a distinct and complementary role in gathering evidence, convening experts, and communicating with the public, decision-makers, and stakeholders. Strong national institutions are essential for supporting and sustaining a robust ecosystem of science advice and public engagement. With science and technology increasingly a part of daily life, the need for science advice will only intensify.

"Science knows no country, because knowledge belongs to humanity," Pasteur told us. This is more relevant today than ever before. In the coming years, scientists will be part of the solutions to global challenges, and they will do this collaboratively. International gatherings, such as those of the G7, the G20, the United Nations, and the Commonwealth Heads of Government, could benefit from advance meetings of science advisers, both to help frame the issues and to signal how science might be harnessed to solve important problems. And Canada, as a member country and one that has made science and innovation a priority, is in a good position to convene and coordinate such efforts. I look forward to working with my colleagues to build on the momentum of next month's summit in Charlevoix.

–Mona Nemer

10.1126/science.aau1744

CREDITS: (INSET) ADAPTED FROM ISTOCK.COM/SIRAANAMWONG; (TOP RIGHT) IVY SINKUNAS/INNOVATION, SCIENCE AND ECONOMIC DEVELOPMENT CANADA

NEWS

> “Large fish, like sturgeons, have been suggested as possible explanations for the monster myth.”
>
> **Neil Gemmell of the University of Otago**, in Dunedin, New Zealand, on a project to study environmental DNA from Scotland’s Loch Ness for signs of unusual creatures.

IN BRIEF

Edited by **Jeffrey Brainard**

DRUG DEVELOPMENT

FDA approves first migraine preventive

ILLUSTRATION: DAVIDE BONAZZI/@SALZMANART

The U.S. Food and Drug Administration (FDA) approved on 17 May the first drug to prevent migraine headaches, an antibody that blocks the receptor of a protein thought to play a central role in causing the condition (*Science*, 8 January 2016, p. 116). The treatment, erenumab (Aimovig), is made by Amgen in Thousand Oaks, California, and Novartis of Basel, Switzerland. The required monthly injections will cost $545 per dose in the United States. Migraines, which can be disabling, are estimated to afflict 6% of men and 17% of women in the United States. Nearly 1% of Americans are chronic sufferers, experiencing headaches on 15 or more days per month. In three clinical trials involving 2199 migraine sufferers, those who received erenumab instead of a placebo experienced 1 to 2.5 fewer migraine days per month. Three other companies have similar drugs awaiting FDA approval.

Sweden cancels Elsevier contract

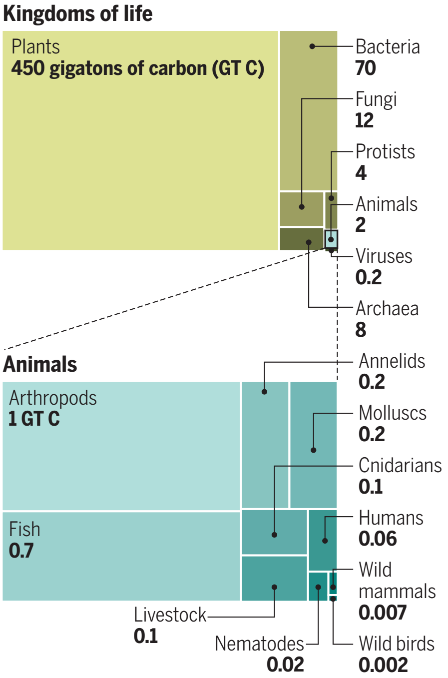
PUBLISHING | Swedish universities and research institutes have canceled their contract with academic publishing giant Elsevier after failing to reach an agreement that aligns with the country’s vision for providing free access to all publicly funded research by 2026. The Bibsam Consortium, which negotiates licensing arrangements on behalf of Sweden’s universities, demanded a more “sustainable price model.” It estimates that Swedish institutes spent €1.3 million in 2017 to publish articles in Elsevier’s open-access journals, on top of €12 million for accessing the publisher’s subscription journals. The consortium sought a deal that required Elsevier to provide immediate open access to all papers by their researchers. The cancellation means Swedish universities will not have access to Elsevier content published after 30 June.

New head for DOE science

SCIENCE POLICY | President Donald Trump has nominated a neuroscientist and former member of the White House staff under President Barack Obama to lead the U.S. Department of Energy’s (DOE’s) science programs. The White House announced on 18 May the nomination of Chris Fall, currently principal deputy director of DOE’s Advanced Research Projects Agency-Energy (ARPA-E), which helps commercialize promising technologies, to head the $6 billion DOE Office of Science. It is the nation’s leading funder of the physical sciences and supports an array of facilities used extensively by researchers. Prior to joining ARPA-E, Fall spent 6 years at the Office of Naval Research in a variety of roles, including a 3-year stint working on national security issues in Obama’s Office of Science and Technology Policy. The Senate must confirm Fall for the role.

Moderate drinking study stopped

BIOMEDICINE | The U.S. National Institutes of Health (NIH) has halted a study of the possible benefits of moderate drinking while it conducts an ethics and scientific review. The Moderate Alcohol



BIOLOGY

Plants weigh the most, but humans, livestock add up

Plants collectively pack more heft than any other kingdom of life, even more than bacteria. That's just one of the surprises in a comprehensive survey of the scientific literature on Earth's biomass. It also shows that the most diverse groups are not necessarily the ones with the most biomass, so this measure provides a different perspective of life on this planet, says Ron Milo, a quantitative biologist at the Weizmann Institute of Science in Rehovot, Israel, whose student Yinon Bar-On combed through hundreds of studies. Humans weigh in at only about 0.06 gigatons of carbon—on par with krill and termites—but the human impact on biomass since the beginning of civilization has been huge, Milo's group reports this week in the *Proceedings of the National Academy of Sciences*. Humans and their livestock (mostly cattle and pigs) outweigh wild mammals by more than 22-fold, and domesticated fowl surpass all other birds in biomass, they found. There has also been a twofold decline in plant biomass since civilization began. Uncertainty in numbers for subsurface microbes made estimating their overall mass difficult, the authors say. The biomass figures set the stage for their next challenge: estimating the planet's most abundant protein.

and Cardiovascular Health Trial has come under scrutiny because the alcoholic beverage industry is putting up two-thirds of the funding for the $10 million, 10-year study. In March, *The New York Times* reported on emails and other documents indicating that NIH employees urged industry officials to fund the trial, a request that apparently violated agency policy. At a Senate hearing last week, NIH Director Francis Collins said "sufficient concerns" led the agency to suspend enrollment in the trial until the investigation is completed in June. The study, led by Beth Israel Deaconess Medical Center in Boston, had already recruited 105 of a planned 7800 participants.

U.S. science eyes spending boost

RESEARCH FUNDING | The process of setting 2019 funding levels for U.S. science agencies is underway, and so far, researchers are mostly pleased with the results. This month, Senate and House of Representatives spending panels began advancing appropriations bills that generally reject President Donald Trump's requests for cuts to science programs and instead specify several increases for the fiscal year that begins 1 October. Both bodies would give the Department of Energy's science office a roughly 5% boost to $6.6 billion. The House bills call for the National Science Foundation's budget to grow by 5.2% to $8.2 billion and NASA's science directorate to rise 7.4% to $6.7 billion. The U.S. Geological Survey would get a 1.6% boost to $1.2 billion, while the Department of Agriculture's competitive grants program grows 3.8% to $415 million. But the bills cut science programs at the Environmental Protection Agency by 8.9%, to $644 million, and trim research at the National Oceanic and Atmospheric Administration (NOAA) by 8.2%, to $503 million, including a 37.6% cut to NOAA climate science, to $99 million. The House has yet to prepare a bill to fund the National Institutes of Health. The two chambers ultimately must reconcile their spending bills.

Reprieve for climate program

CLIMATE SCIENCE | A spending bill approved last week by the appropriations committee of the U.S. House of Representatives tells NASA to resurrect the $10 million Carbon Monitoring System, which President Donald Trump's administration canceled (*Science*, 11 May, p. 586). Climate scientists said the administration's move would jeopardize plans to verify the emission cuts that countries agreed to in the Paris climate accords. The monitoring system has supported the collection of data about carbon sources and sinks and creating high-resolution models of the planet's flows of carbon. The provision is contained in a House bill that sets 2019 spending levels for NASA and other agencies.

CREDITS: (PHOTO) MICHAEL HANSON/AURORA PHOTOS; (GRAPHIC) N. DESAI/*SCIENCE*; (DATA) Y. M. BAR-ON *ET AL. PNAS*, 10.1073/PNAS.1711842115 (2018)

Published by AAAS

THREE QS

Struggling to serve Iran

Kaveh Madani returned to his native Iran last September to serve in the government as deputy vice president for the environment. Now 36, he had spent 14 years overseas and worked since 2013 as a faculty member at Imperial College London, where he studied Iran's dwindling water resources. But in Iran, hardliners had him interrogated and accused him of spying. In April, during a work trip to Bangkok, he resigned his post and moved to an undisclosed location outside Iran. He spoke with *Science* about his experiences. (An extended version of this interview is available at https://scim.ag/IranReturn.)

Q: Why did you return to Iran?

A: I thought if I succeed, there would be more trust in scientists and more trust in people of my generation. And to inspire other Iranians outside to go back and help. … In some ways, my time in Tehran was the best sabbatical one academic could wish for. This was a wonderful learning experience. On the other hand, I realize that I'm lucky I'm not in prison, or dead.

Q: Is there a water crisis in Iran?

A: Our demand and consumption are way higher than the available water. … The government cannot deny the deterioration of Lake Urmia. So speaking out about the environment became a sort of safe space [for Iranian citizens].

Q: What message were hardliners trying to send by discrediting you?

A: [They said I] was trying to paint a dark picture of Iran's water sources to justify shutting down the agricultural sector and importing GMOs [genetically modified organisms]. They called me a bioterrorist. I said, "You're misinterpreting scientific facts." … I knew I wasn't a spy. Getting rid of me was their main goal. … Some of the questions I was getting during my interrogation in Tehran were the same as what I get at the borders of some countries in the West. Hardliners don't like a peaceful environment because conflict is essential to their survival, no matter where they are. I think I was more of a threat because I was breaking down borders. Trying to solve things with a scientific approach, trying to be inclusive.

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

Insider to lead cancer agency

PUBLIC HEALTH | The controversial International Agency for Research on Cancer (IARC) last week elected cancer epidemiologist Elisabete Weiderpass, a member of IARC's scientific council, as its new director. The announcement immediately triggered industry calls for Weiderpass to change the agency's allegedly biased procedures for rating cancer risks. IARC, an arm of the World Health Organization based in Lyon, France, has come under fire for calling the widely used herbicide glyphosate a "possible human carcinogen"—a classification that Weiderpass has defended but other agencies disagreed with (*Science*, 24 June 2016, p. 1504). Weiderpass, the first woman at the agency's helm, is Brazilian and holds Finnish and Swedish passports; she leads the research department of the Cancer Registry of Norway, in Oslo, as well as the genetic epidemiology group at Folkhälsan Research Center in Finland.

New ministry for Russian science

SCIENCE POLICY | In a major restructuring, the Russian government has created a new Ministry for Science and Higher Education with a new leader. The move splits the former Ministry of Education and Science, headquartered in Moscow. The new science ministry will also oversee universities while a new Ministry of Education manages primary and secondary schools. Heading the science ministry will be Mikhail Kotyukov, a former head of the controversial Federal Agency for Scientific Organizations (FASO). Kotyukov, who has a background in finance, has headed FASO since its creation in the fall of 2013. The agency was established as part of a reform of the Russian Academy of Sciences and given control over academy property, a move opposed by scientists and widely seen as a power grab by the Russian government. Rather than become part of RAS, as some members had proposed, FASO will now be absorbed into the science ministry, further strengthening the government's grip on the academy.

Satellites resume water watch

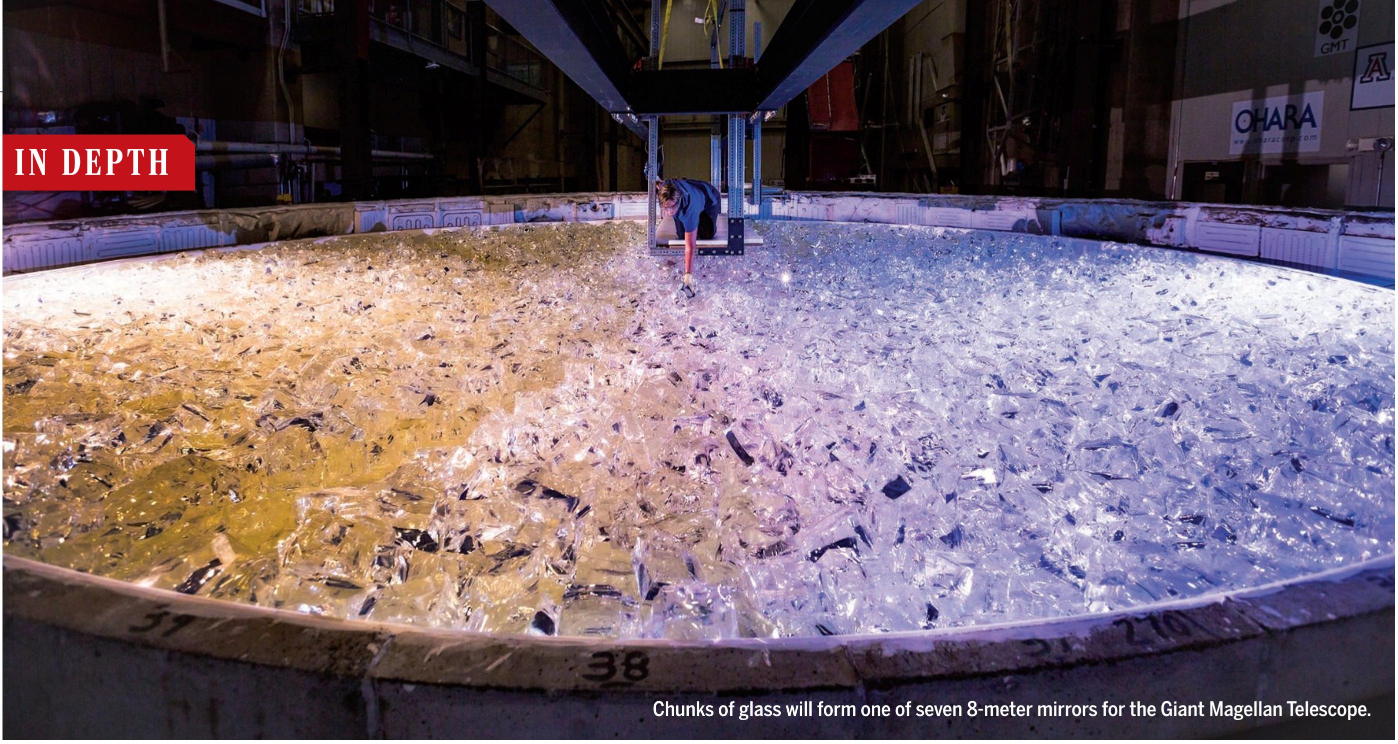
EARTH SCIENCE | NASA and a German partner launched two satellites on 22 May that will measure trace changes in Earth's gravity as a proxy for the movement of water on the planet, replacing a pair that failed in orbit last year. The $550 million Gravity Recovery and Climate Experiment Follow-On (GRACE-FO) mission took off from Vandenberg Air Force Base in California. Like its predecessors, GRACE-FO will gather data by constantly monitoring, via microwave pulses, the changing distance between its satellites as the gravity of large objects below, such as ice sheets, first pulls one spacecraft and then the other. This technique has provided insights into the melting of ice sheets, the emptying of groundwater aquifers, and the motion of deep ocean currents.

BIOETHICS

A forgotten contribution, now honored

By hanging a portrait of her this month, not far from its depictions of presidents, the Smithsonian Institution's National Portrait Gallery in Washington, D.C., honored a woman whose contributions to biomedical research were major, ongoing—and without her knowledge or consent. Henrietta Lacks, a poor woman from rural Virginia, died of cervical cancer in 1951 at age 31. While she was alive, doctors removed and biopsied a tumor and used it to create what became the HeLa line (named for her) of immortal cells, capable of renewing themselves in cultures indefinitely. Researchers used them to develop the polio vaccine and in numerous other medical advances.

IMAGE: SMITHSONIAN INSTITUTION NATIONAL PORTRAIT GALLERY



Chunks of glass will form one of seven 8-meter mirrors for the Giant Magellan Telescope.

PHOTO: GIANT MAGELLAN TELESCOPE/GMTO CORPORATION

ASTRONOMY

Rival giant telescopes join to seek U.S. funding

Projects will offer "public access" in bid for support from astronomy community

By **Eric Hand**

Two U.S.-led giant telescope projects, rivals for nearly 2 decades, announced this week that they have agreed to join forces. The Giant Magellan Telescope (GMT), a 25-meter telescope under construction in Chile, and the Thirty Meter Telescope (TMT), which backers hope to build atop Mauna Kea in Hawaii, are still short of partners and money. They will now work together to win funding from the National Science Foundation (NSF) in Alexandria, Virginia, which could help the projects catch up to a third giant telescope, the 39-meter European Extremely Large Telescope (ELT), due to begin operations in 2024. It is a historic peace accord to end a conflict that has divided funders and delayed both projects.

"This division has set back U.S. astronomy a decade," says Richard Ellis, an astronomer at University College London, and a former leader of the TMT effort. "Let's turn the corner." Patrick McCarthy, a GMT vice president in Pasadena, California, adds, "It's time for these two projects to come together behind a single vision."

The partnership, approved by the GMT board this month and by the TMT board last month, commits the two projects to developing a joint plan that would allow astronomers from any institution to use the telescopes; under previous plans observing time was available only to researchers from nations or institutions that had provided funding. The projects are discussing awarding at least 25% of each telescope's time to nonpartners through a competitive process to be administered by the National Center for Optical-Infrared Astronomy—an umbrella organization that will replace the National Optical Astronomy Observatory (NOAO), based in Tucson, Arizona, sometime in fiscal year 2019 (*Science*, 3 February 2017, p. 442).

Telescope backers hope the public access plan will help persuade the federal government to pay for at least 25% of the total cost of the two facilities, a share that could reach $1 billion. (Cost estimates for the GMT and the TMT are $1 billion and $1.4 billion, respectively, but astronomers expect both numbers to grow.) "There are many science projects that are $1 billion class projects," says David Silva, NOAO's director. "The investment that we would want is of a similar size."

Advocates for the TMT and the GMT "missed opportunities" to unite astronomers, lawmakers, and funding agencies around a single telescope in the early 2000s, Ellis says. Instead, the projects diverged. The TMT began in 2002 under the leadership of the California Institute of Technology in Pasadena and the University of California system. To make a giant mirror that wouldn't deform under its own weight, the TMT's partners looked for inspiration in a fly's eye, fragmenting the main 30-meter mirror into nearly 500 small, 1.4-meter hexagonal pieces, each individually adjustable. The GMT, initially led by the Carnegie Observatories in Pasadena, has seven larger segments, each 8 meters across. They are made of stiff, lightweight glass, cast in molds that spin as the glass cools to bring it into a parabolic shape. In 2005, the GMT cast the first of its mirrors, and the divergent technological paths were set.

More than a decade later, the timing is right for rapprochement. First, the U.S. decadal survey in astronomy, a once-a-decade priority-setting exercise for big projects run by the National Academies of Sciences, Engineering, and Medicine, is set to begin next year, and the public access plan could solidify astronomers' support for putting the giant telescopes high on the list. In making their case, the teams will argue the benefits of having telescopes in both the northern and southern hemispheres. "When you are covering the whole sky, you have greater scientific reach," says Wendy Freedman, an astronomer at The University of Chicago in Illinois who was the founding leader of the GMT. The teams will also argue that the telescopes have complementary strengths. The design of the GMT, for instance, makes it ideal for a high-resolution spectrograph designed to probe the atmospheres of exoplanets. The TMT, which has more light-gathering power, could host a multi-object spectrograph to quickly gather demographic statistics on the universe's first galaxies.

Uniting the projects could also help overcome funding challenges. Each project has already amassed support from countries and deep-pocketed institutions, but neither has enough money to cross the finish line. TMT construction has faced costly delays because of legal challenges and protests by Native Hawaiian groups that object to building atop Mauna Kea, which they consider sacred; planners are considering relocating the TMT to Spain's Canary Islands. And although the GMT hopes to start observing in 2024 with four of its seven mirrors, it, too, needs new partners to see it through. The European Southern Observatory, meanwhile, has made steady progress on the ELT, aided by its regular member state contributions. "The ELT is real," Silva says. "That focused minds."

In the end, success will hinge on NSF and Congress. Silva says NSF Astronomical Sciences Division Director Richard Green, as well as Anne Kinney, an astronomer recently chosen to head the agency's mathematical and physical sciences directorate, "were willing to open the door to exploring this possibility." But Kinney says she and Green first want to see the results of the decadal survey. "Depending on how the GMT and TMT come out, we are very interested in the large telescopes," she says.

A key lawmaker appears to hold similar views. Representative John Culberson (R–TX), who chairs the House of Representatives spending panel that sets funding levels for NSF, told *Science* he could envision NSF helping pay for the telescopes. "But I'm waiting to hear from the next decadal survey" before taking a definitive position, he says. Still, Culberson's panel last week approved a 2019 spending bill that directs NSF to "articulate a plan to ensure public access in future large optical observatories."

Even if NSF ultimately gets money to help build the two telescopes, funding for long-term operating expenses may be harder to find. The agency is already expecting to strain to pay running costs for new telescopes such as the Large Synoptic Survey Telescope, an 8.4-meter survey telescope in Chile—let alone two much bigger ones. Silva says the agency is "really struggling to enable leading-edge facilities while maintaining healthy grants programs." ■

With additional reporting by Jeffrey Mervis.

PHOTO: QILAI SHEN/IN PICTURES LTD./CORBIS/GETTY IMAGES

NEUROSCIENCE

China's ambitious brain science project inches forward

Beijing and Shanghai open new research centers ahead of the launch of a 15-year national effort

By **Dennis Normile**, *in Shanghai, China*

The nascent China Brain Project took another step toward reality last week with the launch of the Shanghai Research Center for Brain Science and Brain-Inspired Intelligence. The new center and its Beijing counterpart, launched 2 months ago, are expected to become part of an ambitious national effort to bring China to the forefront of neuroscience. But details of that 15-year project—expected to rival similar U.S. and EU efforts in scale and ambition—are still being worked out, 2 years after the government made it a priority.

Preparation for the national effort "was taking quite a long time," says Zhang Xu, a neuroscientist and executive director of the new center here. So Beijing and Shanghai got started on their own plans, he says. China's growing research prowess and an increasing societal interest in neuroscience—triggered in part by an aging population—as well as commercial opportunities and government support are all coming together to make this "a good time for China's brain science efforts," Zhang says.

Government planners called for brain research to be a key science and technology project in the nation's 13th Five-Year Plan, adopted in spring 2016. The effort would have three main pillars, according to a November 2016 *Neuron* paper from a group that included Poo Mu-ming, director of Shanghai's Institute of Neuroscience (ION), part of the Chinese Academy of Sciences (CAS). It would focus on basic research on neural mechanisms underlying cognition, translational studies of neurological diseases with an emphasis on early diagnosis and intervention, and brain simulations to advance artificial intelligence and robotics. Support under the 5-year plan was just the start of a 15-year program, the group wrote.

The authors saw an opportunity for basic research in the "enormous gap in our knowledge [of the brain] at the mesoscopic level," where neural circuits connect specific types of brain cells. They suggested China might make an important contribution by mapping those circuits and activity patterns using animal models, a plan now called the International Mesoscopic Connectome Project. China has developed expertise in studying the brain using macaques, and animal research faces less opposition in China than in the West. (Zhang, who is also at ION, advocates for a fourth focus, on how education in mathematics and language, for example, affects neural development.)

But an official green light is still pending. Scientists suggest the complexity of the project and a recent revamp of responsibilities at the Ministry of Science and Technology have delayed a final decision. "We sincerely hope we can see the announcement of the launch of this project very soon," says Yang Xiong-Li, a neuroscientist at Fudan Univer-

Elderly women in Shanghai, China, suffering from dementia, one of the diseases the new effort may tackle.

sity here. The city of Beijing and local institutions took the initiative and launched the Chinese Institute for Brain Research, announced on 22 March; Beijing is putting up much of the funding for the institute. It and the Shanghai center "aim to become world-class research institutes for brain research," says Luo Minmin, a neuroscientist at the National Institute of Biological Sciences in Beijing and co-director of what will be the northern center.

The Shanghai center will "integrate and expand" ongoing efforts, Zhang says. The city's Huashan Hospital, long a center for neurosurgery, is already building a new facility dedicated to clinical neuroscience, he notes; CAS and the city of Shanghai also put up money for a brain mapping project in 2012 and a scheme to get area brain researchers to collaborate in 2014. The region hosts budding commercial efforts to develop brain-inspired technologies. Zhang estimates some 1 billion yuan ($157 million) annually has been flowing into brain-related research here in recent years.

Although the new center will start recruiting its own research teams later this year, Zhang says, many of its principal investigators will have dual appointments and keep labs at their current institutes. The center will provide a mechanism to share pricey equipment and will establish a database to gather research and clinical information not currently being shared. International cooperation will be a priority.

Once the national effort is underway, the two new centers are likely to become its northern and southern hubs. Researchers expect a high-level committee to set priorities, but the two centers will retain a degree of autonomy. In Beijing, for instance, "we are encouraged by the Board of Trustees, which is led by the Beijing municipal government, to do whatever we think is important," Luo says.

As for the overall project, Yang urges a step-by-step approach. "Given the ambitious scope of the project and the relatively limited resources in China, I don't think it is realistic to start the project all at once."

The next step just might be the launch of Poo's proposed connectome project. It was the topic of a 2 May symposium in Beijing organized by the science ministry. But as for when it might be approved, "I am not at liberty to make any statement at this moment," Poo wrote in an email to *Science*. ■

With reporting by Bian Huihui in Shanghai, China.

PHOTO: ISTOCK.COM/D-KEINE

ETHICS

German law allows use of DNA to predict suspects' looks

Bavaria opens the door to DNA phenotyping, provoking debate about the technique's ethics and accuracy

By **Gretchen Vogel**, *in Berlin*

Police in the German state of Bavaria now have the authority to analyze forensic DNA samples to predict the geographical ancestry and physical characteristics—hair color, eye color, skin color, and age—of an unknown suspect who poses an imminent danger. The controversial law, which passed the Landtag, the state parliament in Munich, on 15 May, is the first in Germany to allow what has been dubbed DNA phenotyping, and it has sparked renewed debate here and in other countries about the advantages—and risks—of such methods.

German federal authorities, and police in the country's 15 other states, are only allowed to perform DNA fingerprinting, in which they look for an exact match between crime scene DNA and samples in a database of known criminals or from a suspect. DNA phenotyping and ancestry prediction, however, have been used in the Netherlands, France, the United Kingdom, Canada, and several U.S. states. Some of those jurisdictions have explicit laws permitting and regulating the practice. In others, police have begun to use DNA phenotyping without explicit regulation. Well-publicized success stories—and advances in the science—have fed interest in the method. Switzerland, like Germany, does not allow such use of DNA, but lawmakers there are considering reversing that ban.

Some critics caution that although the underlying science connecting genetic markers to certain physical features is solid, DNA phenotyping and ancestry prediction can still be easily misunderstood by police and the public. Proponents "exaggerate the numerical certainties," says Veronika Lipphardt of the Albert Ludwigs University of Freiburg in Germany, who studies the history and uses of population genetics. "That creates the impression that it's clear-cut what race someone is or where someone comes from, and that's not true." One U.S. company is even using DNA phenotyping to create facial sketches of suspects, which goes beyond what many scientists now consider credible.

"You would need a lot of training of police forces to use [DNA phenotyping] responsibly," says legal scholar Carsten Momsen of the Free University of Berlin, who worries the technique could lead to the targeting of minority groups. Bavaria's new law doesn't cover investigations of crimes that have already been committed; those are still ruled by federal law. Instead, it allows DNA phenotyping in situations of "imminent danger," when police suspect someone is planning a serious crime. For example, authorities

Traces of DNA from items at a crime scene can predict a person's hair, eye, and skin color—and more, some say.

could analyze trace DNA found on stashes of weapons or bombmaking materials. The uncertainties in the technique can be especially problematic in such high-pressure situations, Momsen says.

Geneticist Manfred Kayser of Erasmus University Medical Center in Rotterdam, the Netherlands, who has developed several of the DNA phenotyping techniques permitted under the new Bavarian law (*Science*, 18 February 2011, p. 838), acknowledges that "there has to be education and training involved. You have to be able to work with probabilities." Still, Kayser notes that police constantly have to decide how trustworthy various kinds of evidence are. At least investigators can quantify the uncertainty in results from DNA analyses, he says. "With eyewitness testimony you have no idea if it's right or wrong."

The techniques attempt to wring information out of traces of blood, skin, semen, or other DNA-containing cells found at a crime scene. They read single-nucleotide polymorphisms (SNPs), individual DNA bases that can vary between people and directly influence or correlate with differences in certain physical features. From databases of SNP variants and people's appearances, scientists can build computer models that predict someone's likely hair, eye, and skin color from their DNA. They can also use age-related chemical changes in DNA to estimate a person's age to within about 4 years.

> ***"You would need a lot of training of police forces to use [DNA phenotyping] responsibly."***
>
> **Carsten Momsen**,
> Free University of Berlin

The accuracy of the results depends on several factors, including the quality—and quantity—of the sample tested and whether the genetic variants found in the sample are well-represented in the model databases. Kayser and his colleagues in Rotterdam and at Indiana University-Purdue University Indianapolis recently released their latest iteration of a tool called HIrisPlex-S. The test returns probabilities for three eye colors (blue, brown, or green/hazel); red, blond, brown, or black hair; and five skin shades, ranging from "very pale" to "dark-black."

In April, they reported in *Forensic Science International: Genetics* that the test correctly predicted all three characteristics for 17 out of 19 simulated crime scene samples. (Two samples did not have high enough quality DNA to make a prediction.) The test also correctly identified five out of six other samples as having cells from two or more people, although it could not say what they looked like. Police in the Netherlands and France have used the test for eye and hair color in several cases. Ancient DNA researchers have also used it to predict characteristics of people known from archaeological remains.

Going beyond skin, hair, and eye color to predicting geographical ancestry is harder. Computer programs can compare a suspect's SNP patterns to those in databases from multiple populations. But some populations may be missing from the databases and others may be indistinguishable because of mixing. Telling whether a person's ancestors were East Asian, African, Native American, Western Eurasian, or from Oceania is fairly straightforward. But predicting whether someone has northern European versus Middle Eastern ancestry is sometimes impossible.

Still more controversial is Parabon NanoLabs in Reston, Virginia, a company that generates facial shape profiles in addition to skin, eye, and hair color and geographical ancestry. Ellen Greytak, its director of bioinformatics, says the firm starts with a standard face based on the person's DNA-derived ancestry and sex. Further analysis of genetic markers leads to a face that "emphasizes the features we predict will be distinctive" about the person, she says.

That goes too far, Kayser says. "The science is not there at all." The company has not published its methods or systematic validation tests, he notes. Greytak says the company doesn't claim to produce an exact portrait. "It's not going to point to an individual. It's going to exclude people. … The vast majority of people are not going to fit the description." The company's website describes several cases where its phenotyping helped authorities find a culprit.

Ironically, DNA phenotyping might have been of little use in the case that helped prompt the new Bavarian regulations. In late 2016, a medical student was raped and murdered in Freiburg; an asylum seeker, originally from Afghanistan, was convicted of the crime. But some authorities complained that they could have narrowed their search more quickly if they had been able to use trace DNA to predict what the culprit looked like and where they might be from.

The crucial evidence in the Freiburg murder case, however, was nongenetic: a strand of hair with an unusual dye pattern—dark at the root, blond at the ends. Police spotted the suspect, who had the same dye pattern, in surveillance video from a tram stop near the crime scene. A DNA test would have only indicated dark hair, not enough to pick out the suspect in the video footage. Kayser agrees that the case wasn't well suited to argue for a law to allow DNA phenotyping. "It was the wrong case to make that claim," he says. ■

PLANETARY SCIENCE

B612 plans asteroid hunt with fleet of small satellites

After abandoning plans for a large space telescope, private foundation pins hopes on new technologies

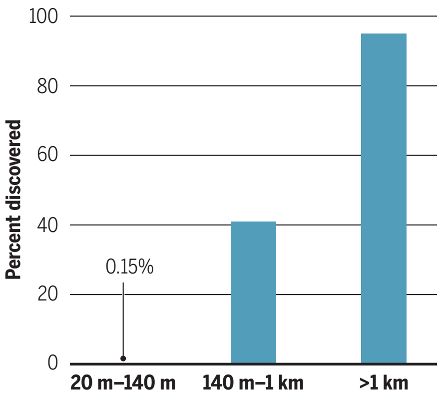
By **Adam Mann**

Last week, an asteroid the size of Egypt's Great Pyramid of Giza whizzed by Earth, missing it by half the distance to the moon. The concern that we may one day not be so lucky has long preoccupied the B612 Foundation, a private organization in Mill Valley, California, dedicated to finding asteroids that cross Earth's orbit and could devastate humanity. B612 itself had a near-death experience 3 years ago, when its bold plans for an asteroid-hunting space telescope fell apart. But now, its ambitions are rising again with a new technique for finding menacing objects.

On 10 May, B612 announced a partnership with York Space Systems, a Denver-

On the hunt

Congress has tasked NASA with finding 90% of near-Earth asteroids bigger than 140 meters. Most of the biggest are known. Organizations like the B612 Foundation are trying to make progress on smaller ones.



CREDITS: (GRAPH) J. YOU/SCIENCE; (DATA) PAUL CHODAS/JET PROPULSION LABORATORY

based maker of standard 85-kilogram satellites, to investigate building a fleet of small asteroid hunters. For many years, B612—which takes its name from the asteroid home of Antoine de Saint-Exupéry's Little Prince—aimed to build and launch a much larger craft, Sentinel, a $450 million space telescope with a 50-centimeter mirror. In 2012, NASA agreed to provide logistical support. But fundraising stalled and, in 2015, the agency ended its agreement with B612 because it wasn't meeting mileposts, essentially killing the telescope. Now, B612 has developed a new technique to do the same thing at a far lower cost with small space telescopes. Ed Lu, B612's co-founder, expects the first telescope to cost about $10 million and believes a full constellation "would be a factor of many, many cheaper" than Sentinel.

Some asteroid astronomers are skeptical of the new approach, saying the technology is far from proven. "To be very, very blunt, what they are proposing and what they've demonstrated is not going to help us find more NEOs [near-Earth objects]," says astronomer Timothy Spahr, CEO of space consultancy firm NEO Sciences in Marlborough, Massachusetts, who does independent work for NASA.

Astronomers believe they have already discovered 95% of the NEOs of civilization-ending size, a kilometer or more across. But ones between 140 meters and 1 kilometer across—still big enough to wipe out a city—are so faint that an estimated 59% remain undiscovered. Finding them, and perhaps developing plans for diverting threats, requires a big telescope, like Sentinel, or long exposures on a smaller one. But if the asteroids are moving fast, they shift from pixel to pixel on the telescope's camera, and the long exposures provide little benefit.

In 2013, astronomer Michael Shao of the Jet Propulsion Laboratory (JPL) in Pasadena, California, and his colleagues outlined a technique called synthetic tracking, which gets around this issue by taking up to 100 snapshots in a single second. A dim asteroid won't appear in any particular image, but by arranging and stacking the photos in thousands of combinations, the system can get lucky and preferentially brighten an unknown fast-moving object.

Doing this for an unknown asteroid of unknown brightness and speed at an unknown point in the sky is computationally demanding. But Shao says advanced processing power makes it feasible. He adds that the technique would help a small telescope do as good a job as a bigger one, even though it gathers less light. It could also help uncover the fastest moving asteroids, such as the object 'Oumuamua, an interstellar asteroid spotted zipping through the solar system last year.

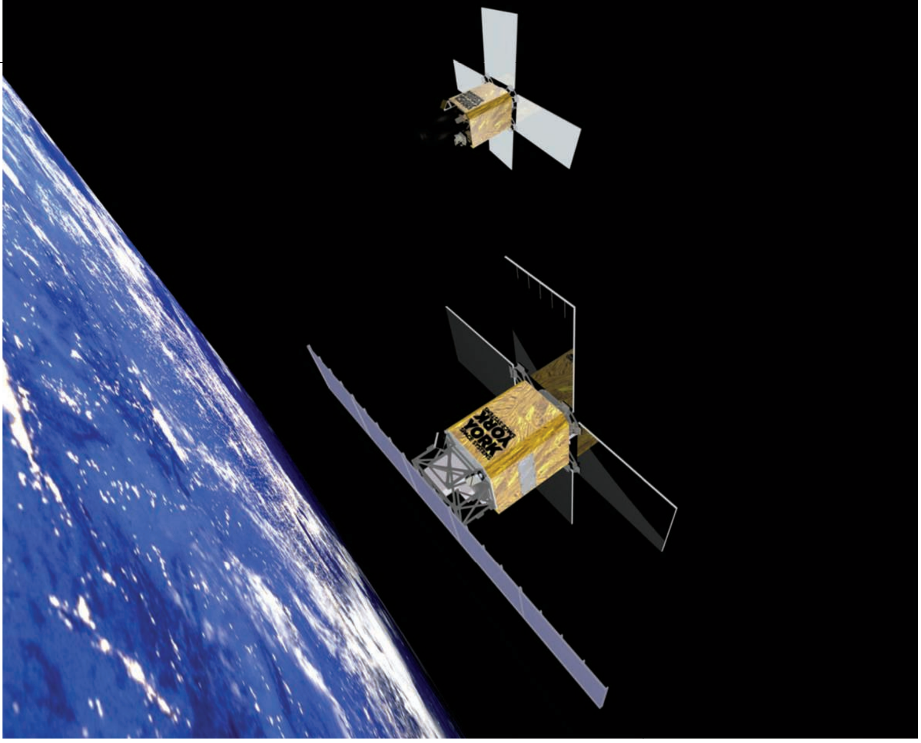

The B612 Foundation is partnering with York Space Systems to build a fleet of small asteroid-spotting satellites.

IMAGE: ARTIST RENDERING OF AN ICEYE SAR AND COMMERCIAL OPTICAL TELESCOPE SPACECRAFT BEING DEVELOPED WITH YORK SPACE SYSTEMS

Shao and his team simulated how nine 20-centimeter telescopes in orbit around the sun would fare with the technique. In a 2015 study, they projected that the constellation could discover 90% of Earth-grazing asteroids 50 meters or larger in diameter in 5 years, three times faster than other techniques and faster than Sentinel itself. Shao says within a week of publishing his paper, B612 contacted him.

Last year, B612 tested the technique in space by asking the Earth-observing company Planet to turn one of its SkySat satellites around and aim its 35-centimeter telescope at already known asteroids. By applying synthetic tracking, the team could tease out images of the faint objects.

Now, B612 needs to raise money for its new plan. But the asteroid-hunting field is becoming more crowded. Starting in 2021, the Large Synoptic Survey Telescope, an enormous 8.4-meter observatory in Chile, will photograph large parts of the sky each night and could increase the number of known NEOs from about 18,000 to more than 100,000. NASA would also like to build the Near-Earth Object Camera (NEOCam), a space-based mission that would replicate many of Sentinel's goals and abilities. Although the mission was not selected in a competition last year, it was given money for continued study, a sign that it might one day fly.

Lindley Johnson, NASA's planetary defense officer in Washington, D.C., says NEOCam could launch 4 years after it's given the go-ahead, and that it would likely be better at finding asteroids than B612's constellation. He says synthetic tracking is good for detecting fast-moving objects near Earth, like debris and satellites, but won't work as well for faint and distant objects. "I think they are overoptimistic in their assessments," he says. Shao, however, says experts at JPL have assured him that small satellites can provide the power, computation, and communications needed for synthetic asteroid tracking to work.

Mark Sykes, CEO and director of the Planetary Science Institute in Tucson, Arizona, says any of these asteroid missions could pay off for research as well as for planetary protection. Spotting asteroids passing closer than the moon, like the 15 May object, would allow researchers to explore at close range objects that hail from the distant solar system. "We would have a one-stop shopping center for sampling material that has formed throughout the solar system," he says, adding that asteroid mining companies would also be interested in such near-Earth targets.

Lu shares those hopes. He wants B612 to compile nearby asteroids in a user-friendly catalog, like a Google Maps for the solar system—a step that he says would help future scientists, explorers, and miners. "If you buy into the idea that humans will someday live and work and economically operate in space, then you need a map," he says. ■

Adam Mann is a journalist in Oakland, California.

Corrected 24 May 2018. See full text.

INFECTIOUS DISEASES

A call to arms against the other retrovirus

HTLV-1, discovered just before HIV but almost forgotten, infects millions and causes cancer

By **Kai Kupferschmidt**

The discovery of the first human retrovirus in 1980 was a minor scientific sensation. Researchers knew retroviruses—which transcribe their RNA genome into DNA and integrate it into a host cell's genome—existed in animals. But until Robert Gallo, then at the National Cancer Institute in Bethesda, Maryland, found the human T-cell leukemia virus-1 (HTLV-1), some doubted that retroviruses infect humans. HTLV-1 was soon eclipsed, however, by another retrovirus that would go on to kill more than 35 million people and keep generations of scientists busy: HIV. "If you were working on retroviruses you switched to HIV," says epidemiologist Antoine Gessain of the Pasteur Institute in Paris.

In some aboriginal communities in Australia, almost half the people over age 50 are infected with HTLV-1.

PHOTO: ANOEK DE GROOT/AFP/GETTY IMAGES

But in a 10 May open letter to World Health Organization (WHO) Director-General Tedros Adhanom Ghebreyesus, Gallo and 59 other virologists, epidemiologists, and patient advocates call for a global effort to eradicate HTLV-1, which infects millions and causes cancer and several other diseases. Just like HIV, HTLV-1 spreads through blood, semen, and breast milk, and the letter argues that the testing and prevention strategies used against HIV should be employed to stop HTLV-1.

HTLV-1 is nowhere to be found among the many dozens of diseases on fact sheets on WHO's website; it's not recognized as a neglected tropical disease and isn't on WHO's list of sexually transmitted diseases either. "I still scratch my head about how we managed to stay on the margins of HIV and of sexual health," says Graham Taylor, who studies HTLV-1 at Imperial College London.

The virus probably jumped from monkeys to humans somewhere in Africa tens of thousands of years ago; the slave trade helped spread it around the world. Today, scientists estimate that 5 million to 10 million people are infected, but the real number is probably much higher, Taylor says. One hot spot is southern Japan, where up to 6% of pregnant women have been found to carry the virus. But poor countries are most affected, which partly explains the neglect, Gallo says. In Jamaica and other Caribbean islands, about 6% of the entire population may carry the virus. Parts of Brazil have high rates, millions are infected in Africa, and in some aboriginal communities in Australia, almost half the people over age 50 are infected, a recent study found. Europe and the United States can't rest easy, Taylor says: "If we are not careful, this could become more prevalent in our communities without us being aware. … As other infections like HIV come under control, people's behavior may change."

Most people infected with HTLV-1 never have symptoms, but about 3% to 5% develop adult T-cell leukemia, a cancer that kills patients after 8 to 10 months on average and for which treatment has not improved in the past quarter-century. Another 4% develop tropical spastic paraparesis, a disease similar to multiple sclerosis. Other inflammatory diseases and immune deficiency have also been reported, and an Australian study published in March showed that those infected with HTLV-1 are more likely to develop bronchiectasis, a disease in which parts of the airways are enlarged, and to die from it.

Yet HTLV-1's modest death toll, and the decades that often go by between infection and disease, have kept it out of the limelight. In response to the letter, a WHO spokesperson pointed out that another virus, hepatitis B, causes more than 400,000 cancers each year worldwide. In contrast, the International Agency for Research on Cancer estimates that only 3000 cancer diagnoses annually are directly linked to HTLV-1. Taylor thinks the real number is higher, but he says, "I really don't think this should turn into a 'my disease is more important than your disease' competition."

Routine testing for HTLV-1 should be available in sexual health clinics everywhere, the researchers write in the letter, and mothers in endemic regions should routinely be screened for HTLV-1 and advised not to breastfeed if they are infected. Since antenatal testing was introduced in Japan's Nagasaki region in 1987, the infection rate in the population has dropped from 7.2% to 1%. The letter also calls for universal HTLV-1 testing of blood and organ donors, because transplants have transmitted the virus, albeit rarely. "I honestly believe that nobody should have a transplant where the donor has not been tested for HTLV-1," Taylor says. (Until 2009, the United States did screen organ donors, but the practice was stopped because false positive results were disqualifying too many healthy organs.)

Gallo believes renewed attention to HTLV-1 could have broader benefits. In the early 1980s, the virus sped up the discovery of HIV, he says, because it alerted scientists to the possibility that a retrovirus might be causing the mysterious new syndrome called AIDS. Today, a better understanding of HTLV-1's powerful carcinogenic properties may lead researchers to new insights about cancer, Gallo says. ■

FEATURES

BASIC INSTINCTS

Some say artificial intelligence needs to learn like a child

By **Matthew Hutson**

Babies are born with instincts that help us learn common sense, so far elusive for AI algorithms.

PHOTO: NATALIADERIABINA/ISTOCKPHOTO

It's a Saturday morning in February, and Chloe, a curious 3-year-old in a striped shirt and leggings, is exploring the possibilities of a new toy. Her father, Gary Marcus, a developmental cognitive scientist at New York University (NYU) in New York City, has brought home some strips of tape designed to adhere Lego bricks to surfaces. Chloe, well-versed in Lego, is intrigued. But she has always built upward. Could she use the tape to build sideways or upside down? Marcus suggests building out from the side of a table. Ten minutes later, Chloe starts sticking the tape to the wall. "We better do it before Mama comes back," Marcus says in a singsong voice. "She won't be happy." (Spoiler: The wall paint suffers.)

Implicit in Marcus's endeavor is an experiment. Could Chloe apply what she had learned about an activity to a new context? Within minutes, she has a Lego sculpture sticking out from the wall. "Papa, I did it!" she exclaims. In her adaptability, Chloe is demonstrating common sense, a kind of intelligence that, so far, computer scientists have struggled to reproduce. Marcus believes the field of artificial intelligence (AI) would do well to learn lessons from young thinkers like her.

Researchers in machine learning argue that computers trained on mountains of data can learn just about anything—including common sense—with few, if any, programmed rules. These experts "have a blind spot, in my opinion," Marcus says. "It's a sociological thing, a form of physics envy, where people think that simpler is better." He says computer scientists are ignoring decades of work in the cognitive sciences and developmental psychology showing that humans have innate abilities—programmed instincts that appear at birth or in early childhood—that help us think abstractly and flexibly, like Chloe. He believes AI researchers ought to include such instincts in their programs.

Yet many computer scientists, riding high on the successes of machine learning, are eagerly exploring the limits of what a naïve AI can do. "Most machine learning people, I think, have a methodological bias against putting in large amounts of background knowledge because in some sense we view that as a failure," says Thomas Dietterich, a computer scientist at Oregon State University in Corvallis. He adds that computer scientists also appreciate simplicity and have an aversion to debugging complex code. Big companies such as Facebook and Google are another factor pushing AI in that direction, says Josh Tenenbaum, a psychologist at the Massachusetts Institute of Technology (MIT) in Cambridge. Those companies are most interested in narrowly defined, near-term problems, such as web search and facial recognition, in which blank-slate AI systems can be trained on vast data sets and work remarkably well.

But in the longer term, computer scientists expect AIs to take on much tougher tasks that require flexibility and common sense. They want to create chatbots that explain the news, autonomous taxis that can handle chaotic city traffic, and robots that nurse the elderly. "If we want to build robots that can actually interact in the full human world like C-3PO," Tenenbaum says, "we're going to need to solve all of these problems in much more general settings."

Some computer scientists are already trying. In February, MIT launched Intelligence Quest, a research initiative now raising hundreds of millions of dollars to understand human intelligence in engineering terms. Such efforts, researchers hope, will result in AIs that sit somewhere between pure machine learning and pure instinct. They will boot up following some embedded rules, but will also learn as they go. "In some sense this is like the age-old nature-nurture debate, now translated into engineering terms," Tenenbaum says.

Part of the quest will be to discover what babies know and when—lessons that can then be applied to machines. That will take time, says Oren Etzioni, CEO of the Allen Institute for Artificial Intelligence (AI2) in Seattle, Washington. AI2 recently announced a $125 million effort to develop and test common sense in AI. "We would love to build on the representational structure innate in the human brain," Etzioni says, "but we don't understand how the brain processes language, reasoning, and knowledge."

Ultimately, Tenenbaum says, "We're trying to take one of the oldest dreams of AI seriously: that you could build a machine that grows into intelligence the way a human does—that starts like a baby and learns like a child.”

IN THE PAST FEW YEARS, AI has shown that it can translate speech, diagnose cancer, and beat humans at poker. But for every win, there is a blunder. Image recognition algorithms can now distinguish dog breeds better than you can, yet they sometimes mistake a chihuahua for a blueberry muffin. AIs can play classic Atari video games such as *Space Invaders* with superhuman skill, but when you remove all the aliens but one, the AI falters inexplicably.

Machine learning—one type of AI—is responsible for those successes and failures. Broadly, AI has moved from software that relies on many programmed rules (also known as Good Old-Fashioned AI, or GOFAI) to systems that learn through trial and error. Machine learning has taken off thanks to powerful computers, big data, and advances in algorithms called neural networks. Those networks are collections of simple computing elements, loosely modeled on neurons in the brain, that create stronger or weaker links as they ingest training data.

With its Alpha programs, Google's DeepMind has pushed deep learning to its apotheosis. Each time rules were subtracted, the software seemed to improve. In 2016, AlphaGo beat a human champion at Go, a classic Chinese strategy game. The next year, AlphaGo Zero easily beat AlphaGo with far fewer guidelines. Months later, an even simpler system called AlphaZero beat AlphaGo Zero—and also mastered chess. In 1997, a classic, rule-based AI, IBM's Deep Blue, had defeated chess champion Garry Kasparov. But it turns out that true chess virtuosity lies in knowing the exceptions to the exceptions to the exceptions—information best gleaned through experience. AlphaZero, which learns by playing itself over and over, can beat Deep Blue, today's best chess programs, and every human champion.

Yet systems such as Alpha clearly are not extracting the lessons that lead to common sense. To play Go on a 21-by-21 board instead of the standard 19-by-19 board, the AI would have to learn the game anew. In the late 1990s, Marcus trained a network to take an input number and spit it back out—about the simplest task imaginable. But he trained it only on even numbers. When tested with odd numbers, the network floundered. It couldn't apply learning from one domain to another, the way Chloe had when she began to build her Lego sideways.

The answer is not to go back to rule-based GOFAIs. A child does not recognize a dog with explicit rules such as “if number of legs=4, and tail=true, and size>cat.” Recognition is more nuanced—a chihuahua with three legs won't slip past a 3-year-old. Humans are not blank slates, nor are we hardwired. Instead, the evidence suggests we have predispositions that help us learn and reason about the world. Nature doesn't give us a library of skills, just the scaffolding to build one.

Harvard University psychologist Elizabeth Spelke has argued that we have at least four “core knowledge” systems giving us a head start on understanding objects, actions, numbers, and space. We are intuitive physicists, for example, quick to understand objects and their interactions. According to one study, infants just 3 days old interpret the two ends of a partially hidden rod as parts of one entity—a sign that our brains might be predisposed to perceive cohesive objects. We're also intuitive psychologists. In a 2017 *Science* study, Shari Liu, a graduate student in Spelke's lab, found that 10-month-old infants could infer that when an animated character climbs a bigger hill to reach one shape than to reach another, the character must prefer the former. Marcus has shown that 7-month-old infants can learn rules; they show surprise when three-word sentences (“wo fe fe”) break the grammatical pattern of previously heard sentences (“ga ti ga”). According to later research, day-old newborns showed similar behavior.

Marcus has composed a minimum list of 10 human instincts that he believes should be baked into AIs, including notions of causality, cost-benefit analysis, and types versus instances (dog versus my dog). Last October at NYU, he argued for his list in a debate on whether AI needs “more innate machinery,” facing Yann LeCun, an NYU computer scientist and Facebook's chief AI scientist. To demonstrate his case for instinct, Marcus showed a slide of baby ibexes descending a cliff. “They don't get to do million-trial learning,” he said. “If they make a mistake, it's a problem.”


In a triumph of machine learning, AlphaGo beat Go champion Ke Jie in 2017.

PHOTO: XU YU/XINHUA/ALAMY STOCK PHOTO

LeCun, disagreeing with many developmental psychologists, argued that babies might be learning such abilities within days, and if so, machine learning algorithms could, too. His faith comes from experience. He works on image recognition, and in the 1980s he began arguing that hand-coded algorithms to identify features in pictures would become unnecessary. Thirty years later, he was proved right. Critics asked him: “Why learn it when you can build it?” His reply: Building is hard, and if you don't fully understand how something works, the rules you devise are likely to be wrong.

But Marcus pointed out that LeCun himself had embedded one of the 10 key instincts into his image-recognition algorithms: translational invariance, the ability to recognize an object no matter where it appears in the visual field. Translational invariance is the principle behind convolutional neural networks, or convnets, LeCun's greatest claim to fame. In the past 5 years they've become central to image recognition and other AI applications, kicking off the current craze for deep learning.

LeCun tells *Science* that translational invariance, too, could eventually emerge on its own with better general learning mechanisms. “A lot of those items will kind of spontaneously pop up as a consequence of learning how the world works,” he says. Geoffrey Hinton, a pioneer of deep learning at the University of Toronto in Canada, agrees. “Most of the people who believe in strong innate knowledge have an unfounded belief that it's hard to learn billions of parameters from scratch,” he says. “I think recent progress in deep learning has shown that it is actually surprisingly easy.”

The debate over where to situate an AI on a spectrum between pure learning and pure instinct will continue. But that issue overlooks a more practical concern: how to design and code such a blended machine. How to combine machine learning—and its billions of neural network parameters—with rules and logic isn't clear. Neither is how to identify the most important instincts and encode them flexibly. But that hasn't stopped some researchers and companies from trying.

A ROBOTICS LABORATORY at The University of New South Wales in Sydney, Australia, is dressed to look like a living room and kitchen—complete with bottles of James Boag's Premium Lager in the fridge. Computer scientist Michael Thielscher explains that the lab is a testbed for a domestic robot. His team is trying to endow a Toyota Human Support Robot (HSR), which has one arm and a screen for a face, with two humanlike instincts. First, they hope to program the HSR to decompose challenges into smaller, easier problems, just as a person would parse a recipe into several steps. Second, they want to give the robot the ability to reason about beliefs and goals, the way humans instinctively think about the minds of others. How would the HSR respond if a person asked it to fetch a red cup, and it found only a blue cup and a red plate?

So far, their software shows some humanlike abilities, including the good sense to fetch the blue cup rather than the red plate. But more rules are programmed into the system than Thielscher would like. His team has had to tell their AI that cup is usually more important than red. Ideally, a robot would have the social instincts to quickly learn people's preferences on its own.

Other researchers are working to inject their AIs with the same intuitive physics that babies seem to be born with. Computer scientists at DeepMind in London have developed what they call interaction networks. They incorporate an assumption about the physical world: that discrete objects exist and have distinctive interactions. Just as infants quickly parse the world into interacting entities, those systems readily learn objects' properties and relationships. Their results suggest that interaction networks can predict the behavior of falling strings and balls bouncing in a box far more accurately than a generic neural network.

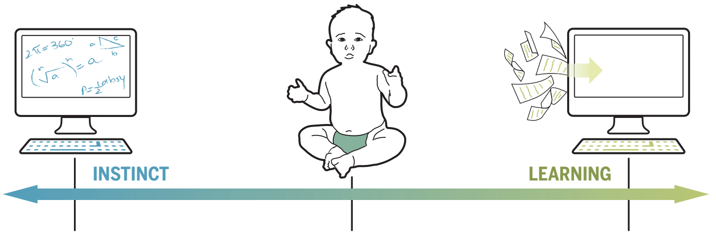
Vicarious, a robotics software company in San Francisco, California, is taking the idea further with what it calls schema networks. Those systems, too, assume the existence of objects and interactions, but they also try to infer the causality that connects them. By learning over time, the company's software can plan backward from desired outcomes, as people do. (I want my nose to stop itching; scratching it will probably help.) The researchers compared their method with a state-of-the-art neural network on the Atari game *Breakout*, in which the player slides a paddle to deflect a ball and knock out bricks. Because the schema network could learn about causal relationships—such as the fact that the ball knocks out bricks on contact no matter its velocity—it didn't need extra training when the game was altered. You could move the target bricks or make the player juggle three balls, and the schema network still aced the game. The other network flailed.

Besides our innate abilities, humans also benefit from something most AIs don't have: a body. To help software reason about the world, Vicarious is "embodying" it so it can explore virtual environments, just as a baby might learn something about gravity by toppling a set of blocks. In February, Vicarious presented a system that looked for bounded regions in 2D scenes by essentially having a tiny virtual character traverse the terrain. As it explored, the system learned the concept of containment, which helps it make sense of new scenes faster than a standard image-recognition convnet that passively surveyed each scene in full. Concepts—knowledge that applies to many situations—are crucial for common sense. "In robotics it's extremely important that the robot be able to reason about new situations," says Dileep George, a co-founder of Vicarious. Later this year, the company will pilot test its software in warehouses and factories, where it will help robots pick up, assemble, and paint objects before packaging and shipping them.

One of the most challenging tasks is to code instincts flexibly, so that AIs can cope with a chaotic world that does not always follow the rules. Autonomous cars, for example, cannot count on other drivers to obey traffic laws. To deal with that unpredictability, Noah Goodman, a psychologist and computer scientist at Stanford University in Palo Alto, California, helps develop probabilistic programming languages (PPLs). He describes them as combining the rigid structures of computer code with the mathematics of probability, echoing the way people can follow logic but also allow for uncertainty: If the grass is wet it probably rained—but maybe someone turned on a sprinkler. Crucially, a PPL can be combined with deep learning networks to incorporate extensive learning. While working at Uber, Goodman and others invented such a "deep PPL," called Pyro. The ride-share company is exploring uses for Pyro such as dispatching drivers and adaptively planning routes amid road construction and game days. Goodman says PPLs can reason not only about physics and logistics, but also about how people communicate, coping with tricky forms of expression such as hyperbole, irony, and sarcasm.

Different minds

Over time, artificial intelligence (AI) has shifted from algorithms that rely on programmed rules and logic—instincts—to machine learning, where algorithms contain few rules and ingest training data to learn by trial and error. Human minds sit somewhere in the middle.



Rule-based AI
IBM's Deep Blue, which bested chess champion Garry Kasparov in 1997, relied on rules and logic.

Humans
Babies learn by trial and error. But developmental cognitive scientists say we also begin life with basic instincts that help us gain a flexible common sense.

Machine learning AI
With few programmed rules, DeepMind's AlphaZero today can beat IBM's Deep Blue at chess. But it doesn't generalize.

GRAPHIC: N. DESAI/SCIENCE

CHLOE MIGHT NOT MASTER sarcasm until her teen years, but her inborn knack for language is already clear. At one point in Marcus's apartment, she holds out a pair of stuck Lego bricks. "Papa, can you untach this for me?" Her father obliges without correcting her coinage. Words and ideas are like Lego pieces, their parts readily mixed and matched, and eagerly tested in the world.

After Chloe tires of building on the wall, an older, slightly more seasoned intelligent system gets a chance to try it: her brother Alexander, age 5. He quickly constructs a Lego building that protrudes farther. "You can see the roots of what he's doing in what she did," Marcus says. When asked, Alexander estimates how far the structure might stick out before it would fall. "He's pretty well-calibrated," Marcus observes. "He hasn't had 10 million trials of glued-on-the-wall Lego things in order to assess the structural integrity. He's taking what he knows about physics, and so forth, and making some inferences."

Marcus is obviously proud, not only of his offspring's capabilities, but also that they uphold his theories of how we learn about the world—and how AIs should be learning, too. Done with their Lego buildings, Chloe and Alexander leap into their father's arms. They squeal with delight as he spins them around and around, offering them another chance to fine-tune their intuitive senses of physics—and fun. ■

Matthew Hutson is a journalist based in New York City.

PHOTO: PAUL TAYLOR/GETTY IMAGES

THE WAR ON

Wheat sensitivity isn't imaginary, most researchers now agree. But what's really behind it?

By **Kelly Servick**

The patients weren't crazy—Knut Lundin was sure of that. But their ailment was a mystery. They were convinced gluten was making them sick. Yet they didn't have celiac disease, an autoimmune reaction to that often-villainized tangle of proteins in wheat, barley, and rye. And they tested negative for a wheat allergy. They occupied a medical no man's land.

About a decade ago, gastroenterologists like Lundin, based at the University of Oslo, came across more and more of those enigmatic cases. "I worked with celiac disease and gluten for so many years," he says, "and then came this wave." Gluten-free choices began appearing on restaurant menus and creeping onto grocery store shelves. By 2014, in the United States alone, an estimated 3 million people without celiac disease had sworn off gluten. It was easy to assume that people claiming to be "gluten sensitive" had just been roped into a food fad.

"Generally, the reaction of the gastroenterologist [was] to say, 'You don't have celiac disease or wheat allergy. Goodbye,'" says Armin Alaedini, an immunologist at Columbia University. "A lot of people thought this is perhaps due to some other [food] sensitivity, or it's in people's heads."

GLUTEN

Stretchy, resilient gluten allows bread to rise. But millions now shun it.

But a small community of researchers started searching for a link between wheat components and patients' symptoms—commonly abdominal pain, bloating, and diarrhea, and sometimes headaches, fatigue, rashes, and joint pain. That wheat really can make nonceliac patients sick is now widely accepted. But that's about as far as the agreement goes.

As data trickle in, entrenched camps have emerged. Some researchers are convinced that many patients have an immune reaction to gluten or another substance in wheat—a nebulous illness sometimes called nonceliac gluten sensitivity (NCGS).

Others believe most patients are actually reacting to an excess of poorly absorbed carbohydrates present in wheat and many other foods. Those carbohydrates—called FODMAPs, for fermentable oligosaccharides, disaccharides, monosaccharides, and polyols—can cause bloating when they ferment in the gut. If FODMAPs are the primary culprit, thousands of people may be on gluten-free diets with the support of their doctors and dietitians but without good reason.

Those competing theories were on display in a session on wheat sensitivity at a celiac disease symposium held at Columbia in March. In back-to-back talks, Lundin made the case for FODMAPs, and Alaedini for an immune reaction. But in an irony that underscores how muddled the field has become, both researchers started their quests believing something completely different.

KNOWN WHEAT-RELATED ILLNESSES have clear mechanisms and markers. People with celiac disease are genetically predisposed to launch a self-destructive immune response when a component of gluten called gliadin penetrates their intestinal lining and sets off inflammatory cells in the tissue below. People with a wheat allergy respond to wheat proteins by churning out a class of antibodies called immunoglobulin E that can set off vomiting, itching, and shortness of breath. The puzzle, for both doctors and researchers, is patients who lack both the telltale antibodies and the visible damage to their intestines but who feel real relief when they cut out gluten-containing food.

Some doctors have begun to approve and even recommend a gluten-free diet. "Ultimately, we're here not to do science, but to improve quality of life," says Alessio Fasano, a pediatric gastroenterologist at Massachusetts General Hospital in Boston who has studied NCGS and written a book on living gluten-free. "If I have to throw bones on the ground and look at the moon to make somebody better, even if I don't understand what that means, I'll do it."

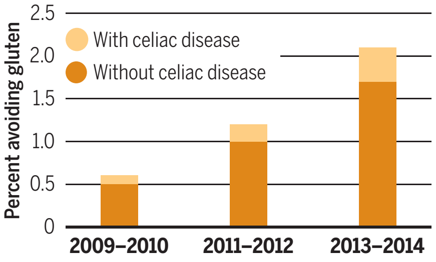
Like many doctors, Lundin believed that (fad dieters and superstitious eaters aside) some patients have a real wheat-related ailment. His group helped dispel the notion that NCGS was purely psychosomatic. They surveyed patients for unusual levels of psychological distress that might express itself as physical symptoms. But the surveys showed no differences between those patients and people with celiac disease, the team reported in 2012. As Lundin bluntly puts it: "We know they are not crazy."

Still, skeptics worried that the field had seized on gluten with shaky evidence that it was the culprit. After all, nobody eats gluten in isolation. "If we did not know about the specific role of gluten in celiac disease, we would never have thought gluten was responsible for [NCGS]," says Stefano Guandalini, a pediatric gastroenterologist at the University of Chicago Medical Center in Illinois. "Why blame gluten?"

Defenders of NCGS generally acknowledge that other components of wheat might contribute to symptoms. In 2012, a group of proteins in wheat, rye, and barley called amylase trypsin inhibitors emerged as a potential offender, for example, after a team led by biochemist Detlef Schuppan of Johannes Gutenberg University Mainz in Germany (then at Harvard Medical School in Boston) reported that those proteins can provoke immune cells.

Against the grain

Data from the National Health and Nutrition Examination Survey show the rising tide of gluten avoidance by people without celiac disease. Celiac diagnoses also rose, but probably not its actual prevalence.



But without biological markers to identify people with NCGS, researchers have relied on self-reported symptoms measured through a "gluten challenge": Patients rate how they feel before and after cutting out gluten. Then doctors reintroduce gluten or a placebo—ideally disguised in indistinguishable pills or snacks—to see whether the symptoms tick back up.

Alaedini has recently hit on a more objective set of possible biological markers—much to his own surprise. "I entered this completely as a skeptic," he says. Over his career, he has gravitated toward studying spectrum disorders, in which diverse symptoms have yet to be united under a clear biological cause—and where public misinformation abounds. His team published a study in 2013, for example, that debunked the popular suggestion that children with autism had high rates of Lyme disease. "I do studies [where] there is a void," he says.

In NCGS, Alaedini saw another poorly defined spectrum disorder. He did accept that patients without celiac disease might somehow be sensitive to wheat, on the basis of several trials that measured symptoms after a blinded challenge. But he was not convinced by previous studies claiming that NCGS patients were more likely than other people to have certain antibodies to gliadin. Many of those studies lacked a healthy control group, he says, and relied on commercial antibody kits that gave murky and inconsistent readings.

In 2012, he contacted researchers at the University of Bologna in Italy to obtain blood samples from 80 patients their team had identified as gluten sensitive on the basis of a gluten challenge. He wanted to test the samples for signs of a unique immune response—a set of signaling molecules different from those in the blood of healthy volunteers and celiac patients. He wasn't optimistic. "I thought if we were going to see something, like with a lot of spectrum conditions that I have looked at, we would see small differences."

The results shocked him. Compared with both healthy people and those with celiac, these patients had significantly higher levels of a certain class of antibodies against gluten that suggest a short-lived, systemic immune response. That didn't mean gluten itself was causing disease, but the finding hinted that the barrier of those patients' intestines might be defective, allowing partially digested gluten to get out of the gut and interact with immune cells in the blood. Other elements—such as immune response–provoking bacteria—also might be escaping. Sure enough, the team found elevated levels of two proteins that indicate an inflammatory response to bacteria. And when 20 of the same patients spent 6 months on a gluten-free diet, their blood levels of those markers declined.

For Alaedini, the beginnings of a mechanism emerged: Some still-unidentified wheat component prompts the intestinal lining to become more permeable. (An imbalance in gut microbes might be a predisposing factor.) Components of bacteria then seem to sneak past immune cells in the underlying intestinal tissue and make their way to the bloodstream and liver, prompting inflammation.

GRAPH: J. YOU/*SCIENCE*

“This is a real condition, and there can be objective, biological markers for it,” Alaedini says. “That study changed a lot of minds, including my own.”

The study also impressed Guandalini, a longtime skeptic about the role of gluten. It “opens the way to finally reach an identifiable marker for this condition,” he says.

BUT OTHERS SEE the immune-response explanation as a red herring. To them, the primary villain is FODMAPs. The term, coined by gastroenterologist Peter Gibson at Monash University in Melbourne, Australia, and his team, encompasses a smorgasbord of common foods. Onions and garlic; legumes; milk and yogurt; and fruits including apples, cherries, and mangoes are all high in FODMAPs. So is wheat: Carbs in wheat called fructans can account for as much as half of a person’s FODMAP intake, dietitians in Gibson’s group have estimated. The team found that those compounds ferment in the gut to cause symptoms of irritable bowel syndrome, such as abdominal pain, bloating, and gas.

Gibson has long been skeptical of studies implicating gluten in such symptoms, arguing that those findings are hopelessly clouded by the nocebo effect, in which the mere expectation of swallowing the dreaded ingredient worsens symptoms. His team found that most patients couldn’t reliably distinguish pure gluten from a placebo in a blinded test. He believes that many people feel better after eliminating wheat not because they have calmed some intricate immune reaction, but because they’ve reduced their intake of FODMAPs.

Lundin, who was firmly in the immune-reaction camp, didn’t believe that FODMAPs could explain away all his patients. “I wanted to show that Peter was wrong,” he says. During a 2-week sabbatical in the Monash lab, he found some quinoa-based snack bars designed to disguise the taste and texture of ingredients. “I said, ‘We’re going to take those muesli bars and we’re going to do the perfect study.’”

His team recruited 59 people on self-instituted gluten-free diets and randomized them to receive one of three indistinguishable snack bars, containing isolated gluten, isolated FODMAP (fructan), or neither. After eating one type of bar daily for a week, they reported any symptoms. Then they waited for symptoms to resolve and started on a different bar until they had tested all three.

Before analyzing patient responses, Lundin was confident that gluten would cause the worst symptoms. But when the study’s blind was lifted, only the FODMAP symptoms even cleared the bar for statistical significance. Twenty-four of the 59 patients had their highest symptom scores after a week of the fructan-laced bars. Twenty-two responded most to the placebo, and just 13 to gluten, Lundin and his collaborators—who included Gibson—reported last November in the journal *Gastroenterology*. Lundin now believes FODMAPs explain the symptoms in most wheat-avoiding patients. “My main reason for doing that study was to find out a good method of finding gluten-sensitive individuals,” he says. “And there were none. And that was quite amazing.”

AT THE COLUMBIA MEETING, Alaedini and Lundin went head to head in consecutive talks titled “It’s the Wheat” and “It’s FODMAPS.” Each has a list of criticisms of the other’s study. Alaedini contends that by recruiting broadly from the gluten-free population, instead of finding patients who reacted to wheat in a challenge, Lundin likely failed to include people with a true wheat sensitivity. Very few of Lundin’s subjects reported symptoms outside the intestines, such as rash or fatigue, that might point to a widespread immune condition, Alaedini says. And he notes that the increase in patients’ symptoms in response to the FODMAP snacks was just barely statistically significant.

Lundin, meanwhile, points out that the patients in Alaedini’s study didn’t go through a blinded challenge to check whether the immune markers he identified really spiked in response to wheat or gluten. The markers may not be specific to people with a wheat sensitivity, Lundin says.

Although consumers focus on gluten, other wheat components could be at the root of symptoms.

PHOTO: MATT RAINEY/*THE NEW YORK TIMES*

Despite the adversarial titles of their talks, the two researchers find a lot of common ground. Alaedini agrees that FODMAPs explain some of the wheat-avoidance phenomenon. And Lundin acknowledges that some small population may really have an immune reaction to gluten or another component of wheat, though he sees no good way to find them.

After the meeting, Elena Verdú, a gastroenterologist at McMaster University in Hamilton, Canada, puzzled over the polarization of the field. “I don’t understand why there is this need to be so dogmatic about ‘it is this, it is not that,’” she says.

She worries that the scientific confusion breeds skepticism toward people who avoid gluten for medical reasons. When she dines with celiac patients, she says, waiters sometimes meet requests for gluten-free food with smirks and questions. Meanwhile, the conflicting messages may send nonceliac patients down a food-avoidance rabbit hole. “Patients are withdrawing gluten first, then lactose, and then FODMAPs—and then they are on a really, really poor diet,” she says.

But Verdú believes careful research will ultimately break through the superstitions. She is president of the North American Society for the Study of Celiac Disease, which this year awarded its first grant to study nonceliac wheat sensitivity. She’s hopeful that the search for biomarkers like those Alaedini has proposed will show that inside the monolith of gluten avoidance lurk multiple, nuanced conditions. “It will be difficult,” she says, “but we are getting closer.” ■

INSIGHTS

PERSPECTIVES

White storks migrate in large flocks between their breeding sites in Europe and their overwintering grounds in Africa.

PHOTO: NATURE PHOTOGRAPHERS LTD/ALAMY STOCK PHOTO

ECOLOGY

Following the leader, for better or worse

Juvenile white storks use different behavioral strategies for thermal soaring

By **Gabrielle A. Nevitt**

White storks (*Ciconia ciconia*) are highly social birds that migrate thousands of kilometers each year between breeding locations in Europe and overwintering grounds in Africa. The birds migrate during the day in large, seemingly disorganized flocks and rely on thermals to power their flight (*1*). On page 911 of this issue, Flack *et al.* (*2*) provide a glimpse into the behavioral interactions that white storks use to exploit atmospheric thermals in a migratory flock.

Thermals form when air over Earth's surface heats up during daylight hours. Warm air rises because it is less dense, and the resulting thermal provides a source of vertical lift that soaring birds can exploit (*1*). Soaring birds like white storks have evolved adaptations, such as broad stiff wings, that enable them to soar on thermals like a glider, thus allowing them to travel more efficiently than they could by flapping flight alone (*3*). However, because thermals are not always predictable, the challenge for white storks is to find appropriate thermals that they can use to gain sufficient altitude to power a long glide.

During their first migration, juveniles and adults migrate together to overwintering sites, but it has been unclear whether individuals work collectively to exploit thermals to migrate these great distances (*4*, *5*). Given the technical challenge of simultaneously tracking individuals within flocks, a first step in addressing this question is to examine how juveniles interact with each other as they take off for their first migration shortly after fledging.

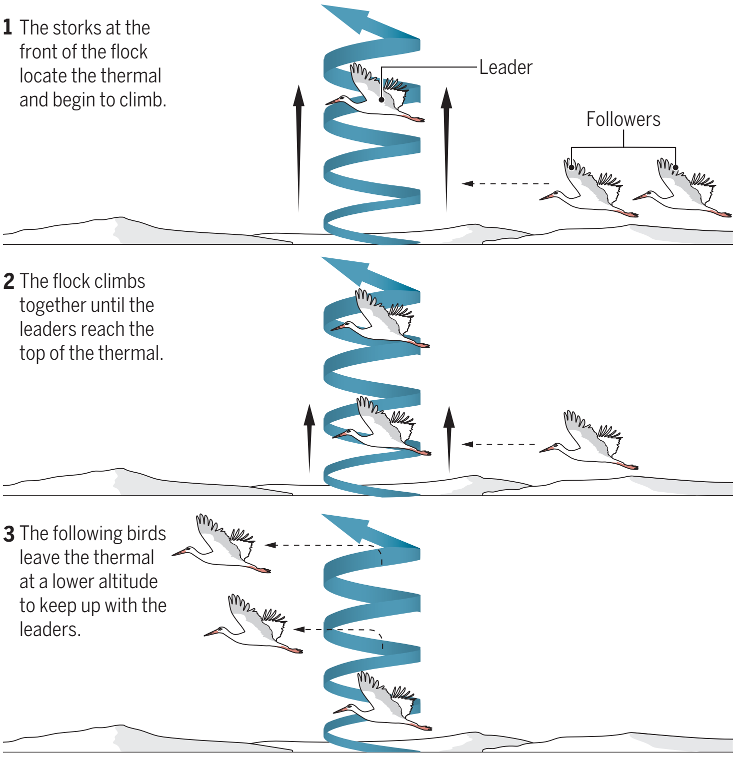
During the summer of 2014, Flack *et al.* followed a flock of 27 fledgling white storks from a small colony of 22 nests as they migrated for the first time from their breeding colony in southern Germany to Spain. Birds were tagged with devices that allowed the researchers to simultaneously monitor each individual's geographical position, elevation, speed, and direction. The birds were also equipped with three axis accelerometers to monitor how frequently they were flapping their wings. The researchers focused on the first 5 days after the fledglings started their migration.

The researchers identified three modes of flight—flapping flight, thermalling, and gliding—and also determined how frequently birds flapped their wings. To further examine interactions between individuals as they located and exploited thermals, the researchers developed a precise method to measure the time required for a tagged bird to reach the position of another tagged bird. Thus, this study provides a high-resolution picture of how white storks interact with each other during their first migration.

The data suggest that how individuals find and exploit thermals is related to an underlying initial variation in flight performance. Juvenile white storks with the lowest proportion of wing-flapping activity tended to fly ahead of birds that flapped their wings more. The individuals in the lead were able to find thermals and circle to higher altitudes than the birds that made up the rear of the flock; instead of searching for thermals on their own, these birds followed the leaders directly to them. Once the followers arrived at the thermal, they could climb rapidly, but they were also inclined to leave the thermals prematurely at a lower altitude, presumably owing to a drive to keep up with the rest of the flock (see the figure).

The leader-follower model for exploiting thermals

Juvenile white storks migrate in flocks, but individuals vary in their flight performance. Leaders search for thermals, whereas followers find thermals by watching the leaders. Leaders start their glide from a higher altitude than followers, presumably because their departure prompts the others to follow them. Followers cannot migrate as far as leaders because they must expend more energy on flapping flight.



Flack *et al.* were also able to examine how far birds traveled during the first 4 weeks of migration. Migratory distance was strongly correlated with the birds' initial wing-flapping behavior. Birds that flew ahead in the flock and needed to flap their wings less migrated farther than birds that followed and that needed to flap their wings more.

The authors suggest that differences in flight performance in juvenile storks lead to different behavioral strategies for exploiting thermals, giving individuals that are leaders an edge in how far they get on their first migration. However, the authors could not explain these differences by body mass, sex, or other demographic parameters that they recorded. This begs the question as to whether the leader-follower strategy might be context dependent. The study focused on a single flock for just 5 days, but there is a natural variation in flight performance among a cohort of fledglings, and a follower in one flock could drop out to be a leader in another flock of more evenly matched individuals.

The researchers' findings illustrate how dynamic flocking behavior is. They initially tagged all 61 fledglings at the colony, but the study group was reduced to just 27 birds because some individuals left the focal flock, stopped migrating altogether, or died. In addition, the tagged juveniles were joined by unmonitored individuals. As the authors point out, interactions with these untagged individuals could not be assessed.

Further research might explore whether leader-follower roles change within the flocks themselves. Other migratory species that are adapted to travel in V-formation to conserve energy (*6*) have been shown to frequently change position in a pairwise manner, such that birds spend similar amounts of time leading and following (*7*). In juvenile Northern bald ibis (*Geronticus eremita*), the V-formation gives birds in the rear an aerodynamic advantage, but, by frequently exchanging positions, that advantage is more evenly distributed among the flock members (*7*).

An added intriguing wrinkle in this story is that juvenile white storks frequently migrate with adults, which would seem to impose a handicap on juveniles. Juvenile white storks rely on flapping flight more than adults, and this higher energetic cost of trying to keep up is thought to contribute to juvenile morbidity during the first migration (*5*). One might also predict that more-experienced birds should position themselves in leader positions in mixed-aged flocks, not only because they are stronger fliers, but also because they know where the flock needs to go (*8*). However, this might put younger birds at a disadvantage because there is an energetic cost to being a follower if they are constantly trying to keep up with the faster leaders. Comparisons of age-related performance in the use of thermals by Eurasian griffon vultures (*Gyps fulvus*) suggest that juveniles do not use thermals as efficiently under intermediate wind shear (*4*). It will be intriguing to explore how thermal soaring contributes to organized movement of individuals within mixed-aged migratory flocks. ■

Department of Neurobiology, Physiology, and Behavior, University of California, Davis, Davis, CA 95616, USA.
Email: ganevitt@ucdavis.edu

REFERENCES AND NOTES

1. C. J. Pennycuick, *Ibis* **114**, 178 (1972).
2. A. Flack, M. Nagy, W. Fiedler, I. D. Couzin, M. Wikelski, *Science* **360**, 911 (2018).
3. C. J. Pennycuick, *J. Avian Biol.* **29**, 33 (1998).
4. R. Harel, N. Horvitz, R. Nathan, *Sci. Rep.* **6**, 27865 (2016).
5. S. Rotics *et al.*, *J. Anim. Ecol.* **85**, 938 (2016).
6. H. Weimerskirch, J. Martin, Y. Clerquin, P. Alexandre, S. Jiraskova, *Nature* **413**, 697 (2001).
7. B. Voelkl *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **112**, 2115 (2015).
8. N. Chernetsov, P. Berthold, U. Querner, *J. Exp. Biol.* **207**, 937 (2004).

ACKNOWLEDGMENTS

The author gratefully acknowledges support from the NSF (grants 1142084 and 1258828).

10.1126/science.aat7920

GRAPHIC: N. DESAI/*SCIENCE*

ORGANIC ELECTRONICS

Enhancing energy transport in conjugated polymers

Scalable synthesis enables long-distance exciton transport in crystalline polymer nanofibers

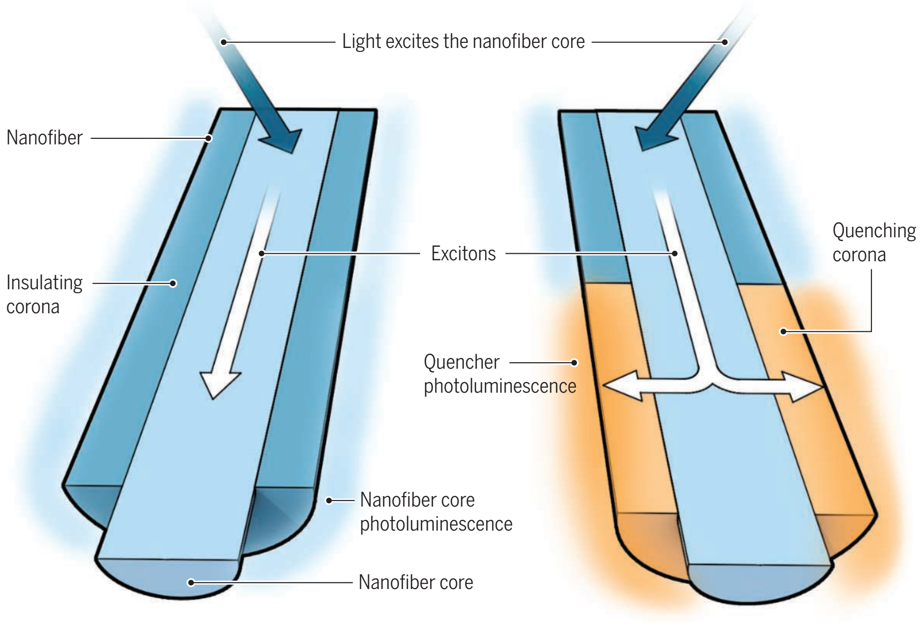
By **Russell J. Holmes**

The conversion of light into usable chemical energy by plants is enabled by the precise spatial arrangement of light-absorbing photosynthetic systems and associated molecular complexes (*1*). In organic solar cells, there is also the need to control intermolecular spacing and molecular orientation, as well as thin-film crystallinity and morphology, so as to enable efficient energy migration and photoconversion (*2*). In an organic solar cell, light absorption creates excitons, tightly bound electron-hole pairs that must be efficiently dissociated into their component charge carriers in order to create an electrical current. Thus, long-range exciton migration must occur from the point of photogeneration to a dissociating site. On page 897 of this issue, Jin *et al.* (*3*) report on a conjugated polymer nanofiber system that yields exciton diffusion lengths greater than 200 nm. In comparison, organic solar cells are typically constructed with materials having exciton diffusion lengths one order of magnitude smaller than this value, which limits device thickness and optical absorption. Their approach exploits a sequential synthesis method that enables measurement of this long exciton diffusion length (see the figure).

Exciton diffusion is typically considered to occur by means of two mechanisms (*4*). The first is through direct electron-exchange between nearest neighbors, also known as Dexter transfer. Because this mechanism involves the physical exchange of electrons, it requires the associated donor and acceptor molecules to have overlapping electronic orbitals. The second mechanism for energy transfer occurs via the nonradiative coupling of electronic dipoles, termed Förster transfer. Förster transfer contrasts with Dexter transfer in that it may occur through occupied space, given that it is mediated by the electromagnetic field. Because Förster transfer fundamentally requires that the donor molecule be capable of light emission, dark states are precluded from this mechanism and instead migrate exclusively via Dexter transfer. In highly luminescent materials, the rate of Förster transfer is typically faster than the rate of Dexter transfer and is responsible for transport.

Probing excitons

Jin *et al*. studied long-range exciton transport in a nanofiber core surrounded by either an insulating, nonquenching corona or a quenching corona that can accept excitons from the core.



With a fully insulating corona, excitons migrate over distances exceeding 200 nm.

The exciton diffusion length can be extracted by reducing the insulating corona segment length.

Jin *et al.* probe exciton migration in crystalline, luminescent nanofibers of poly(di-*n*-hexylfluorene) (PDHF) using photoluminescence quenching. At its core, this method involves probing the photoluminescence intensity or decay kinetics from a material of interest subject to an exciton quencher placed at a known distance from the photogeneration site. Measurement of the photoluminescence as the distance to the quencher is varied allows the exciton diffusion length to be determined. In order to measure the exciton diffusion length of PDHF, nanofibers were synthesized from segments that have coronae consisting of either poly(ethylene glycol) (PEG) or quaternized polythiophene (QPT).

An electrically insulating PEG corona has a large energy gap and hence will not interact with an exciton in PDHF. By contrast, QPT has a narrow energy gap and will actively accept excitons and quench photoluminescence from the PDHF core. Jin *et al.* extracted exciton diffusion lengths by systematically varying the PEG segment length, which in turn varies the number of excitons that reach the corona, and hence, the photoluminescence from QPT. This elegant photoluminescence quenching experiment is facilitated by the use of a scalable, seeded growth method, in which diblock copolymer segments of either PDHF-PEG or PDHF-QPT can be attached to form the subsequent nanofibers with different types of coronae (*5*).

It is interesting to consider the origin of the large exciton diffusion lengths reported by Jin *et al.* Prior work has demonstrated the ability to realize improved exciton transport in crystalline materials by reducing energetic disorder and exciton trapping while also permitting a more optimized orientation of molecular transition dipole moments for excitonic energy transfer (*6*). Micrometer-scale exciton diffusion lengths have been previously reported in

Department of Chemical Engineering and Materials Science, University of Minnesota, Minneapolis, MN 55455, USA. Email: rholmes@umn.edu

GRAPHIC: C. BICKEL/*SCIENCE*

Published by AAAS

discrete single crystals, although transport in these systems can also involve long-lived dark excitons (*7*, *8*). The PDHF nanofibers reported by Jin *et al.* are highly crystalline, with favorable molecule orientation that leads to efficient Förster transfer. Long exciton diffusion lengths are realized for luminescent excitons without a contribution from long-lived dark states.

As researchers strive to exert greater control over energy transport in conjugated systems, it is clear that progress is needed on multiple fronts. The work of Jin *et al.* reinforces the notion that crystalline order plays an important role in facilitating exciton diffusion. More specifically, low disorder and strong π-orbital overlap can enable more effective exciton transport. However, Jin *et al.* note that these factors alone are insufficient to explain the reported long exciton diffusion lengths.

> ***"The PDHF nanofibers reported by Jin et al. are highly crystalline, with favorable molecule orientation that leads to efficient Förster transfer."***

Indeed, they highlight the possible role played by exciton delocalization and coherence. In order to realize paradigm-shifting improvements in exciton transport, it is essential to fully escape the limitations of the diffusive or subdiffusive transport regimes and realize ballistic transport (*9*). Whether through exciton delocalization and coherence, or through the design of architectures that remove the normal isotropy of energy transfer, the realization of ballistic transport in thin film would offer new frontiers of interesting science, as well as new device applications that extend beyond the realm of photoconversion. ■

REFERENCES

1. F. Fassioli, R. Dinshaw, P. Arpin, G. Scholes, *J. R. Soc. Interface* **11**, 20130901 (2013).
2. G. Hedley, A. Ruseckas, I. Samuel, *Chem. Rev.* **117**, 796 (2017).
3. X.-H. Jin *et al.*, *Science* **360**, 897 (2018).
4. S. M. Menke, R. J. Holmes, *Energy Environ. Sci.* **7**, 499 (2014).
5. C. Boott, J. Gwyther, R. Harniman, D. Hayward, I. Manners, *Nat. Chem.* **9**, 785 (2017).
6. R. R. Lunt, J. B. Benziger, S. R. Forrest, *Adv. Mater.* **22**, 1233 (2010).
7. H. Najafov, B. Lee, Q. Zhou, L. C. Feldman, V. Podzorov, *Nat. Mater.* **9**, 938 (2010).
8. G. M. Akselrod *et al.*, *Nat. Commun.* **5**, 3646 (2014).
9. J. Brédas, E. Sargent, G. Scholes, *Nat. Mater.* **16**, 35 (2017).

10.1126/science.aat6009

CHEMISTRY

Cold chemistry with two atoms

Two atoms react to form a molecule in an optical "beaker"

By **Edvardas Narevicius**

For centuries, chemists have written equations representing chemical reactions by using symbols for atoms and molecules; for example, $2H_2O + 2Na \rightarrow 2NaOH + H_2$. This short notation shows only four reacting particles, but even in a classroom demonstration where a small piece of sodium is dropped in water, the total number of reactants will be on the order of Avogadro's number ($\sim 6 \times 10^{23}$). On page 900 of this issue, Liu *et al.* (*1*) instead study a chemical reaction taking place between a minimal number of participants. In their experiment, exactly two atoms collide, absorb a photon, and form a molecule in the excited state. And this time, the reaction equation, $Na + Cs \rightarrow NaCs^*$ (where the asterisk denotes an excited molecule), describes exactly the process that takes place in the laboratory.

During room-temperature liquid-phase reactions, local interactions and conditions vary widely, with many quantum states contributing to the reaction dynamics. Ensemble averaging can be helpful for describing these reactions, and thermodynamic properties such as the Gibbs free energy can be used together with classical dynamics calculations to reliably predict reaction rates and outcomes.

At another extreme are reactions between reactants in interstellar space, where particle densities are 10 orders of magnitude lower than in a liquid and where reactions happen only when two reactants collide. Herschbach, Lee, and co-workers were the first to experimentally explore this regime with low-density cold molecular beams (*2*). Removing interactions with the environment allowed direct observation of reaction mechanisms.

In the past decade, scientists have been able to observe reactions at ever-lower collision energies. At low collision energies, quantum-mechanical wavelike behavior plays a central role, necessitating use of quantum statistics (*3*) and consideration of quantum tunneling (*4*) and quantum scattering resonances (*5*–*7*). Liu *et al.* now measure inelastic collisions between only two trapped atoms, achieve formation of an individual molecule, and perform molecular spectroscopy.

It is surprising that the authors could construct a molecule by a reaction between two atoms. Usually, chemists increase densities until the collision frequency is high enough for products to be formed. Liu *et al.* instead use a clever technique to confine the two atoms in a small (several cubic micrometers) trap that is sufficiently deep to keep atoms with a temperature of 1 millikelvin (mK). They create the trap with tightly focused laser beams that confine particles in the region with the highest laser field intensity. Similar to laboratory tweezers, atoms trapped inside can be moved around by translating the laser beams in space.

To be able to place the atoms in optical tweezers, the authors first had to lower the kinetic temperature down to tens of microkelvin. To achieve this temperature, the authors used one of the most successful tools of modern atomic and optical physics, laser cooling (*8*). They laser-cooled Na and Cs atoms in individual optical tweezer traps and detected them using a fluorescence signal, leading to a ~33% success rate in the preparation of two reactants in a well-defined initial state.

After cooling the atoms and trapping them in individual laser tweezers, the authors adiabatically merged the two traps and formed an optical "beaker" containing both Na and Cs. The next step was to show and prove that collisions between the two atoms are detectable. To do so, the authors prepared a beaker containing two atoms in which excited hyperfine states were populated. Hyperfine states form through interactions between electron and nuclear spins within an atom, with excitation energies on the order of 100 mK. During a collision, the internal excitation energy can be released in the form of kinetic energy, and atoms will be ejected out of the shallow optical trap. The atoms are lost after several milliseconds, indicating a high collision rate. The inelastic collisions can be strongly suppressed by controlling the initial quantum state of the atoms. If atoms are driven to the lowest energy state, their lifetime in the trap increases to several seconds.

The long trapping lifetimes enabled the authors to perform the impressive feat of building a single molecule out of two atoms. In principle, a collision between two ground-state atoms cannot produce a

---

Chemical and Biological Physics Department, Weizmann Institute of Science, 76100 Rehovot, Israel. Email: edvardas.narevicius@weizmann.ac.il

bound molecule. A third body is needed to carry away the excess bonding energy. Usually, interaction with another atom or molecule serves this purpose. In Liu *et al.*'s study, photons take care of energy conservation.

Two photons are needed to create a molecule out of two separate atoms (see the figure). The first photon that starts the reaction is absorbed by an atom. However, the photon energy is tuned just below the atomic transition and can only be absorbed if a molecular state is reached (see the figure). By scanning the laser frequency, the authors could detect several bound molecular levels on the electronically excited potential and assign them to particular molecular vibrations, demonstrating that molecular spectra can be measured with only two atoms. The electronically excited molecule quickly emits a photon and a molecule in a vibrationally excited level. Formation of a molecule is detected as a simultaneous loss of fluorescence of the Na and Cs atoms. Here again, the authors harness the low temperatures that can be achieved with laser cooling. The transition probability between free atoms and bound molecular state depend on the overlap between corresponding wave functions. To form a molecule, two atoms have to spend sufficient time at distances on the order of the length of a molecular bond. This time can be increased by reducing the relative kinetic energies that correspond to temperatures in the range of several microkelvin. This leads to higher excitation probabilities that do not require high-intensity laser fields (*9*).

Liu *et al.*'s method opens many interesting avenues for research. The authors plan to demonstrate that with the help of an additional laser field, molecules can be coherently transferred into a single selected vibrational level. Such a degree of quantum state control would allow studies of quantum state–resolved collisions and reactions between molecules. The ability to work with well-defined numbers of atoms will provide unprecedented precision in unraveling the microscopic mechanisms of complex interactions. Another interesting application for "designer" cold molecules would be the assembly of a lattice of strongly interacting polar molecules that could serve as qubits or as a tool to investigate unusual quantum phases. ■

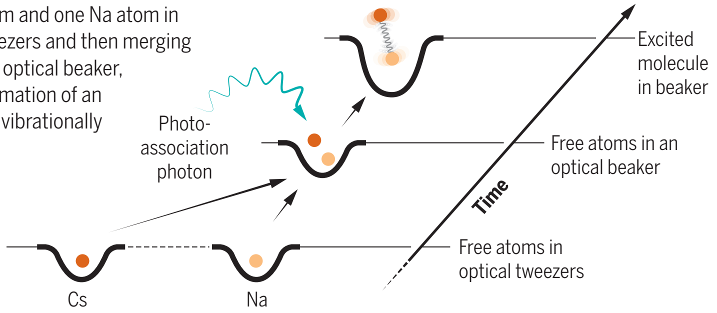
REFERENCES

1. L. R. Liu *et al.*, *Science* **360**, 900 (2018).
2. Y. T. Lee, J. D. McDonald, P. R. LeBreton, D. R. Herschbach, *Rev. Sci. Instrum.* **40**, 1402 (1969).
3. S. Ospelkaus *et al.*, Science **327**, 853 (2010).
4. M. Tizniti *et al.*, *Nat. Chem.* **6**, 141 (2014).
5. A. B. Henson, S. Gersten, Y. Shagam, J. Narevicius, E. Narevicius, *Science* **338**, 234 (2012).
6. S. Chefdeville *et al.*, *Science* **341**, 1094 (2013).
7. S. N. Vogels *et al.*, *Science* **350**, 787 (2015).
8. C. Cohen-Tannoudji, D. Guéry-Odelin, *Advances in Atomic Physics: An Overview* (World Scientific, 2011).
9. K. M. Jones, E. Tiesinga, P. D. Lett, P. S. Julienne, *Rev. Mod. Phys.* **78**, 483 (2006).

10.1126/science.aat7917

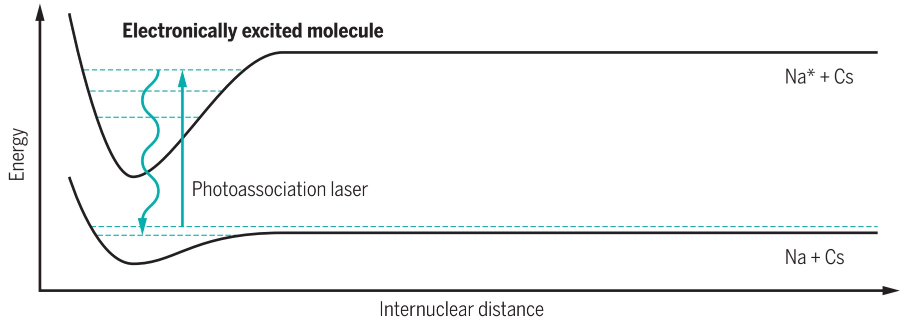
Making exactly one molecule

By confining one Cs atom and one Na atom in separate molecular tweezers and then merging the tweezers to form an optical beaker, Liu *et al.* achieve the formation of an individual molecule in a vibrationally excited state.



Photoassociation process

Free Na and Cs atoms collide, absorb a photon, and form an electronically excited molecule, which emits another photon and forms a stable molecule.



GRAPHIC: C. BICKEL/SCIENCE

SYNTHETIC BIOLOGY

Illuminating dark depths

Microelectronic processing and engineered bacteria provide real-time insights into the gut

By **Peter R. Gibson** *and* **Rebecca E. Burgell**

The inner workings of human physiology remain one of the final frontiers of science. In particular, the gastrointestinal (GI) microenvironment is incompletely understood. On page 915 of this issue, Mimee *et al.* (*1*) have developed an ultralow-power ingestible sensor of various compounds (ingestible micro-bio-electronic device, IMBED) by combining engineered bioluminescent bacteria and microelectronic processing. This has the potential to unlock a wealth of information about the body's structure and function, its relationship with the environment, and the impact of disease and therapeutic interventions.

Tiny sensors can now be made through advances in semiconductor microelectronics and microfabrication, and their signals can be computed and communicated wirelessly to remote receivers (*2*). Furthermore, the increasingly low power demands of such devices, developments in battery technology, and advances in wireless signal transmission have permitted deeper, more durable exploration within the body (*2*). Such technology can be used to study the luminal content of the GI tract, as understanding its physiology and pathology is limited, largely owing to difficulties in access. Although devices can be inserted into regions close to the exterior, few provide measurements under truly physiological conditions. For example, access to the colon generally involves bowel cleansing, which disrupts normal physiology. Inaccessibility almost precludes such approaches to most of the small intestine. This negatively affects our comprehension of inflammatory intestinal disorders such as Crohn's disease, celiac disease, and irritable bowel syndrome, which are prevalent worldwide (*3*, *4*).

The design of biosensors in the form of capsules, no greater in size than a large vitamin pill, has revolutionized the ability

Department of Gastroenterology, Alfred Hospital and Monash University, Melbourne, Victoria 3004, Australia. Email: peter.gibson@monash.edu

to access the GI tract without the need for physiological disruption (*5*). The addition of a miniaturized camera on such devices has enabled endoscopic visualization of the entire GI tract and opened the small intestine to medical scrutiny and scientific research (*2*, *5*). Such devices are now in routine use worldwide for diagnosing inflammatory, ulcerative, and proliferative lesions and identifying sites and causes of bleeding (*6*).

With their success has come the demand for new investigative technology that not only provides an assessment of bowel structure but that also expands our understanding of normal GI physiology and the function of nutrients and microbiota in health and disease. Microsensors have been developed to map temperature, pH, and gas concentrations as the capsule moves down the gut (*7*–*10*). By utilizing a semipermeable membrane, gaseous small molecules are sampled by microsensors in the capsule, and their concentrations are transmitted in real time to an external wireless receiver and recording device. For example, the observed concentrations of hydrogen, carbon dioxide, and methane give unparalleled insights into the inner workings of the microbiota of the intestine (*8*, *9*). GI microbiota are key for gut health and play critical, although incompletely understood, roles in overall health (*11*).

Mimee *et al.* offer a system for detecting biomolecules in the GI tract. A probiotic bacterium (*12*) was genetically engineered to detect and metabolize heme (a component of red blood cells) and to convert heme metabolites into a luminescent signal. A capsule enclosed the bacteria within a semipermeable membrane, and electronic photodetectors and transmission platforms detected the luminescent signal. After capsule ingestion, this information was transmitted wirelessly in real time to an external receiver and recorder, such as a smartphone (see the figure). When tested in vivo in pigs, the ingested capsule detected heme (that is, lysed red blood cells) with high specificity.

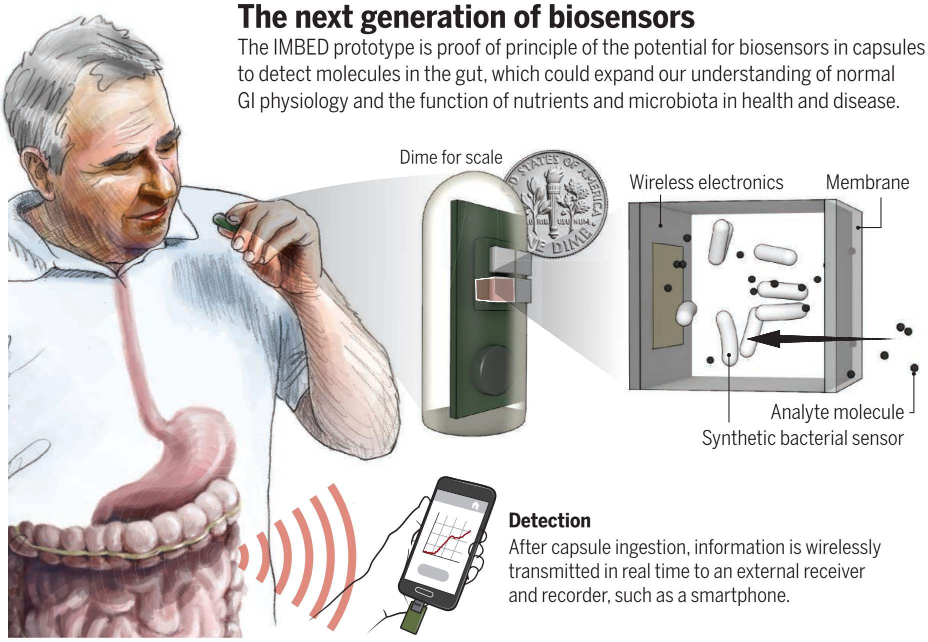
The current application of this prototype has limited utility in clinical medicine: Bleeding from the stomach is readily detected and treated in clinical practice (*13*). However, the IMBED prototype provides proof of concept that a specific molecule can be detected in vivo and this information wirelessly transmitted outside the body. Indeed, Mimee *et al.* show that this device can detect other compounds within the gut, including a biomarker of inflammation and specific bacteria.

The future challenge is to find an appropriate application for the technology. The approach of Mimee *et al.* was to examine a molecule (heme) not normally found in the GI lumen. A similar situation might apply to molecules associated with cancers and inflammation, as well as infection (such as a specific microbe). Assessing multiple molecules simultaneously might also be advantageous, although technical issues such as whether the bioluminescent system can be adapted to different wavelengths need to be addressed. There is also potential for robotic capabilities to be included in such devices, enabling, for example, drug release when the analyte is detected (*2*).

However, there are limitations. In the case of the GI tract, sometimes the greatest barrier is that we do not know what we do not know. As such, the clinical utility of the device is likely to be limited by the rate at which biomarkers are characterized. Furthermore, for most molecules, it is not their presence but rather where and in what concentration they are found that will be key to understanding GI physiology. The microenvironment of each segment of the GI tract varies, and thus a device that responds to a change in analyte concentration is crucial. Whether the IMBED system is able to respond to change over time is uncertain—it is not clear whether the bioluminescence signal can be reset once activated.

Critical to next-generation devices will be location determination. For optical imaging of the gut, anatomical landmarks can be used (*5*, *6*). However, as inclusion of a camera results in excessive power needs, nonoptical means are necessary. Changes in pH appear to be reliable at estimating passage through different GI regions (*14*), and a drop in temperature indicates capsule excretion (*7*). The development and validation of a gas-sensing capsule (*8*, *9*) alternatively provides the theoretical use of other molecules for localization. Whether such a biosensor can be successfully incorporated into the IMBED requires further study.

Despite these limitations, it is exciting to watch where this technology ultimately takes us. Although the GI tract is ideal for this technology, its application to other parts of the body, such as solid organs or the circulation, is also theoretically possible. The former would require implantation and removal. The latter would necessitate further miniaturization and the application of different structural components, such as carbon nanotubes, to provide flexibility, strength, and safety (*15*).


The next generation of biosensors

The IMBED prototype is proof of principle of the potential for biosensors in capsules to detect molecules in the gut, which could expand our understanding of normal GI physiology and the function of nutrients and microbiota in health and disease.

Detection

After capsule ingestion, information is wirelessly transmitted in real time to an external receiver and recorder, such as a smartphone.

The IMBED system described by Mimee *et al.* provides another step in delving into the dark depths of the GI tract and could lead to real-time identification, detection, and quantification of biomolecules in humans, with great potential benefit to clinical medicine and research. ■

REFERENCES

1. M. Mimee *et al.*, *Science* **360**, 915 (2018).
2. S. S. Mapara, V. B. Patravale, *J. Controlled Rel.* **261**, 337 (2017).
3. C. Canavan *et al.*, *Aliment. Pharmacol. Ther.* **40**, 1023 (2014).
4. C. N. Bernstein, *Aliment. Pharmacol. Ther.* **46**, 911 (2017).
5. E. Rondonotti *et al.*, *Endoscopy* **50**, 423 (2018).
6. M. F. Hale *et al.*, *World J. Gastroenterol.* **20**, 7752 (2014).
7. A. D. Farmer *et al.*, *United Eur. Gastroenterol. J.* **1**, 413 (2013).
8. K. Kalantar-zadeh *et al.*, *Gastroenterol.* **150**, 37 (2016).
9. K. Kalantar-zadeh *et al.*, *Nat. Electron.* **1**, 79 (2018).
10. J. Z. Ou *et al.*, *Sci. Rep.* **6**, 33387 (2016).
11. J. C. Clemente *et al.*, *BMJ* **360**, j5145 (2018).
12. B. Ou, *Appl. Microbiol. Biotechnol.* **100**, 8693 (2016).
13. A. Lanas *et al.*, *Nat. Rev. Dis. Primers* **4**, 18020 (2018).
14. N. Zarate *et al.*, *Am. J. Physiol. Gastroint. Liver Physiol.* **299**, G1276 (2010).
15. A. Jain *et al.*, *Biotechnol. J.* **10**, 447 (2015).

10.1126/science.aat8658

GRAPHIC: V. ALTOUNIAN/*SCIENCE*

CANCER MICROBIOME

Cancer immunity thwarted by the microbiome

Microbial bile acid metabolites promote liver metastasis

By **Nadine Hartmann**[1] and **Mitchell Kronenberg**[1,2]

Liver cancer rates have tripled since 1980, making it one of the leading causes of cancer deaths worldwide (*1*). The efficacy of a variety of cancer treatments, as well as carcinogenesis itself, can be influenced by the microbiome (*2*). Microbiome-cancer interactions are complex, as the microbiome can promote chronic inflammation or directly affect cancer cells. Furthermore, by influencing the immune system, the microbiome may either promote or inhibit antitumor immune responses (*3*). On page 876 of this issue, Ma *et al.* (*4*) show that a type of commensal bacteria, *Clostridium* species, prevents an effective immune response to both primary liver tumors and liver metastases by a previously unknown mechanism: not by means of its own metabolites, but by modifying a host product—the bile acids that liver cells are exposed to.

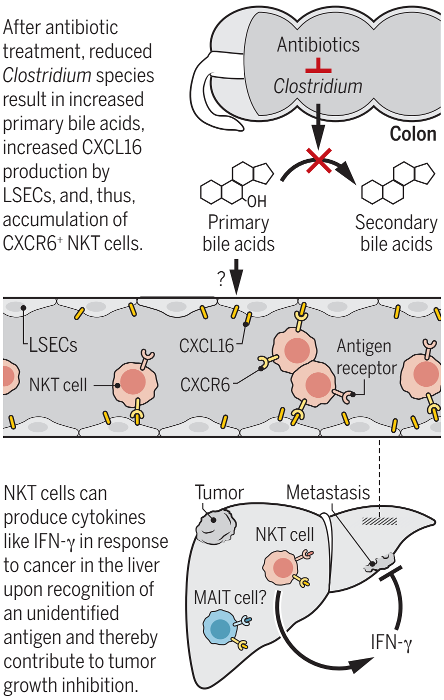
About 75% of the blood that passes through the liver is supplied from the intestine by the portal vein, thus constantly exposing the liver to components and metabolites from the intestinal microbiota. To avoid excessive immune stimulation, at steady state, hepatic tissue promotes immune tolerance rather than immune activation (*5*). The liver has a mostly noncirculating immune system consisting of T lymphocytes, B lymphocytes, and macrophages, among others (*6*). In the absence of infection or injury, these immune cells are predominantly in the sinusoidal blood vessels rather than the tissue parenchyma. The liver-resident lymphocyte populations have different properties from their circulating counterparts. For example, natural killer T (NKT) cells constitute up to 40% of the total intrahepatic T lymphocytes in mice, but they are much less frequent in other sites.

Although most T lymphocytes recognize peptides bound to or presented by antigen-presenting molecules encoded by the highly polymorphic major histocompatibility complex (MHC) genes, NKT cells recognize lipids presented by the monomorphic MHC class I–like protein CD1d (*7*). NKT cells are distinguished by their rapid and intense cytokine responses: Intrahepatic NKT cells produce copious amounts of interferon-γ (IFN-γ) when activated, which has been associated with their ability to respond to tumors (*8*).

To understand the role of the microbiome in hepatic cancers, Ma *et al.* analyzed primary liver tumors in transgenic mice as well as liver metastases from lymphoid or melanoma tumor cell lines engrafted outside the liver. All mice received a cocktail of three antibiotics to minimize the microbiome. Antibiotic treatment reduced tumor growth or liver metastases, but not metastases to other sites, such as the lungs. This tissue-specific effect correlated with an increase in activated liver NKT cells. Furthermore, tumor-bearing CD1d-deficient mice were unaffected by antibiotic treatment. These mice lack not only NKT cells but also other CD1d-reactive cells in the liver. The accumulation of NKT cells in the liver requires expression of the chemokine receptor CXCR6 (C-X-C chemokine receptor type 6), which binds the ligand CXCL16 expressed on the surface of liver sinusoidal endothelial cells (LSECs) (see the figure). Antibiotic treatment increased LSEC expression of CXCL16 by an unknown mechanism. In CXCR6-deficient mice, antibiotic treatment did not augment tumor growth, suggesting that the increased antitumor response following antibiotics was due to enhanced NKT cell accumulation in the liver mediated by CXCL16-CXCR6 interactions.

Primary bile acids promote liver tumor immunity

Clostridium species modify bile acids, which prevents antitumor immune responses in the liver.



After antibiotic treatment, reduced *Clostridium* species result in increased primary bile acids, increased CXCL16 production by LSECs, and, thus, accumulation of CXCR6$^+$ NKT cells.

NKT cells can produce cytokines like IFN-γ in response to cancer in the liver upon recognition of an unidentified antigen and thereby contribute to tumor growth inhibition.

GRAPHIC: V. ALTOUNIAN/SCIENCE

How does the gut microbiome signal to LSECs? Bile acids—which are synthesized in the liver, stored in the gall bladder, and secreted into the intestine—are involved in dietary fat absorption. Recently, bile acids have been found to shape the intestinal microbiome composition, but microbiome metabolism also affects bile acid profiles (*9*). Primary bile acids that are synthesized in the liver stimulate hepatocytes to produce proinflammatory cytokines and chemokines like CXCL16, thus contributing to the recruitment of CXCR6$^+$ lymphocytes to the liver (*10*). However, primary bile acids are metabolized to secondary bile acids by intestinal bacteria and are returned to the liver by the enterohepatic circulation (*9*). Ma *et al.* showed that LSECs up-regulate *Cxcl16* messenger RNA (mRNA) production when exposed to primary bile acids, whereas *Cxcl16* mRNA production was decreased upon secondary bile acid exposure. Also, treatment with antibiotics depleted the main producers of secondary bile acids and increased CXCL16 expression in LSECs. Importantly, increasing primary bile acids (for example, through feeding) increased intrahepatic NKT cells and reduced tumor metastases. Because the selective use of antibiotics implicated Gram-positive bacteria, and because *Clostridium* species carry out a major enzymatic step in secondary bile acid formation, Ma *et al.* determined that colonization with a commensal *Clostridium* species led not only to decreased intrahepatic NKT cells but also increased liver tumor metastases. Interestingly, it has been shown that a high-fat diet-induced increase of intestinal *Clostridium* species increases levels of secondary bile acids in the liver, which in turn promotes liver cancer (*11*).

---

[1] *Division of Developmental Immunology, La Jolla Institute for Allergy and Immunology, La Jolla, CA 92037, USA.* [2] *Division of Biology, University of California, San Diego, La Jolla, CA 92037, USA. Email: mitch@lji.org*

Why is a specialized T cell subset critical for immune protection from liver metastases? Although NKT cells recognize lipids containing fatty acids, bile acids have an entirely different structure and are not known to be antigens. Therefore, the role of primary bile acids in NKT cell activation is indirect, by promoting NKT cell accumulation but not their activation. Ma *et al.* analyzed CD1d-expressing tumors, and they suggest that CD1d-presented lipid antigens from the tumors could be directly recognized by NKT cells. It is also possible that NKT cells could be activated by lipid antigens from $CD1d^-$ tumors presented by $CD1d^+$ macrophages or dendritic cells or even by inflammatory cytokines such as interleukin-12 (IL-12) (*12*). Regardless, the prevalence of NKT cells in mouse liver and their strong effector function allow them to be a dominant player in the anticancer response of this organ. The situation is different in humans, as intrahepatic NKT cells are much less frequent. However, mucosal-associated invariant T (MAIT) cells, which recognize microbial vitamin B metabolites presented by MHC-related protein 1 (MR1), are prevalent in human liver, and they also exert rapid effector functions, such as IFN-γ secretion (*7*). Interestingly, MAIT cells also express CXCR6, and, although their antitumor potential is uncertain, they do infiltrate some metastatic lesions in the liver of colorectal carcinoma patients (*13*).

Although there are bile acid profiles that are associated with cancer, several studies advocate modifying bile acid production to improve cancer therapy, for example, through agonists of bile acid receptors or simply by changes in diet (*14*). Thus, as for most biological systems, bile function has different facets. It is designed to emulsify fats and help absorb dietary lipids, but after modification by the microbiome, secondary bile acids alter immune function to promote liver cancer and liver metastases. ■

REFERENCES

1. R. L. Siegel *et al.*, *CA Cancer J. Clin.* **68**, 7 (2018).
2. L. Zitvogel *et al.*, *Science* **359**, 1366 (2018).
3. R. F. Schwabe, C. Jobin, *Nat. Rev. Cancer* **13**, 800 (2013).
4. C. Ma *et al. Science* **360**, eaan5931 (2018).
5. C. Mehrfeld *et al.*, *Front. Immunol.* **9**, 635 (2018).
6. D. G. Doherty, *J. Autoimmun.* **66**, 60 (2016).
7. S. Chandra, M. Kronenberg, *Adv. Immunol.* **127**, 145 (2015).
8. N. Y. Crowe *et al.*, *J. Exp. Med.* **196**, 119 (2002).
9. J. M. Ridlon *et al.*, *Curr. Opin. Gastroenterol.* **30**, 332 (2014).
10. K. Allen *et al.*, *Am. J. Pathol.* **178**, 175 (2011).
11. S. Yoshimoto *et al.*, *Nature* **499**, 97 (2013).
12. N. A. Nagarajan, M. Kronenberg, *J. Immunol.* **178**, 2706 (2007).
13. C. R. Shaler *et al.*, *Cancer Immunol. Immun.* **66**, 1563 (2017).
14. F. G. Schaap *et al.*, *Nat. Rev. Gastroenterol. Hepatol.* **11**, 55 (2014).

10.1126/science.aat8289

CELL BIOLOGY

The RNA face of phase separation

RNA regulates the formation, identity, and localization of phase-separated granules

By **Magdalini Polymenidou**

Phase separation or liquid unmixing—a phenomenon resembling the formation of oil droplets in vinegar—has emerged as a major driver of functional compartmentalization within cells, allowing the rapid and dynamic isolation of specific activities from the surrounding cellular environment, without the need of a membrane (*1*). A flurry of exciting recent studies demonstrated the importance of RNA binding proteins (RBPs) with low complexity regions (LCRs) in this process [for example, (*2*, *3*)], yet the role of incorporated RNAs is less well understood. On pages 918 and 922 of this issue, Maharana *et al.* (*4*) and Langdon *et al.* (*5*) propose that cellular RNAs regulate the formation, subcellular localization, and identity of these granules, which has implications for stress responses and pathologic protein aggregation in neurodegenerative diseases.

Cellular stressors induce the formation of perhaps the best studied of these so-called membraneless organelles, stress granules (SGs), which consist of multiple RBPs in complex with cytoplasmic RNAs. SGs stall the translation of incorporated RNAs and contain RBPs with LCRs—also called prion-like domains—which can phase-separate, both in vitro and within cells (*1*–*3*). Several of these SG-incorporating prion-like RBPs form pathologic protein aggregates in affected neurons of patients with neurodegenerative diseases. The same proteins form phase-separated liquid droplets in vitro, which are dynamic in a similar manner to SGs in cells. These liquid droplets transform over time into solid-like structures (*1*–*3*), somewhat resembling the pathologic protein aggregates in patients with neurodegeneration. These observations have reinforced the hypothesis that SGs act as precursors of pathologic aggregates (*6*, *7*). How this transition occurs in the degenerating brain remains unclear, and a direct in vivo demonstration is still missing.

A main driver of protein phase separation is concentration, so Maharana *et al.* asked whether the intrinsic cellular concentrations of RBPs known to form membraneless organelles could explain this behavior. Surprisingly, they found that in the absence of any other components, several RBPs implicated in neurodegeneration phase-separate in vitro in concentrations similar to those physiologically found in the nucleus. Yet inside the cell, the same proteins remain diffused in the nucleus and are typically excluded from phase-separated compartments, even under cellular stress conditions, which drive their condensation into cytoplasmic SGs. Searching for which nuclear component may prevent the phase separation of RBPs, they found that high RNA:protein ratios—like those typically found in the nucleus—prevent phase separation, whereas low ratios, similar to those in the cytoplasm, promote it (see the figure). This effect seems to be independent of sequence specificity for RBP binding. How these RNAs keep RBPs from phase-separating in the absence of specific binding and what the role of LCRs is in this process remain unresolved. Interestingly, Maharana *et al.* show that in vitro high amounts of short, nonspecific RNAs keep prion-like RBPs soluble. Conversely, longer RNAs that can form higher-order secondary structures and which specifically bind RBPs can nucleate phase separation. This is potentially the reason that *Neat1* (nuclear paraspeckle assembly transcript 1), a long noncoding RNA with specific FUS (fused in sarcoma) binding sites (*8*), can capture FUS out of the diffused nuclear pool to form paraspeckles (phase-separated nuclear compartments that regulate gene expression).

> ***"…high RNA:protein ratios… prevent phase separation… low ratios…promote it…"***

Maharana *et al.* propose that these observations may explain the somewhat preferential phase separation of nuclear proteins into cytoplasmic SGs, which in turn may transform into aggregates in the degenerating brain. Cytoplasmic aggregates are toxic

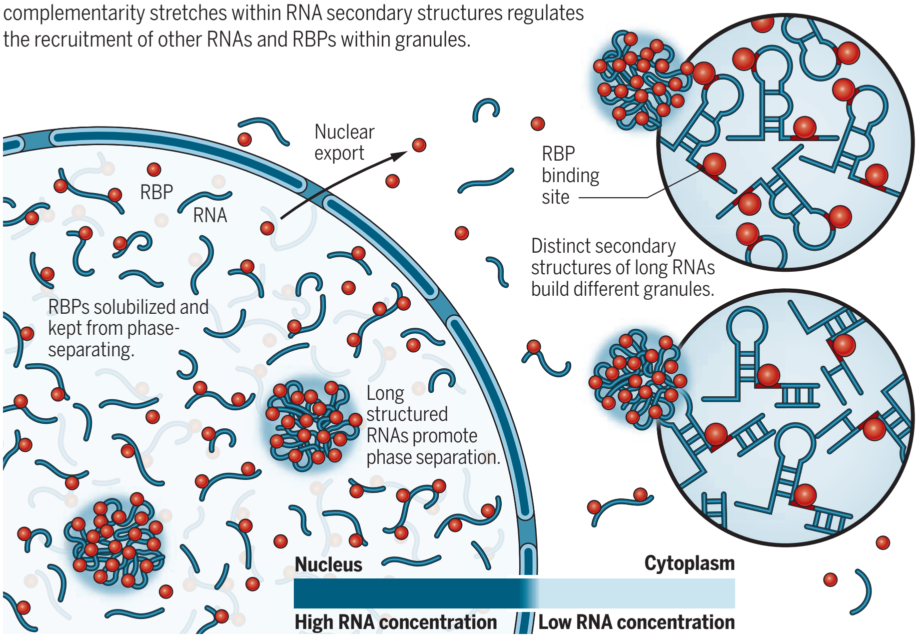
---

Institute of Molecular Life Sciences, University of Zurich, Switzerland. Email: magdalini.polymenidou@imls.uzh.ch

to cells because they interfere with major cellular functions, including nucleocytoplasmic transport (*9*), so promoting their removal is a focus of drug development for neurodegenerative diseases. Because many of the key protein components of cytoplasmic aggregates play essential cellular functions, therapeutic approaches aimed at their transcriptional silencing are not suitable. Instead, regulating their transition to pathologic aggregates might be a viable strategy, and a recent study showed that interfering with the kinetics of SG formation may reverse neurodegenerative processes associated with TDP-43 (TAR-DNA-binding protein of 43 kDa) aggregation (*10*). The observation that small nonspecific RNAs solubilize FUS and prevent its phase separation both in vitro and in cells (*4*) now opens up possibilities for exploring RNA mimetics as an alternative strategy to counteract the pathologic solidification of RBPs.

RNA controls phase separation

High nuclear RNA concentrations prevent phase separation of RBPs in a nonspecific manner. Long RNAs with stable secondary structure can capture RBPs out of solution, leading to phase-separated nuclear compartments such as paraspeckles and cytosolic granules. The exposure of specific RBP binding sites and complementarity stretches within RNA secondary structures regulates the recruitment of other RNAs and RBPs within granules.



The role of RNA secondary structure and RBP binding specificity in defining the identity and subcellular localization of phase-separated granules is the focus of the Langdon *et al.* study. They investigated two distinct RNA granules of yeast cells containing the Whi3 protein, the phase separation of which was shown to control cellular memory. Whi3 protein phase-separates via its RNA binding motif (RRM) and its polyglutamine (polyQ) tract and forms granules localized either near the nucleus or toward cellular growth tips. These two types of granules contain different RNA species and levels of Whi3 protein (*5*, *11*). Langdon *et al.* show that RNA secondary structure determines their interaction with other RNAs through exposure or masking of complementary sequences, forming double-stranded RNA, as well as specific Whi3 RNA binding sites. Importantly, the two types of Whi3-containing granules do not mix, either in vitro or in cells, supporting their distinct structures. These findings suggest that RNA secondary structure delineates the complex composition of granules through the formation of specific RNA-RNA and RNA-RBP interactions.

Both studies point to RNA secondary structure—*Neat1* or Whi3-interacting RNAs—as a major navigator of the formation, identity, and composition of phase-separated granules, and both groups speculated that these principles are likely to be generally applicable for other granules, including cytoplasmic SGs. In line with this, RNA-RNA interactions were recently shown to contribute to the formation of SGs and define their overall RNA composition (*12*). Moreover, the protein composition of SGs varies among different cell types, with neurons demonstrating the highest complexity (*13*). In this RNA-centric view of phase separation, a key question now is whether the specific cytoplasmic RNA pool of each cell type drives the disparity in the protein composition of SGs.

Another related outstanding question is what defines the specificity of protein components in pathologic aggregates? Even though many different proteins are localized to physiologic SGs, only a few of them persist in pathologic aggregates. Moreover, different RBPs comprise the pathologic aggregates in specific disease subtypes, even in the absence of disease-causing mutations. For example, TDP-43 is the main aggregated protein in a large subset of patients with frontotemporal dementia (FTD), whereas FUS seems undisturbed in the same neurons. Conversely, FUS forms cytoplasmic aggregates in a distinct group of FTD patients that do not show TDP-43 aggregates. In view of the studies from Maharana *et al.* and Langdon *et al.*, the distinct RNA binding specificities of TDP-43 (*14*) and FUS (*8*) may account for their preferential incorporation into diverse granules before their conversion into pathologic aggregates. Which are these specific RNAs and whether they remain trapped within the pathologic assemblies are key questions.

It will be important to further dissect the role of RNA in the formation and identity of different phase-separated RBPs and to understand the rules that regulate their differential composition within a cell and among different cell types. Recently, long RNAs associated with repeat expansion disorders, including Huntington's disease, muscular dystrophy, and amyotrophic lateral sclerosis, were shown to phase-separate independently of RBPs (*15*). In view of the studies by Maharana *et al.* and Langdon *et al.*, the accumulation of foci containing repetitive RNA in cells from patients with neurodegeneration may trigger a broad imbalance in the overall RBP:RNA ratios and interactions and a more general misregulation of phase separation in nuclear and cytoplasmic compartments than previously understood. ■

REFERENCES AND NOTES

1. S. Boeynaems *et al.*, *Trends Cell Biol.* **28**, 420 (2018).
2. A. Molliex *et al.*, *Cell* **163**, 123 (2015).
3. D. T. Murray *et al.*, *Cell* **171**, 615 (2017).
4. S. Maharana *et al.*, *Science* **360**, 918 (2018).
5. E. M. Langdon *et al.*, *Science* **360**, 922 (2018).
6. M. Polymenidou, D. W. Cleveland, *Cell* **147**, 498 (2011).
7. M. Ramaswami, J. P. Taylor, R. Parker, *Cell* **154**, 727 (2013).
8. C. Lagier-Tourenne *et al.*, *Nat. Neurosci.* **15**, 1488 (2012).
9. A. C. Woerner *et al.*, *Science* **351**, 173 (2016).
10. L. A. Becker *et al.*, *Nature* **544**, 367 (2017).
11. H. Zhang *et al.*, *Mol. Cell* **60**, 220 (2015).
12. B. Van Treeck *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **115**, 2734 (2018).
13. S. Markmiller *et al.*, *Cell* **172**, 590 (2018).
14. M. Polymenidou *et al.*, *Nat. Neurosci.* **14**, 459 (2011).
15. A. Jain, R. D. Vale, *Nature* **546**, 243 (2017).

ACKNOWLEDGMENTS

M.P. apologizes to all authors whose relevant papers could not be cited due to space limitations. M.P. is grateful to F. Allain, M. Perez Berlanga, and M. Jambeau for critical reading of the manuscript. M.P. is supported by the Swiss National Science Foundation and holds funding from Target ALS, the ALS association, and the U.S. Department of Defense. She is an EMBO Young Investigator.

10.1126/science.aat8028

GRAPHIC: KELLIE HOLOSKI/*SCIENCE*

POLICY FORUM

SCIENCE EDUCATION

Disparities in science literacy

Cognitive and socioeconomic factors don't fully explain gaps

By **Nick Allum**[1], **John Besley**[2], **Louis Gomez**[3,4], **Ian Brunton-Smith**[5]

Much is known about how adult science literacy varies internationally and over time, and about its association with attitudes and beliefs. However, less is known about disparities in science literacy across racial and ethnic groups (*1*). This is particularly surprising in light of substantial research on racial and ethnic disparities in related areas such as educational achievement, math and reading ability (*2*), representation in science, technology, engineering, and math (STEM) occupations (*3*), and health literacy (*4*). Given the importance of science literacy to securing and sustaining many jobs, to understanding key health concepts to enhance quality of life, and to increasing public engagement in societal decision-making (*5*), it is concerning if the distribution of science literacy is unequally stratified, particularly if this stratification reflects broader patterns of disadvantage and cultural dominance as experienced by minorities and educationally underserved populations. We describe here such disparities in science literacy in the United States and attempt to explain underlying drivers, concluding that the science literacy disadvantage among black and Hispanic adults relative to whites is only partially explained by measures of broader, foundational literacies and socioeconomic status (SES).

The main source of evidence about U.S. patterns and trends in science literacy is the National Science Board's *Science and Engineering Indicators* (SEI) survey module (*3*), administered biennially since 2006 to a subsample of respondents for the General Social Survey (GSS), a high-quality biennial survey that seeks to provide a representative picture of American adults (aged 18 or older). Science literacy is captured by questions covering basic scientific facts and processes, but subgroup analysis is only presented in SEI for gender, age, education, and income because sample sizes do not permit more granular analysis. Research using different questions (*6*) found that white Americans score more highly than blacks and Hispanics, although sample sizes for black and Hispanic groups were relatively low and only bivariate analysis was presented. Other work (*7*) found that black Americans reported lower confidence in science, even after adjusting for attitudinal and demographic factors, though this study did not look at science literacy.

In contrast to the sparse research on race and adult science literacy, there is voluminous evidence of racial inequalities in educational measures of children's science knowledge (*8*). Moving beyond narrow science literacy to health literacy and foundational reading literacy, we see similar stratified patterns where white Americans do better than blacks and Latinos, with substantial variation across SES groups (*9*).

> **"...educational interventions need to measure...not just the quantity of instruction... but also quality."**

In the present study, our first objective is to examine racial and ethnic disparities in science literacy among adults in the United States. We estimate these by combining data from six waves of the GSS between 2006 and 2016 ($n = 2339$). We take it as axiomatic that the explanation for such disparities must be found in socially determined factors that fall differentially on different groups. Our second objective is therefore to investigate plausible factors, including demographics, foundational literacy, attitudes, and access to information that could account for such disparities [see supplementary materials (SM) for details on all data and analyses].

The GSS science survey module includes multiple choice (mostly true/false) questions about science content and process, along with open-ended questions. We regard these questions as indicators of the broader construct of science literacy. To measure foundational literacy, we rely on a well-established measure of verbal ability, Wordsum, that has been included in the GSS since 1974 (*10*). We use a standard set of demographic controls, including gender, birth cohort, geographical region, education, income, and religion. We also employ the following covariates that we hypothesize could account for between-group literacy differences.

We use a particularly rich measure of SES, the Cambridge Social Interaction and Stratification (CAMSIS) scale (*11*), that represents differences in status, prestige, and economic advantage, based on respondents' occupational groups. This measure is useful as it reflects the kinds of personal networks, social class, and cultural milieu, in which views about science develop and which may overlap with racial inequality. Further, we might surmise that a largely white teaching force has often failed to understand contexts of social life and interests of black and Hispanic students, to connect those interests to scientific phenomena, and to support scientific literacy about the phenomena. This neglect may lead to different levels of science literacy even given equivalent formal qualifications (*12*).

Some minority groups express less trust and confidence in science compared to whites (*7*). Low confidence in science arguably could lead to lack of science engagement in various settings, thus lower knowledge scores. We therefore include measures that ask people how much "confidence" they have in the "scientific community" and how positive they are about science.

One potential benefit of the internet would be in reducing the gap between the information-rich and information-poor. Yet, knowledgeable individuals are often able to acquire information more effectively, so the internet may exacerbate knowledge gaps (*13*). We include a question on whether respondents have sought science information on the internet.

Pooling the samples across all years yields a mean science literacy quiz score for whites of 8.6 (out of a maximum possible 13), Hispanics 6.8, and blacks 6.5. A one-way analysis of variance shows statistically significant differences between groups ($F = 283$; df = 3). The overall mean for all groups combined is 8.0 with a standard deviation of 2.7; the average difference between whites and the two racial and ethnic groups is quite substantial, at around two-thirds of a standard deviation.

We tried to gain better understanding of these disparities by adjusting for potential confounding factors. We fit several multivariate ordinary least squares regression models with science knowledge as the dependent variable. The first model included indicators

---

[1]Department of Sociology, University of Essex, Colchester, UK. [2]Department of Advertising and Public Relations, Michigan State University, East Lansing, MI, USA. [3]Department of Education, University of California, Los Angeles, CA, USA. [4]Carnegie Foundation for the Advancement of Teaching, Stanford, CA, USA. [5]Department of Sociology, University of Surrey, Guildford, UK. Email: nallum@essex.ac.uk

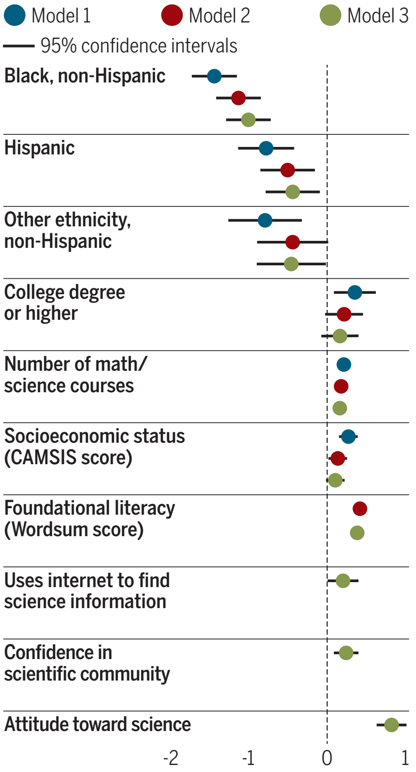
for race and ethnicity alongside demographic variables: sex, birth cohort, region of residence, college education, and income. We also include SES and number of science courses taken at high school or college level. If science literacy disparities are due to these factors, then statistically adjusting for them should attenuate or remove any residual differences among the groups. In the second model, we added our measure of foundational literacy. If this eliminates race or ethnicity differences, it would support the hypothesis that broader literacy disparities lie behind science literacy disparities. In the third model, we added attitude toward science, confidence in science, and science internet use, to see if these more proximal features of orientation to science and technology might lie behind science literacy differences.

Even after adjusting for demographic variables, science knowledge disparities are only partially reduced. There is still around a 1.5-point difference in the average scores of black and white Americans [see the figure, model 1 (blue)]. The gap for Hispanics is narrower, but their mean is still about a point lower than that for whites. Most of the other included predictors are significantly associated with science literacy and consistent with extant research (fig. S1 and table S1). Model 1 accounts for about 20% of the variance in science literacy.

Adding foundational literacy to the model shrinks the coefficients for both minority groups. Residual gaps are just over one point for blacks and a half point for Hispanics [see the figure, model 2 (red)]. This model accounts for just under 22% of the variance. The hypothesis that disparities in foundational literacy account for gaps in science literacy receives some support, but there is much left to explain. Adding behavioral and psychological variables increases the amount of variance explained by the model to 30%, but we see little change in the race and ethnicity coefficients [see the figure, model 3 (green)]. Race and ethnicity continue to matter even when comparing the science literacy of people with similar science attitudes.

Change in science knowledge score per unit change in independent variables

Unstandardized regression coefficients. Model 1 = demographic variables; model 2 = model 1 + foundational literacy; model 3 = model 2 + behavioral/psychological variables. Effect of black/Hispanic/other is compared to white. Effect of college or above is compared to less than college. The zero line means no effect. See full results in the SM.



GRAPHIC: ADAPTED BY J. YOU/SCIENCE

Overall, disparities in science literacy cannot be straightforwardly "explained" by intergroup differences in the levels of our measured characteristics. We performed a decomposition analysis and found that whereas around one-third of the variation in knowledge scores is explained by the independent variables, only about half of the total race and ethnicity gaps are explained by these observables. For both blacks and Hispanics, differences in foundational literacy compared to whites are the most important of the observable influences on the size of the disparity. The remaining, unexplained, portion of the gap must in large part be driven by variables not in our models. Translating confidence in, and positive attitudes toward, science into higher science literacy appears to be less common for blacks and Hispanics than for whites, for unclear reasons.

A principal question we wanted to address was the extent to which ethnic and racial inequalities in science literacy are simply reflections of well-established disparities in more fundamental axes of disadvantage, including broader foundational literacies. They are not. When we compared whites with black and Hispanic respondents who hold similar attitudes toward science and have the same degree of confidence in its institutions, we still find persistent disparities in science literacy. We do not claim to have captured all of these disadvantages in our analysis, as our variables are measured with error and are relatively broad-brush, but we have at least included key dimensions.

This analysis invites the question as to what could be responsible for the remaining gaps. It may be that the specific knowledge questions asked or the survey response context favors whites, but, more important, we suspect that our education measures mask considerable heterogeneity in the experiences of children, young people, and adults of different races and ethnicities. Graduating from high school, earning a college degree, or taking a science class can consist in a wide variety of experiences, some of which are likely correlated with race and ethnicity. As recently as the early 1970s, black and Hispanic children were much more likely to attend schools funded at a lower than average rate and intentionally segregated by ethnicity, and such segregation has continued de facto to varying degrees (*14*). Microsocial experiences of nondominant groups in any learning environment—for example, stereotype threat and racial microaggressions—can shape learning experiences (*15*). Although we adjust for educational qualifications, we do not capture differences in the quality of education experienced by blacks and Hispanics.

This suggests that educational interventions need to measure, and target, not just the quantity of instruction and formal qualifications, as we do here, but also quality. We may also be able to craft training and public awareness campaigns to help scientists, teachers, and employers to be more sensitive to the subtle manifestations of bias. Whatever the remedy, ignoring science literacy disparities among underserved groups does not serve science or society well. ■

REFERENCES

1. National Academies of Sciences, Engineering, and Medicine, *Science Literacy: Concepts, Contexts, and Consequences* (National Academies Press, Washington, DC, 2016).
2. G. Kena *et al.*, *The Condition of Education 2014* (NCES 2014-083, U.S. Department of Education, National Center for Education Statistics); http://nces.ed.gov/pubs2014/2014083.pdf (2014).
3. National Science Board, "Science and Engineering Indicators 2016" (National Science Foundation, Arlington, VA, 2016).
4. R. V. Rikard *et al.*, *BMC Public Health* **16**, 975 (2016).
5. W. K. Hallman, in *Handbook of the Science of Science Communication* (Oxford Univ. Press, 2017), pp. 61–72.
6. Pew Research Center, *A Look at What the Public Knows and Does Not Know About Science* (Washington, DC, 2015).
7. E. Plutzer, *Bull. Sci. Technol. Soc.* **33**, 146 (2013).
8. National Center for Education Statistics (2012), The Nation's Report Card: Science 2011 (NCES 2012–465). Institute of Education Sciences, U.S. Department of Education (Washington, DC, 2011).
9. C. C. Cutilli, I. M. Bennett, *Orthop. Nurs.* **28**, 27 (2009).
10. N. Malhotra, J. A. Krosnick, E. Haertel, The psychometric properties of the GSS Wordsum vocabulary test. *GSS Methodological Report* 11 (NORC, 2007).
11. K. Prandy, *Sociology* **24**, 629 (1990).
12. S. D. Museus, R. T. Palmer, R. J. Davis, D. Maramba, *Racial and Ethnic Minority Student Success in STEM Education: ASHE Higher Education Report, Volume 36, Number 6* (Wiley, 2011).
13. L. Su, M. A. Cacciatore, D. A. Scheufele, D. Brossard, M. A. Xenos, *Sci. Commun.* **36**, 352 (2014).
14. L. Darling-Hammond, in *Handbook of Research on Multicultural Education*, J. A. Banks, Ed. (Jossey-Bass, San Francisco, 2001), pp. 465–583.
15. S. M. Quintana, L. Mahgoub, *Theory Into Practice* **55**, 94 (2016).

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/360/6391/861/suppl/DC1

10.1126/science.aar8480

PHOTO: ISTOCK.COM/LIDERINA

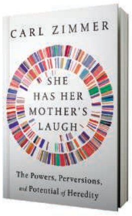
HUMAN HEREDITY

Our inheritance

A fascinating history of heredity research reveals the field's highs and lows

By **Axel Meyer**

With hundreds of ancient human genome sequences at our fingertips and millions of contemporary samples provided by customers of consumer genetics companies, now—more than ever before—we are able to discover, decipher, and interpret mixing, migration events, and genetic variants in human populations. Into this zeitgeist enters Carl Zimmer's most enjoyable new book, *She Has Her Mother's Laugh*, with a sweeping overview of the history of our understanding of heredity.



She Has Her Mother's Laugh
The Powers, Perversions, and Potential of Heredity
Carl Zimmer
Dutton, 2018. 672 pp.

From Mendel's peas and Soviet-era Lysenkoism to chimeras, CRISPR, and the microbiome, the book traces the scientific discovery of genes, chromosomes, and epigenetics, as well as recent advances in gene therapy and paleogenomics. It is ambitious in scope, but Zimmer is one of the best science journalists of our times, with a long history of setting the bar for beautiful, clear, and scholarly writing. He is true to form in this book.

The reviewer is at the Department of Biology, University of Konstanz, 78457 Konstanz, Germany, and the Radcliffe Institute for Advanced Study, Harvard University, Cambridge, MA 02138, USA. Email: axel.meyer@uni-konstanz.de

Zimmer begins with a personal tale of a visit to a genetic counselor that preceded the birth of his first child. This experience reveals how little he actually knew about heredity at the time. Later, he seems to "hear heredity"—his daughter's laugh is so similar to his wife's—but why? And what makes his second daughter so different from the first?

Interesting historical tidbits are peppered throughout the book. The term "hereditas," Zimmer reveals, for example—employed by the Romans—referred not to biological inheritance but to a legal concept concerning who would inherit a relative's assets. Genealogies, we learn, evolved in their depiction of biological connections from vertical lines to treelike structures, which, to the French, resembled the forked *pé de grue* ("crane's foot," later rendered into "pedigree" in English).

In the 14th through the 16th centuries, Europeans began to conceive of the link between generations as being related to the blood, although other cultures maintained very different ideas. Native populations on the Malaysian island of Langkawi, for example, continued to believe that kinship was achieved by sharing foods.

In his daughter's laugh—so similar to her mother's—Zimmer wondered if he was "hearing heredity."

Evolutionary biology is Zimmer's specialty: He coauthored one of the best textbooks in the field, and it shows (*1*). But evolution and heredity are not the same, both constituting huge fields that only partially overlap.

Charles Darwin famously misunderstood inheritance, proposing that tiny particles shed from an organism's many cells migrated to the gonads, where they were transferred to offspring. His cousin Francis Galton did further disservice to the field of heredity, having conceived of the notion of "eugenics": a program whereby strategically arranged marriages would result in "better" humans from generation to generation. Ironically, Nazi bombs destroyed the Galton laboratory during World War II. However, the lab's research continued, becoming part of University College London's Department of Genetics, Evolution, and Environment in 2013.

"Individual 'Z,'" as Zimmer refers to himself after having his own genome sequenced, consults experts in New Haven and New York to help him interpret the results. His Y-chromosome haplotypes, he learns, are shared with Napoleon's (a lock of hair was saved after the general's death, allowing French researchers to extract some genetic information in 2011). Luckily for Zimmer, his results were otherwise rather boring.

The book reminds one how little information is typically transmitted between generations in many families. Often, we don't know where our ancestors were born, what diseases plagued them, or how long they lived. Genome sequencing stands to shed light on some of these genealogical questions. Now we can say, for example, that we are 43% Ashkenazi Jewish or put a name to a mysterious malady that plagued our mother's sister.

When one writes about genetics and heredity in humans, one runs the danger of igniting societal short fuses. Zimmer does not shy away from controversies, but he does, artfully, ship around some of the shallows. He points out, for example, how historically transient, arbitrary, and scientifically wrong it is to classify humans based on skin colors. "[S]kin color," he writes, "is not a timeless hallmark of human races. It has changed in different places and at different times."

She Has Her Mother's Laugh is more than a historical account of the field of genetics. It's a treasure trove of curious facts, contextual tidbits, and up-to-date reports on the trials and tribulations of heredity told in a most entertaining way. ■

REFERENCE

1. C. Zimmer, D. J. Emlen, *Evolution: Making Sense of Life* (Roberts and Company, Greenwood Village, CO, 2012).

10.1126/science.aat3786

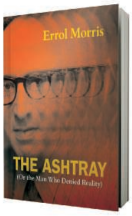
PHILOSOPHY OF SCIENCE

Truth with a vengeance

A filmmaker with an ax to grind takes aim at Thomas Kuhn's legacy

By **Joseph D. Martin**

Thomas Kuhn, an American physicist, historian, and philosopher of science, achieved prominence with the 1962 publication of his influential book, *The Structure of Scientific Revolutions*. Science, it contended, does not generate incrementally truer descriptions of reality but develops through radical paradigm shifts. Kuhn's theory of scientific revolutions motivated expansions and reactions that reconfigured the history and philosophy of science.



The Ashtray (Or the Man Who Denied Reality)
Errol Morris
University of Chicago Press, 2018. 223 pp.

Errol Morris, an acclaimed documentary filmmaker who briefly studied under Kuhn at Princeton, instead remembers the acerbic philosopher for flinging the titular ashtray at him during an argument in 1972. Morris hurls *The Ashtray* back, targeting not Kuhn's person but his legacy. "Our posttruth moment is all Thomas Kuhn's fault" would be an unfair summation of the book—but only just.

Structure, Morris writes, brutalized the truth, and Kuhn was an avatar of postmodernism who rejected the external world's very existence. Deep personal animosity—so overt that he compares Kuhn to Hitler in his acknowledgments—fuels Morris's critiques.

A filmmaker's sensibility informs the book's cadence; rather than demonstrating Kuhn's errors, Morris conjures an image of his wrongness. He interweaves philosophical explication, cultural allusion that careens between the highbrow and the low, interviews with authoritative figures, and copious illustrations. Incommensurability is his principal target.

Kuhn's notion that adherents of one paradigm inhabit a different mental world from, and cannot communicate meaningfully with, those of another is the most radical element of his philosophy. Morris, who favors Saul Kripke's causal theory of reference, considers incommensurability to be epistemic anarchy.

For Kripke, theoretical commitments do not bedevil our linguistic reference to things. Morris adjudges Kripke's philosophy a firm foundation for scientific knowledge, setting up a choice: "Does science progress toward a more truthful apperception of the physical world? Or is it all a matter of opinion, a sociological phenomenon that reflects consensus, not truth?" Hard-nosed scientific certitude or frothy-lipped relativism?

This dyad, too neat to encapsulate Kuhn's legacy, recalls Morris Zapp, the bumptious protagonist of David Lodge's *Changing Places*. Zapp had a mission: a series of Jane Austen commentaries so panoptic "that when each commentary was written there would be simply *nothing further to say* about the novel in question."

The joke, of course, is the perfect totality with which any such endeavor escapes the point. We don't study literature in search of finality but to enable new ways of thinking. Daniel Dennett makes a similar suggestion about philosophy, asking whether it would be preferable to be so definitively correct that you foreclose an area of inquiry or to err so spectacularly that people debate your blunders for centuries. Scientists tend to choose the former, philosophers the latter.

Kuhn had many fecund failures. *Structure*, for all its flaws—many of Morris's specific critiques are apt—forced new approaches to questions about how scientific truth claims arise in practice. Morris might even have been inclined to agree, if not for the depth of his animus. "The role of a good documentary," he once told the *Boston Globe*, "is not to convince you about what happened, but to force you to *think* about what happened."

Theories of reason and reference describe the clean, ideal ways that our minds might latch onto the world. But the practice of science is far messier. Kuhn sensitized historians and philosophers to that messiness. He wondered how philosophy would have to adapt if it countenanced the historical processes by which scientists actually achieved understanding. Close scrutiny reveals those to be profoundly dissimilar to the tidy philosophical frameworks Morris favors, which are little informed by, and have scarcely influenced, the practices that provide the sturdiest accounts of the natural world.

Morris charges Kuhn with undermining the truth and authority of science. But that authority, the crux of cultural battles over climate change and evolution, has less to do with abstract, transcendent truth than with reliability, consensus, and the social processes that enable them.

The ironclad truth Morris seeks, so integral to the popular image of science, is one of science's greatest vulnerabilities. A responsible scientist will always admit doubt. Combined with the widespread perception that science seeks irreproachable truth, that is a powerful weapon for those who would undermine scientific authority.

Thomas Kuhn was unkind to Errol Morris. By all accounts, he was often unkind. That kindness and influence so frequently fail to intersect is a bitter injustice. But so, too, is vilification in the service of vendetta. ■

10.1126/science.aat0237

The reviewer is at the Department of History and Philosophy of Science, University of Cambridge, Cambridge CB2 1TN, UK. Email: jdm205@cam.ac.uk

PODCAST

Inventing Ourselves
The Secret Life of the Teenage Brain
Sarah-Jayne Blakemore
PublicAffairs, 2018, 256 pp.

The social risk of being rejected by peers may lead adolescents to take physical risks.

Elusive, evasive, and uncommunicative, the human adolescent is among the most enigmatic subjects ever to be studied. This week on the *Science* podcast, Sarah-Jayne Blakemore delves into the teenage brain, revealing the extraordinary features that define this transitional state in human development. www.sciencemag.org/podcasts

10.1126/science.aat9678

PHOTO: ISTOCK.COM/CASTENOID

The pumpkin toadlet is just one example of the diverse fauna endemic to Brazil.

PHOTO: TARAN GRANT

Edited by **Jennifer Sills**

Brazil's government attacks biodiversity

To address the rapid and massive loss of biodiversity worldwide, scientific research must inform agile decision-making. The political leaders of Brazil, the country with the planet's greatest biological wealth (*1*), continue to undermine this goal. In 2001, citing the laudable objective of preventing biopiracy, the government created the Genetic Heritage Management Council (CGen) (*2*), making established biodiversity research activities illegal overnight by imposing severe limitations on access and shipment of specimens and samples. After 5 years of pressure from the academic community, the government finally resolved to exempt several areas of basic research from the restrictive policy (*3*). Now, after severe budget cuts (*4*, *5*) and legal maneuvers directly affecting biodiversity (*6*), the Brazilian Federal Government has reversed its position once again with the so-called New Law on Biodiversity (*7*, *8*), striking a potentially fatal blow to biodiversity.

Disguised as a milestone of progress for scientific research and endorsed by some scientific sectors (*9*), this law constitutes a monumental setback. According to the law, which is retroactive to 2000 or 2015, depending on the area of research (*8*), and technically applies to researchers at both Brazilian and foreign institutions (*10*), virtually every research activity on Brazilian biodiversity—even nongenetic studies in taxonomy, ecology, physiology, and behavior—must be registered in the National System of Genetic Resource Management and Associated Traditional Knowledge (SisGen), which was created to assist the CGen. The deadline for meeting these draconian rules is 5 November 2018 (*9*, *11*). Failure to comply, including the unregistered dissemination of scientific results, even if based only on public databases like GenBank or previously published data or results, will result in heavy fines (*7*). A lecture given by the coordinator of the Academic Sectoral Board of the CGen, titled "How not to be fined," highlights the punitive spirit of the law (*10*). Paradoxically, commercial activities involving Brazilian biota, such as export of ornamental fishes, plants, grains, and other marketable products, remain unaffected by the law (*7*).

If not repealed or substantially overhauled, this Byzantine labyrinth of unnecessary demands and threats will decimate scientific research on Brazilian biodiversity by requiring scientists to divert an inordinate amount of already limited resources from research to the time-consuming process of registering every specimen, DNA sequence, photograph, and any other observation of Brazilian biodiversity before publication, presentation at scientific meetings, or dissemination to media outlets. Scientists must take back the reins of their own activities by demanding that the Brazilian government implement laws that facilitate international collaboration and encourage biodiversity research instead of stifling it. Otherwise, a substantial part of the world's biodiversity and its benefits may silently vanish behind a wall of bureaucracy.

Flávio Alicino Bockmann,[1]* Miguel Trefaut Rodrigues,[2] Tiana Kohlsdorf,[1] Lorian Cobra Straker,[1] Taran Grant,[2] Mário César Cardoso de Pinna,[3] Fernando Luis Medina Mantelatto,[1] Aléssio Datovo,[3] José Perez Pombal Jr.,[4] John Campbell McNamara,[1] Eduardo Andrade Botelho de Almeida,[1] Wilfried Klein,[1] Annie Schmaltz Hsiou,[1] Milton Groppo,[1] Ricardo Macedo Corrêa e Castro,[1] Dalton de Souza Amorim[1]

[1]Department of Biology, Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto, SP 14040-901, Brazil. [2]Department of Zoology, Instituto de Biociências, Universidade de São Paulo, São Paulo, SP 05508-090, Brazil. [3]Museu de Zoologia da Universidade de São Paulo, São Paulo, SP 04263-000, Brazil. [4]Department of Vertebrates, Museu Nacional, Universidade Federal do Rio de Janeiro, Rio de Janeiro, RJ 20940-040, Brazil.
*Corresponding author.
Email: fabockmann@ffclrp.usp.br

REFERENCES

1. R. A. Mittermeier, C. G. Mittermeier, P. R. Gil, E. O. Wilson, *Megadiversity: Earth's Biologically Wealthiest Nations* (CEMEX, 1997).
2. Medida Provisória No 2.186-16, (23 August 2001); www.planalto.gov.br/ccivil_03/mpv/2186-16.htm [in Portuguese].
3. Resolução Nº 21 (31 August 2006); www.mma.gov.br/estruturas/sbf_dpg/_arquivos/res21cons.pdf [in Portuguese].
4. H. Escobar, *Science*, 10.1126/science.aal0276 (2016).
5. C. Angelo, *Nature*, 10.1038/nature.2017.21766 (2017).
6. H. T. Pinheiro *et al.*, *Science* **350**, 1043 (2015).
7. Lei Nº 13.123 (20 May 2015); www.planalto.gov.br/ccivil_03/_ato2015-2018/2015/lei/l13123.htm [in Portuguese].
8. Decreto Nº 8.772 (11 May 2016); www.planalto.gov.br/ccivil_03/_ato2015-2018/2016/decreto/D8772.htm [in Portuguese].
9. M. da Silva, D. R. de Oliveira, *Braz. J. Microbiol*. **49**, 1 (2018).
10. Audiovisual IB-USP, "Lei da Biodiversidade e Plataforma SisGen: Como não ser multado. - Dra. Manuela da Silva" (2018); https://youtu.be/x6t4rXARkEs [in Portuguese].
11. D. R. Oliveira, M. da Silva, F. do Carmo, R. Angeli, *J. Ciênc*. **5772**, 1 (2017).

10.1126/science.aat7540

NEXTGEN VOICES: SUBMIT NOW

Broad interests: Benefits for science

Add your voice to *Science*! Our new NextGen VOICES survey is now open:

How do broad interests benefit your science? Describe one non-science interest or hobby, and explain how it has made you a better scientist.

To submit, go to www.sciencemag.org/nextgen-voices

Deadline for submissions is 1 June. A selection of the best responses will be published in the 6 July issue of *Science*. Submissions should be 150 words or less. Anonymous submissions will not be considered.

The road to wild yak protection in China

China is home to about 22,000 wild yaks, which account for 90% of the global wild yak population (*1*). The International Union for Conservation of Nature (IUCN) categorizes the wild yak, a cold-tolerant herbivore, as a vulnerable species (*2*), mainly attributable to excessive hunting for food and trade. China's wild yak is also threatened by land-use change, disease, environmental pollution, genetic contamination, climate change, and resource competition (*1*, *3*, *4*). In recent decades, infrastructure construction in China has grown increasingly disruptive to remaining wild yak populations.

Most wild yaks live in or near the Tibetan Plateau (*1*). These regions are located in or adjacent to areas zoned for the Western Development Strategy (*5*), an ambitious plan proposed in 1999 to increase the economic level and quality of life of China's rural citizens. The plan's implementation has accelerated the construction of railways and roads. Railways in the central western region of China, which account for 76.6% of China's current rail traffic (*6*), expanded from 70,000 km in 2014 (*7*) to 95,000 km in 2016 (*6*). Meanwhile, highway density increased from 7.7 km/100 km^2 in 1999 to 20.6 km/100 km^2 in 2008 (*8*). The increasing density of railways and roads has fragmented the habitats of wild yaks and forced them to migrate to resource-limited areas to escape from predation and conflict with humans.

China's wild yaks are an important genetic resource for breeding new yak species in an effort to sustainably develop animal husbandry in the Tibetan area and enrich the region's biodiversity (*1*, *9*). Wild yaks are also an important component of biodiversity in nature (*9*). To protect China's wild yaks, the Chinese government has built several nature reserves (*10*, *11*), but even there, the yaks are threatened by illegal hunting (*10*). To ensure the safety of China's wild yaks, China must further expand the nature reserves and effectively enforce existing hunting bans. The government should also implement scientific management and protection policies that minimize habitat fragmentation, resource plunder, and predation.

Ming Chen,[1,2]* Yingzhu Sun,[1,2] Chunping Yang,[1,2]* Guangming Zeng,[1,2] Zhongwu Li,[1,2] Jiachao Zhang[3]

[1]College of Environmental Science and Engineering, Hunan University, Changsha 410082, China. [2]Key Laboratory of Environmental Biology and Pollution Control (Hunan University), Ministry of Education, Changsha 410082, China. [3]College of Resources and Environment, Hunan Agricultural University, Changsha 410128, China.

*Corresponding author. E-mail: mchensn@hnu.edu.cn (M.C.); yangc@hnu.edu.cn (C.Y.)

REFERENCES

1. Q. Shi *et al.*, *J. Nat. Conserv.* **32**, 35 (2016).
2. IUCN, "Red list of threatened species" (2017); www.iucnredlist.org/details/2892/2890.
3. X. Liang, A. Kang, N. Pettorelli, *Oryx* **51**, 361 (2017).
4. J. Berger *et al.*, *Sci. Rep.* **5**, 8676 (2015).
5. H. H. Lai, *Modern China* **28**, 432 (2002).
6. Xinhuanet, "Length of railroad lines in service extends to 124,000 km where high-speed railway is 22,000 km long" (2017); www.xinhuanet.com/politics/2017-01/03/c_129430123.htm [in Chinese].
7. Xinhuanet, "Length of railroad lines in service extends to 112,000 km" (2015); www.xinhuanet.com/politics/2015-01/29/c_127436479.htm [in Chinese].
8. "The 10th anniversary of Western Development: The achievement of traffic development" (2009); www.chinahighway.com/news/2009/376530.php [in Chinese].
9. L. Chunnian *et al.*, *Mitochondrial DNA Part A* **27**, 4266 (2016).
10. P. J. Buzzard *et al.*, *Oryx* **44**, 577 (2010).
11. G. B. Schaller, W. Liu, *Biol. Conserv.* **76**, 1 (1996).

10.1126/science.aat6749

Trout in hot water: A call for global action

Trout are one of the most culturally, economically, and ecologically important taxonomic groups of freshwater fishes worldwide (*1*). Native to all continents in the Northern Hemisphere, trout belong to seven genera, which are distributed across 52 countries. These cold-water specialists provide recreation and food to millions of people and play important roles in ecosystem functioning and health (*2*). They are also extremely sensitive to human disturbances because they require cold, clean, complex, and connected habitats for survival and persistence (*3*)—all attributes that humans have substantially altered and degraded (*4*, *5*). Despite their importance as societal icons and as indicators of biodiversity, many of the world's trout species and lineages are endangered and some require immediate conservation efforts to reverse their precarious decline.

Of the 124 recognized species of trout (*6*, *7*), only 67 (54%) have been assessed by the International Union for Conservation of Nature (IUCN) (*7*). Alarmingly, 73% of these species are currently threatened with global extinction, and four are now extinct (*7*). Although some of these species are likely subspecies, lineages, or distinct ecotypes, this level of threat is exceptionally high compared with other vertebrate groups assessed by the IUCN (*8*). Widespread imperilment of trout reflects numerous human activities identified in the IUCN assessments, including invasive species, overfishing, pollution, dams, deforestation, agriculture, grazing, and mining (*7*). Climate change is further stressing trout populations by warming water temperatures, shifting streamflow regimes, increasing extreme events (such as floods, drought, and wildfire), and facilitating species invasions (*9*, *10*). Worse, climate change often acts in synergy with other stressors to further endanger trout (*11*, *12*), a pattern that will be intensified in coming decades as global temperatures continue to rise, with important consequences for trout, aquatic ecosystem functioning, and human well-being.

Reversing these declines will require progressive conservation efforts to protect native trout diversity and ameliorate ongoing and future threats at local and global scales. To preserve these unique fishes, we must protect ecological and genetic diversity, which are critical for long-term resiliency, viability, and adaptation in the face of rapid environmental change. Innovative conservation approaches include reconnecting rivers with floodplains, establishing native fish refuges, restoring habitat diversity, and reducing invasive species, including nonnative trout stocking programs. Moreover, comprehensive, coordinated, and comparable approaches are needed immediately to assess conservation status and to delineate conservation units across the globe, particularly for data-poor species. Only by

Many trout species, such as these native bull trout, are endangered.

PHOTO: JOEL SARTORE

addressing threats at their root causes can we accomplish these conservation goals.

Clint C. Muhlfeld,[1,2]* Daniel C. Dauwalter,[3] Ryan P. Kovach,[1] Jeffrey L. Kershner,[4] Jack E. Williams,[5] John Epifanio[6]
[1]Northern Rocky Mountain Science Center, U.S. Geological Survey, West Glacier, MT 59936, USA. [2]Flathead Lake Biological Station, Division of Biological Sciences, University of Montana, Polson, MT, 59860 USA. [3]Trout Unlimited, Boise, ID 83702, USA. [4]Northern Rocky Mountain Science Center, U.S. Geological Survey, Bozeman, MT 59936, USA. [5]Trout Unlimited, Medford, OR 97501, USA. [6]Illinois Natural History Survey, University of Illinois, Urbana, IL 61801, USA.
*Corresponding author. Email: cmuhlfeld@usgs.gov

REFERENCES AND NOTES

1. J. Prosek, *Trout of the World* (Stewart, Tabori, & Chang, Revised edition, 2013), pp. 224.
2. C. M. Holmlund, M. Hammer, *Ecol. Econ.* **29**, 253 (1999).
3. R. J. Behnke, *Trout and Salmon of North America* (The Free Press, 2002).
4. F. R. Hauer *et al.*, *Sci. Adv.* **2**, e1600026 (2016).
5. A. L. Haak, J. E. Williams, *J. Conserv. Plan.* **9**, 38 (2013).
6. We searched the FishBase database (www.fishbase.org) for all described freshwater species of trout belonging to the genera *Oncorhynchus*, *Salvelinus*, *Salmo*, *Hucho*, *Parahucho*, *Brachymystax*, and *Salvethymus*.
7. We searched for each species on The IUCN Red List (www.iucnredlist.org). Of the 124 species, IUCN has assessed 67, and of those, 49 are listed as Vulnerable, Endangered, or Critically Endangered, and 4 are extinct.
8. IUCN Red List, "Table 1: Numbers of threatened species by major groups of organisms (1996–2017)" (http://cmsdocs.s3.amazonaws.com/summarystats/2017-3_Summary_Stats_Page_Documents/2017_3_RL_Stats_Table_1.pdf).
9. R. P. Kovach *et al.*, *Rev. Fish Biol. Fish.* **26**, 135 (2016).
10. C. C. Muhlfeld *et al.*, *Nat. Clim. Change* **4**, 620 (2014).
11. R. P. Kovach *et al.*, *J. Appl. Ecol.* **54**, 638 (2017).
12. C. C. Muhlfeld *et al.*, *Glob. Change Biol.* **23**, 4663 (2017).

10.1126/science.aat8455

TECHNICAL COMMENT ABSTRACTS

Comment on "Plant diversity increases with the strength of negative density dependence at the global scale"

Lisa Hülsmann and Florian Hartig

LaManna *et al.* (Reports, 30 June 2017, p. 1389) claim that subadult trees are proportionally less common at high conspecific adult density (CNDD), and that this effect increases toward the tropics and for rare species. We show that the CNDD-abundance correlation may have arisen from a methodological artifact and that a range of processes can explain the reported latitudinal patterns.

Full text: dx.doi.org/10.1126/science.aar2435

Response to Comment on "Plant diversity increases with the strength of negative density dependence at the global scale"

Joseph A. LaManna, Scott A. Mangan, Alfonso Alonso, Norman A. Bourg, Warren Y. Brockelman, Sarayudh Bunyavejchewin, Li-Wan Chang, Jyh-Min Chiang, George B. Chuyong, Keith Clay, Susan Cordell, Stuart J. Davies, Tucker J. Furniss, Christian P. Giardina, I. A. U. Nimal Gunatilleke, C. V. Savitri Gunatilleke, Fangliang He, Robert W. Howe, Stephen P. Hubbell, Chang-Fu Hsieh, Faith M. Inman-Narahari, David Janík, Daniel J. Johnson, David Kenfack, Lisa Korte, Kamil Král, Andrew J. Larson, James A. Lutz, Sean M. McMahon, William J. McShea, Hervé R. Memiaghe, Anuttara Nathalang, Vojtech Novotny, Perry S. Ong, David A. Orwig, Rebecca Ostertag, Geoffrey G. Parker, Richard P. Phillips, Lawren Sack, I-Fang Sun, J. Sebastián Tello, Duncan W. Thomas, Benjamin L. Turner, Dilys M. Vela Díaz, Tomáš Vrška, George D. Weiblen, Amy Wolf, Sandra Yap, Jonathan A. Myers

Hülsmann and Hartig suggest that ecological mechanisms other than specialized natural enemies or intraspecific competition contribute to our estimates of conspecific negative density dependence (CNDD). To address their concern, we show that our results are not the result of a methodological artifact and present a null-model analysis that demonstrates our original findings—(i) stronger CNDD at tropical relative to temperate latitudes, and (ii) a latitudinal shift in the relationship between CNDD and species abundance—persist even after controlling for other processes that might influence spatial relationships between adults and recruits.

Full text: dx.doi.org/10.1126/science.aar3824

Comment on "Plant diversity increases with the strength of negative density dependence at the global scale"

Ryan A. Chisholm and Tak Fung

LaManna *et al.* (Reports, 30 June 2017, p. 1389) found higher conspecific negative density dependence in tree communities at lower latitudes, yielding a possible mechanistic explanation for the latitudinal diversity gradient. We show that their results are artifacts of a selective data transformation and a forced zero intercept in their fitted model. A corrected analysis shows no latitudinal trend.

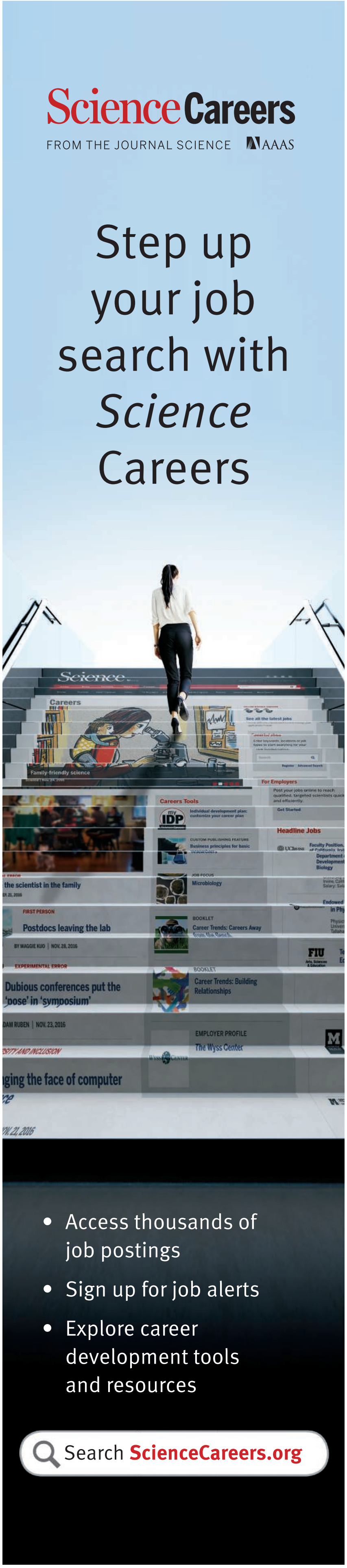
Full text: dx.doi.org/10.1126/science.aar4685

Response to Comment on "Plant diversity increases with the strength of negative density dependence at the global scale"

Joseph A. LaManna *et al.*

Chisholm and Fung claim that our method of estimating conspecific negative density dependence (CNDD) in recruitment is systematically biased, and present an alternative method that shows no latitudinal pattern in CNDD. We demonstrate that their approach produces strongly biased estimates of CNDD, explaining why they do not detect a latitudinal pattern. We also address their methodological concerns using an alternative distance-weighted approach, which supports our original findings of a latitudinal gradient in CNDD and a latitudinal shift in the relationship between CNDD and species abundance.

Full text: dx.doi.org/10.1126/science.aar5245



TECHNICAL COMMENT

FOREST ECOLOGY

Comment on “Plant diversity increases with the strength of negative density dependence at the global scale”

Lisa Hülsmann* and Florian Hartig

LaManna *et al*. (Reports, 30 June 2017, p. 1389) claim that subadult trees are proportionally less common at high conspecific adult density (CNDD) and that this effect increases toward the tropics and for rare species. We show that the CNDD-abundance correlation may have arisen from a methodological artifact and that a range of processes can explain the reported latitudinal pattern.

Conspecific negative density dependence (CNDD) has long been discussed as a key mechanism for maintaining local species richness and global biodiversity patterns. LaManna *et al.* (*1*) analyzed CNDD in 24 plots of the global CTFS-ForestGEO network, defining density dependence as an effect of local adult density on the recruit/adult ratio in 10 m × 10 m and 20 m × 20 m quadrats. Fitting this ratio against conspecific adult density with Ricker and offset-power models, they found that CNDD increases toward the tropics and with species rarity.

After carefully examining their statistical analysis, we believe that these patterns, intriguing as they are, may have arisen from a methodological artifact. We simulated synthetic data under a range of processes with and without CNDD (*2*) and found that both Ricker and offset-power models produce biased CNDD estimates as soon as a process breaks the spatial coupling between adults and recruits. For example, dispersal, adult mortality, or niche processes (*3*–*5*) can produce CNDD and CNDD-abundance correlations compatible with values reported in LaManna *et al.*, despite no CNDD being present whatsoever (Fig. 1). Moreover, the strength of the bias depends on community characteristics such as species richness or adult proportion, which suggests an alternative explanation for the reported latitudinal pattern.

One can understand the reason for this bias. If recruitment is not entirely local, the assumption that local counts of adults (A) and recruits (S) should be directly proportional does not hold, even without any CNDD (*6*). A special example of this is the frequent quadrats with recruits but no adults in the data. To explain these observations, the authors introduced a background adult density of $A = 0.1$ in the Ricker model. This decision, however, creates S/A ratios of ≥ 10 for the corrected observations with $A = 0.1$ and $S > 0$, much higher than the mean ratio of approximately 4 in the data where $A > 0$. To explain these seemingly high reproduction ratios in quadrats with only background adult density, the Ricker model must wrongly assume high intrinsic growth paired with strong density dependence (Fig. 2C). And because $S > 0$ and $A = 0$ occur more frequently for rare species and in the tropics,

Theoretical Ecology, Faculty of Biology and Preclinical Medicine, University of Regensburg, 93053 Regensburg, Germany.
*Corresponding author. Email: lisa.huelsmann@ur.de

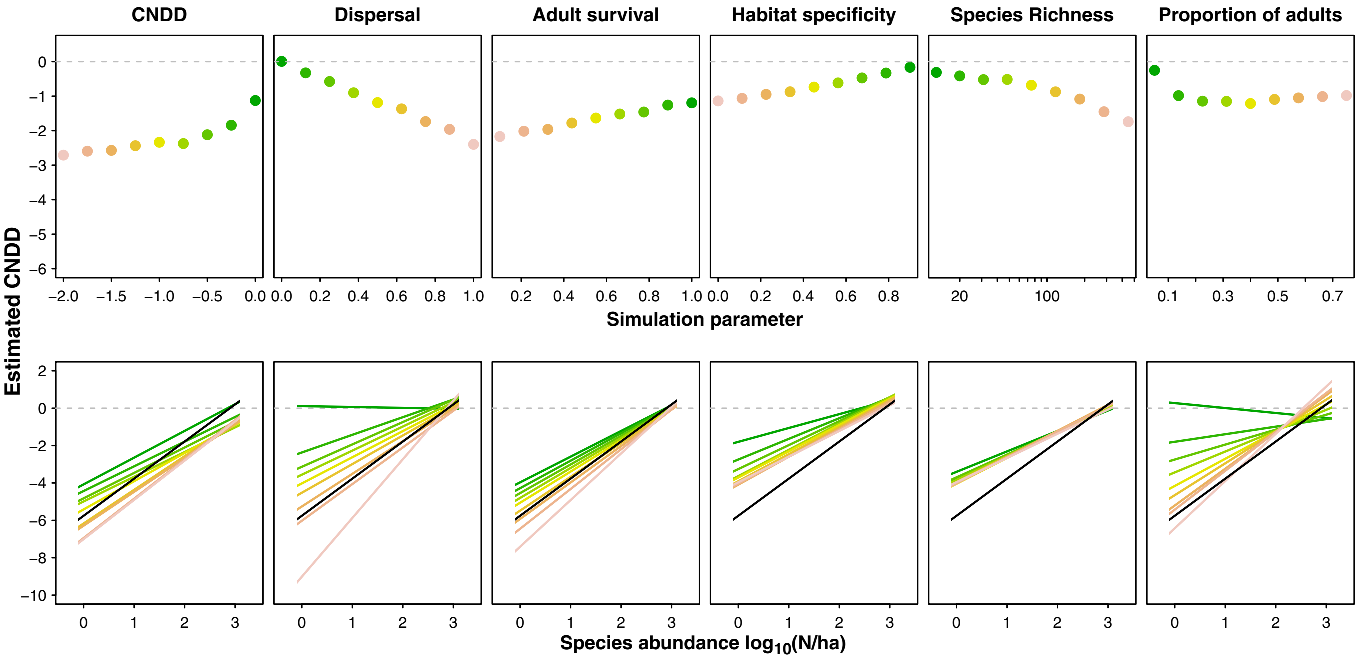


Fig. 1. Processes and community characteristics that create artifacts in estimated CNDD patterns. (**A** and **B**) Weighted mean CNDD (A) and CNDD-abundance correlations (B) estimated by the Ricker model with simulated data, varying ecological parameters (CNDD, dispersal, adult survival, habitat specificity), and community characteristics (species richness, proportion of adults) at the 10 m × 10 m scale. CNDD is zero except for the CNDD subplots where the same value was used for all species. See (*2*) for details of the simulation settings. Line colors correspond to parameter values in the upper subpanels; black lines depict CNDD-abundance correlations for the tropical Barro Colorado Island (BCI) (*16*) forest plot. Parameters estimation follows LaManna *et al.*

CNDD values and CNDD-abundance correlations emerge that are compatible with the patterns reported in LaManna *et al.*, even when no CNDD is present (Fig. 2A).

In the offset-power model, zeros were dealt with differently by adding a value of 1 to all observations. Again, this distorts recruit/adult ratios, as $S/A \neq (S + 1)/(A + 1)$, and more so for rare species with smaller values of S and A. When LaManna *et al.* then fitted a mixed-effects model with a random species effect on the CNDD slope but no random intercept (*7*), the rarity-dependent S/A distortion is compensated by species-specific CNDD (Fig. 2B). When we fitted the more natural standard random slope and intercept model that accounts for species-specific S/A ratios, the CNDD-abundance correlation vanished in simulated and real data (Fig. 2D).

LaManna *et al.* realized that CNDD estimates could be biased and therefore ran a number of simulations and null models to explore and potentially correct this bias. However, our results (Figs. 1 and 2) clearly contradict their assertion that their "results are generally robust to these potential biases." Examining their null models, we find several shortcomings: (i) Their analyses and corrections often only pertained to mean values and not to CNDD-abundance correlations, and where they did, simulations were not run with those critically low abundances that distort CNDD estimates; (ii) their null simulations included only dispersal, but none of the other factors that we show to create bias (Fig. 1); and (iii) their assumption that only 10% of the recruits derive from outside the quadrats produces unrealistically low dispersal relative to empirical values for primary dispersal (*8*). Even a uniform kernel with a 10-m radius would predict 52% of all seeds to disperse outside a 10 m × 10 m quadrat. With our null simulations, we tend to find clear CNDD signals and CNDD-abundance correlations similar to what LaManna *et al.* reported under perfectly neutral (CNDD = 0) conditions. Accordingly, we find their low CNDD estimates for rare temperate species far more surprising than strong CNDD and CNDD-abundance slopes in the tropics.

A reliable correction of CNDD estimates, even with more realistic null models, seems very difficult. One would require species-specific values for dispersal and all other spatial processes that we have shown to influence CNDD estimates (*9*). Species-specific values are important because a correlation of traits such as dispersal or habitat specificity with species abundance or latitude could create additional artifacts (*10*), let alone the effect of intraspecific trait variation. Based on this, we believe it is virtually impossible to devise appropriate corrections or null models for the analysis presented by LaManna *et al.* Directly analyzing CNDD in demographic rates such as growth and mortality of saplings (*11*), which are much less affected by dispersal, seems far more promising to us.

The biased estimator questions the evidence for the reported CNDD patterns, but does not constitute proof of their absence. Yet, when combining our analysis with general ecological knowledge, we find it rather unlikely that the patterns reported in LaManna *et al.* are primarily caused by CNDD. A previous study found no support for a latitudinal CNDD pattern in tree mortality (*12*), and although some studies have reported CNDD-abundance correlations in the tropics (*11*, *13*), effect sizes tended to be much smaller than what LaManna *et al.* reported, particularly as they inappropriately contrasted CNDD and heterospecific density effects. Strong CNDD-abundance correlations also seem unlikely from theoretical considerations. It is frequently assumed that CNDD is mediated by specialized pathogens (*14*). It would be surprising if such strict mutualisms could persist and exert effective control for exceedingly rare species. Moreover, if CNDD applied only to rare species, these would be stabilized against each other, but the question remains how infrequent species could escape competition from common ones that do not suffer from strong CNDD.

If not real CNDD, what then causes the latitudinal patterns in the (biased) CNDD estimates? Diversity differences alone could create a latitudinal pattern in mean CNDD but would not lead to a latitudinal shift of the CNDD-abundance correlation (Fig. 1). However, either latitudinal differences in spatial processes [e.g., stronger habitat effects or reduced dispersal in temperate regions (*9*)] or latitudinal differences in the size structure could create the pattern. Size structure is a highly sensitive parameter for the bias because LaManna *et al.* distinguished adults and recruits according to their diameter (Fig. 1, adult proportion). Because the original data were not fully available to us, we could not examine these possibilities in more detail.

We believe that LaManna *et al.*'s results should be interpreted far more carefully. Given the instability of the statistical methods, we do not see strong evidence for any of the reported CNDD patterns. To us, the findings mainly suggest that the spatial association between adults and recruits in the tropics is surprisingly weak, as already reported by (*15*), and that either this spatial association becomes stronger toward the temperate zone, or another parameter (possibly size structure) creates the latitudinal pattern in the biased estimator. A latitudinal gradient of abundance-dependent CNDD would also be compatible with our findings, but overall, we find it far more likely that a range of ecological processes—possibly in combination with CNDD, but not primarily driven by it—are responsible for the reported patterns.

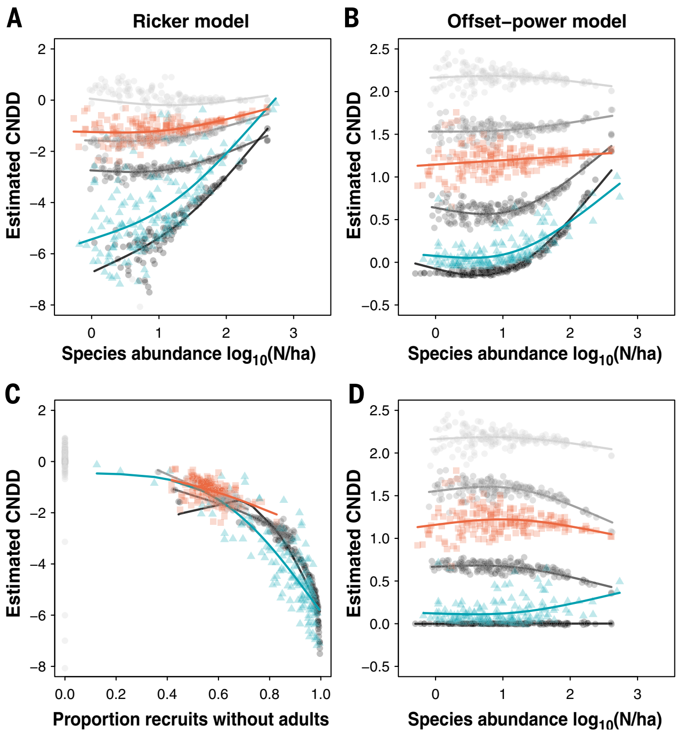


Fig. 2. CNDD bias in Ricker and offset-power models. (**A** and **B**) Estimated CNDD versus abundance [$\log_{10}$(N/ha)] per species from the Ricker model (A) and the offset-power model (B) at the 10 m × 10 m scale. Grayscale circles result from data simulated randomly without CNDD and changing spatial association between adults and recruits, ranging from perfect spatial coupling (lightest gray) to no spatial association (black). Blue triangles depict CNDD estimates for the tropical BCI (*16*) forest plot; orange squares denote the simulation model without CNDD used in the appendix of LaManna *et al*. (**C**) For the Ricker model, CNDD bias is highly correlated with the proportion of corrected adult counts. (**D**) Fitting a species-specific recruit/adult ratio in the offset-power model removes the CNDD-abundance correlation. See (*2*) for details of the simulation settings.

REFERENCES AND NOTES

1. J. A. LaManna *et al.*, *Science* **356**, 1389–1392 (2017).
2. Code and technical details of the data simulations and analyses are archived at https://github.com/LisaHuelsmann/CommentTo-LaMannaEtAl-Science.
3. R. John *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **104**, 864–869 (2007).
4. M. R. Metz, *J. Ecol.* **100**, 969–979 (2012).
5. T. G. Seidler, J. B. Plotkin, *PLOS Biol.* **4**, e344 (2006).
6. R. P. Freckleton, O. T. Lewis, *Proc. R. Soc. B* **273**, 2909–2916 (2006).

7. Note that LaManna *et al*. stated that they applied a random intercept and slope for CNDD (supplementary material, p. 5), but the code we obtained from the authors suppresses the random intercept, which is in line with the results they reported.
8. H. C. Muller-Landau, S. J. Wright, O. Calderón, R. Condit, S. P. Hubbell, *J. Ecol.* **96**, 653–667 (2008).
9. F. May, T. Wiegand, S. Lehmann, A. Huth, M. J. Fortin, *Glob. Ecol. Biogeogr.* **25**, 575–585 (2016).
10. L. S. Comita, *Science* **356**, 1328–1329 (2017).
11. L. S. Comita, H. C. Muller-Landau, S. Aguilar, S. P. Hubbell, *Science* **329**, 330–332 (2010).
12. J. Hille Ris Lambers, J. S. Clark, B. Beckage, *Nature* **417**, 732–735 (2002).
13. S. A. Mangan *et al*., *Nature* **466**, 752–755 (2010).
14. J. H. Connell, in *Dynamics of Populations*, P. J. de Boer, G. R. Gradwell, Eds. (Centre for Agricultural Publishing and Documentation, Wageningen, Netherlands, 1971), pp. 298–312.
15. S. Getzin, T. Wiegand, S. P. Hubbell, *Proc. R. Soc. B* **281**, 20140922 (2014).
16. R. Condit, S. Lao, R. Pérez, S. B. Dolins, R. B. Foster, S. P. Hubbell, *Barro Colorado Forest Census Plot Data, 2012 Version* (Center for Tropical Forest Science Databases, 2012).

ACKNOWLEDGMENTS

We thank R. Condit and S. P. Hubbell for providing the BCI data and for discussion that improved this comment. The BCI forest dynamics research project was founded by S. P. Hubbell and R. B. Foster and is now managed by R. Condit, S. Lao, and R. Perez under the Center for Tropical Forest Science and the Smithsonian Tropical Research Institute in Panama. Numerous organizations have provided funding, principally the U.S. National Science Foundation, and hundreds of field workers have contributed.

31 October 2017; accepted 18 April 2018
10.1126/science.aar2435

TECHNICAL RESPONSE

FOREST ECOLOGY

Response to Comment on "Plant diversity increases with the strength of negative density dependence at the global scale"

Joseph A. LaManna,[1,2]* Scott A. Mangan,[2] Alfonso Alonso,[3] Norman A. Bourg,[4] Warren Y. Brockelman,[5,6] Sarayudh Bunyavejchewin,[7] Li-Wan Chang,[8] Jyh-Min Chiang,[9] George B. Chuyong,[10] Keith Clay,[11] Susan Cordell,[12] Stuart J. Davies,[13,14] Tucker J. Furniss,[15] Christian P. Giardina,[12] I. A. U. Nimal Gunatilleke,[16] C. V. Savitri Gunatilleke,[16] Fangliang He,[17,18] Robert W. Howe,[19] Stephen P. Hubbell,[20] Chang-Fu Hsieh,[21] Faith M. Inman-Narahari,[12] David Janík,[22] Daniel J. Johnson,[23] David Kenfack,[13,14] Lisa Korte,[3] Kamil Král,[22] Andrew J. Larson,[24] James A. Lutz,[15] Sean M. McMahon,[25,26] William J. McShea,[27] Hervé R. Memiaghe,[28] Anuttara Nathalang,[5] Vojtech Novotny,[29,30,31] Perry S. Ong,[32] David A. Orwig,[33] Rebecca Ostertag,[34] Geoffrey G. Parker,[26] Richard P. Phillips,[11] Lawren Sack,[20] I-Fang Sun,[35] J. Sebastián Tello,[36] Duncan W. Thomas,[37] Benjamin L. Turner,[38] Dilys M. Vela Díaz,[2] Tomáš Vrška,[22] George D. Weiblen,[39] Amy Wolf,[19,40] Sandra Yap,[41] Jonathan A. Myers[1,2]

Hülsmann and Hartig suggest that ecological mechanisms other than specialized natural enemies or intraspecific competition contribute to our estimates of conspecific negative density dependence (CNDD). To address their concern, we show that our results are not the result of a methodological artifact and present a null-model analysis that demonstrates that our original findings—(i) stronger CNDD at tropical relative to temperate latitudes and (ii) a latitudinal shift in the relationship between CNDD and species abundance—persist even after controlling for other processes that might influence spatial relationships between adults and recruits.

To explore potential bias in our estimates of conspecific negative density dependence (CNDD) (*1*), Hülsmann and Hartig (*2*) present simulations that vary several processes, including dispersal, habitat specificity, adult/recruit ratios, and species richness. Some of these simulations produce spuriously strong CNDD for rare species, leading them to suggest that our methods might be biased. If this were correct, then our estimates of CNDD would be biased toward stronger effects for rare species at any latitude. However, this was not the case, because our original estimates of CNDD varied substantially among rare species when matched for abundance [figure 2D in (*1*)]. Furthermore, median CNDD for rare species differed across latitudes, with rare species having stronger median CNDD in tropical than in temperate forests [figure 2C in (*1*)].

The strong negative biases in CNDD that Hülsmann and Hartig observe for rare species are erroneously generated because their simulation models assume that a certain proportion of recruits are globally dispersed across an entire 50-ha forest plot on Barro Colorado Island (BCI). This assumption is biologically unrealistic because the vast majority of species at BCI exhibit dispersal limitation (*3*), spatially clumped distributions (*4*, *5*), and an average mean dispersal distance of 28 m (*6*). Although such distances allow some recruits to disperse quite far from parent trees, recruit density is still greatest around the parent (*7*, *8*). That is why analyses that assume global dispersal, as in Hülsmann and Hartig, underestimate or fail to detect CNDD when it is actually present (*8*) (Fig. 1). Moreover, the assumption of global dispersal results in recruits of rare species being farther from adults than recruits of common species ($r = -0.43$, $P < 0.0001$, $N = 187$ species in the BCI data set), mimicking the same pattern that would be produced if CNDD were stronger for rare species. Even where Hülsmann and Hartig relax the assumption of completely global dispersal, a large proportion of recruits are still randomly dispersed across the entire plot (e.g., 50% directly under parent tree, 50% globally dispersed; Fig. 1). Moreover, these empirically unjustifiable assumptions about dispersal permeate their simulations of other processes (i.e., the authors use some degree of global dispersal for all other panels in their figure 1). Thus, the unrealistic assumption of global dispersal largely generates the biases that Hülsmann and Hartig claim to observe.

Hülsmann and Hartig suggest that our use of an offset value introduced bias in our CNDD estimates. We applied an offset value to retain recruits located in 20 m × 20 m quadrats without a conspecific adult in the calculation of CNDD (*1*). Retention of all recruits was important because, in many cases, recruits in quadrats without adults

---

[1]Tyson Research Center, Washington University, St. Louis, MO, USA. [2]Department of Biology, Washington University, St. Louis, MO, USA. [3]Center for Conservation and Sustainability, Smithsonian Conservation Biology Institute, National Zoological Park, Washington, DC, USA. [4]Hydrological-Ecological Interactions Branch, Earth System Processes Division, Water Mission Area, U.S. Geological Survey, Reston, VA, USA. [5]Ecology Laboratory, BIOTEC, National Science and Technology Development Agency, Science Park, Pathum Thani, Thailand. [6]Institute of Molecular Biosciences, Mahidol University, Salaya, Nakhon, Pathom, Thailand. [7]Research Office, Department of National Parks, Wildlife and Plant Conservation, Bangkok, Thailand. [8]Taiwan Forestry Research Institute, Taipei 10066, Taiwan. [9]Department of Life Science, Tunghai University, Taichung, Taiwan. [10]Department of Botany and Plant Physiology, University of Buea, Buea, Cameroon. [11]Department of Biology, Indiana University, Bloomington, IN, USA. [12]Institute of Pacific Islands Forestry, U.S. Department of Agriculture Forest Service, Hilo, HI, USA. [13]Center for Tropical Forest Science–Forest Global Earth Observatory, Smithsonian Tropical Research Institute, Panama City, Republic of Panama. [14]Department of Botany, National Museum of Natural History, Washington, DC, USA. [15]Wildland Resources Department, Utah State University, Logan, UT, USA. [16]Department of Botany, Faculty of Science, University of Peradeniya, Peradeniya, Sri Lanka. [17]Joint Lab for Biodiversity Conservation, Sun Yat-sen University (SYSU)–University of Alberta, State Key Laboratory of Biocontrol, School of Life Sciences, SYSU, Guangzhou 510275, China. [18]Department of Renewable Resources, University of Alberta, Edmonton, Alberta, Canada. [19]Department of Natural and Applied Sciences, University of Wisconsin, Green Bay, WI, USA. [20]Department of Ecology and Evolutionary Biology, University of California, Los Angeles, CA, USA. [21]Institute of Ecology and Evolutionary Biology, National Taiwan University, Taipei, Taiwan. [22]Department of Forest Ecology, Silva Tarouca Research Institute, Brno, Czech Republic. [23]Biology Department, Utah State University, Logan, UT, USA. [24]Department of Forest Management, College of Forestry and Conservation, University of Montana, Missoula, MT, USA. [25]Center for Tropical Forest Science–Forest Global Earth Observatory, Smithsonian Environmental Research Center, Edgewater, MD, USA. [26]Forest Ecology Group, Smithsonian Environmental Research Center, Edgewater, MD, USA. [27]Conservation Ecology Center, Smithsonian Conservation Biology Institute, National Zoological Park, Front Royal, VA, USA. [28]Institut de Recherche en Ecologie Tropicale/Centre National de la Recherche Scientifique et Technologique, Libreville, Gabon. [29]New Guinea Binatang Research Centre, P.O. Box 604, Madang, Papua New Guinea. [30]Biology Centre, Academy of Sciences of the Czech Republic, Prague, Czech Republic. [31]Faculty of Science, University of South Bohemia, Branisovska 31, Ceske Budejovice 370 05, Czech Republic. [32]Institute of Biology, University of the Philippines Diliman, Quezon City, Philippines. [33]Harvard Forest, Harvard University, Petersham, MA, USA. [34]Department of Biology, University of Hawaii, Hilo, HI, USA. [35]Department of Natural Resources and Environmental Studies, National Dong Hwa University, Hualian, Taiwan. [36]Center for Conservation and Sustainable Development, Missouri Botanical Gardens, St. Louis, MO, USA. [37]School of Biological Sciences, Washington State University, Vancouver, WA, USA. [38]Smithsonian Tropical Research Institute, Balboa, Ancon, Republic of Panama. [39]Department of Plant and Microbial Biology, University of Minnesota, St. Paul, MN, USA. [40]Department of Biology, University of Wisconsin, Green Bay, WI, USA. [41]Institute of Arts and Sciences, Far Eastern University Manila, Manila, Philippines.
***Corresponding author. Email: joe.a.lamanna@gmail.com**

likely came from adults in adjacent quadrats. The mean distance (±SE) between recruits in these quadrats and the nearest conspecific adult was 26.95 ± 0.04 m (24.16 ± 0.19 m and 27.08 ± 0.04 m for temperate and tropical trees, respectively), and 80% were within 36 m of a conspecific adult (31 and 36 m for temperate and tropical trees, respectively). These distances are well within average dispersal kernels for tree species (*6*, *9*). Therefore, we applied an offset value to quadrats with recruits but no conspecific adults, so as to ensure that these recruits remained in the calculation of CNDD and to avoid bias that results from excluding these recruits (*1*, *10*, *11*). However, adding the offset value to all quadrats did not qualitatively change either the relationship between species rarefied richness and CNDD across latitudes ($r = -0.877$, $P < 0.001$) or the latitudinal shift in the relationship between CNDD and species abundance ($r = -0.552$, $P = 0.006$). Moreover, these findings persisted when we used an alternative distance-weighted approach to estimate CNDD that avoids the use of an offset altogether (*11*). Therefore, the main findings of our original paper are robust to the statistical approach used to estimate CNDD.

To further verify that our results reflect changes in CNDD across species and latitudes, we used a null model similar to the model used by Hülsmann and Hartig that simply modifies the assumption of global dispersal. The model, recommended by Wiegand and Moloney (*12*), fixes adult locations and disperses recruits away from adults given some dispersal kernel. Instead of globally dispersing all or a fraction of recruits (as done by Hülsmann and Hartig), we use the best-fitting model for seed dispersal across tropical and temperate forests (Clark's 2dT dispersal kernel) (*6*, *7*, *9*, *12*). As in Hülsmann and Hartig's model, our null model also preserves habitat specificity (adult locations), adult/recruit ratios, and abundances for each species to test whether these processes could have generated artificial CNDD patterns. As suggested by Wiegand and Moloney, we also modeled immigration by treating each forest plot as a torus, allowing recruits dispersing off the plot to immigrate back into the plot from the other side (*12*). We performed two versions of this null model: (i) with every species having a fixed mean dispersal distance of 30 m [approximating the mean across >60 tropical and temperate species (*6*, *7*)] and (ii) with species having allometrically scaled dispersal distances that incorporate intraspecific variation. For the allometric-dispersal model, we modeled inter- and intraspecific variation in dispersal using a recent meta-analysis, which showed that more than half of the total variation in mean dispersal distance across >200 plant species scaled allometrically with their maximum height (*9*). We used this relationship, along with observed variation around this relationship, to simulate a wide range of empirical dispersal distances for each species given its maximum height and to calculate null-expected values of CNDD for each species in our data set (*12*).

The results of this null-model analysis using the distance-weighted measure of adult abundance [described briefly above and presented in detail in (*11*)] are shown in Figs. 2 and 3. The expected values of CNDD were near zero and did not show evidence of strong bias across rare and common species. Results from these null-model analyses supported the conclusions from our original paper (*1*). Moreover, similar results were obtained when we applied the same null-model analysis using the Ricker model with offset approach presented in (*1*). Therefore, although we acknowledge that other processes such as dispersal and habitat specificity may contribute to our estimates of CNDD, null-model results strongly suggest that these processes alone cannot account for observed patterns in CNDD across latitudes and species. Instead, our results are consistent with the idea that density dependence, caused by specialized natural enemies and/or intraspecific competition, strongly contributes to the latitudinal diversity gradient.

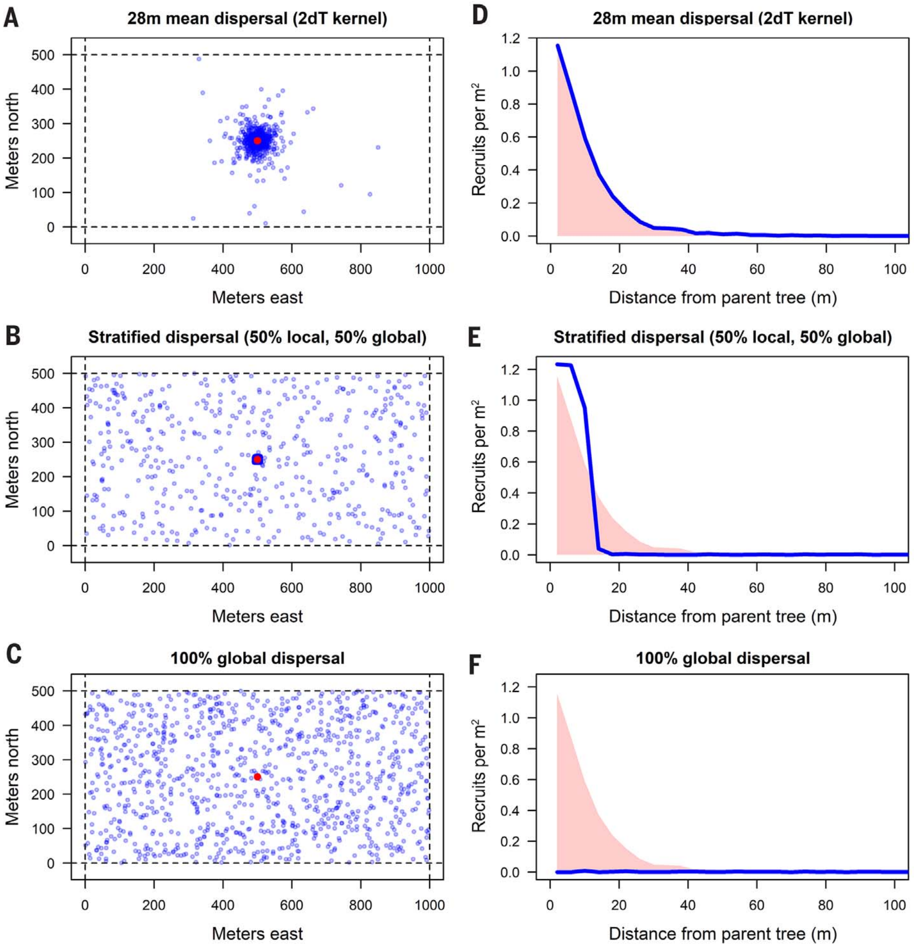


Fig. 1. Simulated spatial patterns of dispersal. (**A** to **C**) Simulated spatial patterns of 1000 recruits (blue points) relative to a parent tree (red point) in a 50-ha plot the same size as the plot at BCI. (**D** to **F**) Recruits per m^2 as a function of distance from the parent. Dispersal patterns were produced by Clark's 2dT dispersal kernel (*7*) at the average observed mean dispersal distance across more than 60 tropical and temperate species (*6*, *7*) [(A) and (D)], by stratified dispersal (used in Hülsmann and Hartig) whereby a certain proportion of recruits are dispersed into the same quadrat as the adult and the rest are globally dispersed across the entire forest plot [(B) and (E)], and by complete global dispersal [(C) and (F)]. Blue lines in (D) to (F) show the expected recruit density for each dispersal kernel [(D), observed mean dispersal from the literature; (E), stratified; (F), global]. The pink shaded areas show the expected recruit density using the observed mean dispersal from the literature [same as in (D)] (*6*, *7*). The assumption of global dispersal is inappropriate because it vastly underestimates expected recruit density near the parent tree [compare blue lines with pink shaded areas in (E) and (F)] and produces a pattern identical to that expected under strong CNDD (*8*). Even stratified dispersal (E) underestimates the density of recruits between 10 and 30 m. Recruits dispersing off the plot were retained by treating the plot as a torus (*12*).

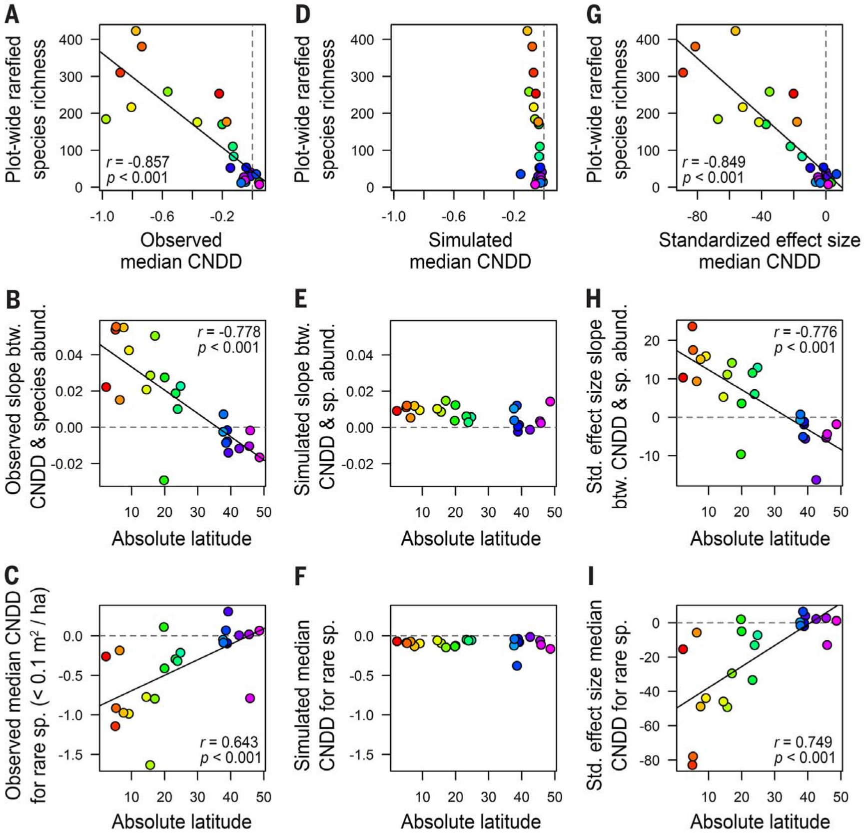


Fig. 2. Results from a null model that incorporates empirically supported values of dispersal and preserves habitat specificity, adult/recruit ratios, and abundances for each species. The null model, recommended by Wiegand and Moloney (*12*), fixes adult locations to preserve habitat specificity of a species and then disperses recruits away from adults. This is the same model used by Hülsmann and Hartig, except that it incorporates empirically supported dispersal values. In this version of the null model, all species were given the same mean dispersal distance of 30 m, based on empirical estimates for more than 60 tropical and temperate species (*6*, *7*). (**A** to **F**) A distance-weighted measure of adult abundance was used to calculate both observed [(A) to (C)] and simulated values [(D) to (F)] of CNDD (*11*). Observed and simulated values are presented on the same scale for comparison. The analysis was performed at the 20 m × 20 m scale. (**G** to **I**) Standardized effect sizes are the observed value minus the mean simulated value from 100 iterations of the null model divided by the standard deviation of the simulated values. These results support the main results in our original paper [figures 1 and 2 in (*1*)]. Colors reflect distance from the equator [see legend of figures 1 and 2 in (*1*)]. Linear fits are shown, along with Spearman rank correlation coefficients and their *P* values. Gray dashed lines represent a median CNDD of zero [(A), (C), (D), (F), (G), and (I)] or a zero slope between CNDD and species abundance [(B), (E), and (H)].

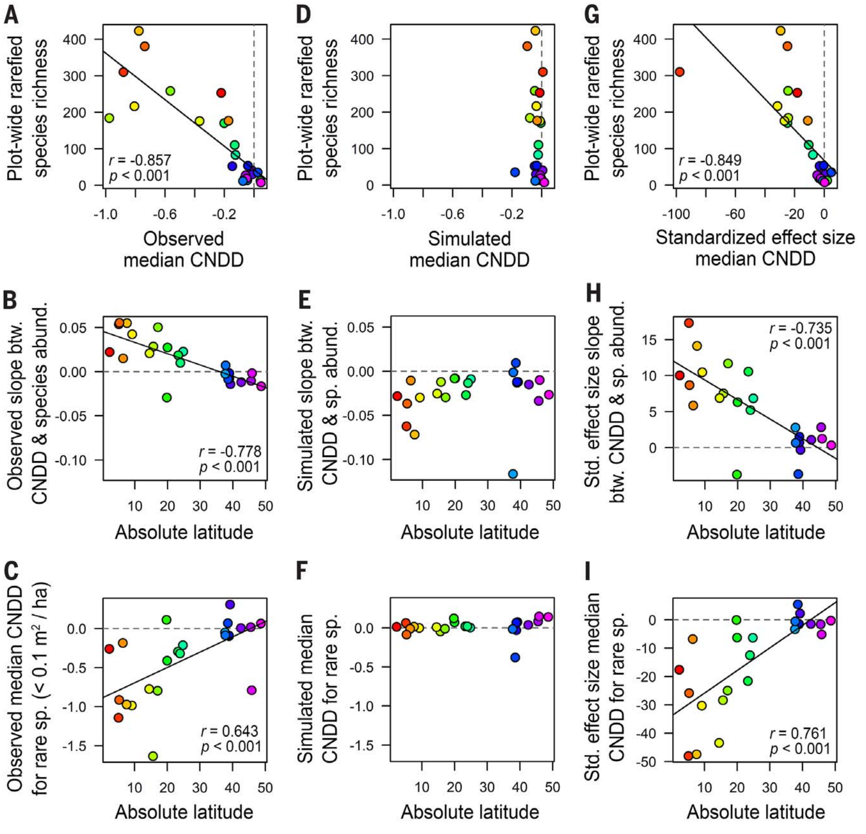


Fig. 3. Results from a null model that incorporates empirically supported interspecific differences in dispersal and preserves habitat specificity, adult/recruit ratios, and abundances for each species. The null model is the same as the model shown in Fig. 2, except that all species are given a mean dispersal distance that scales allometrically with their maximum height and incorporates variation in dispersal within species (*9*).

The null-model results we report here support the main conclusions of our original paper: CNDD was (i) stronger in tropical than in temperate forests and (ii) stronger for rare species than for common species in the tropics, but equivalent or weaker for rare species than for common species in temperate latitudes (*1*). The latter finding is supported by both observational and experimental studies of tropical seedling growth and survival as well as temperate sapling recruitment (*13*–*15*). Hülsmann and Hartig raise the common misconception that strong negative density dependence for rare species should always increase their likelihood of local extinction. This is true if stabilizing forces such as CNDD are lacking in rare species. However, theory shows that strong CNDD can promote the persistence of rare species in communities, even when CNDD is relatively weaker for common species (*16*), as we observed in tropical forests (*1*). Under these conditions, rare species that are strongly stabilized by interactions with specialized natural enemies or other conspecifics (i.e., strong CNDD) should be less susceptible to local extinction from ecological drift than rare species that are not stabilized (*1*, *16*). Our results not only are consistent with this idea, they suggest fundamental differences in the nature of local stabilizing interactions that contribute to the maintenance of species diversity across temperate and tropical latitudes.

REFERENCES AND NOTES

1. J. A. LaManna *et al.*, *Science* **356**, 1389–1392 (2017).
2. L. Hülsmann, F. Hartig, *Science* **360**, eaar2435 (2018).
3. S. P. Hubbell *et al.*, *Science* **283**, 554–557 (1999).
4. R. Condit *et al.*, *Science* **288**, 1414–1418 (2000).
5. K. E. Harms, R. Condit, S. P. Hubbell, R. B. Foster, *J. Ecol.* **89**, 947–959 (2001).
6. H. C. Muller-Landau, S. J. Wright, O. Calderón, R. Condit, S. P. Hubbell, *J. Ecol.* **96**, 653–667 (2008).
7. J. S. Clark, M. Silman, R. Kern, E. Macklin, J. HilleRisLambers, *Ecology* **80**, 1475–1494 (1999).
8. J. H. Lambers, J. S. Clark, B. Beckage, J. HilleRisLambers, *Nature* **417**, 732–735 (2002).
9. F. J. Thomson, A. T. Moles, T. D. Auld, R. T. Kingsford, *J. Ecol.* **99**, 1299–1307 (2011).
10. I. A. Dickie, J. M. Hurst, P. J. Bellingham, *Science* **338**, 469 (2012).
11. J. A. LaManna *et al.*, *Science* **360**, eaar5245 (2018).
12. T. Wiegand, K. A. Moloney, *Handbook of Spatial Point-Pattern Analysis in Ecology* (CRC Press, 2014).
13. L. S. Comita, H. C. Muller-Landau, S. Aguilar, S. P. Hubbell, *Science* **329**, 330–332 (2010).
14. S. A. Mangan *et al.*, *Nature* **466**, 752–755 (2010).
15. K. Zhu, C. W. Woodall, J. V. Monteiro, J. S. Clark, *Ecology* **96**, 2319–2327 (2015).
16. G. Yenni, P. B. Adler, S. K. Ernest, *Ecology* **93**, 456–461 (2012).

ACKNOWLEDGMENTS

We thank an anonymous reviewer for suggesting the distance-weighted approach. We also thank everyone involved in the collection of the vast quantity of data in the CTFS-ForestGEO network; see table S20 in our original paper for site-specific acknowledgments. Supported by NSF grant DEB-1545761 (S.J.D.), NSF grants DEB-1256788 and DEB-1557094 (J.A.M.), NSF grant DEB-1257989 (S.A.M.), and the Tyson Research Center. We declare no conflicts of interest. The data and R code for analyses are available at the Smithsonian Institution's ForestGEO database portal: https://forestgeo.si.edu/explore-data and https://forestgeo.si.edu/plant-diversity-increases-strength-negative-density-dependence-global-scale.

15 November 2017; accepted 18 April 2018
10.1126/science.aar3824

TECHNICAL COMMENT

FOREST ECOLOGY

Comment on "Plant diversity increases with the strength of negative density dependence at the global scale"

Ryan A. Chisholm[1,2]* **and Tak Fung**[1]

LaManna *et al*. (Reports, 30 June 2017, p. 1389) found higher conspecific negative density dependence in tree communities at lower latitudes, yielding a possible mechanistic explanation for the latitudinal diversity gradient. We show that their results are artifacts of a selective data transformation and a forced zero intercept in their fitted model. A corrected analysis shows no latitudinal trend.

In ecological systems, conspecific negative density dependence (CNDD) occurs when individuals are more negatively affected by neighboring conspecific individuals than by neighboring heterospecific individuals (*1*–*4*). Communities with stronger CNDD should, all else equal, have higher diversity, because CNDD prevents any one species from becoming too abundant. Accordingly, it is theoretically possible that global variation in CNDD drives global variation in diversity (*5*). In particular, the latitudinal diversity gradient, whereby diversity decreases with increasing latitude, could be driven by a concomitant weakening in CNDD with increasing latitude. Thus, there is scientific interest in testing the hypothesis that species in tropical high-diversity sites are more susceptible to conspecific competition than are their temperate counterparts.

LaManna *et al*. (*6*) tested this hypothesis by analyzing the relationship between sapling density and adult density for nearly 2.4 million trees of more than 3000 species across 24 forest plots worldwide. For each species at each site, they fitted the Ricker population dynamics model (*7*) to static counts of adults and saplings within 10 m × 10 m or 20 m × 20 m quadrats. They found that the fitted CNDD parameter of the Ricker model was, on average, substantially more negative in high-diversity, low-latitude forests, providing apparent support for the hypothesis that CNDD explains high tropical tree diversity.

However, the analysis of LaManna *et al*. has two flaws. First, they transformed a subset of their data prior to analysis while leaving the remainder of the data untransformed, which biased the model results. For each species, they added +0.1 to the adult abundance in any quadrat that had nonzero sapling abundance and zero adult abundance, but did not transform adult abundance values in other quadrats. The problem LaManna *et al*. were trying to address with this selective transformation was that the Ricker model has a zero intercept and therefore cannot account for observations that have nonzero sapling abundance and zero adult abundance. However, the selective transformation they used is not valid (and renders goodness-of-fit statistics meaningless). Indeed, we cannot think of any circumstances under which it is valid to transform only a subset of data prior to analysis.

How did the selective transformation used by LaManna *et al*. introduce a specific bias toward

---

[1]Department of Biological Sciences, National University of Singapore, 117558 Singapore. [2]Smithsonian Tropical Research Institute, Balboa, Ancón, Republic of Panamá.
*Corresponding author. Email: ryan.chis@gmail.com

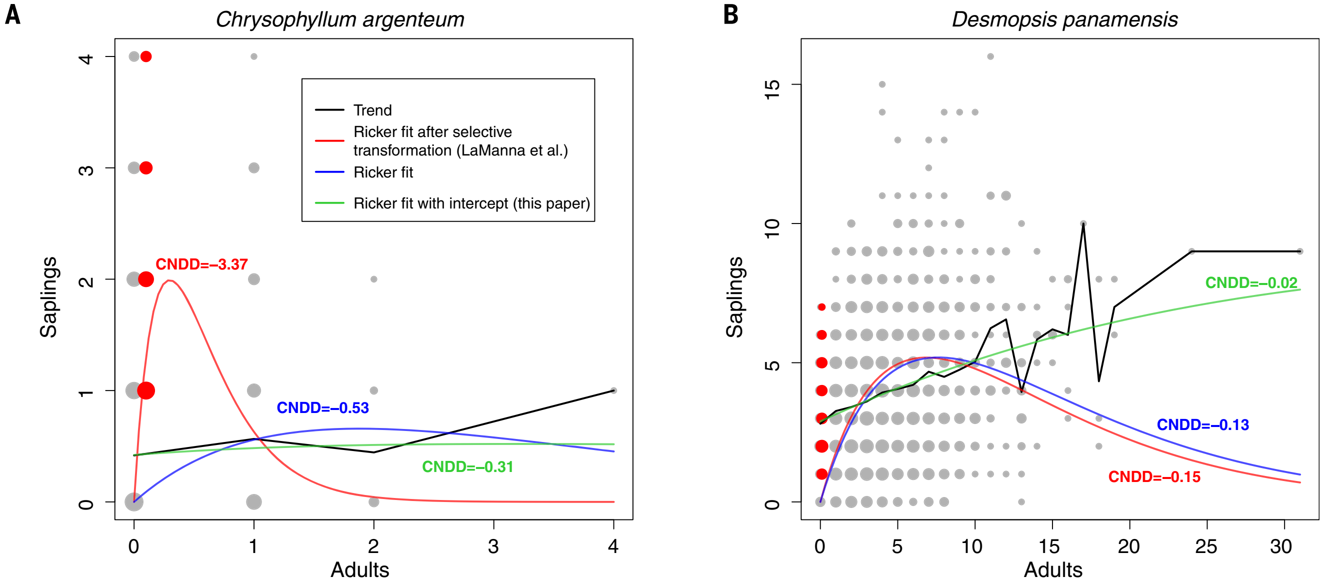


Fig. 1. The selective transformation and forced zero intercept used by LaManna *et al*. inflate CNDD, especially for rare species. (**A** and **B**) Sapling abundance versus adult abundance for two illustrative species at the Barro Colorado Island plot in Panama, *Chrysophyllum argenteum* (a rare species) and *Desmopsis panamensis* (a common species), across quadrats [gray points, with sizes proportional to log(number of quadrats + 1)]. The selective transformation applied by LaManna *et al*. (red points) leads to a fitted model (red curves) that has an inflated intrinsic growth rate and inflated CNDD relative to the case with no selective transformation (blue curves). The bias is more acute for rarer species (A). Even without the selective transformation, the Ricker model (blue curves) provides a poor fit to the average trend (black lines) because of the forced zero intercept; a Ricker model with an intercept term characterizes the data more accurately (green curves; $F_{1,1245} = 5372.5$, $P = 6 \times 10^{-454}$ for *C. argenteum*; $F_{1,1245} = 1800.1$, $P = 5 \times 10^{-244}$ for *D. panamensis*).

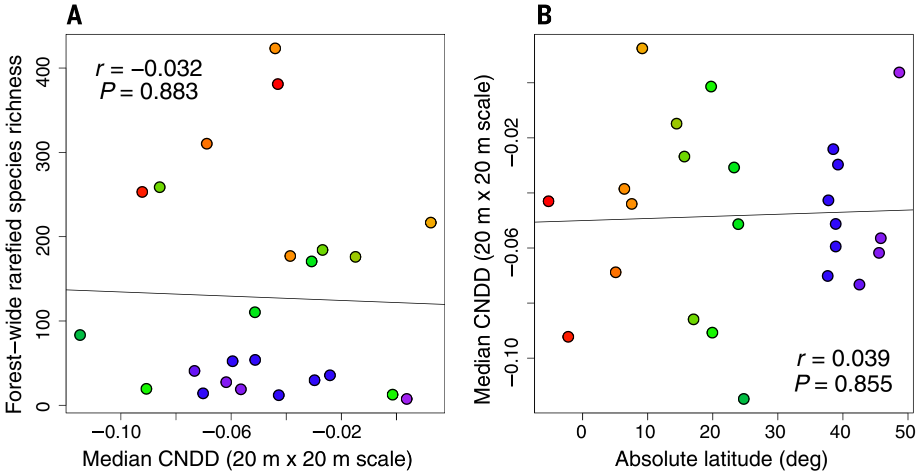


Fig. 2. Species richness exhibits no systematic relationship with the strength of CNDD across tropical and temperate forests. (**A**) Species richness versus estimated CNDD (median over all species in a plot) for the 24 forest plots of LaManna *et al.* (points) with fitted least-squares regression line. (**B**) Estimated CNDD versus latitude. The methods and data here are identical to those used to produce the results of LaManna *et al.*, except that we have not applied a selective data transformation prior to analysis and we have used an intercept in the fitted Ricker model (see text and Fig. 1; the graphs here should be directly compared with figures 1E and 1C, respectively, in LaManna *et al.*). Spearman rank correlation coefficients (r) with P values are shown.

higher CNDD in tropical species-rich forests? The selective transformation moved data points with nonzero sapling abundance and zero adult abundance off the vertical axis, which caused the fitted Ricker model to estimate an artificially high intrinsic growth rate and consequently artificially strong negative density dependence to compensate (Fig. 1). This effect was more pronounced for rare species, which have a higher proportion of presence quadrats with nonzero sapling abundance and zero adult abundance, and thus a higher proportion of transformed data points (e.g., Fig. 1A; Spearman $r = -0.25$, $P = 2 \times 10^{-52}$ for the relationship between species abundance and fraction of presence quadrats that were transformed). This then led to more strongly biased CNDD estimates for rare species. Because species-rich sites have more rare species than species-poor sites, the fraction of data points affected by the selective transformation of LaManna *et al.* was positively correlated with site richness (Spearman $r = 0.95$, $P = 2 \times 10^{-6}$). For example, at the most species-poor site, Zofin, the average abundance is 212.7 trees/ha and LaManna *et al.* transformed 7% of presence quadrats, whereas at the most species-rich site, Khao Chong, the average abundance is 6.4 trees/ha and they transformed 50% of presence quadrats. The higher fraction of selectively transformed points at species-rich sites led to a stronger bias in CNDD estimates at these sites and ultimately inflated the latitudinal gradient in estimated CNDD. LaManna *et al.* also investigated shifts of +0.01 and +0.001 (their table S19), but these analyses suffer from the same rare-species bias.

The second problem in the analysis of LaManna *et al.* was that the Ricker model was forced to have a zero intercept, when the data do not suggest this (e.g., Fig. 1) and the resulting residual plots are skewed (not shown by LaManna *et al.*). We fitted a Ricker model with a nonzero intercept to the same forest data and obtained better fits than with a zero intercept [as judged by F tests: $P < 0.05$ for 99.4% of species; $P < 3 \times 10^{-5}$ (Bonferroni-adjusted threshold) for 98.3% of species]. Statistically, the nonzero intercept is a valid solution (unlike the selective transformation of LaManna *et al.*) to the problem of quadrats with nonzero sapling abundance and zero adult abundance. Biologically, the nonzero intercept for each species can be interpreted as the central tendency of a stochastic immigration process that does not vary systematically with the number of adults in a quadrat. With our improved model, we found that the median estimated CNDD over all species at all plots was lower by about a factor of 30 than in LaManna *et al.*, the relationship between plot richness and estimated CNDD disappeared, and the latitudinal pattern in estimated CNDD also disappeared (Fig. 2).

LaManna *et al.* claimed that their results were robust because qualitatively similar patterns arose when they excluded quadrats with nonzero sapling abundance and zero adult abundance. But doing this remedied only the problem of selective transformation and not the problem of using a zero intercept. LaManna *et al.* also claimed that their results were robust because qualitatively similar patterns arose when they fitted a model with a different functional form: the "offset-power model" presented in their supplementary materials. However, the offset-power model also suffers from fatal statistical flaws. The offset-power model of LaManna *et al.*, as shown by their own simulations, estimated strong CNDD even when there was none (figure S11B in LaManna *et al.* shows that the parameter r is estimated at less than 0.4 on average when the true value is 1.0).

Overall, we found that the main results of LaManna *et al.* are artifacts of a selective transformation of the data prior to analysis and a forced zero intercept not suggested by the data. Our corrected analysis yields much weaker CNDD overall and no global CNDD gradient with species richness or latitude. We recommend that future studies test the robustness of our conclusion by using models that are fully spatial, dynamic, and stochastic.

REFERENCES AND NOTES

1. J. N. Klironomos, *Nature* **417**, 67–70 (2002).
2. S. A. Mangan *et al.*, *Nature* **466**, 752–755 (2010).
3. L. S. Comita, H. C. Muller-Landau, S. Aguilar, S. P. Hubbell, *Science* **329**, 330–332 (2010).
4. R. A. Chisholm, H. Muller-Landau, *Theor. Ecol.* **4**, 241–253 (2011).
5. L. S. Comita *et al.*, *J. Ecol.* **102**, 845–856 (2014).
6. J. A. LaManna *et al.*, *Science* **356**, 1389–1392 (2017).
7. W. E. Ricker, *J. Fish. Res. Board Can.* **11**, 559–623 (1954).

ACKNOWLEDGMENTS

We thank M. Detto, M. Visser, H. Muller-Landau, D. Wardle, M. Luskin, and I. Dickie for helpful discussion and comments on the manuscript. **Funding:** Supported by McDonnell Foundation grant 220020470. **Author contributions:** R.A.C. conceived the study. Both authors performed analyses and wrote the paper. **Competing interests:** None. **Data and materials availability:** The data used here are the same as in LaManna *et al.* (6) and are available from those authors upon request.

10 November 2017; accepted 18 April 2018
10.1126/science.aar4685

TECHNICAL RESPONSE

FOREST ECOLOGY

Response to Comment on “Plant diversity increases with the strength of negative density dependence at the global scale”

**Joseph A. LaManna,[1,2]* Scott A. Mangan,[2] Alfonso Alonso,[3] Norman A. Bourg,[4]
Warren Y. Brockelman,[5,6] Sarayudh Bunyavejchewin,[7] Li-Wan Chang,[8]
Jyh-Min Chiang,[9] George B. Chuyong,[10] Keith Clay,[11] Susan Cordell,[12]
Stuart J. Davies,[13,14] Tucker J. Furniss,[15] Christian P. Giardina,[12]
I. A. U. Nimal Gunatilleke,[16] C. V. Savitri Gunatilleke,[16] Fangliang He,[17,18]
Robert W. Howe,[19] Stephen P. Hubbell,[20] Chang-Fu Hsieh,[21]
Faith M. Inman-Narahari,[12] David Janík,[22] Daniel J. Johnson,[23] David Kenfack,[13,14]
Lisa Korte,[3] Kamil Král,[22] Andrew J. Larson,[24] James A. Lutz,[15]
Sean M. McMahon,[25,26] William J. McShea,[27] Hervé R. Memiaghe,[28]
Anuttara Nathalang,[5] Vojtech Novotny,[29,30,31] Perry S. Ong,[32] David A. Orwig,[33]
Rebecca Ostertag,[34] Geoffrey G. Parker,[26] Richard P. Phillips,[11] Lawren Sack,[20]
I-Fang Sun,[35] J. Sebastián Tello,[36] Duncan W. Thomas,[37] Benjamin L. Turner,[38]
Dilys M. Vela Díaz,[2] Tomáš Vrška,[22] George D. Weiblen,[39] Amy Wolf,[19,40]
Sandra Yap,[41] Jonathan A. Myers[1,2]**

Chisholm and Fung claim that our method of estimating conspecific negative density dependence (CNDD) in recruitment is systematically biased, and present an alternative method that shows no latitudinal pattern in CNDD. We demonstrate that their approach produces strongly biased estimates of CNDD, explaining why they do not detect a latitudinal pattern. We also address their methodological concerns using an alternative distance-weighted approach, which supports our original findings of a latitudinal gradient in CNDD and a latitudinal shift in the relationship between CNDD and species abundance.

Chisholm and Fung (*1*) claim that our statistical approach (*2*) produces CNDD estimates that are systematically biased, causing rare species to always exhibit strong CNDD. In particular, they argue that our CNDD estimates for rare species are biased because an offset value was added to retain recruits ($y > 0$ sapling density) located in 20 m × 20 m quadrats without a conspecific adult ($x = 0$ adult density). Retention of all recruits was important because our data demonstrate that recruits in quadrats without adults likely came from adults in adjacent quadrats (*3*). To avoid bias that results from excluding these recruits (*4*), we therefore added a minimum adult density of 0.1 (offset value) to quadrats containing recruits but no conspecific adults. This approach allowed us to include all quadrats and meet the assumptions of a conventional Ricker model that fixes the intercept at the origin (*5*). This is what Chisholm and Fung refer to as “selective data transformation.”

Three lines of evidence refute Chisholm and Fung’s claim. First, if their claim is correct, then our approach should have always produced strong estimates of CNDD for rare species regardless of their latitude. Yet our estimates of CNDD vary substantially among rare species when they are matched for abundance [figure 2D in (*2*); see below]. Moreover, median CNDD for rare species differed across latitudes, with rare tropical species having stronger median CNDD than rare temperate species [figure 2C in (*2*)]. Second, adding the offset value to all quadrats does not qualitatively change either the relationship between species rarefied richness and CNDD across latitudes ($r = -0.877$, $P < 0.001$) or the latitudinal shift in the relationship between CNDD and species abundance ($r = -0.552$, $P = 0.006$). Third, these findings persist using an alternative distance-weighted approach to estimate CNDD that avoids the use of an offset altogether (described in detail below).

In contrast, Chisholm and Fung offer an alternative approach to retain all recruits by including an additional additive intercept in the Ricker model (Fig. 1). However, their modification of the Ricker model and inferences made from it are flawed for two reasons. First, the extra intercept in Chisholm and Fung’s model is confounded with the CNDD parameter (and recruitment parameter, or r) such that the same reduction in recruitment with increasing adult densities (i.e., CNDD) can be described by a wide range of CNDD-parameter values (Fig. 1C). This problem is similar to another flawed attempt to add a third parameter to the Ricker model, which has been demonstrated to produce strongly biased estimates of CNDD (*6*). Second, the Ricker model already has an intercept that measures density-independent recruitment (r) (Fig. 1B). The addition of a second intercept removes a constant proportion of recruits from the calculation of the CNDD parameter in the Ricker model, thereby assuming that these recruits are immune to the influence of neighboring adults. This assumption is flawed because all recruits (even those that immigrate from outside the forest plot) are subject to potential effects from intraspecific competition and shared natural enemies associated with neighboring adult trees (*7*, *8*). Thus, the addition of an extra-intercept parameter suggested by Chisholm and Fung fundamentally alters the Ricker model so that it no longer measures the biological process of interest: CNDD.

Benchmark tests, which quantify a model’s ability to recover known parameter values, further demonstrate that Chisholm and Fung’s extra-intercept Ricker model produces severely biased estimates of CNDD. Building on benchmark tests provided in our original paper (*2*), we simulated data with known values of CNDD across the range of recruitment and mean adult abundances observed in the data while incorporating uncertainty due to immigration and random error (*2*, *9*). We then tested the ability of the extra-intercept model to recover these known CNDD values. CNDD estimates from the extra-intercept model were severely biased, imprecise, and centered at zero, regardless of the known CNDD value (Fig. 2, A and C). Indeed, Chisholm and Fung state that their approach resulted in CNDD being generally weak across species, which contradicts previous experimental demonstrations of strong CNDD for several tropical and temperate species in our study (*10*–*12*). This explains why they show no patterns in CNDD across species or latitudes. Moreover, estimated values of Chisholm and Fung’s extra-intercept parameter were not associated with known values of immigration in the benchmark tests ($r = 0.010$, $P = 0.089$), which suggests that the extra intercept does not reflect immigration as they claim.

Inspired by Chisholm and Fung’s concern, we used an alternative approach to estimate CNDD that avoids the use of an offset. This approach yields relationships between CNDD and latitude, and between CNDD and species abundance, nearly identical to those in our original paper. This approach uses a distance-weighted measure of adult abundance for each quadrat, heavily weighting the number of conspecific adults in that quadrat but also incorporating (with lesser weights) the number of conspecific adults in neighboring quadrats (Fig. 3). To weight adult abundances, we used Clark’s 2dT dispersal kernel because it is well supported in previous studies of seed dispersal and reflects the probability that a sapling recruit in a focal quadrat came from an adult at a given distance (*13*, *14*). This kernel also reflects the decreasing influence with distance that an adult may have on recruits in a focal quadrat via intraspecific competition or shared natural enemies (*7*, *8*). We used a mean dispersal-distance parameter of 30 m for the 2dT kernel, which is the mean dispersal distance estimated across more than 60 species in two of the largest studies of seed dispersal in tropical and temperate forests (*13*, *14*) (Fig. 3A). The distance-weighted approach yields only nonzero values of adult abundance, reflecting the nonzero probability that recruits from any adult can potentially disperse to any quadrat in these 4- to 50-ha forest plots, and thus negates the need for an offset value. Moreover, because there are only nonzero values of adult abundance (no data at $x = 0$), this approach meets the assumptions of a conventional Ricker model that fixes the intercept at the origin (*5*).

Results using this distance-weighted approach support our original conclusions that CNDD is stronger in tropical than in temperate forests and that the relationship between CNDD and species abundance changes with latitude (Fig. 4). Null-model analyses conducted using the distance-weighted approach further demonstrate that these results are robust to the influence of other processes that might affect the spatial distribution of recruits relative to adults (e.g., habitat specificity, adult/recruit ratios, species abundances) (*3*). Moreover, benchmark tests that incorporate uncertainty due to immigration and random error (described above) demonstrate that the distance-weighted approach accurately recovers known values of CNDD (Fig. 2, B and D). In summary, regardless of whether we use (i) the distance-weighted approach (which requires no offset value), (ii) the Ricker model with an offset applied to quadrats with recruits but no adults (*2*), (iii) the Ricker model with an offset applied to all quadrats (described above), or (iv) an offset-power model (*2*), we find the same qualitative results as those presented in our original paper (*2*).

In conclusion, we demonstrate that our key findings are not an artifact of systematic bias. We also demonstrate that the alternative approach proposed by Chisholm and Fung fundamentally alters the Ricker model so that it no longer measures CNDD accurately or precisely, explaining

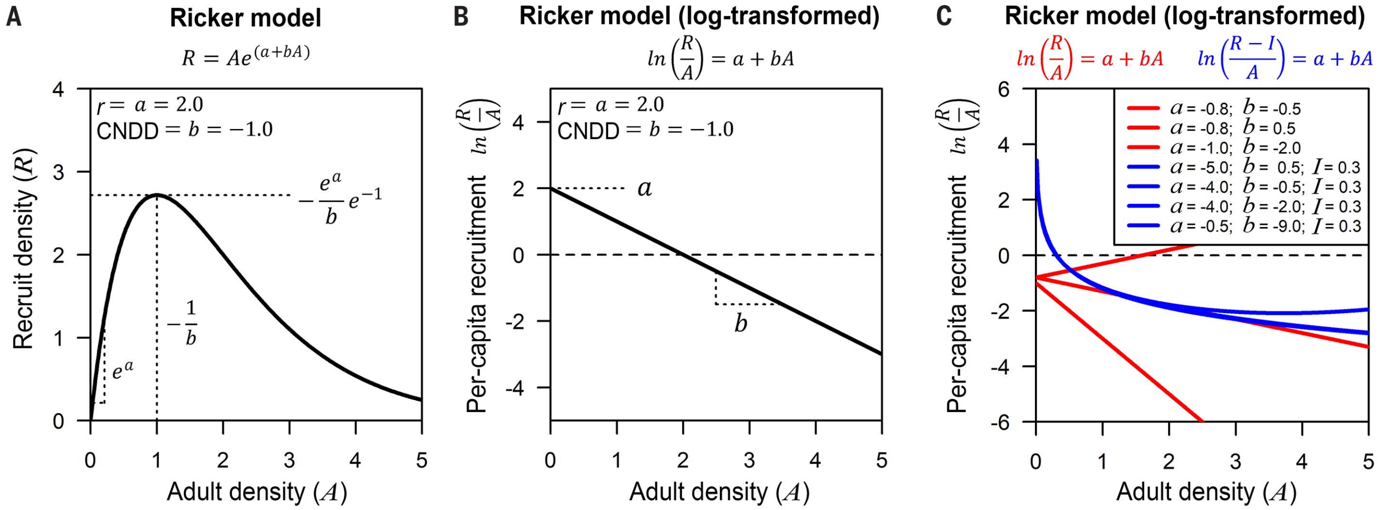


Fig. 1. Basic parameterization of the Ricker model and comparison to the model proposed by Chisholm and Fung. (**A** and **B**) The Ricker model measures linear changes in log-transformed per capita recruitment with increasing adult densities (*5*). The intercept, *a*, measures density-independent per capita recruitment (*r*), and the slope, *b*, measures density-dependent recruitment (CNDD). (**C**) Comparison of the Ricker model (red) to the model suggested by Chisholm and Fung (blue), which adds an intercept, *I*, to the right side of the equation in (A). *I* is mathematically confounded with *a* and *b*, similar to other attempts that add a third parameter to the Ricker model (*6*), such that a wide range values of *a* and *b* can describe the same density dependence in recruitment. Hence, the five overlapping blue curves shown in (C) can be described by a wide range of *a* and *b* values. That is not the case with the conventional Ricker model (red lines).

---

[1]Tyson Research Center, Washington University, St. Louis, MO, USA. [2]Department of Biology, Washington University, St. Louis, MO, USA. [3]Center for Conservation and Sustainability, Smithsonian Conservation Biology Institute, National Zoological Park, Washington, DC, USA. [4]Hydrological-Ecological Interactions Branch, Earth System Processes Division, Water Mission Area, U.S. Geological Survey, Reston, VA, USA. [5]Ecology Laboratory, BIOTEC, National Science and Technology Development Agency, Science Park, Pathum Thani, Thailand. [6]Institute of Molecular Biosciences, Mahidol University, Salaya, Nakhon, Pathom, Thailand. [7]Research Office, Department of National Parks, Wildlife and Plant Conservation, Bangkok, Thailand. [8]Taiwan Forestry Research Institute, Taipei 10066, Taiwan. [9]Department of Life Science, Tunghai University, Taichung, Taiwan. [10]Department of Botany and Plant Physiology, University of Buea, Buea, Cameroon. [11]Department of Biology, Indiana University, Bloomington, IN, USA. [12]Institute of Pacific Islands Forestry, U.S. Department of Agriculture Forest Service, Hilo, HI, USA. [13]Center for Tropical Forest Science–Forest Global Earth Observatory, Smithsonian Tropical Research Institute, Panama City, Republic of Panama. [14]Department of Botany, National Museum of Natural History, Washington, DC, USA. [15]Wildland Resources Department, Utah State University, Logan, UT, USA. [16]Department of Botany, Faculty of Science, University of Peradeniya, Peradeniya, Sri Lanka. [17]Joint Lab for Biodiversity Conservation, Sun Yat-sen University (SYSU)–University of Alberta, State Key Laboratory of Biocontrol, School of Life Sciences, SYSU, Guangzhou 510275, China. [18]Department of Renewable Resources, University of Alberta, Edmonton, Alberta, Canada. [19]Department of Natural and Applied Sciences, University of Wisconsin, Green Bay, WI, USA. [20]Department of Ecology and Evolutionary Biology, University of California, Los Angeles, CA, USA. [21]Institute of Ecology and Evolutionary Biology, National Taiwan University, Taipei, Taiwan. [22]Department of Forest Ecology, Silva Tarouca Research Institute, Brno, Czech Republic. [23]Biology Department, Utah State University, Logan, UT, USA. [24]Department of Forest Management, College of Forestry and Conservation, University of Montana, Missoula, MT, USA. [25]Center for Tropical Forest Science–Forest Global Earth Observatory, Smithsonian Environmental Research Center, Edgewater, MD, USA. [26]Forest Ecology Group, Smithsonian Environmental Research Center, Edgewater, MD, USA. [27]Conservation Ecology Center, Smithsonian Conservation Biology Institute, National Zoological Park, Front Royal, VA, USA. [28]Institut de Recherche en Ecologie Tropicale/Centre National de la Recherche Scientifique et Technologique, Libreville, Gabon. [29]New Guinea Binatang Research Centre, P.O. Box 604, Madang, Papua New Guinea. [30]Biology Centre, Academy of Sciences of the Czech Republic, Prague, Czech Republic. [31]Faculty of Science, University of South Bohemia, Branisovska 31, Ceske Budejovice 370 05, Czech Republic. [32]Institute of Biology, University of the Philippines Diliman, Quezon City, Philippines. [33]Harvard Forest, Harvard University, Petersham, MA, USA. [34]Department of Biology, University of Hawaii, Hilo, HI, USA. [35]Department of Natural Resources and Environmental Studies, National Dong Hwa University, Hualian, Taiwan. [36]Center for Conservation and Sustainable Development, Missouri Botanical Gardens, St. Louis, MO, USA. [37]School of Biological Sciences, Washington State University, Vancouver, WA, USA. [38]Smithsonian Tropical Research Institute, Balboa, Ancon, Republic of Panama. [39]Department of Plant and Microbial Biology, University of Minnesota, St. Paul, MN, USA. [40]Department of Biology, University of Wisconsin, Green Bay, WI, USA. [41]Institute of Arts and Sciences, Far Eastern University Manila, Manila, Philippines.

***Corresponding author. Email: joe.a.lamanna@gmail.com**

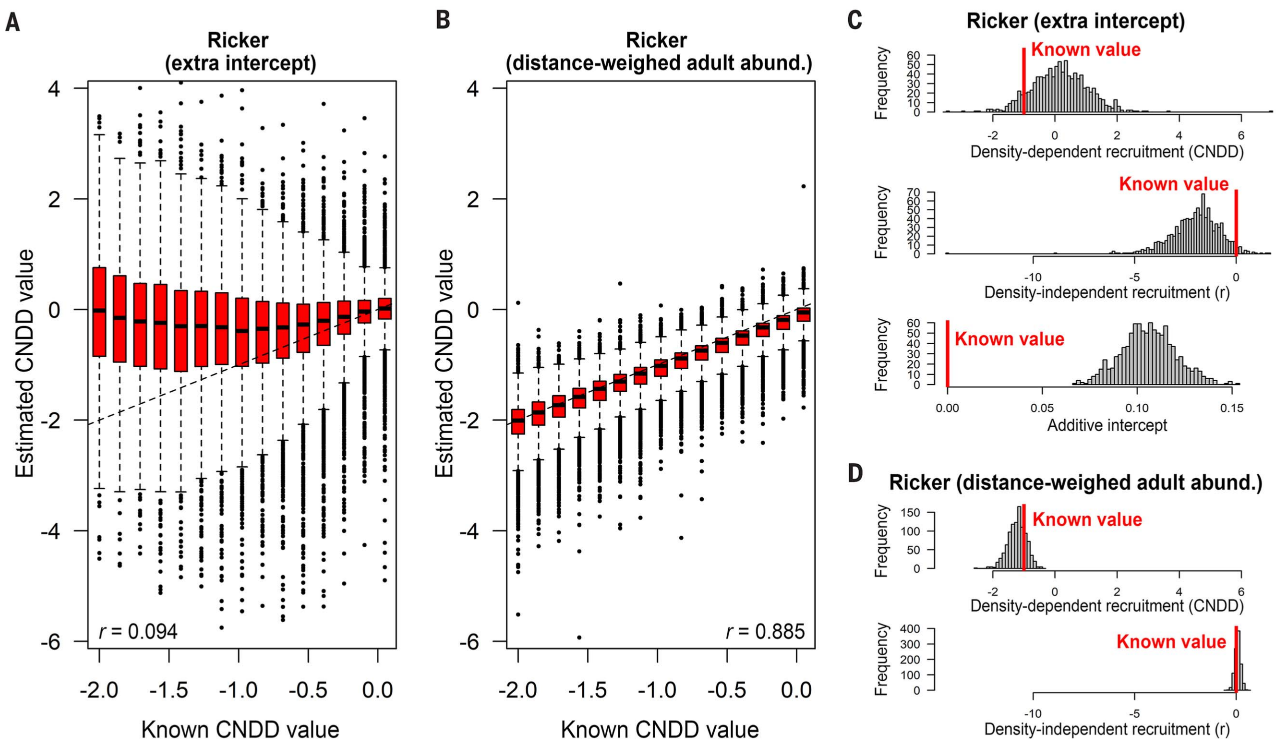


Fig. 2. Benchmark tests that attempt to recover known values of CNDD from simulated data with error. (**A** and **B**) Tests of the ability of the extra-intercept approach proposed by Chisholm and Fung (A) and the distance-weighted adult abundance approach (B) to recover known values of CNDD across a wide range of CNDD, *r*, mean adult abundance, random error, and immigration values. An unbiased approach should produce estimates that fall on the identity line (diagonal black dashed lines). The greater spread of estimated values in (A) relative to (B) reflects the confounded nature of the CNDD and extra-intercept parameters in the model proposed by Chisholm and Fung. (**C** and **D**) One thousand iterations of the same benchmark test at one set of known parameter values (i.e., known values in red). Histograms show the distribution of estimated values from the Ricker model using the extra-intercept approach (C) and the distance-weighted approach (D). The distribution of estimated values should fall around the known value if the model is not biased. The extra-intercept model tends to bias CNDD weak (i.e., closer to zero) and tends to bias *r* low. The extra-intercept model also estimates a wide range of nonzero intercepts, even though these simulated data were generated without an intercept.

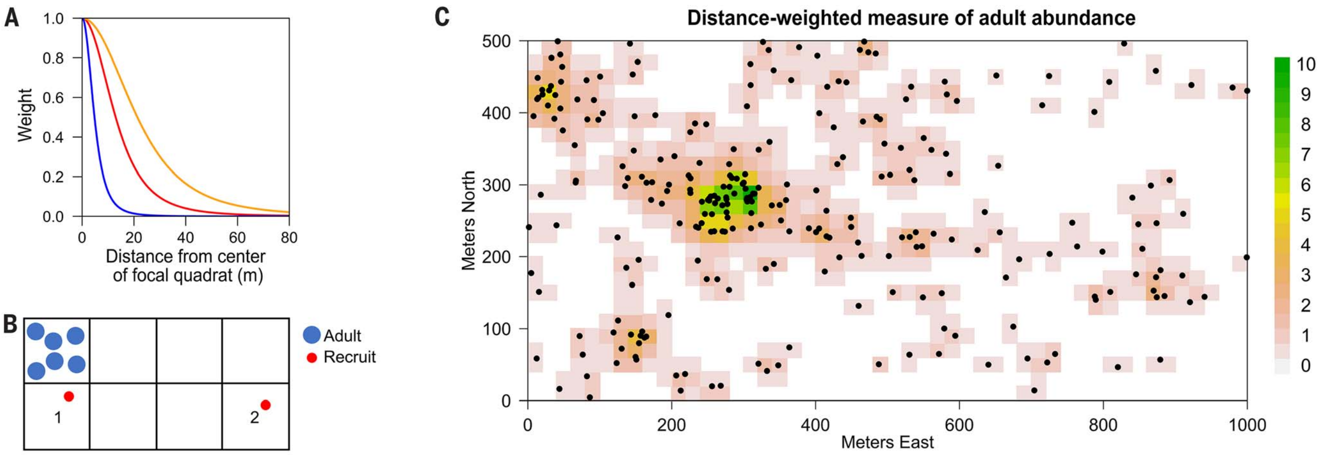


Fig. 3. Distance-weighted measure of adult abundance. (**A**) Distance kernel used to weight adult abundances as a function of distance from a focal quadrat. We used the mean dispersal distance (30 m, red curve) across more than 60 tropical and temperate species estimated from Clark's 2dT kernel (*13*, *14*), but inferences were similar using more extreme values [i.e., 10 m (blue curve), 60 m (orange curve)]. (**B**) The distance-weighted approach is advantageous because it provides information on adults in neighboring quadrats. Whereas both quadrats 1 and 2 have naïve adult abundances of zero, the distance-weighted adult abundance of quadrat 1 would be greater than quadrat 2 because of the presence of six adults in the neighboring quadrat. (**C**) Example of distance-weighted adult abundance values (color scale) for *Cassipourea elliptica* at Barro Colorado Island, Panama (range of distance-weighted adult abundance values is 0.003 to 10.48). Black dots are adult individuals of this species.

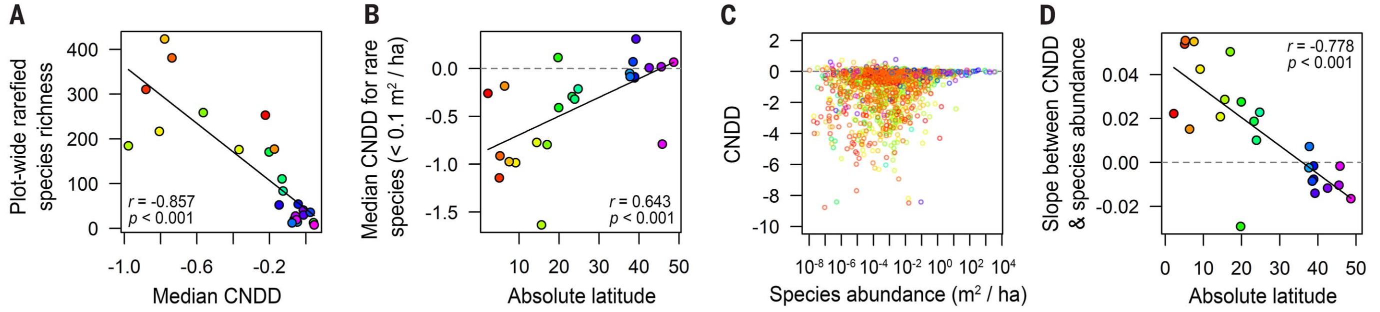


Fig. 4. Results using the Ricker model with distance-weighted measure of adult abundance. (**A**) Species rarefied richness increased with the median strength of CNDD across tropical and temperate forest plots. (**B**) Median CNDD was stronger for rare species in tropical than in temperate forests. (**C**) Estimates of CNDD plotted against species abundance for all species and forests (tropical species in red/orange; temperate species in blue/purple; mid-latitude species in green). (**D**) The slope between CNDD and species abundance decreased with increasing latitude. These results support the main results in our original paper [figures 1 and 2 in (*2*)]. Colors reflect distance from the equator [see captions of figures 1 and 2 in (*2*)]. Results are shown at the 20 m × 20 m scale. Linear fits are shown, along with Spearman rank correlation coefficients and their *P* values.

why their method shows no latitudinal pattern in CNDD. Finally, we show that our key findings—(i) tree species diversity increases with the strength of CNDD across temperate and tropical forests, and (ii) there is a latitudinal shift in the relationship between CNDD and species abundance—are supported using a distance-weighted approach that addresses the methodological concerns raised by Chisholm and Fung.

REFERENCES AND NOTES

1. R. A. Chisholm, T. Fung, *Science* **360**, eaar4685 (2018).
2. J. A. LaManna *et al.*, *Science* **356**, 1389–1392 (2017).
3. J. A. LaManna *et al.*, *Science* **360**, eaar3824 (2018).
4. I. A. Dickie, J. M. Hurst, P. J. Bellingham, *Science* **338**, 469 (2012).
5. B. M. Bolker, *Ecological Models and Data in R* (Princeton Univ. Press, 2008).
6. F. Clark, B. W. Brook, S. Delean, H. R. Akçakaya, C. J. Bradshaw, *Methods Ecol. Evol.* **1**, 253–262 (2010).
7. D. H. Janzen, *Am. Nat.* **104**, 501–528 (1970).
8. J. H. Connell, in *Dynamics of Populations*, P. J. den Boer, G. R. Gradwell, Eds. (Centre for Agricultural Publishing and Documentation, Wageningen, Netherlands, 1971), pp. 298–312.
9. R. P. Freckleton, A. R. Watkinson, R. E. Green, W. J. Sutherland, *J. Anim. Ecol.* **75**, 837–851 (2006).
10. S. A. Mangan *et al.*, *Nature* **466**, 752–755 (2010).
11. L. S. Comita *et al.*, *J. Ecol.* **102**, 845–856 (2014).
12. J. A. Bennett *et al.*, *Science* **355**, 181–184 (2017).
13. J. S. Clark, M. Silman, R. Kern, E. Macklin, J. HilleRisLambers, *Ecology* **80**, 1475–1494 (1999).
14. H. C. Muller-Landau, S. J. Wright, O. Calderón, R. Condit, S. P. Hubbell, *J. Ecol.* **96**, 653–667 (2008).

ACKNOWLEDGMENTS

We thank an anonymous reviewer for suggesting the distance-weighted approach. We also thank everyone involved in the collection of the vast quantity of data in the CTFS-ForestGEO network; see table S20 in our original paper for site-specific acknowledgments. Supported by NSF grant DEB-1545761 (S.J.D.), NSF grants DEB-1256788 and DEB-1557094 (J.A.M.), NSF grant DEB-1257989 (S.A.M.), and the Tyson Research Center. We declare no conflicts of interest.
The data and R code for analyses are available at the Smithsonian Institution's ForestGEO database portal: https://forestgeo.si.edu/explore-data and https://forestgeo.si.edu/plant-diversity-increases-strength-negative-density-dependence-global-scale.

8 December 2017; accepted 18 April 2018
10.1126/science.aar5245

Global health shifts to local experts with global partners

Local health specialists seek international collaborations to fight health emergencies

By **Anne Q. Hoy**

A decade and a half ago, severe acute respiratory syndrome (SARS) took hold in the coastal Chinese province once known as Canton. Within months, the respiratory virus had spread across four continents. The U.S. Centers for Disease Control and Prevention issued its first travel warning. In time, 774 people had died and 8,098 were sickened by the outbreak.

The response of governments and global public health organizations to the world's first known instance of atypical pneumonia has informed responses to more recent global health epidemics, including the Ebola outbreak in West Africa from 2014 to 2016 and the Zika virus in Brazil in 2015.

At the center of the U.S. response to the outbreaks was Jimmy Kolker, now a visiting scholar at the American Association for the Advancement of Science's Center for Science Diplomacy. At the time, he was the assistant secretary for global affairs at the U.S. Health and Human Services Department, a position to which he brought experience as the former ambassador to Burkina Faso from 1999 to 2002 and later to Uganda from 2002 to 2005.

Kolker traced the evolution of global health on 4 May as part of the AAAS-Hitachi Lecture on Science and Society, at AAAS's Washington, D.C., headquarters. The annual lecture series is a decade-long collaboration between AAAS and Hitachi Ltd., the global technology and innovation company.

As the U.S. ambassador to Uganda seeking to implement the President's Emergency Plan for AIDS Relief, Kolker said diplomacy was "a crucial but ill-defined" requirement. Applying policies based on science and evidence in Uganda required weighing uncertainties with limited data, a topic Kolker explored in an essay recently published in *Science & Diplomacy*.

Global health crises require international health institutions, governments, scientists, and diplomats to work together. Yet, this growing but still tiny field of health diplomacy increasingly needs scientists skilled in translating research into actionable policy options—skills AAAS has long worked to hone through multiple science policy, engagement, and diplomacy programs designed to help scientists effectively place research findings into context to ensure that science informs policy. "The data rarely speaks for itself," Kolker said.

The interplay of science and diplomacy takes on many forms. The SARS epidemic in 2003, for instance, underscored the necessity of open public and international communications, a practice the Chinese government initially discounted. After SARS, the CDC embedded disease surveillance and detection experts with Chinese counterparts.

"The system that was set up, the protocols that were put in place meant that when the next potential pandemic, H7N9, broke out, the Chinese response was timely and enabled governments and experts around the world to bring to bear the knowledge needed to control the outbreak," Kolker said.

At the outset of the Zika outbreak in Brazil, the government hesitated to share samples necessary to develop blood bank screening, diagnostics, and vaccines. A U.S. team from the CDC, the National Institutes of Health, and other agencies responded to the Brazilian government's request for an experts' meeting. Kolker, who led the group, said the meeting "changed the dynamic" and produced a plan that gave political approval for Brazil's top health institutions to deal directly with American counterparts to streamline cooperation on Zika research, countermeasures, and field studies.

The unexpected 2014 Ebola outbreak in West Africa highlighted the benefit of an earlier collaboration between the CDC and Ugandan scientists dating from what was then the world's largest Ebola outbreak in 2000. The CDC helped Uganda develop its own world-class Ebola laboratory and response center at the Uganda Virus Research Institute. During the West African outbreak that killed thousands, a single case of Ebola in Uganda was quickly diagnosed, and contacts were traced and isolated.

"There was one death in Uganda," said Kolker. "The methods were in place to deal with the outbreak. We didn't hear about it because Uganda had world-class capacity and did not require outside or emergency help."

Emerging economies, as the Ugandan outcome demonstrates, are not looking for "donors to provide aid;" instead, Kolker said, they want to work as partners with experts to help build First World capacity in their own health systems and institutions.

Against this backdrop, U.S. funding for international health security is being squeezed, Kolker added, including programs designed to assist emerging economies to meet World Health Organization standards to prevent and respond to global health emergencies.

The White House proposed deep cuts to global health programs in

PHOTO: CHRISTIAN KEENAN/STRINGER/GETTY IMAGES

Hong Kong passengers don masks during the 2003 SARS outbreak.

its fiscal 2019 budget proposal, including programs that address HIV/AIDS, malaria, immunizations, and parasitic diseases, said David Parkes, program associate of AAAS's R&D Budget and Policy Program. The president's budget plan would cut these CDC programs alone by a total of $80 million, taking funding levels 16.3% below enacted fiscal 2018 levels.

The fiscal 2019 budget proposal also called for a 36% reduction below enacted fiscal 2018 levels for the U.S. Agency for International Development. It would cut 17% from a CDC infectious disease program that develops tools to stop diseases spread between animals and people, reported the Global Health Technologies Coalition, a group dedicated to advancing deadly disease treatments and diagnostic tools.

Budget constraints and global political trends heighten the need for organizations like AAAS, academic institutions, philanthropies, and the private sector to forge partnerships with emerging economies, Kolker said. Technical partnerships are needed to further expand health care infrastructure and medical expertise. Several AAAS science diplomacy training and public engagement programs contribute to meeting this goal but will be hard put to fill gaps if government commitment is reduced, Kolker added.

In discussing his experience with epidemics, Kolker explained how multidisciplinary and multinational approaches are especially beneficial as researchers in emerging economies increasingly want to select the health experts they partner with and the research they want to pursue. Such collaborations have led research teams to no longer focus exclusively on infectious diseases. Increasingly, they also study chronic diseases such as diabetes, heart disease, and cancer that present growing health burdens in their countries.

"Our health research system has led U.S. institutions to use African institutions as research platforms. This underestimates the ability of Third World institutions to actually identify their own priorities," said Kolker. "In the 21st century, the new paradigm is that lower- and middle-income countries actually want partnerships with the world's best experts to bring their own capacity to First World standards."

AAAS extends science in theological education program

Science in the Seminaries enhances science education across the religious community

By **Anne Q. Hoy**

Building on broad interest generated by a three-year pilot project integrating science into theological education, the American Association for the Advancement of Science is now expanding the initiative to advance understanding of science and technology across the religious community to as many as 35 seminaries over the next 5 years.

A set of seven seminaries in Michigan, Missouri, Indiana, Illinois, and Wisconsin have been selected as the first of four groups to participate in the expanded program over the next 18 months. Three additional groups of seminaries will be chosen to join the program in stages.

"The hope is that the seminary students exposed to enriched classes will find science relevant and interesting to their vocations, and in the future, help them make science a positive force for their congregations and favorably impact the everyday lives of a broad swath of Americans," said Jennifer Wiseman, director of the AAAS Dialogue on Science, Ethics, and Religion (DoSER) program.

Fr. John Kartje, a trained astrophysicist, rector, and president of Mundelein Seminary, a Roman Catholic theological seminary in a northern Chicago suburb, expressed hope that his school's participation in the expansion phase will give students and faculty a "deeper knowledge and better appreciation for the findings and methodologies of scientific research."

"As Mundelein Seminary prepares priests who will be serving in parishes across the country, it is important that they possess a basic foundation in the fundamentals of science and are up to date on current trends in research and discovery," Kartje said. "Such 'scientific fluency' will help them better connect with their parishioners and better integrate their theological expertise into the scientifically savvy culture in which they will minister."

Anna Case-Winters, professor of theology at McCormick Theological Seminary, said the chance to participate in the initiative drew immediate faculty support from the U.S. Presbyterian theological seminary located in Chicago. "We believe that for religious leaders today, capacity for thoughtful interaction with the sciences is not optional—it is essential," Case-Winters said.

Participating seminaries have pledged to incorporate science into at least two of their core courses and to hold at least one campus-wide event over the next 18 months. "We provide science resources and they plug that into the larger context of their programs," said Curtis Baxter, a DoSER program associate. "The seminaries decide on their own how to incorporate the science into courses they already teach."

AAAS will recruit science advisers from nearby research and academic institutions to share knowledge and experience in designing engaging science coursework, assisting theological educators at each seminary to sort out how best to integrate science into courses the seminaries have selected. The program makes available information on advances in science and technology and provides the institutions with access to the *Science* family of journals.

Coursework that seminaries plan to fold into their core classes covers a broad sweep of science, ranging from evolution of the cosmos to genetics and neuroscience. One seminary plans to explore topics such as anxiety, addiction, and dementia to prepare church leaders to better minister to congregants. Some seminaries plan hands-on activities, including a pilgrimage to Israel to study archaeological sites and field trips to laboratories and science museums.

The five other seminaries participating are the Seventh-Day Adventist Theological Seminary of Andrews University in Berrien Springs, Michigan; Nazarene Theological Seminary in Kansas City, Missouri; Bethany Theological Seminary in Richmond, Indiana; Kenrick-Glennon Seminary in St. Louis, Missouri; and the Sacred Heart Seminary and School of Theology in Hales Corners, Wisconsin.

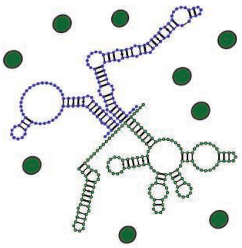
Screeners needed for journalism awards

Scientists from the United States and abroad who will be in the Washington, D.C., area between late August and late September are needed to review the scientific accuracy of entries in the prestigious AAAS Kavli Science Journalism Awards competition. If you can volunteer, please contact Nkongho Beteck (nbeteck@aaas.org) for screening dates and categories.

RESEARCH

Secondary structures of RNA condensates

Langdon et al., *p. 922*



IN *SCIENCE* JOURNALS

Edited by **Stella Hurtley**

A freshwater turtle known as the red-eared slider (*Trachemys scripta elegans*)

NEURODEVELOPMENT

Evolution of the brain

Just how related are reptilian and mammalian brains? Tosches *et al*. used single-cell transcriptomics to study turtle, lizard, mouse, and human brain samples. They assessed how the mammalian six-layered cortex might be derived from the reptilian three-layered cortex. Despite a lack of correspondence between layers, mammalian astrocytes and adult neural stem cells shared evolutionary origins. General classes of interneuron types were represented across the evolutionary span, although subtypes were species-specific. Pieces of the much-folded mammalian hippocampus were represented as adjacent fields in the reptile brains.—PJH

Science, this issue p. 881

PALEOCLIMATE

Forcing the East Asian summer monsoon

What factors have controlled the intensity of the East Asian summer monsoon over the recent geological past? To answer this key question requires a robust proxy for rainfall amounts. Beck *et al*. measured the beryllium isotopic content of loess from China, from which they reconstructed a 550,000-year-long record of rainfall. Rainfall correlated with orbital precession and global variations in ice volume. This finding suggests that the monsoon is governed by low-latitude interhemispheric gradients in solar radiation levels, rather than by high-northern-latitude solar radiation levels as previously suggested. —HJS

Science, this issue p. 877

VALLEYTRONICS

Tracking the spin-valley current

Taking advantage of the electron's spin and valley degrees of freedom requires a method for generating currents of carriers that have a particular spin or come from a particular valley in the electronic structure. Jin *et al*. used a hetero-structure made out of adjacent layers of WSe_2 and WS_2 to create a spin-valley diffusion current without applying an external electric field. Instead, they used circularly polarized laser light to initiate the diffusion and a second laser pulse to image the propagation of the carriers. With long lifetimes and diffusion lengths, the method may be of practical use in future valleytronic devices. —JS

Science, this issue p. 893

COLD MOLECULE PHYSICS

Lighting the way to molecules, one by one

When chemists run reactions, what they are really doing is mixing up an enormous number of reacting partners and then hoping that they collide productively. It is possible to manipulate atoms more deliberately with a scanning tunneling microscope tip, but the process is then confined to a surface. Liu *et al*. directly manipulated individual atoms with light to form single molecules in isolation (see the Perspective by Narevicius). They used optical tweezers of two different colors to selectively steer ultracold sodium (Na) and cesium (Cs) atoms together. A subsequent optical excitation formed NaCs. —JSY

Science, this issue p. 900; see also p. 855

CREDITS: (TOP TO BOTTOM) E. M. LANGDON *ET AL*.; STEPHAN JUNEK/MPI BRAIN RESEARCH

SYNTHETIC BIOLOGY

Using bugs in the gut to detect blood

Bacteria are environmentally resilient and can be engineered to sense various biomolecules. Mimee *et al*. combined biosensor bacteria with a miniaturized wireless readout capsule to produce a minimally invasive device capable of in vivo biosensing in harsh, difficult-to-access environments (see the Perspective by Gibson and Burgell). The device successfully measured gastrointestinal bleeding in pigs. —SYM

Science, this issue p. 915;
see also p. 856

DIABETES

A stimulating therapy for diabetes

In type 2 diabetes, insulin resistance leads to elevated blood glucose and increased risk of cardiovascular disorders. The brain participates in glucose metabolism, but whether and how modulation of brain activity affects systemic blood concentrations of glucose is poorly understood. In diabetic and nondiabetic patients, ter Horst *et al.* found that dopamine release induced by deep brain electrical stimulation of the nucleus accumbens decreased systemic glucose concentrations. Conversely, pharmacological systemic dopamine depletion reduced insulin-mediated blood glucose uptake. —MM

Sci. Transl. Med. **10**, eaar3752 (2018).

EPIDEMIOLOGY

A faster way to detect Zika in mosquitoes

A major challenge in preventing mosquito-borne diseases is providing rapid and affordable tests to identify infected insects. Conventional techniques are often time-consuming and too expensive when analyzing large numbers of mosquitoes. Leveraging a technique known as near-infrared spectroscopy (NIRS), Fernandes *et al.* tested for Zika virus in female *Aedes aegypti* mosquitoes. NIRS distinguished infected from uninfected mosquitoes with up to 99% accuracy 7 days after infection. NIRS is faster by a factor of 18 and cheaper by a factor of 110 than RT-qPCR (quantitative reverse transcription polymerase chain reaction), a technique commonly used for pathogen screening in mosquitoes. —PJB

Sci. Adv. 10.1126/
sciadv.aat0496 (2018).

PHOTO: ISTOCK.COM/NINAMALYNA

MICROBIOME

Bile acids and liver cancer

Liver cancer is a leading cause of cancer-related deaths in the United States. The composition of the gut microbiome influences many human diseases, including liver inflammatory disorders. Ma *et al*. found that commensal gut bacteria can recruit the immune system to control the growth of liver tumors in mice (see the Perspective by Hartmann and Kronenberg). *Clostridium* species modified bile acids to signal liver sinusoidal endothelial cells to produce the chemokine CXCL16. This recruited natural killer T (NKT) immune cells to perform antitumor surveillance of the liver. Growth of both primary and metastatic cancer was reduced by NKT cell–driven killing. —PNK

Science, this issue p. 876;
see also p. 858

MIGRATORY BEHAVIOR

Follow the leader

What role do social dynamics play in guiding collective migrations? Identifying such dynamics requires following individual animals across long migratory distances. Flack *et al*. used GPS tags to follow individual juvenile white storks on their southern migration (see the Perspective by Nevitt). Birds generally fell into two categories: leaders and followers. Leaders sought out areas of thermal uplift, flapped less in transit, and flew farther. Followers followed leaders into thermals but had different trajectories, exhibited greater flapping effort, and flew shorter total distances. —SNV

Science, this issue p. 911;
see also p. 852

IN OTHER JOURNALS

Edited by **Sacha Vignieri**
and **Jesse Smith**

Premature graying of hair is related to immune function.

HAIR COLOR

The roots of gray hair

Gray hair is an inevitable part of aging. Melanocytes are the culprit cells that slow production of the color pigments called melanin, but how and why this occurs with age largely remains a mystery. Harris *et al.* make a link between the immune system and premature graying. They find that the protein MITF (microphthalmia-associated transcription factor), which controls melanocyte stem cell function, also works to trigger melanocyte immune responses. Interferons normally kickstart the immune response to viral and bacterial infection, but when MITF cannot regulate interferon, hair turns gray in mouse models. These findings may shed light on why chronic illness or certain autoimmune disorders can accelerate the graying process. —PNK

PLOS Biol. 10.1371/journal.pbio.2003648 (2018).

ORGANOMETALLICS

Aluminum's breakup with fluoroalkenes

Carbon-fluorine bonds are hard to break. As a result, remediation of fluorocarbon waste streams is an enduring challenge. Bakewell *et al.* explored C–F scission in a variety of fluorinated propene derivatives by using an unusual synthetic aluminum compound. The electron-rich compound, with Al in the +1 oxidation state, inserted into both olefinic and allylic C–F bonds to form Al(III) products that were characterized crystallographically. Theory implicated two simultaneous competing mechanisms, respectively involving stereoretentive direct oxidative addition and stereoinverting intermediacy of a metallocyclopropane. —JSY

Angew. Chem. Int. Ed. 10.1002/
anie.201802321 (2018).

GENE THERAPY

Better to transfer than transfuse?

β-Thalassemia is a blood disease caused by mutations

A gene controlling *Wolbachia* proliferation in *Nasonia* wasps is under positive selection.

GENETICS

How hosts can defeat selfish elements

Wolbachia bacterial infections are horizontally passed through the eggs of nematodes and insects and can selfishly affect reproductive outcomes, resulting in an increased number of female offspring. Because *Wolbachia* can affect reproduction, it has been eyed as a potential gene drive system to eliminate disease-carrying vectors, such as mosquitos. Crossing species of *Nasonia* wasps that maintain differing levels of *Wolbachia*, Funkhouser-Jones *et al.* mapped a gene named *Wolbachia density suppressor* that controls *Wolbachia* proliferation within hosts. Further investigation suggested that this gene is under positive selection. This adds to our understanding of the dynamics between hosts and selfish parasites such as *Wolbachia* and may provide information of interest for the design of gene drive systems. —LMZ

Curr. Biol. 10.1016/j.cub.2018.04.010 (2018).

in the β-globin gene. β-Globin is a subunit of hemoglobin, an oxygen-carrying protein in red blood cells. Patients with β-thalassemia are severely anemic and require lifelong transfusions of red blood cells. In two clinical trials involving a total of 22 patients, Thompson *et al.* tested a gene therapy–based treatment that might eliminate the need for repeated transfusions. They isolated hematopoietic stem cells from the patients, used a lentiviral vector to transfer a "normal" copy of the β-globin gene into the cells, and then infused the modified cells back into the patients. After 26 months, the patients showed sustained expression of the transgenic hemoglobin, and nearly 70% of them no longer required transfusions. —PAK

New Engl. J. Med. **378**, 1479 (2018).

PHYSIOLOGY

How drinking (alcohol) affects drinking (water)

The hormone FGF21 (fibroblast growth factor 21) is a major regulator of drinking in mice. FGF21 is produced in the liver and links metabolic status to behavioral responses through actions in the brain. It has many effects on metabolism, including regulation of preference for sugar consumption. Song *et al.* found that various stresses, such as alcohol consumption or a ketogenic diet, caused increased production of FGF21 and stimulation of water drinking. Although alcohol consumption has well-known acute effects on water balance through inhibition of antidiuretic hormone action, this work shows an important role for FGF21 in stimulating water drinking after alcohol consumption. FGF21 can even suppress preference for alcohol consumption in favor of pure water. —LBR

Cell Metab. 10.1016/j.cmet.2018.04.001 (2018).

ENZYME EVOLUTION

Something from nothing

Specialized enzymes often originate from the refinement of a promiscuous enzyme by evolution, rather than through the emergence of activity from an inactive protein. Chalcone isomerase (CHI) catalyzes a simple chemical reaction but emerged within a protein family whose other extant members have no known enzymatic function. Kaltenbach *et al.* used phylogenetic reconstruction to predict the sequence of ancestors along the course of CHI evolution. The distant ancestor of the protein family had no CHI activity, despite presenting catalytic residues within an active site–like pocket. A small number of peripheral mutations could induce CHI activity in the ancestral protein, unlocking the catalytic potential already present in the binding site. —MAF

Nat. Chem. Biol. 10.1038/s41589-018-0042-3 (2018).

COSMOLOGY

Simulating the future of our Universe

The energy densities of matter and dark energy in our Universe are approximately equal at the current epoch. Testing whether this is a coincidence requires running cosmological simulations into the future, something that has generally been avoided because only the past can be observed. Salcido *et al.* simulated the futures of universes with and without dark energy, assessing its impact on the total star formation within each universe. The presence of dark energy has surprisingly little effect, because most stars have formed before it becomes dominant. Eighty-eight percent of all stars that will ever form in our Universe have done so already. The results suggest that the current epoch is not particularly special. —KTS

Mon. Not. R. Astron. Soc. 10.1093/mnras/sty879 (2018).

PHOTO: SIMON VAN NOORT/IZIKO MUSEUMS OF SOUTH AFRICA

ALSO IN *SCIENCE* JOURNALS

Edited by **Stella Hurtley**

INORGANIC CHEMISTRY

Transforming nitrogen without carbon

How much carbon does it take to make nitric acid? The counterintuitive answer nowadays is quite a lot. Nitric acid is manufactured by ammonia oxidation, and all the hydrogen to make ammonia via the Haber-Bosch process comes from methane. That's without even accounting for the fossil fuels burned to power the process. Chen *et al*. review research prospects for more sustainable routes to nitrogen commodity chemicals, considering developments in enzymatic, homogeneous, and heterogeneous catalysis, as well as electrochemical, photochemical, and plasma-based approaches. —JSY

Science, this issue p. 873

SINGLE-CELL ANALYSIS

Mapping the planarian transcriptome

A cell type's transcriptome defines the active genes that control its biology. Two groups used single-cell RNA sequencing to define the transcriptomes for essentially all cell types of a complete animal, the regenerative planarian *Schmidtea mediterranea*. Because pluripotent stem cells constantly differentiate to rejuvenate any part of the body of this species, all developmental lineages are active in adult animals. Fincher *et al*. determined the transcriptomes for most, if not all, planarian cell types, including some that were previously unknown. They also identified transition states and genes governing positional information. Plass *et al*. used single-cell transcriptomics and computational algorithms to reconstruct a lineage tree capturing the developmental progressions from stem to differentiated cells. They could then predict gene programs that are specifically turned on and off along the tree, and they used this approach to study how the cell types behaved during regeneration. These whole-animal transcriptome "atlases" are a powerful way to study metazoan biology. —BAP

Science, this issue p. 874, p. 875

CATALYSIS

Reduction can make cobalt act precious

Enzymes rely on abundant metals such as iron and nickel to manipulate hydrogen. Chemists, on the other hand, have largely had to rely on precious metals such as platinum and rhodium for the task. Friedfeld *et al*. now report a simple trick to make cobalt act more like rhodium. Reduction of Co(II) to Co(I) by zinc reinforced binding of phosphine ligands to the metal to facilitate its use in asymmetric hydrogenation of alkenes. The cobalt catalysts tolerated alcohol solvents, unlike their rhodium congeners, and could be applied to a 200-gram-scale reduction at 0.08% loading. —JSY

Science, this issue p. 888

PHOTOVOLTAICS

Poking a semiconductor

Noncentrosymmetric crystal structure can lead to a peculiar kind of charge separation under illumination called the bulk photovoltaic (BPV) effect. Solar cells made of such materials, however, typically have low efficiency. Yang *et al*. expanded the class of materials capable of exhibiting the BPV effect by making ordinarily centrosymmetric materials, such as $SrTiO_3$ and TiO_2, lose their inversion symmetry. The authors accomplished this by applying a point force on the surface of the material. This induced a strain gradient and the loss of inversion symmetry, resulting in large photovoltaic currents under illumination. The mechanism, dubbed the flexo-photovoltaic effect, is expected to apply to most semiconductors. —JS

Science, this issue p. 904

ORGANIC ELECTRONICS

A longer exciton pathway

Organic semiconductors typically exhibit exciton diffusion lengths on the order of tens of nanometers. Jin *et al*. prepared nanofibers from block polymers consisting of emissive polyfluorene cores surrounded by coronas of polyethylene glycol and polythiophene (see the Perspective by Holmes). Excitons generated in the polyfluorene cannot enter the polyethylene glycol layer and so diffuse more than 200 nm. This distance can be tuned by varying the length of the polyethylene glycol—a feature that could potentially be exploited in the development of organic devices such as photovoltaics. —MSL

Science, this issue p. 897; see also p. 854

MOLECULAR BIOLOGY

RNA and membraneless organelles

Membraneless compartments can form in cells through liquid-liquid phase separation (see the Perspective by Polymenidou). But what prevents these cellular condensates from randomly fusing together? Using the RNA-binding protein (RBP) Whi3, Langdon *et al*. demonstrated that the secondary structure of different RNA components determines the distinct biophysical and biological properties of the two types of condensates that Whi3 forms. Several RBPs, such as FUS and TDP43, contain prion-like domains and are linked to neurodegenerative diseases. These RBPs are usually soluble in the nucleus but can form pathological aggregates in the cytoplasm. Maharana *et al*. showed that local RNA concentrations determine distinct phase separation behaviors in different subcellular locations. The higher RNA concentrations in the nucleus act as a buffer to prevent phase separation of RBPs; when mislocalized to the cytoplasm, lower RNA concentrations trigger aggregation. —SYM

Science, this issue p. 922, p. 918; see also p. 859

CANCER IMMUNOLOGY

Killing without poking holes

Given the success of T cell–centric cancer immunotherapies, there is considerable interest in understanding exactly how tumors sometimes evade this form of treatment. Kearney *et al*. carried out a series of genome-wide CRISPR screens to identify mechanisms of tumor immune evasion from cytotoxic T lymphocytes (CTLs) and natural killer (NK) cells. Interferon-γ signaling and antigen presentation were critical for CTL-mediated killing of cancer cells, and TNF (tumor necrosis factor) signaling was a key effector mechanism for both CTL and NK cell antitumor activity. The same immune evasion mechanisms arose upon screening with perforin-deficient CTLs, suggesting that tumors evade the immune system by dampening the effects of cytokines, rather than reducing direct killing via perforin. —AB

Sci. Immunol. **3**, eaar3451 (2018).

EVOLUTION

Multiple parasites speed host evolution

Virtually all organisms are parasitized by multiple species, but our current understanding of host-parasite interactions is based on pairwise species interactions. Betts *et al*. address this by using the bacterium *Pseudomonas aeruginosa* and five different phage virus parasites. Increasing parasite diversity accelerated the rate of host evolution, driving both faster genomic evolution within populations and greater divergence between populations. Thus, different parasite loads prompt different evolutionary dynamics and profoundly shape host evolution by different mechanisms. —CA

Science, this issue p. 907

CAR T CELL SIGNALING

More ITAMs for more potent receptors

A form of cancer immunotherapy uses the patient's own T cells, which are engineered to express a chimeric antigen receptor (CAR) that recognizes a cancer cell antigen. Increasing signal transduction efficiency in these CAR T cells would enhance treatment efficacy. James engineered synthetic T cell receptors with differing numbers of ITAMs, a protein motif involved in immune cell signaling. Increasing ITAM number enhanced the fraction of T cells that became activated by antigen, suggesting a strategy to improve the potency of CAR T cells. —ERW

Sci. Signal. **11**, eaan1088 (2018).

REVIEW SUMMARY

INORGANIC CHEMISTRY

Beyond fossil fuel–driven nitrogen transformations

Jingguang G. Chen,* Richard M. Crooks,* Lance C. Seefeldt,* Kara L. Bren, R. Morris Bullock, Marcetta Y. Darensbourg, Patrick L. Holland, Brian Hoffman, Michael J. Janik, Anne K. Jones, Mercouri G. Kanatzidis, Paul King, Kyle M. Lancaster, Sergei V. Lymar, Peter Pfromm, William F. Schneider, Richard R. Schrock

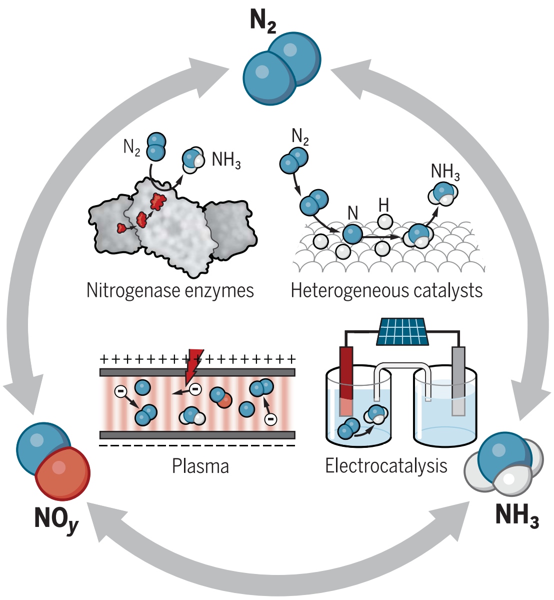
BACKGROUND: The invention of the Haber-Bosch (H-B) process in the early 1900s to produce ammonia industrially from nitrogen and hydrogen revolutionized the manufacture of fertilizer and led to fundamental changes in the way food is produced. Its impact is underscored by the fact that about 50% of the nitrogen atoms in humans today originate from this single industrial process. In the century after the H-B process was invented, the chemistry of carbon moved to center stage, resulting in remarkable discoveries and a vast array of products including plastics and pharmaceuticals. In contrast, little has changed in industrial nitrogen chemistry. This scenario reflects both the inherent efficiency of the H-B process and the particular challenge of breaking the strong dinitrogen bond. Nonetheless, the reliance of the H-B process on fossil fuels and its associated high CO_2 emissions have spurred recent interest in finding more sustainable and environmentally benign alternatives. Nitrogen in its more oxidized forms is also industrially, biologically, and environmentally important, and synergies in new combinations of oxidative and reductive transformations across the nitrogen cycle could lead to improved efficiencies.

ADVANCES: Major effort has been devoted to developing alternative and environmentally friendly processes that would allow NH_3 production at distributed sources under more benign conditions, rather than through the large-scale centralized H-B process. Hydrocarbons (particularly methane) and water are the only two sources of hydrogen atoms that can sustain long-term, large-scale NH_3 production. The use of water as the hydrogen source for NH_3 production requires substantially more energy than using methane, but it is also more environmentally benign, does not contribute to the accumulation of greenhouse gases, and does not compete for valuable and limited hydrocarbon resources. Microbes living in all major ecosystems are able to reduce N_2 to NH_3 by using the enzyme nitrogenase. A deeper understanding of this enzyme could lead to more efficient catalysts for nitrogen reduction under ambient conditions. Model molecular catalysts have been designed that mimic some of the functions of the active site of nitrogenase. Some modest success has also been achieved in designing electrocatalysts for dinitrogen reduction. Electrochemistry avoids the expense and environmental damage of steam reforming of methane (which accounts for most of the cost of the H-B process), and it may provide a means for distributed production of ammonia. On the oxidative side, nitric acid is the principal commodity chemical containing oxidized nitrogen. Nearly all nitric acid is manufactured by oxidation of NH_3 through the Ostwald process, but a more direct reaction of N_2 with O_2 might be practically feasible through further development of nonthermal plasma technology. Heterogeneous NH_3 oxidation with O_2 is at the heart of the Ostwald process and is practiced in a variety of environmental protection applications as well. Precious metals remain the workhorse catalysts, and opportunities therefore exist to develop lower-cost materials with equivalent or better activity and selectivity. Nitrogen oxides are also environmentally hazardous pollutants generated by industrial and transportation activities, and extensive research has gone into developing and applying reduction catalysts. Three-way catalytic converters are operating on hundreds of millions of vehicles worldwide. However, increasingly stringent emissions regulations, coupled with the low exhaust temperatures of high-efficiency engines, present challenges for future combustion emissions control. Bacterial denitrification is the natural analog of this chemistry and another source of study and inspiration for catalyst design.

OUTLOOK: Demands for greater energy efficiency, smaller-scale and more flexible processes, and environmental protection provide growing impetus for expanding the scope of nitrogen chemistry. Nitrogenase, as well as nitrifying and denitrifying enzymes, will eventually be understood in sufficient detail that robust molecular catalytic mimics will emerge. Electrochemical and photochemical methods also demand more study. Other intriguing areas of research that have provided tantalizing results include chemical looping and plasma-driven processes. The grand challenge in the field of nitrogen chemistry is the development of catalysts and processes that provide simple, low-energy routes to the manipulation of the redox states of nitrogen. ■

ON OUR WEBSITE

Read the full article at http://dx.doi.org/10.1126/science.aar6611



Possible routes for nitrogen transformations that eliminate or minimize the need for fossil fuels. A more thorough understanding of nitrogenase may lead to more efficient homogeneous catalysts for reduction of N_2 to NH_3. Coupling of theory and experiment will lead to more effective and stable heterogeneous and electrocatalysts. Innovative energy sources, such as plasmas, which involve nonequilibrium chemistry, may lead to new nitrogen conversion mechanisms.

ILLUSTRATION: K. HOLOSKI

The list of author affiliations is available in the full article online.
*Corresponding author. Email: jgchen@columbia.edu (J.G.C.); crooks@cm.utexas.edu (R.M.C.); lance.seefeldt@usu.edu (L.C.S.)

Cite this article as J. G. Chen *et al.*, *Science* **360**, eaar6611 (2018). DOI: 10.1126/science.aar6611

Chen *et al.*, *Science* **360**, 873 (2018) 25 May 2018 **1 of 1**

REVIEW

INORGANIC CHEMISTRY

Beyond fossil fuel–driven nitrogen transformations

Jingguang G. Chen,[1,2]* Richard M. Crooks,[3]* Lance C. Seefeldt,[4]* Kara L. Bren,[5] R. Morris Bullock,[6] Marcetta Y. Darensbourg,[7] Patrick L. Holland,[8] Brian Hoffman,[9] Michael J. Janik,[10] Anne K. Jones,[11] Mercouri G. Kanatzidis,[9] Paul King,[12] Kyle M. Lancaster,[13] Sergei V. Lymar,[2] Peter Pfromm,[14] William F. Schneider,[15] Richard R. Schrock[16]

Nitrogen is fundamental to all of life and many industrial processes. The interchange of nitrogen oxidation states in the industrial production of ammonia, nitric acid, and other commodity chemicals is largely powered by fossil fuels. A key goal of contemporary research in the field of nitrogen chemistry is to minimize the use of fossil fuels by developing more efficient heterogeneous, homogeneous, photo-, and electrocatalytic processes or by adapting the enzymatic processes underlying the natural nitrogen cycle. These approaches, as well as the challenges involved, are discussed in this Review.

Reduced forms of nitrogen, particularly ammonia (NH_3), are vital to life. Before the early 1900s, all reduced N came from biological nitrogen fixation by microbes. This landscape changed in the early 1900s with the invention by Haber and Bosch of an industrial process to reduce N_2 to NH_3 ($N_2 + 3H_2 \rightleftarrows 2NH_3$). The Haber-Bosch (H-B) process has led to fundamental changes in the way food is produced, and its impact is underscored by the fact that about 50% of the N atoms in humans today originate from it (*1*). For these reasons and others, the H-B process has been called the most impactful invention of the 20th century (*2*). However, it is energy-intensive, consuming 1 to 2% of the world's annual energy output (*3*). The main energy requirements arise from the high reaction temperatures (~700 K) and pressures (~100 atm) and, most importantly, from the need for large quantities of H_2. Nevertheless, at scale, the H-B process is surprisingly energy-efficient. In addition to the high energy demand, however, there is a second problem with the process that is potentially more serious. Specifically, the source of H_2 is usually natural gas (*4*), and, consequently, the H-B process generates about 1.9 metric tons of CO_2 per metric ton of NH_3 produced (3 CO_2 per 8 NH_3) (*5*).

Nitrogen in its more oxidized forms is also industrially, biologically, and environmentally important. Oxidation of N_2 to nitric oxide (NO) and nitrogen dioxide (NO_2), which are collectively referred to as NO_x (the representation of neutral forms of oxidized nitrogen), occurs naturally in lightning and also during combustion of fuels in air (*6*). In fact, plasma-driven oxidation of N_2 to NO (the Birkeland-Eyde, or B-E, process) preceded the H-B process as the first commercial approach to nitrogen fixation (*7*). However, the B-E process is not economically competitive with the H-B process, and, as a result, the demand for oxidized nitrogen, primarily in the form of nitric acid (HNO_3), is satisfied by catalytic oxidation of H-B–generated NH_3 at elevated temperatures through the Ostwald process (*8*).

Nitrogen oxides (NO_y, which represents both neutral and anionic forms of oxidized nitrogen) are also produced on a large scale by bacteria. These processes include oxidation of NH_3 to nitrate (NO_3^-) by nitrifying bacteria and reduction of NO_3^- to nitrite (NO_2^-), NO, nitrous oxide (N_2O), and ultimately N_2 by denitrifying bacteria. The conversion of NH_3 and NO_3^- to N_2 by the action of these bacteria is used in wastewater treatment and is the primary pathway for loss of applied fertilizers in agriculture. Nitrogen oxides are common environmental pollutants, and thus catalysts to reduce these oxides to N_2 are also of considerable practical importance. For example, NO_x generated during combustion is a primary component of photochemical smog (*9*). The development of active catalysts for reducing NO_x to N_2 (three-way catalytic converters, or TWCs) led to considerable improvements in air quality in the late 1970s (*10*), and TWCs now operate on hundreds of millions of vehicles worldwide. However, increasingly stringent emissions regulations, coupled with the low exhaust temperatures from high-efficiency engines, present challenges for the future control of combustion emissions (*11*). The soluble forms of the nitrogen oxides, primarily NO_2^- and NO_3^-, are similarly hazardous to human health and contribute to the eutrophication of waterways (*12*). Sustainable approaches for catalytically converting these pollutants to N_2 remain an unmet challenge (*13*).

In this Review, we report on the results of a U.S. Department of Energy workshop held in October 2016 that focused on the challenges and opportunities associated with fundamental aspects of nitrogen chemistry. We begin by discussing the thermochemistry of nitrogen transformations; we then focus on the reduced forms of nitrogen, primarily NH_3, followed by the oxidized forms.

Thermochemistry of nitrogen fixation

The energy expenditures for various reductive and oxidative N_2 fixation pathways are compared in Fig. 1 (*14*–*16*). Hydrocarbons, including biomass and fossil fuels (particularly methane), and water are the only two sources of hydrogen atoms that can sustain long-term, large-scale NH_3 production. The use of water as the hydrogen source for NH_3 production requires substantially more energy than using methane (by 613.4 kJ/mol N_2), but it is also more environmentally benign, does not contribute to the accumulation of greenhouse gases, and does not compete for valuable and limited hydrocarbon resources. However, these advantages can be fully realized only if the energy comes from a clean, renewable source such as the Sun or wind.

When water is used as the feedstock, there is only a small (32.9 kJ/mol N_2) thermodynamic advantage for the direct synthesis of NH_3, compared with a two-stage synthesis involving water splitting followed by the H-B process. The overall energy saving may, however, become more substantial if the H-B process is avoided altogether and direct reductive N_2 fixation is conducted under milder conditions.

The two most important N-containing commodity chemicals are NH_3 and HNO_3. In 2017, their worldwide production was estimated to be 150 (*17*) and 50 (*18*) million metric tons, respectively. Presently, nearly all HNO_3 is manufactured using a three-stage approach—steam reforming of methane, the H-B process, and then the Ostwald process (Fig. 1)—but this approach is energetically wasteful compared with direct oxidative N_2 fixation.

An attractive alternative route to HNO_3, which avoids the NH_3 intermediate, is direct N_2 oxidation to aqueous HNO_3 (Eq. 1).

---

[1]Department of Chemical Engineering, Columbia University, New York, NY 10027, USA. [2]Chemistry Division, Brookhaven National Laboratory, Upton, NY 11973, USA. [3]Department of Chemistry, The University of Texas at Austin, Austin, TX 78712, USA. [4]Department of Chemistry and Biochemistry, Utah State University, Logan, UT 84332, USA. [5]Department of Chemistry, University of Rochester, Rochester, NY 14627, USA. [6]Pacific Northwest National Laboratory, Richland, WA 99352, USA. [7]Department of Chemistry, Texas A&M University, College Station, TX 77843, USA. [8]Department of Chemistry, Yale University, New Haven, CT 06511, USA. [9]Department of Chemistry, Northwestern University, Evanston, IL 60208, USA. [10]Department of Chemical Engineering, Pennsylvania State University, University Park, PA 16802, USA. [11]School of Molecular Sciences, Arizona State University, Tempe, AZ 85282, USA. [12]National Renewable Energy Laboratory, Golden, CO 80401, USA. [13]Department of Chemistry and Chemical Biology, Cornell University, Baker Laboratory, Ithaca, NY 14853, USA. [14]Voiland School of Chemical Engineering and Bioengineering, Washington State University, Pullman, WA 99164-6515, USA. [15]Department of Chemical and Biomolecular Engineering, University of Notre Dame, Notre Dame, IN 46556, USA. [16]Department of Chemistry, MIT, 6-331, Cambridge, MA 02139, USA.

***Corresponding author. Email: jgchen@columbia.edu (J.G.C.); crooks@cm.utexas.edu (R.M.C.); lance.seefeldt@usu.edu (L.C.S.)**

$$N_{2(g)} + 2.5O_{2(g)} + H_2O_{(liq)} \rightleftarrows 2H^+_{(aq)} + 2NO_3{}^-{}_{(aq)} \qquad (1)$$

The equilibrium constant for this reaction is ~2.7×10^{-3} $M^4/bar^{3.5}$ (standard Gibbs free energy $\Delta G^0 = 14.6$ kJ/mol N_2) (*14–16*, *19*), which is large enough to drive spontaneous formation of ~0.1 M HNO_3 in any pool of water on the surface of Earth. Moreover, if equilibrium were to be achieved between Earth's atmosphere and oceans, the oceans would contain ~0.02 M HNO_3, and ~90% of atmospheric O_2 would be consumed. Fortunately for life on Earth, under ordinary conditions and in the absence of a catalyst, Eq. 1 is unmeasurably slow in both directions.

Reduction of N_2 to NH_3

Overview

As discussed in the previous two sections, the most important N_2 reduction reaction is NH_3 synthesis, specifically $N_2 + 3H_2 \rightleftarrows 2NH_3$. Given that this reaction is slightly thermodynamically favorable under ambient conditions, major effort has been devoted to developing alternative and environmentally friendly processes that would allow NH_3 to be produced at distributed sources under more benign conditions, rather than through the large-scale centralized H-B process. This goal is particularly important in developing countries where the population, and hence the need for food, is increasing rapidly. In these countries, access to fertilizer is hampered by poor transportation infrastructure and insufficient capital to build large chemical plants. For example, there are no large-scale NH_3 production facilities anywhere in East Africa. To address this need, substantial effort has been expended to understand N_2 reduction using a variety of catalysts, including heterogeneous, enzymatic, and homogeneous catalysts, as well as electro- and photocatalysts.

Heterogeneous catalysis

The commercial H-B process is carried out using a heterogeneous catalyst based on iron and promoted with Al_2O_3 and potassium. It is now well accepted that the relative activity of metallic catalysts can be correlated to their binding energies with N-containing species in terms of a volcano-shaped relationship. Metals that bind nitrogen too strongly or too weakly are on either side of the volcano (*20*). For metals that have low binding energies, N_2 dissociation is rate-limiting. For metals with high binding energies, N_2 dissociation occurs, but desorption of the resulting atomic N (and other N-containing intermediates) is slow, limiting the number of available binding sites and thereby slowing the catalytic rate (*21*).

Recent calculations suggest that the energies of all intermediates and transition states involved in a conventional NH_3 synthesis reaction network scale with the N-binding energy. As shown in Fig. 2, these correlations limit the rate of NH_3 synthesis on transition metals to much lower values than would be possible on a material with both a low N_2 dissociation barrier and more moderate binding energies (*22*). An important challenge is to break this type of scaling relationship, which should lead to the development of catalysts yielding reaction rates that are potentially orders of magnitude higher than the current state of the art.

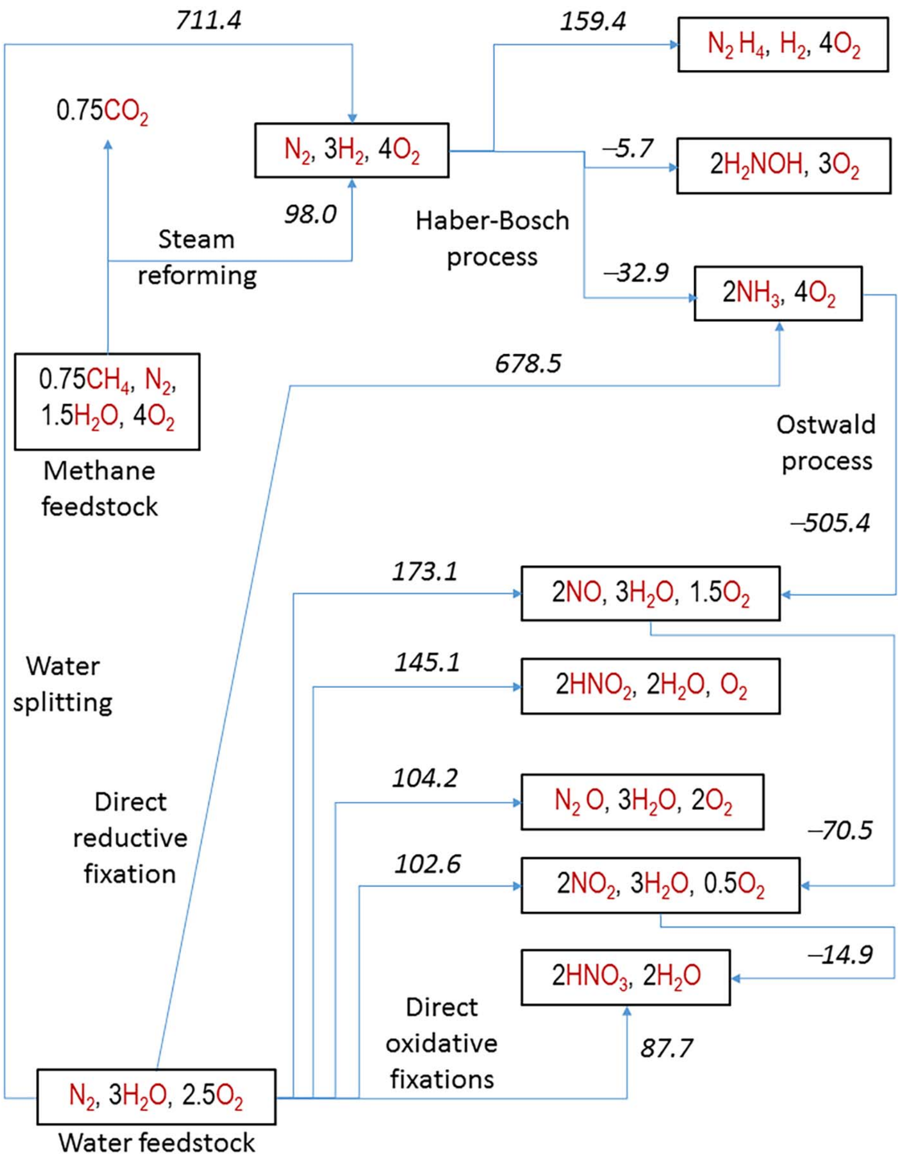


Fig. 1. Atom and energy economy of nitrogen fixation. The numerical values are the standard (*14*) Gibbs free energies in kilojoules per mole of fixed N_2 in the direction of the arrows. All thermochemical data are from (*15*), except for H_2NOH (*16*).

Advances in the synthesis, characterization, and modeling of nontraditional heterogeneous catalysts—such as intermetallics, alloys containing a single heteroelement, and shape-controlled materials—offer the promise for more active catalysts. These combined efforts should lead to catalysts that could potentially operate at lower temperatures and thus reduce the operating pressures of the H-B process. Another opportunity to reduce the CO_2 footprint of the H-B process is to replace the methane reformer with a water electrolyzer to provide CO_2-free H_2, but this would require cost-effective, large-scale electrolyzers.

Enzyme catalysis

A range of microbes (bacteria and archaea) living in all major ecosystems are able to reduce N_2 to NH_3 using the enzyme nitrogenase. Nitrogen fixation by these microbes contributes to the global supply of NH_3, but it is insufficient to support modern intensive agriculture. Exciting frontier areas for the application of microbial N_2 fixation include engineering nitrogen fixation genes into eukaryotes (*23*), such as plants, and using nitrogen-fixing microbes in electrocatalysis cells that permit sustained N_2 reduction to NH_3 (*24*). Such systems offer the promise of providing localized solutions to N_2 reduction but are not likely to displace the need for large-scale N_2 reduction in the near term.

One important lesson to learn from biological nitrogen fixation is how the enzyme nitrogenase accomplishes reduction of N_2 to NH_3 at ambient temperatures and pressures and without H_2. The first nitrogenase to be purified (and the most widely distributed) is molybdenum-dependent and, as shown in Fig. 3, has a metal-cluster active site containing Mo, Fe, S, C, and homocitrate (FeMo cofactor).

An important step toward a mechanistic understanding of nitrogenase emerged from trapping an intermediate with $4H^+$ and $4e^-$ accumulated immediately before binding N_2. Characterization of this intermediate (*25*) (termed the E_4 state) has revealed the presence of bridging hydrides (Fe-H-Fe). The E_4 state undergoes reductive elimination of the two hydrides with release of H_2, with concomitant activation of the metal core by $2e^-$ for binding and reduction of N_2 to a metal-bound diazenido species (HN=N–Fe). Although these studies have provided insights into the central catalytic step, which is the cleavage of the N_2 triple bond, many challenges remain before the entire mechanism can be elucidated. These include a molecular-level understanding of both the early and late stages of the N_2 reduction pathways, an understanding of the roles of adenosine triphosphate (ATP) hydrolysis in driving electron transfer, and then reconciliation of the empirical mechanism with computational studies.

These foregoing challenges primarily involve substrate moieties bound as catalytic intermediates. In parallel, important questions remain about the FeMo cofactor active site of Mo-dependent nitrogenase. For example, what changes in redox and spin states of the seven Fe centers and the Mo center accompany the stepwise transformations of substrates? Additionally, the FeMo cofactor contains the first known example of a carbide (C^{4-}) in biology. The importance and function of this carbide with respect to mechanism are unknown. Understanding the structure of FeMo cofactor intermediates and synthetic analogs could provide important clues to guide the design of new homogeneous and heterogeneous catalysts for N_2 reduction (*26*). Biomimetic approaches could take advantage of not only the coordination chemistry in the FeMo cofactor, but also the cooperative interactions with nearby acidic amino acid residues in the nitrogenase protein.

Electrochemical techniques offer a powerful alternative method of delivering electrons to nitrogenase without a requirement for ATP hydrolysis, and they also offer new avenues for addressing aspects of the nitrogenase mechanism. Electron transfer from an electrode, previously demonstrated for a number of redox enzymes, has been applied to nitrogenase, but only via mediators (no direct electron transfer) (*27*–*29*). As illustrated in a recent study, this approach, in combination with computations (*30*), provides insights into catalytic nitrogen transformations.

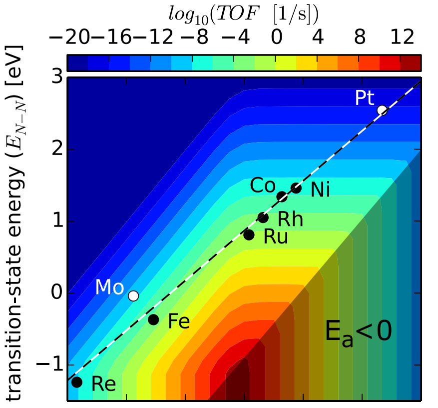


Fig. 2. Linear scaling between nitrogen-binding energy and activation barrier for N_2 dissociation. TOF, turnover frequency; E_a, activation energy. [Reproduced with permission from (*22*)]

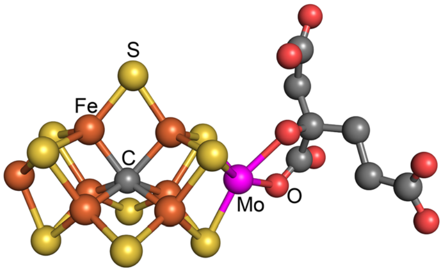


Fig. 3. FeMo cofactor. Fe, rust-colored; S, yellow; C, gray; Mo, magenta; O, red. The depiction is based on metrical data from PDB (Protein Data Bank) ID 1M1N.

A related frontier area is using nanomaterials to deliver electrons to nitrogenase. For example, it has been recently demonstrated that a CdS nanorod–nitrogenase MoFe protein hybrid achieves light-dependent N_2 reduction to NH_3 by using a sacrificial electron donor (*31*). Such systems offer a means to use light as the exclusive energy source to drive N_2 reduction.

Homogeneous catalysis

Synthetic homogeneous catalysts designed for N_2 reduction provide well-defined molecular precatalysts and intermediates, which can be thoroughly characterized by diverse spectroscopic techniques and thereby provide excellent opportunities for determining mechanistic information about N_2 reduction to NH_3. The reactivity of synthetic inorganic complexes offers insights into molecular reactivity and individual bond-making processes that occur in nitrogenases. The discovery of nitrogenase, along with a transition metal compound that contains N_2 bound to ruthenium (*32*), initiated a race to prepare NH_3 catalytically under mild conditions using a transition metal complex. Entire institutes were created in England and Russia in the late 1960s for this purpose. The stoichiometric N_2-to-NH_3 conversion using H_2SO_4 was demonstrated with low-valent Mo-N_2 and W-N_2 complexes in the 1970s (*33*); typically, the electrons required came from the metal complex. In 1985, Pickett and Talarmin accomplished the synthesis of NH_3 by using an electrochemical route starting from the tungsten complex $[W(N_2)_2(PMe_2Ph)_4]$; Me, methyl, Ph, phenyl], with the electrons supplied by the electrode and protons from an added acid (*34*). The catalytic reduction of N_2 to NH_3 by H^+ and e^- was reported in subsequent studies, using, for example, a well-defined Mo-containing catalyst (*35*). Up to eight equivalents of NH_3 per Mo were formed, along with H_2. Experiments and calculations both support a mechanism involving the addition of $6H^+$ and $6e^-$ to the N_2 initially bound to a Mo^{3+} center. Altogether, eight of these proposed intermediates were prepared and characterized.

Additional Mo precatalysts for reduction of N_2 using chemical reductants and an acid—a dimolybdenum-dinitrogen complex (*36*) and a molybdenum nitride (*37*)—were subsequently reported in 2011 and 2015, respectively. In 2013, Fe complexes were reported that catalyze reduction of N_2 to NH_3 using a strong reducing agent (KC_8) and a strong acid at 200 K (*38*); improvements to the Fe system recently produced as many as 84 equivalents of NH_3 per Fe (*39*). These Fe complexes were designed to reflect the trigonal symmetry of the Fe in the catalytic face of the FeMo cofactor (Fig. 3) and to allow variation of an axial ligand, so as to explore tuning of N_2 binding and reduction by the unusual carbide at the center of the FeMo cofactor.

In all these cases, it appears that catalytic N_2 reduction is a six-electron process, involving many intermediate chemical species ($M\text{-}N_xH_y^{n+}$), but, unlike nitrogenase, without a mechanistic requirement for production of H_2. Nonetheless, in all cases, N_2 reduction is accompanied by production of H_2, and at least with the Fe complexes, this has been shown to involve a catalytic process, with H^+ reduction competing with N_2 reduction.

Much work remains to move this field forward. For example, rational design of potential catalysts will be guided by accurate determination of the bond dissociation free energies (BDFEs) of N–H bonds, including likely intermediates such as M–N=NH, M=N–NH_2, and M=N–NH_3. Initial experimental and theoretical estimates have shown that N–H BDFEs of coordinated NH_3 can vary considerably, and they demonstrate the importance of metal identity, metal oxidation state, and the attached ligands, which together tune the stability of potential intermediates. In addition, determination of the energies of these bonds will allow evaluation of energetic requirements of the individual chemical steps that ultimately lead to NH_3. It is also important to identify features that tune the reduction potentials of relevant complexes to ensure that the multiple reductions required for catalysis are energetically feasible. Thus, research should focus on exploring complexes designed to simultaneously promote N_2 binding and multiple reductions at the metal site.

Electrocatalysis and photocatalysis

The potential-pH predominance (Pourbaix) diagram for the N_2-H_2O system shown in Fig. 4 reveals that reduction of N_2 to NH_4^+ or NH_3 (Eq. 2 and the N_2 reduction line in Fig. 4) is thermodynamically possible under moderately reducing conditions.

$$N_{2(g)} + 8H^+_{(aq)}\ [\text{or } 6H^+_{(aq)}] + 6e^- \rightleftarrows 2NH_4{}^+_{(aq)}\ [\text{or } 2NH_{3(aq)}] \qquad (2)$$

Moreover, with a sufficiently active electrocatalyst, the N_2 reduction reaction can, in principle, proceed in a narrow range of negative potentials without interference from water reduction (line *a* in Fig. 4) over the entire pH range. However, at potentials below line *a*, an extremely selective electrocatalyst would be required to suppress parasitic water reduction. The nitrogen reduction reaction could be paired with a water oxidation half-reaction to supply the H^+ and e^- required for a complete electrolysis cell. Such a cell would essentially combine water splitting with NH_3 formation as depicted by the direct reductive fixation route in Fig. 1. Despite the inviting prospect of a scalable NH_3 electrolysis cell running on renewable electricity at or near room temperature, the difficulty of selectively catalyzing the nitrogen reduction reaction has prevented development of an efficient electrochemical system.

At present, only a few examples of pairing the nitrogen reduction reaction with water oxidation at ambient temperature and pressure have been reported, and in all cases, owing to competing H_2 evolution, the faradaic efficiencies for NH_3 are so low that even its reliable detection is challenging (*40*, *41*). For example, an efficiency of 0.83% and a production rate of 9.4×10^{-10} mol NH_3 per square centimeter per second were measured at 353 K at an applied cell voltage of 1.2 V using Pt electrodes and a Nafion electrolyte (*42*). Another electrochemical approach involved an electrochemical cycling process for producing NH_3 (*43*). The cycling included separate steps of LiOH electrolysis, Li nitridation, and Li_3N hydrolysis, thereby circumventing the parasitic H_2 evolution reaction and leading to higher faradaic efficiencies for NH_3.

Proton-conducting ceramic oxides have been explored as electrodes for nitrogen reduction, and when they operate at higher temperatures, both the rate and the faradaic efficiency increase (*44*). For example, efficiencies in excess of 50% and rates in the range of 10^{-9} to 10^{-8} mol NH_3 per square centimeter per second have been attained above 673 K by using oxide or molten chloride electrolytes and a variety of metal or metal oxide cathodes (*44*, *45*). However, higher temperatures introduce an additional energy burden and hence compromise the advantage of this approach compared with the H-B process.

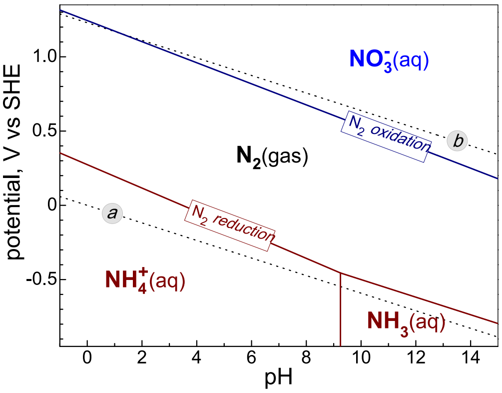
Photochemical reduction of N_2 to NH_3 has also been reported, and these studies may contain guidance relevant to electrochemistry (*46*, *47*). Photocatalytic routes to NH_3 have been summarized recently (*48*). One of the biggest challenges for photocatalytic NH_3 synthesis is the discovery of active and stable photoelectrode materials that are responsive to visible light. Synthetic nitrogenase mimics in the form of chalcogels, composed of Mo- and Fe-containing biomimetic clusters, can potentially accomplish photocatalytic N_2 fixation and conversion to NH_3 at ambient temperature and pressure. Both Fe–S and Mo-Fe–S chalcogels have displayed promising activity toward N_2 reduction (*49*).

Plasma-driven transformations

The feasibility of using plasmas consisting of ions, electrons, and excited molecular species to drive chemical reactions that are thermally inaccessible has been known for many years (*7*). For example, NH_3 was synthesized in laboratory-scale, nonthermal plasmas of N_2 and H_2 at temperatures and pressures lower than those used in the H-B process as early as the 1970s, and more recently, the synthesis has been demonstrated under much milder conditions (1 atm and 140°C) (*50*). The plasma synthesis of NH_3 has been carried out both homogeneously and in combination with heterogeneous catalysts (*7*, *50*, *51*).

Several mechanisms have been proposed to account for enhancements in catalytic activity in the presence of nonthermal plasmas (*52*–*54*). However, the fundamental science and practical engineering of plasma-driven NH_3 syntheses lag behind thermal and electrochemical routes and are therefore ripe for additional research. Further progress in this field will draw on the insights gained from traditional heterogeneous and electrochemical reductions.

Fig. 4. Partial Pourbaix diagram for the N_2-H_2O system. Solid lines correspond to N_2 reduction to NH_4^+ or NH_3 and N_2 oxidation to NO_3^-. Dotted lines *a* and *b* straddle the region of water stability (reduction to H_2 and oxidation to O_2, respectively). Details are given in (*14*), and primary data are in (*15*, *19*).



Chemical looping

Chemical looping is perhaps most familiar from combustion, where two connected fluidized beds are used to cycle solid particles between an oxidized and a reduced state to combust a carbon-containing fuel with oxygen (*55*). Chemical looping can also be performed with two fixed beds instead of fluidized beds. Nitrogen can be used instead of oxygen for a similar looping approach in NH_3 synthesis (*56*). First, N_2 is contacted with a suitable solid-state transition metal to yield a nitride (activation). Second, NH_3 is harvested by contacting the nitride with steam or H_2 (*57*). Recently, lithium has been used as the parent metal for the nitride, but this approach requires liquid-phase electrolysis to recycle lithium, and this step is usually not present in chemical looping (*43*).

Potential advantages of chemical looping include the ability to independently control process conditions for N_2 activation and product harvest. In this sense, chemical looping can break the scaling relationship alluded to in Fig. 2. Operation at atmospheric pressure is another major advantage of chemical looping for nitrogen activation (*58*). On the basis of results from a pilot-scale plant, chemical looping for combustion, including CO_2 capture, is comparable to conventional combustion and appears to be economically attractive (*59*). This may be an indication of the possible economic feasibility of chemical looping for nitrogen activation.

Oxidation of N_2

Overview

The dominance of the H-B process in commercial N_2 fixation is arguably a direct consequence of the ready availability of H_2 from fossil hydrocarbons. As shown in Fig. 1, oxidative fixation of N_2 can, in principle, be achieved with a lower-energy input than reductive fixation if N_2, H_2O, and O_2 are used as reactants. Other nitrogen oxides are accessible at even lower free energies, especially when coupled with water to form the corresponding oxoacids.

Although H-B and natural processes have motivated substantial research into reductive N_2 fixation, direct N_2 oxidation remains largely unexplored despite its great practical value as a replacement for the Ostwald process. The fact that living organisms have not evolved to consume N_2 and O_2 and produce aqueous HNO_3 suggests fundamental chemical challenges that merit investigation. As a result, there is very little literature on the homogeneous oxidative chemistry of N_2.

Direct oxidation of N_2 with O_2

NO_x is unavoidably generated at the high temperatures that prevail during combustion in air. NO is the primary combustion-generated component of NO_x, and its formation can be viewed in the context of its equilibrium with N_2 and O_2 (Eq. 3).

$$N_{2(g)} + O_{2(g)} \rightleftarrows 2NO_{(g)} \qquad (3)$$

The standard enthalpy of this reaction is 90.3 kJ/mol NO, and the entropy is 12.4 J/K per mol NO (*14*, *15*), indicating that equilibrium is unfavorable for NO production under ambient conditions. Owing to the positive entropy, however, the equilibrium NO concentration rises rapidly with temperature, reaching 0.8% in air at 2000 K (*60*).

The direct bimolecular reaction of N_2 and O_2 to produce NO is symmetry-forbidden and has a negligible rate even at combustion temperatures. Rather, NO is produced during combustion via Eqs. 4 and 5, where O (and similarly OH) radicals attack N_2, as originally proposed by Zeldovich (*61*).

$$O + N_2 \rightarrow NO + N \qquad (4)$$

$$N + O_2 \rightarrow NO + O \qquad (5)$$

Analogous reactions are at play in high-temperature thermal plasmas, as in the original B-E process. Thermal plasmas are unlikely to be energy-competitive with the H-B process, even if the latter uses a renewable H_2 source (*62*), because the overall energy efficiencies of N_2 activation are low, and rapid thermal quenching is required to suppress NO decomposition back to N_2 and O_2.

The aforementioned fundamental limitations can be mitigated using low-temperature, nonthermal plasmas generated by either a dielectric barrier discharge or microwave absorption at or below atmospheric pressure, potentially coupled with appropriate catalysts (*7*). According to current estimates, theoretical energy consumption for N_2 fixation via Eq. 3 in a nonthermal plasma (~400 kJ/mol N_2) is lower by more than a factor of 2.5 than that for the H-B process with methane-derived H_2 (*62*–*64*), and the energy efficiencies already attained in the laboratory [600 to 1200 kJ/mol N_2, assuming 100%-efficient plasma generation (*63*, *64*)] are better than the H-B process using H_2 from water electrolysis [~3000 kJ/mol N_2 (*62*)], 40% of which would go to electrolysis (*65*). Experiments and kinetic models show that the primary effect of the nonthermal plasma is to accelerate Zeldovich-like reactions involving nonthermal, vibrationally excited N_2. Practical implementation of nonthermal plasmas is hampered by their relatively low power per unit reactant mass and thus low throughput.

A thermal, heterogeneous N_2 oxidation catalyst would have to activate N_2 and O_2 and be stable at the extreme temperatures at which NO production becomes thermodynamically favored. Because of these stringent demands, nonthermal electro-, photo-, or plasma-assisted oxidation routes likely offer the greatest opportunities for progress. The development of catalytic systems that can effectively couple with these external energy sources will be key to practical advances. Preliminary results with nonthermal plasmas, especially coupled to catalysts, are promising. There is a pressing need to uncover the fundamental mechanisms that unite plasma chemistry, surface chemistry, catalysts, and theory to guide optimization of plasma and rational development of appropriate catalytic materials.

Electrochemical oxidation of N_2 to HNO_3

The oxidation line in Fig. 4 shows that N_2 is electrochemically unstable toward NO_3^- under moderately oxidizing conditions (Eq. 6).

$$N_{2(g)} + 6H_2O_{(liq)} \rightleftarrows 2NO_3{}^-{}_{(aq)} + 12H^+{}_{(aq)} + 10e^- \qquad (6)$$

At pH > 1.3, this 10-electron reaction is more thermodynamically favorable than the parasitic four-electron water oxidation to O_2. Thus, it is possible for NO_3^- to be the only product of an anodic process, particularly in neutral and alkaline solutions, if a sufficiently active and selective electrocatalyst for Eq. 6 can be discovered. The corresponding cell-completing cathodic reaction can be either water reduction to H_2 or O_2 reduction to water. Little apparent progress has yet been made in developing such catalysts, but there is no reason to believe that they cannot eventually be discovered.

By coupling cathodic N_2 reduction (Eq. 2) with anodic N_2 oxidation (Eq. 6), an electrochemical cell consuming only atmospheric N_2 and water as feedstock to yield aqueous NH_4NO_3 (Eq. 7) can be envisioned.

$$N_{2(g)} + 3H_2O_{(liq)} \rightleftarrows NH_4{}^+{}_{(aq)} + NO_3{}^-{}_{(aq)} + H_{2(g)} \qquad (7)$$

The additional H_2 production is necessary for matching cathodic and anodic currents and maintaining the electrolyte pH constant. Such a cell would require a minimum of 1.08 V to operate in neutral or alkaline electrolyte, which is slightly lower than the minimum water electrolysis voltage of 1.23 V. Advances relating to electrochemical N_2 oxidation will require development of catalysts that are sufficiently active (capable of good current densities at low overpotentials), selective (with respect to water oxidation), and stable (toward deactivation under harsh redox and pH operating conditions).

NO_y reduction and NH_3 oxidation

Overview

The redox chemistry of forms of nitrogen other than N_2 is immense and rich, and it is extensively incorporated into biological systems. Both NO_y and NH_3 are environmentally hazardous pollutants generated by industrial and transportation activity, and thus much research has gone into developing and applying NO_y reduction and NH_3 oxidation catalysts. As is often the case, the understanding of these catalysts lags their applications.

Ammonia oxidation and NO_y reduction are critical to environmental protection and are much less well understood or optimized than the synthesis of NH_3. There are appealing opportunities to apply the experimental and computational tools of modern heterogeneous catalysis to these reactions, both to develop a fundamental understanding of the processes and to identify superior materials and catalytic transformations—for example, catalysts that can selectively reduce NO_y to N_2 using hydrocarbon reductants. Environmental applications generally demand high longevity, durability, and tolerance to poisons. Key gaps in understanding include how activity and selectivity can be achieved in relatively cool gas streams, how catalyst structure evolves over long periods of time, how sulfur and other common poisons interfere with reactions, and especially how these effects can be mitigated.

NO_y reduction

NO decomposition, the reverse of Eq. 3, is thermodynamically downhill but difficult to catalyze under the oxygen-rich conditions of interest for environmental protection (*66*). Nevertheless, NO decomposition has been demonstrated using both homogeneous and heterogeneous catalysts. Decomposition appears to involve a two-electron reduction of two NO molecules to a hyponitrite (Eq. 8).

$$2NO + e^- \rightarrow N_2O_2{}^- \qquad (8A)$$

$$N_2O_2{}^- + e^- \rightarrow N_2O_2{}^{2-} \qquad (8B)$$

$$N_2O_2{}^{2-} + 2H^+ \rightarrow N_2O + H_2O \qquad (8C)$$

Coordination complexes of hyponitrites are not uncommon, and a tri-Cu hyponitrite coordination complex has been isolated as an intermediate in the catalytic decomposition of NO to N_2 (*67*). A Cu dimer has similarly been implicated as the active site in a Cu-exchanged zeolite catalyst for NO decomposition (*68*).

NO_x can be catalytically reduced to N_2 via several routes. Precious metals are active for NO_x reduction by CO and hydrocarbons, but only under conditions in which O_2 concentrations are low (*10*). Vanadia/titania catalysts and metal-exchanged zeolites catalyze selective reduction of NO_x by NH_3 (*69*). Nitrosamine appears to be the key N_2-forming intermediate in the catalytic pathway (Eq. 9).

$$NH_3 + NO \rightarrow H_2NNO + H^+ + e^- \qquad (9A)$$

$$H_2NNO \rightarrow N_2 + H_2O \qquad (9B)$$

This N_2-forming reaction is exothermic, implying substantial, but perhaps not insurmountable, kinetic challenges to reversing the chemistry.

NO_x can also be catalytically reduced with H_2 over Pd and Pt catalysts (*13*), and N_2 is typically the desired product (*66*). By suitable selection of promoters and control of the NO/H_2 ratio, however, NH_3 can be produced over Pt catalysts (*70*). Likewise, NH_3 competes with N_2 as the product of the catalytic reduction of NO_2^- and NO_3^- by H_2 over Pd catalysts (*13*).

Certain bacteria analogously reduce NO_3^- and NO_2^- through a series of gaseous intermediates, leading ultimately to N_2 in a process called denitrification (Eq. 10). In these bacteria, NO_y acts as a terminal electron acceptor in an anaerobic respiration process. Electrons flowing through membrane-bound proteins produce a proton gradient that in turn is used to make ATP. Each step in the denitrification process is catalyzed by well-studied metalloenzymes that use unusual metal clusters to achieve their chemistry (71, 72). From a global nitrogen perspective, denitrification by these bacteria functions to convert fixed forms of N (NO_y) into N_2. In some cases, this process is beneficial, such as during wastewater treatment to remove nitrates. In other cases, such as in agriculture, denitrification results in loss of N that was applied to crops as fertilizer.

$$2NO_3^- + 10e^- + 12H^+ \rightarrow N_2 + 6H_2O \qquad (10)$$

NH_3 oxidation

Heterogeneous NH_3 oxidation with O_2 is at the heart of the Ostwald process and is practiced in a variety of environmental protection applications. Precious metals remain the workhorse catalysts, and therefore opportunities exist to develop lower-cost materials having equivalent or better activity and selectivity. The preferred NH_3 oxidation products shift from N_2 at low temperatures to NO at higher temperatures, with N_2O being a minor but highly undesirable product. At present, a full mechanistic understanding of this process and of the kinetic origins of selectivity is lacking.

NH_3 and/or NO_2^- are fuel for bacteria and archaea that use multi-electron oxidative reactions to generate reducing equivalents for cellular respiration and to establish a proton gradient for ATP synthesis (73, 74). Ammonia-oxidizing bacteria (AOB) and archaea (AOA) stoichiometrically convert NH_3 to NO_2^-. AOA and AOB first carry out the selective conversion of NH_3 to hydroxylamine (NH_2OH) by using an integral membrane Cu monooxygenase called ammonia monooxygenase (AMO) (Eq. 11). Active AMO has never been purified, and thus this challenging transformation awaits mechanistic elucidation.

$$NH_3 + \tfrac{1}{2}O_2 \rightarrow NH_2OH \qquad (11)$$

The NH_2OH generated by AMO is oxidized to provide a net electron flow for cellular respiration. How AOA processes NH_2OH is unknown (75). In AOB, this process is carried out by hydroxylamine oxidoreductase (HAO), which contains a rare, covalently modified c-heme, heme P460, to which NH_2OH binds at Fe and undergoes proton-coupled oxidation (73). Recent experiments strongly support NO as the enzymatic product of HAO reactivity, contrary to decades of dogma asserting NO_2^- as the enzymatic product (76). Assuming a two-electron turnover of AMO, the three-electron oxidation of NH_2OH by HAO furnishes one net electron for cellular respiration. Stoichiometric production of NO_2^- from NH_3 by AOB likely involves an additional, but as yet unidentified, third enzyme in the pathway. Cytotoxicity of both NO and NH_2OH may be managed by cytochrome P460, a monoheme enzyme that selectively produces N_2O from the reaction of NH_2OH with Fe-bound NO (77). NO_2^- is itself cellular fuel for nitrite-oxidizing bacteria that effect the two-electron oxidation of NO_2^- to NO_3^- and for the recently discovered "complete ammonia oxidation" (or comammox) bacteria that effect the eight-electron oxidation of NH_3 to NO_3^- (78).

Opportunity abounds to extract broadly applicable insights through the study of the oxidative enzymes operative in nitrification. Establishing the mechanism of NH_3 oxidation by AMO remains a grand challenge, whose difficulty is exacerbated by the integral membrane nature of the enzyme and the sluggish growth of its host organisms. The end goal is tantalizing technology: selective hydroxylation of relatively inert bonds that is generalizable beyond N species—especially considering that enzymes related to AMO, such as particulate methane monooxygenase, effect similar transformations of unactivated alkanes (79).

Mechanistic understanding of NH_2OH and NO_2^- oxidation offers insights into proton-coupled, multi-electron transformations that are crucial to redox catalysis. NH_2OH is an energetic, cytotoxic metabolite that AOB and AOA harness as a source of reducing equivalents. The absence of requisite c-heme biosynthesis pathways in AOA necessitates substitution of HAO for either a nonheme or a Cu-based enzyme for NH_2OH oxidation (80). Thus, opportunities exist to fill in a missing link in the biogeochemical nitrogen cycle while expanding understanding of how transition metals can transduce energy from redox-active small molecules.

Concluding outlook

Since the beginning of the 20th century, both the ability to produce and the need to remediate nitrogen-containing compounds have been coupled to fossil fuels. Demands for greater energy efficiency, smaller and more flexible distributed processes, and environmental protection provide growing impetus to expand the scope of practically viable oxidative and reductive transformations of nitrogen that are not driven by fossil fuels. Opportunities exist to identify new and radically improved pathways, but progress in this regard will require a molecular-level understanding of nitrogen transformation reactions, as well as the translation of these core insights to the discovery of new catalytic systems and alternative means of delivering the energy needed to drive those reactions. These advances will emerge through the collective knowledge and insights to be gained from fundamental research that integrates experiments and theory in hetero- and homogeneous catalysis, photon- and electron-driven processes, and biology.

REFERENCES AND NOTES

1. J. W. Erisman, M. A. Sutton, J. Galloway, Z. Klimont, W. Winiwarter, How a Century of Ammonia Synthesis Changed the World. *Nat. Geosci.* **1**, 636–639 (2008). doi: 10.1038/ngeo325
2. V. Smil, Detonator of the population explosion. *Nature* **400**, 415 (1999). doi: 10.1038/22672
3. International Fertilizer Industry Association, "Ammonia production: Moving towards maximum efficiency and lower GHG emissions," 2014; www.fertilizer.org/images/Library_Downloads/2014_ifa_ff_ammonia_emissions_july.pdf.
4. G. Booth, H. Zollinger, K. McLaren, W. Sharples, A. E. Westwall, in *Ullmanns Encyclopedia of Industrial Chemistry*, B. Elvers, S. Hawkins, G. Schulz, Eds. (2002), vol. 17.
5. I. Rafiqul, C. Weber, B. Lehmann, A. Voss, Energy efficiency improvements in ammonia production - Perspectives and uncertainties. *Energy* **30**, 2487–2504 (2005). doi: 10.1016/j.energy.2004.12.004
6. J. A. Miller, C. T. Bowman, Mechanism and modeling of nitrogen chemistry in combustion. *Pror. Energy Combust. Sci.* **15**, 287–338 (1989). doi: 10.1016/0360-1285(89)90017-8
7. B. S. Patil, Q. Wang, V. Hessel, J. Lang, Plasma N_2-fixation: 1900–2014. *Catal. Today* **256**, 49–66 (2015). doi: 10.1016/j.cattod.2015.05.005
8. C. H. Bartholomew, R. J. Farrauto, *Fundamentals of Industrial Catalytic Processes* (Wiley, 2006).
9. P. A. Leighton, *Photochemistry of Air Pollution* (Academic Press, 1961).
10. M. Shelef, R. W. McCabe, Twenty-five years after introduction of automotive catalysts: What next? *Catal. Today* **62**, 35–50 (2000). doi: 10.1016/S0920-5861(00)00407-7
11. M. Zammit, C. L. DiMaggio, C. H. Kim, C. Lambert, G. G. Muntean, C. H. F. Peden, J. E. Parks, K. Howden, "Future automotive aftertreatment solutions: The 150°C challenge workshop report" (PNNL-22815, Pacific Northwest National Laboratory, 2013).
12. National Research Council, *Clean Coastal Waters: Understanding and Reducing the Effects of Nutrient Pollution* (National Academies Press, 2000).
13. B. P. Chaplin *et al.*, Critical review of Pd-based catalytic treatment of priority contaminants in water. *Environ. Sci. Technol.* **46**, 3655–3670 (2012). doi: 10.1021/es204087q; pmid: 22369144
14. The following standard states at 298 K and 1 bar are adopted throughout this paper: pure liquid for water, 1 mol/liter for aqueous solutes, and gaseous state for all other compounds. These states are designated by the subscripts (liq), (aq), and (g), respectively. Aqueous nitric acid is considered to be fully ionized.
15. D. D. Wagman *et al.*, The NBS tables of chemical thermodynamic properties. *J. Phys. Chem. Ref. Data* **11**, suppl. 2 (1982).
16. L. V. Gurvich, I. V. Veyts, C. B. Alcock, Eds., *Thermochemical Properties of Individual Substances* (Hemisphere Publishing, 1989), vol. 1, part 2.
17. U.S. Geological Survey, Mineral Commodity Summaries (2018); https://minerals.usgs.gov/minerals/pubs/mcs/2018/mcs2018.pdf.
18. Center for Industry Education Collaboration, "Nitric acid," The Essential Chemical Industry–Online (2016); http://essentialchemicalindustry.org/chemicals/nitric-acid.html.
19. D. A. Armstrong *et al.*, Standard electrode potentials involving radicals in aqueous solution: Inorganic radicals (IUPAC Technical Report). *Pure Appl. Chem.* **87**, 1139–1150 (2015). doi: 10.1515/pac-2014-0502
20. A. Logadottir *et al.*, The Brønsted–Evans–Polanyi Relation and the Volcano Plot for Ammonia Synthesis over Transition Metal Catalysts. *J. Catal.* **197**, 229–231 (2001). doi: 10.1006/jcat.2000.3087
21. C. J. H. Jacobsen *et al.*, Catalyst design by interpolation in the periodic table: Bimetallic ammonia synthesis catalysts. *J. Am. Chem. Soc.* **123**, 8404–8405 (2001). doi: 10.1021/ja010963d; pmid: 11516293
22. A. J. Medford *et al.*, From the Sabatier Principle to a Predictive Theory of Transition-Metal Heterogeneous Catalysis. *J. Catal.* **328**, 36–42 (2015). doi: 10.1016/j.jcat.2014.12.033
23. E. J. Vicente, D. R. Dean, Keeping the nitrogen-fixation dream alive. *Proc. Natl. Acad. Sci. U.S.A.* **114**, 3009–3011 (2017). doi: 10.1073/pnas.1701560114; pmid: 28283657
24. R. D. Milton *et al.*, The In Vivo Potential-Regulated Protective Protein of Nitrogenase in Azotobacter vinelandii Supports Aerobic Bioelectrochemical Dinitrogen Reduction In Vitro. *J. Am. Chem. Soc.* **139**, 9044–9052 (2017). doi: 10.1021/jacs.7b04893; pmid: 28595003
25. B. M. Hoffman, D. Lukoyanov, Z.-Y. Yang, D. R. Dean, L. C. Seefeldt, Mechanism of nitrogen fixation by nitrogenase: The next stage. *Chem. Rev.* **114**, 4041–4062 (2014). doi: 10.1021/cr400641x; pmid: 24467365
26. I. Čorić, P. L. Holland, Insight into the Iron-Molybdenum Cofactor of Nitrogenase from Synthetic Iron Complexes with Sulfur, Carbon, and Hydride Ligands. *J. Am. Chem. Soc.* **138**, 7200–7211 (2016). doi: 10.1021/jacs.6b00747; pmid: 27171599

27. P. Paengnakorn *et al*., Infrared spectroscopy of the nitrogenase MoFe protein under electrochemical control: Potential-triggered CO binding. *Chem. Sci.* **8**, 1500–1505 (2017). doi: 10.1039/C6SC02860H; pmid: 28616146
28. R. D. Milton *et al*., Bioelectrochemical Haber-Bosch Process: An Ammonia-Producing H_2 /N_2 Fuel Cell. *Angew. Chem. Int. Ed.* **56**, 2680–2683 (2017). doi: 10.1002/anie.201612500; pmid: 28156040
29. K. Danyal *et al*., Uncoupling nitrogenase: Catalytic reduction of hydrazine to ammonia by a MoFe protein in the absence of Fe protein-ATP. *J. Am. Chem. Soc.* **132**, 13197–13199 (2010). doi: 10.1021/ja1067178; pmid: 20812745
30. N. Khadka *et al*., Mechanism of Nitrogenase H_2 Formation by Metal-Hydride Protonation Probed by Mediated Electrocatalysis and H/D Isotope Effects. *J. Am. Chem. Soc.* **139**, 13518–13524 (2017). doi: 10.1021/jacs.7b07311; pmid: 28851217
31. K. A. Brown *et al*., Light-driven dinitrogen reduction catalyzed by a CdS:nitrogenase MoFe protein biohybrid. *Science* **352**, 448–450 (2016). doi: 10.1126/science.aaf2091; pmid: 27102481
32. A. D. Allen, C. V. Senoff, Nitrogenopentammineruthenium(II) complexes. *Chem. Commun.* **1965**, 621–622 (1965). doi: 10.1039/c19650000621
33. J. Chatt, J. R. Dilworth, R. L. Richards, Recent Advances in the Chemistry of Nitrogen Fixation. *Chem. Rev.* **78**, 589–625 (1978). doi: 10.1021/cr60316a001
34. C. J. Pickett, J. Talarmin, Electrosynthesis of ammonia. *Nature* **317**, 652–653 (1985). doi: 10.1038/317652a0
35. D. V. Yandulov, R. R. Schrock, Catalytic reduction of dinitrogen to ammonia at a single molybdenum center. *Science* **301**, 76–78 (2003). doi: 10.1126/science.1085326; pmid: 12843387
36. K. Arashiba, Y. Miyake, Y. Nishibayashi, A molybdenum complex bearing PNP-type pincer ligands leads to the catalytic reduction of dinitrogen into ammonia. *Nat. Chem.* **3**, 120–125 (2011). doi: 10.1038/nchem.906; pmid: 21258384
37. K. Arashiba *et al*., Catalytic reduction of dinitrogen to ammonia by use of molybdenum-nitride complexes bearing a tridentate triphosphine as catalysts. *J. Am. Chem. Soc.* **137**, 5666–5669 (2015). doi: 10.1021/jacs.5b02579; pmid: 25879994
38. J. S. Anderson, J. Rittle, J. C. Peters, Catalytic conversion of nitrogen to ammonia by an iron model complex. *Nature* **501**, 84–87 (2013). doi: 10.1038/nature12435; pmid: 24005414
39. M. J. Chalkley, T. J. Del Castillo, B. D. Matson, J. P. Roddy, J. C. Peters, Catalytic N_2-to-NH_3 Conversion by Fe at Lower Driving Force: A Proposed Role for Metallocene-Mediated PCET. *ACS Cent. Sci.* **3**, 217–223 (2017). doi: 10.1021/acscentsci.7b00014; pmid: 28386599
40. V. Rosca, M. Duca, M. T. de Groot, M. T. M. Koper, Nitrogen cycle electrocatalysis. *Chem. Rev.* **109**, 2209–2244 (2009). doi: 10.1021/cr8003696; pmid: 19438198
41. F. Zhou *et al*., Electro-synthesis of ammonia from nitrogen at ambient temperature and pressure in ionic liquids. *Energy Environ. Sci.* **10**, 2516–2520 (2017). doi: 10.1039/C7EE02716H
42. R. Lan, S. Tao, Electrochemical synthesis of ammonia directly from air and water using a $Li^+/H^+/NH_4^+$ mixed conducting electrolyte. *RSC Adv.* **3**, 18016–18021 (2013).
43. J. M. McEnaney *et al*., Ammonia synthesis from N2 and H2O using a lithium cycling electrification strategy at atmospheric pressure. *Energy Environ. Sci.* **10**, 1621–1630 (2017). doi: 10.1039/C7EE01126A
44. S. Licht *et al*., Ammonia synthesis by N_2 and steam electrolysis in molten hydroxide suspensions of nanoscale Fe_2O_3. *Science* **345**, 637–640 (2014). doi: 10.1126/science.1254234; pmid: 25104378
45. V. Kyriakou, I. Garagounis, E. Vasileiou, A. Vourros, M. Stoukides, Progress in the electrochemical synthesis of ammonia. *Catal. Today* **286**, 2–13 (2017). doi: 10.1016/j.cattod.2016.06.014
46. R. J. Hamers, J. A. Bandy, D. Zhu, L. Zhang, Photoemission from diamond films and substrates into water: Dynamics of solvated electrons and implications for diamond photoelectrochemistry. *Faraday Discuss.* **172**, 397–411 (2014). doi: 10.1039/C4FD00039K; pmid: 25413482
47. C. Guo, J. Ran, A. Vasileff, S.-Z. Qiao, Rational design of electrocatalysts and photo(electro)catalysts for nitrogen reduction to ammonia (NH3) under ambient conditions. *Energy Environ. Sci.* **11**, 45–56 (2018). doi: 10.1039/C7EE02220D
48. A. J. Medford, M. C. Hatzell, Photon-Driven Nitrogen Fixation: Current Progress, Thermodynamic Considerations, and Future Outlook. *ACS Catal.* **7**, 2624–2643 (2017). doi: 10.1021/acscatal.7b00439
49. J. Liu *et al*., Nitrogenase-mimic iron-containing chalcogels for photochemical reduction of dinitrogen to ammonia. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 5530–5535 (2016). doi: 10.1073/pnas.1605512113; pmid: 27140630
50. J. H. van Helden *et al*., Detailed study of the plasma-activated catalytic generation of ammonia in N_2-H_2 plasmas. *J. Appl. Phys.* **101**, 043305 (2007). doi: 10.1063/1.2645828
51. J. Hong, S. Prawer, A. B. Murphy, Plasma Catalysis as an Alternative Route for Ammonia Production: Status, Mechanisms, and Prospects for Progress. *ACS Sustain. Chem. Eng.* **6**, 15–31 (2018). doi: 10.1021/acssuschemeng.7b02381
52. E. C. Neyts, K. K. Ostrikov, M. K. Sunkara, A. Bogaerts, Plasma Catalysis: Synergistic Effects at the Nanoscale. *Chem. Rev.* **115**, 13408–13446 (2015). doi: 10.1021/acs.chemrev.5b00362; pmid: 26619209
53. J. C. Whitehead, Plasma-catalysis: The known knowns, the known unknowns and the unknown unknowns. *J. Phys. D Appl. Phys.* **49**, 243001 (2016). doi: 10.1088/0022-3727/49/24/243001
54. P. Mehta *et al*., Overcoming Ammonia Synthesis Scaling Relations with Plasma-enabled Catalysis. *Nature Catal.* **1**, 269–275 (2018). doi: 10.1038/s41929-018-0045-1
55. J. Adanez, A. Abad, F. Garcia-Labiano, P. Gayan, L. F. de Diego, Progress in Chemical-Looping Combustion and Reforming Technologies. *Pror. Energy Combust. Sci.* **38**, 215–282 (2012). doi: 10.1016/j.pecs.2011.09.001
56. R. Michalsky, A. M. Avram, B. A. Peterson, P. H. Pfromm, A. A. Peterson, Chemical looping of metal nitride catalysts: Low-pressure ammonia synthesis for energy storage. *Chem. Sci.* **6**, 3965–3974 (2015). doi: 10.1039/C5SC00789E; pmid: 29218166
57. M. E. Gálvez, M. Halmann, A. Steinfeld, Ammonia production via a two-step Al2O3/AlN thermochemical cycle. 1. Thermodynamic, environmental, and economic analyses. *Ind. Eng. Chem. Res.* **46**, 2042–2046 (2007). doi: 10.1021/ie061550u
58. R. Michalsky, P. Pfromm, An ionicity rationale to design solid phase metal nitride reactants for solar ammonia production. *J. Phys. Chem. C* **116**, 23243–23251 (2012). doi: 10.1021/jp307382r
59. A. Lyngfelt, B. Leckner, A 1000 MW_{th} boiler for chemical-looping combustion of solid fuels - Discussion of design and cost. *Appl. Energy* **157**, 475–487 (2015). doi: 10.1016/j.apenergy.2015.04.057
60. W. F. Schneider, in *Environmental Catalysis*, V. H. Grassian, Ed. (Wiley-VCH, 2004), pp. 233–268.
61. A. M. Dean, J. W. Bozzelli, in *Gas-Phase Combustion Chemistry*, W. C. Gardiner, Ed. (Springer, 2000), pp. 125–341.
62. N. Cherkasov, A. O. Ibhadon, P. Fitzpatrick, A review of the existing and alternative methods for greener nitrogen fixation. *Chem. Eng. Process. Process Intensif.* **90**, 24–33 (2015). doi: 10.1016/j.cep.2015.02.004
63. A. Fridman, *Plasma Chemistry* (Cambridge Univ. Press, 2008).
64. V. D. Rusanov, A. A. Fridman, G. V. Sholin, The physics of a chemically active plasma with nonequilibrium vibrational excitation of molecules. *Sov. Phys. Usp.* **24**, 447–474 (1981). doi: 10.1070/PU1981v024n06ABEH004884
65. M. Carmo, D. Fritz, J. Mergel, D. Stolten, A comprehensive review on PEM water electrolysis. *Int. J. Hydrogen Energy* **38**, 4901–4934 (2013).
66. V. I. Pârvulescu, P. Grange, B. Delmon, Catalytic removal of NO. *Catal. Today* **46**, 233–316 (1998). doi: 10.1016/S0920-5861(98)00399-X
67. D. Lionetti, G. de Ruiter, T. Agapie, A trans-Hyponitrite Intermediate in the Reductive Coupling and Deoxygenation of Nitric Oxide by a Tricopper-Lewis Acid Complex. *J. Am. Chem. Soc.* **138**, 5008–5011 (2016). doi: 10.1021/jacs.6b01083; pmid: 27028157
68. M. H. Groothaert, J. A. van Bokhoven, A. A. Battiston, B. M. Weckhuysen, R. A. Schoonheydt, Bis(μ-oxo)dicopper in Cu-ZSM-5 and its role in the decomposition of NO: A combined in situ XAFS, UV-vis-near-IR, and kinetic study. *J. Am. Chem. Soc.* **125**, 7629–7640 (2003). doi: 10.1021/ja029684w; pmid: 12812505
69. C. Paolucci *et al*., Dynamic multinuclear sites formed by mobilized copper ions in NO_x selective catalytic reduction. *Science* **357**, 898–903 (2017). doi: 10.1126/science.aan5630; pmid: 28818971
70. Y. J. Mergler, B. E. Nieuwenhuys, NO reduction by H_2 over promoted Pt catalysts. *Appl. Catal. B* **12**, 95–110 (1997). doi: 10.1016/S0926-3373(96)00068-9
71. R. M. Martínez-Espinosa, J. A. Cole, D. J. Richardson, N. J. Watmough, Enzymology and ecology of the nitrogen cycle. *Biochem. Soc. Trans.* **39**, 175–178 (2011). doi: 10.1042/BST0390175; pmid: 21265768
72. P. Tavares, A. S. Pereira, J. J. G. Moura, I. Moura, Metalloenzymes of the denitrification pathway. *J. Inorg. Biochem.* **100**, 2087–2100 (2006). doi: 10.1016/j.jinorgbio.2006.09.003; pmid: 17070915
73. A. B. Hooper, D. Arciero, D. Bergmann, M. P. Hendrich, in *Respiration in Archaea and Bacteria: Diversity of Prokaryotic Respiratory Systems*, D. Zannoni, Ed., vol. 16 of *Advances in Photosynthesis and Respiration* (Springer, 2004), pp. 121–147.
74. M. Könneke *et al*., Isolation of an autotrophic ammonia-oxidizing marine archaeon. *Nature* **437**, 543–546 (2005). doi: 10.1038/nature03911; pmid: 16177789
75. N. Vajrala *et al*., Hydroxylamine as an intermediate in ammonia oxidation by globally abundant marine archaea. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 1006–1011 (2013). doi: 10.1073/pnas.1214272110; pmid: 23277575
76. J. D. Caranto, K. M. Lancaster, Nitric oxide is an obligate bacterial nitrification intermediate produced by hydroxylamine oxidoreductase. *Proc. Natl. Acad. Sci. U.S.A.* **114**, 8217–8222 (2017). doi: 10.1073/pnas.1704504114; pmid: 28716929
77. J. D. Caranto, A. C. Vilbert, K. M. Lancaster, Nitrosomonas europaea cytochrome P460 is a direct link between nitrification and nitrous oxide emission. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 14704–14709 (2016). doi: 10.1073/pnas.1611051113; pmid: 27856762
78. H. Daims *et al*., Complete nitrification by Nitrospira bacteria. *Nature* **528**, 504–509 (2015). doi: 10.1038/nature16461; pmid: 26610024
79. T. J. Lawton, J. Ham, T. Sun, A. C. Rosenzweig, Structural conservation of the B subunit in the ammonia monooxygenase/particulate methane monooxygenase superfamily. *Proteins* **82**, 2263–2267 (2014). doi: 10.1002/prot.24535; pmid: 24523098
80. C. B. Walker *et al*., Nitrosopumilus maritimus genome reveals unique mechanisms for nitrification and autotrophy in globally distributed marine crenarchaea. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 8818–8823 (2010). doi: 10.1073/pnas.0913533107; pmid: 20421470

ACKNOWLEDGMENTS

Funding: This article evolved from presentations and discussions at the workshop "Frontiers, Opportunities, and Challenges in Biochemical and Chemical N_2 Activation" held in October 2016 in Gaithersburg, Maryland. The workshop was sponsored by the Council on Chemical Sciences, Geosciences and Biosciences of the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences. The authors thank the members of the Council for their encouragement and assistance in developing this workshop. In addition, the authors are indebted to the agencies responsible for funding their individual research efforts, without which this article would not have been possible. **Competing interests:** The authors have no competing interests.

10.1126/science.aar6611

RESEARCH ARTICLE SUMMARY

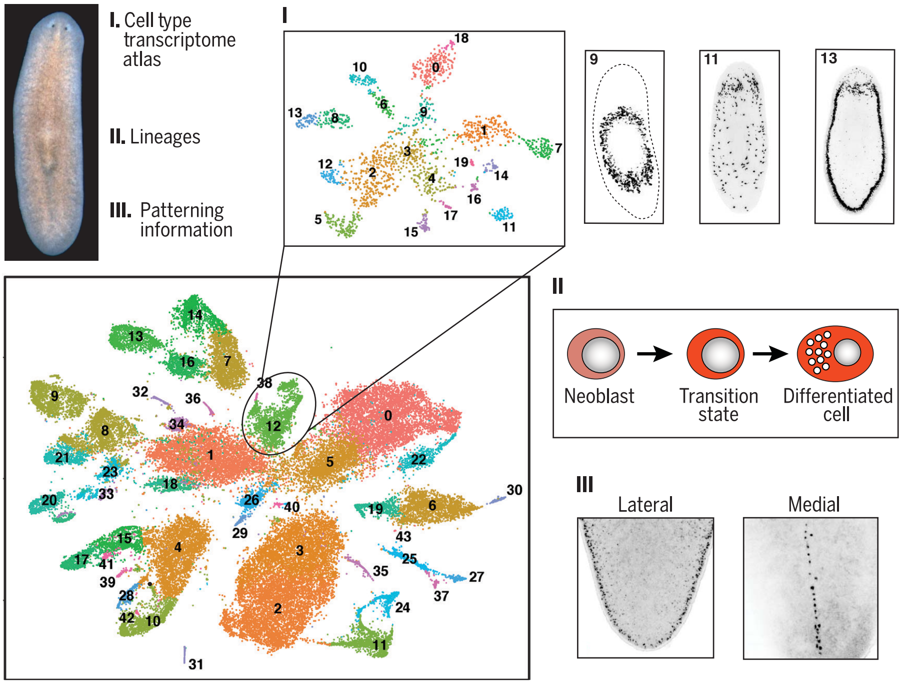
SINGLE-CELL ANALYSIS

Cell type transcriptome atlas for the planarian *Schmidtea mediterranea*

Christopher T. Fincher, Omri Wurtzel, Thom de Hoog, Kellie M. Kravarik, Peter W. Reddien*

INTRODUCTION: The complete sequence of animal genomes has had a transformative impact on biological research. Whereas the genome sequence of an organism contains the information for its development and physiology, the transcriptomes (the sets of actively transcribed genes) of the cell types in an organism define how the genome is used for the unique functions of its cells. Cell number and complexity have historically made the identification of all cell types, much less their transcriptomes, an extreme challenge for most multicellular organisms. Recent advances in single-cell RNA sequencing (SCS) have greatly enhanced the ability to determine cell type transcriptomes, with SCS of thousands of cells readily achievable.

RATIONALE: We reasoned that it might be possible, given these advances, to determine the transcriptomes of essentially every cell type of a complete organism possessing an unknown number of cell types. The planarian *Schmidtea mediterranea*, famous for its regeneration ability, is an attractive case study for such an undertaking. Planarians possess a complex anatomy with diverse differentiated cell types, including many found across animals. Furthermore, planarians contain a proliferating cell population called neoblasts that includes pluripotent stem cells. Neoblasts mediate regeneration and constitutive tissue turnover. Consequently, lineage precursors for essentially all differentiated cells types are also present in adults. Finally, planarians constitutively express positional information guiding tissue turnover. Therefore, comprehensive SCS at a single time point (the adult) could allow transcriptome determination for all differentiated cell types and for lineage precursors, and could identify patterning information that guides new cell production and organization. Capturing this information in most organisms would require sampling adults and many transient embryonic stages.



An atlas of planarian cell type transcriptomes. High-throughput single-cell RNA sequencing of adult planarians reveals a cell type transcriptome atlas that includes rare cell types and many novel cell populations, cellular transition states, and patterning information, as demonstrated by two regionally expressed genes in muscle.

RESULTS: We used the SCS method Drop-seq to determine the transcriptomes for 66,783 individual cells from adult planarians. We locally saturated cell type coverage by iteratively sequencing distinct body regions and assessing the frequency of known rare cell types in the data. Clustering the cells by shared gene expression grouped cells into broad tissue classes. Subclustering of each broad tissue type in isolation enabled separation of cells into the cell populations constituting each tissue. These analyses enabled the identification of a previously unidentified tissue group and the classification of poorly characterized tissues into their constituent cell types, including numerous previously unknown cell types. Transcriptomes were identified for many rare cell types, including those that exist as rarely as ~10 cells in an animal that has 10^5 to 10^6 cells, which suggests that near-to-complete cellular saturation was reached. In addition, transcriptomes for known and novel lineage precursors, from pluripotent stem cell to differentiated cell types, were generated. Precursor transcriptomes identified transcription factors required for maintenance of associated differentiated cells during homeostatic cell turnover. Finally, the data were used to identify genes regionally expressed in muscle, which is the site of planarian patterning gene expression.

ON OUR WEBSITE

Read the full article at http://dx.doi.org/10.1126/science.aaq1736

CONCLUSION: We successfully used SCS to generate transcriptomes for most to all cells of a complete organism. This resource provides a wealth of data regarding the cellular site of expression of thousands of conserved genes and the transcriptomes for cell types widely used in animals. These data will inform studies of these genes and cell types broadly, and will provide a resource for the fields of planarian biology and comparative evolutionary biology. This work also provides a template for the generation of cell type transcriptome atlases, which can be applied to a large array of organisms. ■

The list of author affiliations is available in the full article online.
*Corresponding author. Email: reddien@wi.mit.edu

Cite this article as C. T. Fincher *et al.*, *Science* **360**, eaaq1736 (2018). DOI: 10.1126/science.aaq1736

Fincher *et al.*, *Science* **360**, 874 (2018) 25 May 2018 **1 of 1**

RESEARCH ARTICLE

SINGLE-CELL ANALYSIS

Cell type transcriptome atlas for the planarian *Schmidtea mediterranea*

Christopher T. Fincher,[1,2,3] **Omri Wurtzel,**[1,2] **Thom de Hoog,**[1,2]
Kellie M. Kravarik,[1,2,3] **Peter W. Reddien**[1,2,3]*

The transcriptome of a cell dictates its unique cell type biology. We used single-cell RNA sequencing to determine the transcriptomes for essentially every cell type of a complete animal: the regenerative planarian *Schmidtea mediterranea.* Planarians contain a diverse array of cell types, possess lineage progenitors for differentiated cells (including pluripotent stem cells), and constitutively express positional information, making them ideal for this undertaking. We generated data for 66,783 cells, defining transcriptomes for known and many previously unknown planarian cell types and for putative transition states between stem and differentiated cells. We also uncovered regionally expressed genes in muscle, which harbors positional information. Identifying the transcriptomes for potentially all cell types for many organisms should be readily attainable and represents a powerful approach to metazoan biology.

The complete sequence of animal genomes, such as that of *Caenorhabditis elegans* reported in 1998 and humans in 2001, has had an immeasurable impact on research (*1*–*3*). Whereas the genome sequence of an organism contains the information for its development and physiology, the transcriptomes (the sets of actively transcribed genes) of the cell types in an organism define how the genome is used for the unique functions of its cells. Recent advances in RNA sequencing of individual cells have greatly enhanced the ability to determine cell type transcriptomes (*4*, *5*), and single-cell RNA sequencing (SCS) of thousands of cells has become readily achievable (*6*). For example, the transcriptomes of most cell types of complete *C. elegans* L2 larvae and numerous mouse cells were recently reported with this approach (*7*, *8*). We reasoned that it might be possible, given these advances, to determine the transcriptomes of essentially every cell type of a complete adult organism possessing an unknown number of cell types.

Multicellular organisms can have many millions of cells and hundreds of different cell types, and the cellular composition of organisms varies markedly over the course of development. This complexity has historically made the identification of all cell types, much less their transcriptomes, for most multicellular organisms an extreme challenge. The planarian *Schmidtea mediterranea* is an attractive case study organism for which to generate the transcriptomes for all cells in an animal. Planarians are famous for their ability to regenerate essentially any missing body part, and they possess a complex body plan containing many characterized cell types (*9*, *10*). Despite this complexity, with an average planarian possessing ~10^5 to 10^6 cells (*11*), planarians are smaller with simpler anatomy than humans and many other model systems such as mice. Planarians are also easily dissociated into single-cell suspensions, allowing potential characterization of all cells. Because some planarian cell types, such as glia (*12*, *13*), have only recently been defined with molecular markers, it is probable that undescribed planarian cell types exist. The combination of known and potentially unknown cell types is attractive for developing approaches that can apply to diverse organisms with varying amounts of available cell type information. Planarians possess a population of proliferative cells called neoblasts that contain pluripotent stem cells, enabling their ability to regenerate and replace aged cells in tissue turnover (*14*). Neoblasts are the only cycling somatic cells and the source of all new tissue. Neoblasts contain multiple classes of specialized cells, with transcription factors expressed to specify cell fate (*15*, *16*). Because of the constant turnover of planarian tissues, essentially all stages of all cell lineages, from pluripotent stem cell to differentiated cell, are anticipated to be present in the adult (*9*, *17*).

Planarians also constitutively and regionally express dozens of genes that have roles in positional information (*18*). These genes, referred to as positional control genes (PCGs), are expressed in a complex spatial map spanning anterior-posterior (AP), medial-lateral (ML), and dorsal-ventral (DV) axes (*18*), and their expression is largely restricted to muscle (*19*). PCGs are hypothesized to constitute instructions for the maintenance and regeneration of the body plan. Because of these features, comprehensive SCS at a single time point (the adult) could allow transcriptome identification for all differentiated cell types, lineage precursors for these cells, and the patterning information that guides new cell production and organization. To capture this information in most organisms would require sampling the adult and many transient stages of embryogenesis.

Single-cell RNA sequencing of 50,562 planarian cells

Planarians have a complex internal anatomy including a brain, ventral nerve cords, peripheral nervous system, epidermis, intestine, muscle, an excretory system (the protonephridia), and a centrally located pharynx (*10*). These major tissues are composed of multiple different cell types that, together with other gland and accessory cells, constitute the planarian anatomy.

To detect planarian cell types and states in an unbiased manner, including rare cell types, we used the SCS method Drop-seq (*6*) to determine the transcriptomes for 50,562 individual cells from adults (Fig. 1A, fig. S1A, and table S1). Planarians contain 10^5 to 10^6 cells (*11*), and yet some cell types are extremely rare, such as the ~100 photoreceptor neurons of eyes (*20*). Given such rarity, sequencing random cells from entire animals might not reach cell type saturation with even 10^5 cells sequenced. Therefore, we divided animals into five sections (head, prepharyngeal region, trunk with pharynx removed, tail, and the pharynx itself) and cells from each region were dissociated, sorted by flow cytometry, and sequenced (Fig. 1A, fig. S1A, and table S1). Sequences were aligned to a previously assembled transcriptome (*21*). We targeted cell type saturation by assessing coverage of known, rare cell types during iterative rounds of cell isolation and sequencing in a region-by-region approach. In total, 25 separate Drop-seq runs were completed, yielding cells with an average of 3020 unique molecular identifiers (UMIs) and 1404 genes (~13% of the estimated detection limit) (fig. S1, A to C, table S1, and supplementary materials).

Genes with high variance and expression across cells were used to generate informative principal components using Seurat (*6*, *22*). Cells were clustered using Seurat into 44 distinct major clusters using a graph-based clustering approach and were visualized by applying t-distributed stochastic neighbor embedding on transcriptomes (t-SNE) (Fig. 1B and fig. S1D). Cells from different regions were largely interspersed in the t-SNE plots, except for cells from the pharynx, which contains many unique cell types (fig. S2A). Cell doublets were scarce within the data and did not affect clustering results (fig. S2, B to D). To determine the identity of each cluster, we identified cluster-specific genes by means of a receiver operating characteristic curve analysis and a likelihood ratio test based on zero-inflated data (table S2) (*23*). Expression of established cell type markers within each cluster and fluorescence in situ hybridization (FISH) with cluster-specific markers enabled

---

[1]Whitehead Institute for Biomedical Research, Cambridge, MA 02142, USA. [2]Howard Hughes Medical Institute, Massachusetts Institute of Technology, Cambridge, MA 02139, USA. [3]Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139, USA.
*Corresponding author. Email: reddien@wi.mit.edu

cluster assignment to one of eight previously identified planarian tissue classes: protonephridia, neural, epidermis, intestine, pharynx, muscle, neoblast, and parenchymal (Fig. 1C). The parenchymal class was previously termed "parapharyngeal" because of localization of some enriched markers around the pharynx (24). However, most cell populations within this class exhibit broader localization in the planarian parenchyma. We also identified a ninth group of clusters marked by *CTSL2* (dd175) expression (Fig. 1D). *CTSL2* (dd175) FISH revealed cells with long processes distributed broadly. We designated this group of clusters the *cathepsin*$^+$ class. Hierarchical clustering of a subset of 5000 cells by Euclidean distance, independently of Seurat, recapitulated assignment of cells into these nine tissue classes (fig. S3).

Clusters representing the major planarian tissue classes were generally heterogeneous in terms of gene expression. For example, neural clusters contained a large number of known neuronal cell types, which suggests that multiple distinct cell types could be identified within each major cluster (fig. S4). Therefore, we systematically subclustered each major cluster group (Figs. 2 to 6), identifying >150 subclusters, and determined genes with enriched expression in cells of each subcluster (table S2). Subclustering proved a powerful approach to defining the collection of cell types that constituted each major cluster and identified candidate transition states between stem cells and differentiated cells.

Progenitors in planarian cell lineages

Neoblasts are abundant and express canonical marker genes such as *smedwi-1* (25), *vasa* (26), and *bruli* (27) (Fig. 1C and fig. S5, A and B). Neoblasts are cycling cells and consequently show enrichment in expression of S/G_2/M cell cycle markers (fig. S5C). To identify the transcriptomes

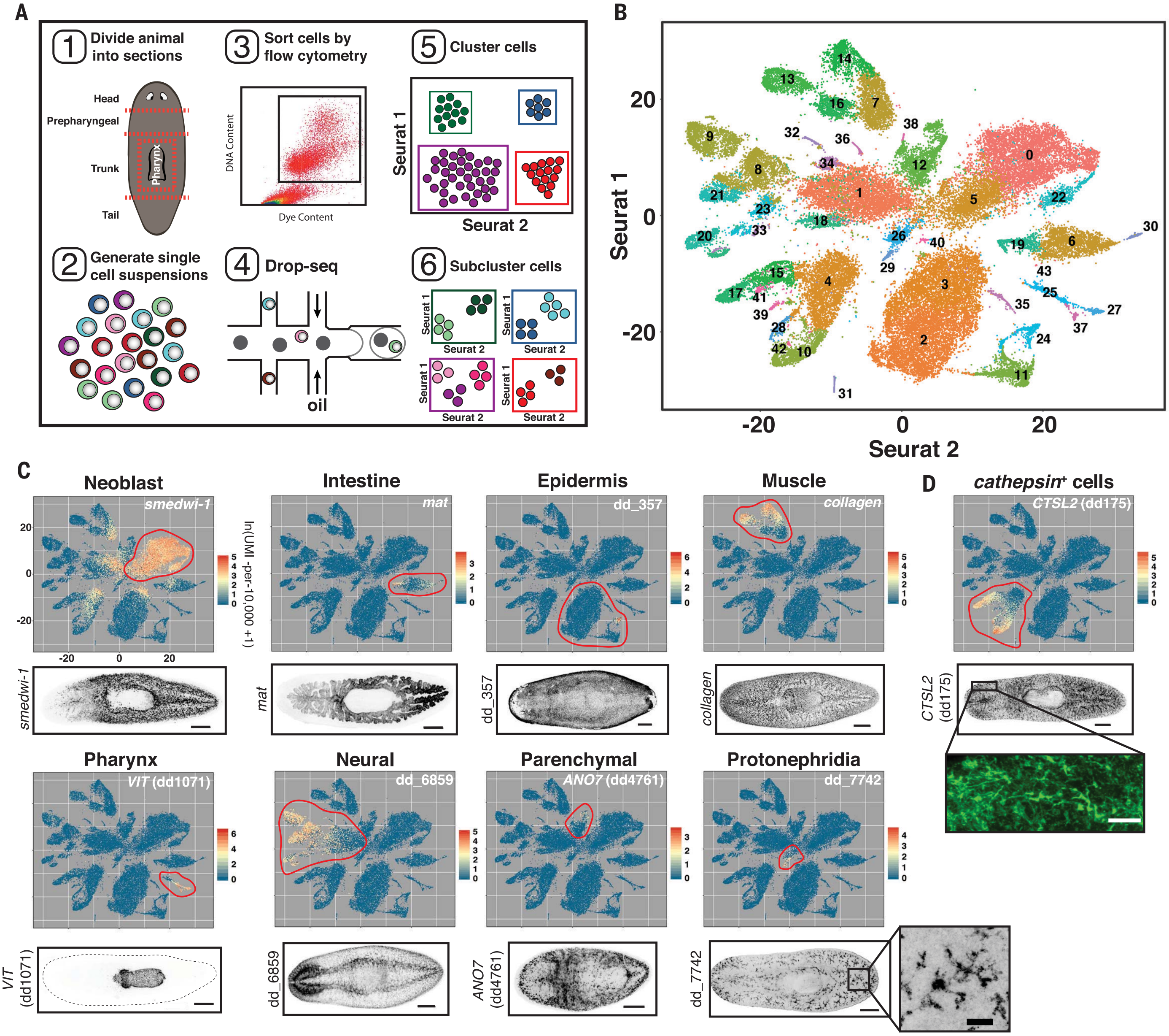


Fig. 1. Drop-seq of 50,562 planarian cells. (**A**) Schematic illustrating the workflow used to isolate and cluster single cells. (**B**) t-SNE representation of 44 clusters generated from the data. (**C** and **D**) Upper panels: t-SNE plots colored according to gene expression (red, high; blue, low) for highly enriched genes from nine planarian tissue classes. Red outlines denote clusters assigned to that tissue class. Lower panels: FISH images for tissue-enriched genes. Scale bars: whole-animal images, 200 μm; insets, 50 μm.

of potential neoblast subpopulations, we selected in silico and subclustered 12,212 cells with *smedwi-1* expression of ≥2.5 [ln(UMI-per-10,000 + 1)] (Fig. 2A and fig. S6A). Resulting clusters on the left of the plot were enriched in S/G_2/M cell cycle markers (fig. S6B). These clusters included the previously characterized major specialized neoblast classes, including γ-neoblasts (intestine progenitors) and ζ-neoblasts (epidermis progenitors) ([28]) (Fig. 2B). A number of other subclusters were also identified, including one marked by expression of the contig dd_10988 (fig. S6, C and D). FISH confirmed that dd_10988 was expressed in a neoblast subset as well as in a number of *smedwi-1*$^-$ cells (Fig. 2C).

The large number of subclustered neoblasts facilitated transcriptome determination for candidate progenitors for many planarian tissues. Clusters to the right of the plot were marked by a G_1/G_0 cell cycle status and displayed expression of various tissue markers (Fig. 2D and fig. S6B). These included a population defined by expression of *POU2/3*, a marker for protonephridia-specialized neoblasts ([29]), and a number of subclusters expressing markers also expressed in specific differentiated tissues or their post-mitotic precursors, such as *ChAT* for the nervous system, *prog-1* for the epidermis, *ASCL4* (dd1854) for parenchymal cells, and *COL4A6A* (dd2337) for muscle (Fig. 2D and fig. S7A). Expression of these markers in *smedwi-1*$^+$ cells suggests that these cells could be transition states for those lineages. Several markers enriched in the dd_10988$^+$ subcluster, including dd_10988, were also expressed in cells of the two *smedwi-1*$^+$ neural subclusters, as well as in neural cells of the initial clustering (figs. S6C and S7, B and C), which suggests that the dd_10988$^+$ subcluster is enriched in neural progenitors. Likewise, many markers enriched in the *PLOD1* (dd3457)$^+$ subcluster were also expressed in the *smedwi-1*$^+$ muscle subcluster, which suggests that the *PLOD1* (dd3457)$^+$ subcluster is enriched in muscle progenitors (fig. S7D). *prox-1*, *hnf-4*, *nkx2.2*, and *gata4/5/6* encode transcription factors expressed in intestinal progenitors ([28]), and in these data all four genes were expressed in γ-neoblasts (Fig. 2, B and E, and fig. S7E), with *hnf-4*, *nkx2.2*, and *gata4/5/6* also expressed in intestinal clusters (Fig. 2E and fig. S7F). *hnf-4*, but not *prox-1*, *nkx2.2*, and *gata4/5/6*, was also expressed in a *smedwi-1*$^+$ cell cluster enriched in *CTSL2* (dd175) expression (the *cathepsin*$^+$ cell marker) and in differentiated *cathepsin*$^+$ cells (Fig. 2E and fig. S7G). The additional transcription factor–encoding genes *ETS1* (dd2092) and *FOXF1* (dd6910) were expressed with *hnf-4* in these cells and also displayed expression patterns similar to that of *CTSL2* (dd175) in the animal (Fig. 2F and fig. S8, A and B) and have recently been shown to regulate the planarian pigment cell lineage ([30]). Pigment cells clustered within the *cathepsin*$^+$ cell class in our data (see below). By FISH, *hnf-4* was indeed coexpressed with *nkx2.2* and *gata4/5/6* in the intestine, but was also coexpressed with *cathepsin*$^+$ cell markers (fig. S8, C to E), which suggests that *hnf-4* is expressed in two distinct lineages. These data demonstrate the utility of this approach for identifying potentially novel neoblast progenitor populations and the transcription factors that define them.

Some planarian neoblasts display pluripotency in clonal assays and are hypothesized to generate all lineage-committed neoblast subpopulations, and are called clonogenic neoblasts ([14]). We selected cells expressing high levels of *smedwi-1* but that excluded ζ- and γ-neoblasts [including subclusters 2, 9, dd_10988+, dd_6998+, dd_17796+, *SAMD15* (dd19710)+, dd_11221+, dd_13666+, and *PLOD1* (dd3457)+] and subclustered this set of neoblasts in isolation (fig. S9, A and B). A remnant ζ-neoblast population (clusters 4 and 6), as well as protonephridia progenitors (cluster 10) and the putative neural (clusters 2 and 5) and muscle (clusters 1 and 9) progenitor populations described above in the *smedwi-1*$^+$ cell subclustering, were identified (fig. S9C and table S2). Clusters 0, 3, 7, and 8 were largely devoid of specifically enriched markers (table S2). It is therefore possible that clonogenic neoblasts are defined by an absence of any tissue-specific markers, as opposed to the unique expression of specific genes.

When all cells were clustered together, numerous *smedwi-1*$^+$ cells were present regionally within each of the other eight major planarian tissue clusters (Fig. 1C). We reasoned that these *smedwi-1*$^+$ cells could represent progenitors for the cell types within each associated tissue cluster. We therefore examined these *smedwi-1*$^+$ cells after taking each tissue class in isolation and subclustering the data.

The planarian epidermis contains ciliated and nonciliated cells as well as dorsal-ventral boundary epidermis ([10], [28], [31]), and the lineage from ζ-neoblasts to epidermal cells is well characterized ([31–33]) (Fig. 2G). SCS reveals gene expression transitions during neoblast epidermal differentiation ([31]); subclustering 11,021 epidermal lineage cells (Fig. 1C) produced subclusters associated with each epidermal lineage stage (Fig. 2H and fig. S10, A and B). Plotting gene expression onto this t-SNE map showed a continuous progression from ζ-neoblast to differentiated cells (Fig. 2I and fig. S10C).

The gene dd_554 [SmedASXL_059179 in ([34])] is expressed in candidate pharynx progenitors ([34]) (*smedwi-1*$^+$ cells at the pharynx base) and in *smedwi-1*$^-$ cells within the pharynx ([34]) (Fig. 2J and fig. S11). Subclustering the 1083 nonmuscle, non-neuronal pharynx cluster cells (Fig. 1C) revealed that *smedwi-1*$^+$ cells sequenced from non-pharynx midbody tissue clustered with pharynx cells, despite not being part of the pharynx itself (Fig. 2, K and L). Because the pharynx lacks neoblasts, pharynx-specialized neoblasts must be outside of the pharynx. This clustering of neoblasts with pharynx cells clearly demonstrates the ability of SCS data clustering to associate lineage precursors with differentiated cells. Similarly, many dd_554$^+$ cells sequenced from outside of the pharynx clustered with pharynx cells (Fig. 2, K and L). Plotting *smedwi-1*/dd_554 expression onto pharyngeal subclusters revealed a progression from *smedwi-1*$^+$ cells isolated outside the pharynx to dd_554$^+$ cells isolated outside the pharynx to dd_554$^+$ cells isolated inside the pharynx to pharyngeal cells (Fig. 2, K and L). These epidermis and pharynx examples demonstrate how precursor stages within cell lineages can be identified from subclustering cells within a major tissue class. Because planarians constantly generate new differentiated cells for essentially all tissue types ([17], [20]), transcriptomes for lineage precursors for essentially every cell type in the body could in principle be studied with this approach.

Cell lineages for many planarian cell types are largely uncharacterized. After tissue type subclustering, *smedwi-1*$^+$ cells were present with locally high expression in resultant t-SNE plots; *smedwi-1* expression level gradually declined in cells across subclusters (Fig. 2, M and N). These *smedwi-1*$^+$ cells, similar to the epidermis and pharynx cases, could represent transition states between pluripotent neoblasts and differentiated cells for the various cells of the protonephridia, intestine, muscle, nervous system, parenchymal, and *cathepsin*$^+$ cells (Fig. 2, M and N). The *smedwi-1*$^+$ cells found within subclusters of the major tissue type classes generally displayed enriched expression of at least one transcription factor. For example, *smedwi-1* expression was high within cells at the center of the parenchymal cell t-SNE plot and displayed a graded decrease projecting in all directions into seven major parenchymal subclusters (Fig. 2N). Each projection was associated with enriched expression of one or more distinct transcription factors, identifying candidate transcription factors associated with the specification of different parenchymal cell types (Fig. 2O).

Subclustering cells by tissue type uncovers rare cell types

The protonephridia, the planarian excretory and osmoregulatory system, contains flame cells for filtering fluids, proximal and distal tubule cells, and a collecting duct ([29], [35], [36]). The protonephridia is a model tissue for studying organ regeneration and the evolution of kidney-like excretory systems. Subclustering of 890 protonephridia cells (Fig. 1C) identified each known protonephridia cell type as a separate subcluster, revealing the complete transcriptomes of these cells (Fig. 3A and fig. S12, A to C). Furthermore, two protonephridia subclusters with *smedwi-1*$^+$ cells were identified (Fig. 2M). One was enriched in flame cell gene expression (e.g., dd_2920) and the other in a proximal tubule marker (dd_10830), which suggests that they might be flame and tubule cell precursors, respectively (fig. S12, D and E).

Less is known regarding the full complement of cell types in other planarian tissues. Ultrastructural studies suggested that the planarian intestine contains two cell types: absorptive enterocytes and secretory goblet cells ([37], [38]). Subclustering of 3025 intestinal cells (Fig. 1C) revealed three distinct cell populations (clusters 4, 5, and 8) (Fig. 3B and fig. S13, A and B). FISH

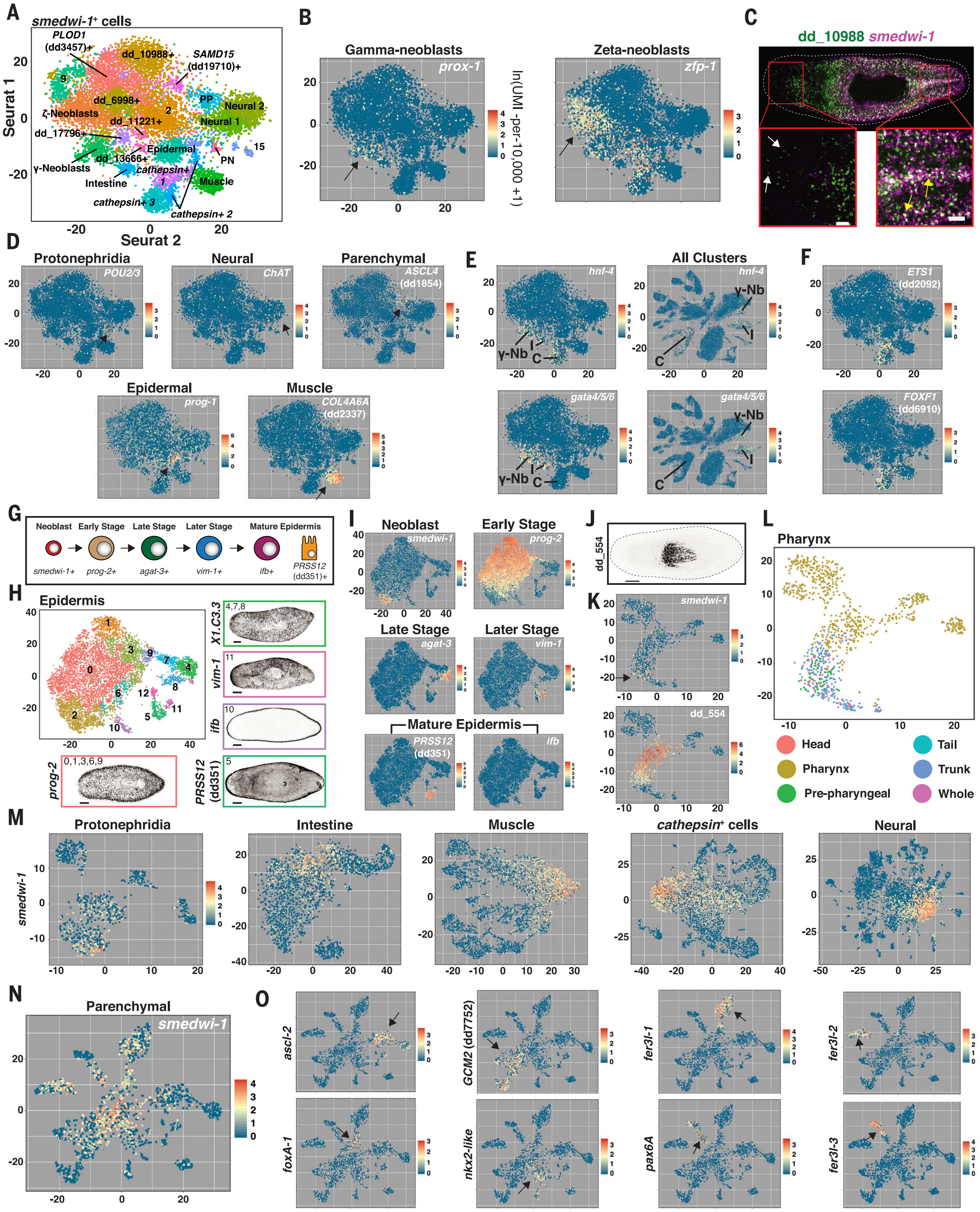


Fig. 2. Subclustering identifies neoblast subpopulations. (**A**) t-SNE representation of 22 clusters generated from subclustering cells with *smedwi-1* expression ≥ 2.5 [ln(UMI-per-10,000 + 1)]. Identity of numbered clusters unknown. PP, parenchymal; PN, protonephridia. Intestine cluster is indicated by lower expression of *smedwi-1* and enriched *gata4/5/6* and *hnf-4* expression. (**B**) *smedwi-1*$^+$ t-SNE plots colored by *prox-1* and *zfp-1* expression. (**C**) Double FISH image for dd_10988 and *smedwi-1*. Yellow arrows highlight coexpression; white arrows denote absence of coexpression. (**D**) *smedwi-1*$^+$ t-SNE plots colored by expression of differentiated tissue-enriched genes. Arrows indicate gene expression sites. (**E**) Left: *smedwi-1*$^+$ t-SNE plots colored by *gata4/5/6* and *hnf-4* expression. Right: All cluster t-SNE plots colored by *gata4/5/6 and hnf-4* expression. C, *cathepsin*$^+$ cells; I, intestine; γ-Nb, γ-neoblasts. (**F**) *smedwi-1*$^+$ t-SNE plots colored by *ETS1* (dd2092) and *FOXF1* (dd6910) expression. (**G**) Epidermal cell maturation stages. (**H**) t-SNE representation of epidermal subclusters. FISH images labeled by their associated cluster(s) are shown. (**I**) Epidermal t-SNE plots colored by epidermal lineage marker expression from (G). (**J**) dd_554 FISH. (**K**) Pharynx t-SNE plots colored by *smedwi-1* and dd_554 expression. (**L**) Pharynx t-SNE plot colored by the body region from which each cell was isolated. (**M** and **N**) t-SNE plots, colored by *smedwi-1* expression, generated by subclustering cells identified as (M) protonephridia, intestine, muscle, *cathepsin*$^+$, neural, and (N) parenchymal. (**O**) Parenchymal t-SNE plots colored by expression of eight transcription factor–encoding genes enriched in (N). Arrows indicate gene expression sites. Scale bars, 50 μm (C), 200 μm [(H) and (J)].

with subcluster-enriched markers (table S2) revealed distinct intestine components. Cluster 4 represented an inner intestine cell layer (Fig. 3C) and was enriched for absorptive enterocyte markers (*39*). Cluster 8 cells were largely present within the primary intestine branches, resembling the pattern of goblet cells (*40*). A third group (cluster 5) represented an outer intestine cell layer and displayed a set of enriched genes different from that of clusters 4 and 8 (Fig. 3C and table S2). In addition to these three main intestine components, clusters representing putative transition states were also identified. Clusters 1, 3, and 6 included many *smedwi-1*$^+$ cells (Fig. 2M). Genes with enriched expression in clusters 0 and 7 displayed expression spanning into the enterocyte cluster (cluster 4), suggesting these might be enterocyte transition states (fig. S14A). Genes with enriched expression in clusters 2 and 3 displayed expression spanning into the outer intestine cluster (cluster 5) and might reflect transition or variant states of these cells (fig. S14B). The Monocle toolkit can be used to predict cellular transitions in lineages (*41*) and was used to build single-cell trajectories for the enterocyte and outer intestine cell lineages, closely recapitulating the candidate transition states identified by Seurat (Fig. 3D, fig. S15, and table S3).

Several transcription factors required for the specification of various planarian cell types have been identified with RNA interference (RNAi) and gene expression studies. Because of constant tissue turnover, RNAi of transcription factor–encoding genes expressed in specific classes of specialized neoblasts in adult planarians can lead to steady depletion of the cell type generated by that specialized neoblast class (*29*, *42*, *43*). The transcriptomes identified here generate a resource of enriched gene expression for different cell types, including transcription factor–encoding genes. Accordingly, inhibition of the transcription factor–encoding *PTF1A* (dd6869) gene, which had enriched expression in candidate transition states for the outer intestine cluster, strongly reduced this cell population while not affecting absorptive enterocytes of the intestine (Fig. 3E).

The nervous system displays by far the greatest known cell type composition complexity of the major planarian tissues. By subclustering 11,907 neuronal cells (Fig. 1C), we identified 61 distinct subclusters representing a diversity of cell types and states (Fig. 4A and fig. S16A). Twelve subclusters had high *smedwi-1* expression, which suggested that they represent neuronal precursors (fig. S16B). Cluster 10 contained cells of the brain branches, as determined by expression of *gpas* (*44*) and *pds* (*45*) (Fig. 4B and fig. S16C). Three subclusters (clusters 3, 7, and 8) were defined by expression of *pc2* (encoding a neuropeptide-processing proprotein convertase) as well as an assortment of markers for rare neuron classes in the cephalic ganglia and ventral nerve cords (Fig. 4C and fig. S16, D and E). We also sequenced

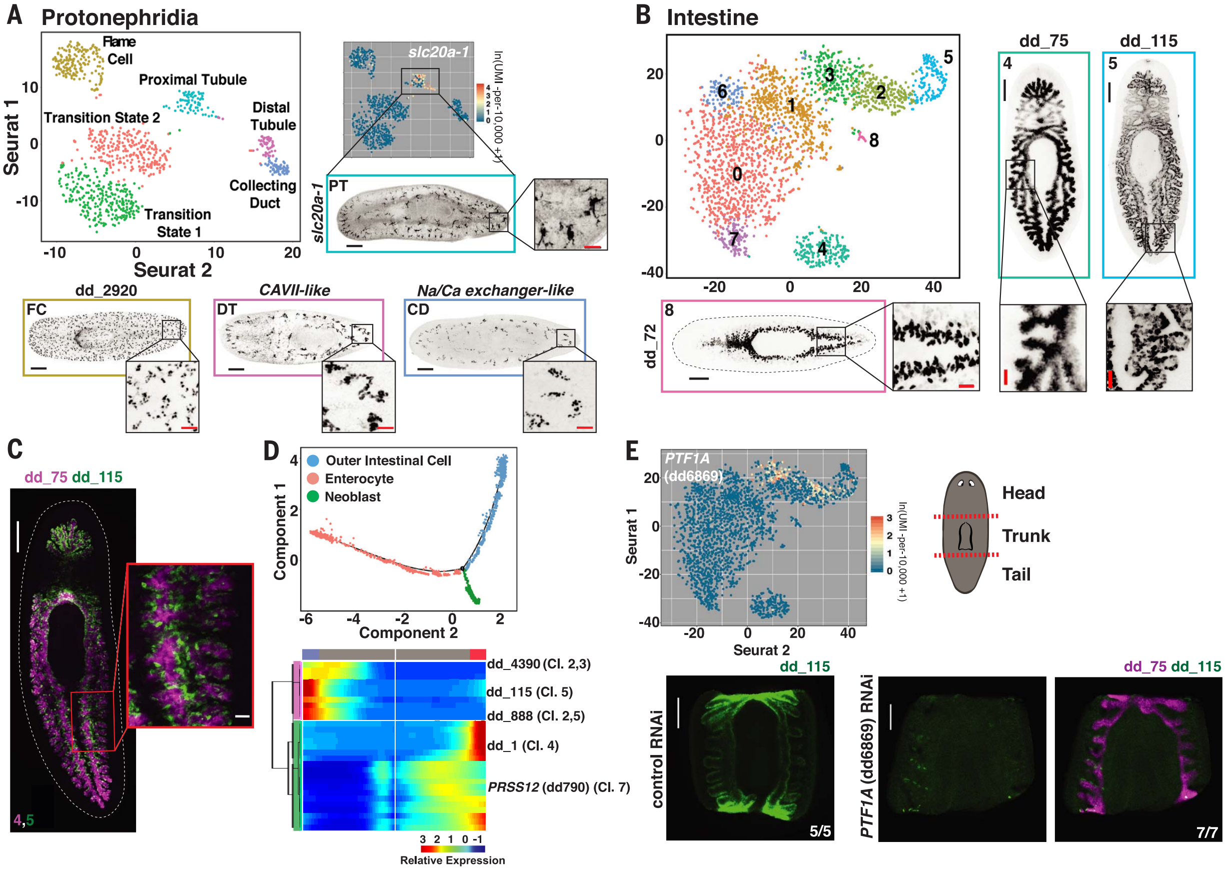


Fig. 3. Subclustering of tissues reveals transcriptomes for known and novel cell populations. (**A**) t-SNE representation of the protonephridial subcluster. FISH images are labeled by their associated cluster. (**B**) t-SNE representation of intestinal subclusters. (**C**) Double FISH images of genes enriched in separate intestinal subclusters. Numbers indicate the associated subcluster for each marker. (**D**) Top: Cell trajectory of enterocyte and outer intestinal cell lineages produced by Monocle. Cells are colored by identity. Bottom: Heat map of branch dependent genes (q value < 10^{-145}) across cells plotted in pseudo-time (*41*). Cells, columns; genes, rows. Beginning of pseudo-time is at center of heat map. "Cl." annotation indicates a log-fold enrichment ≥ 1 of the gene in that intestine Seurat cluster. (**E**) Top left: Intestine t-SNE plot colored by expression of *PTF1A* (dd6869). Top right: Illustration of cutting scheme used to generate fragments. Bottom: dd_115 and dd_75 FISH of control and *PTF1A* (dd6869) RNAi animals. Animals were cut and fixed 23 days after the start of double-stranded RNA (dsRNA) feedings. Scale bars: whole-animal/fragment images, 200 µm; insets, 50 µm.

an additional 7766 cells from the brain region to expand the number of cells in these clusters (fig. S16, F to H). In addition to these large clusters, there existed a number of smaller, compact, and well-separated subclusters. These could be further divided into ciliated and nonciliated neurons according to the expression of *rootletin* (dd6573), which encodes a ciliary rootlet component (Fig. 4D and fig. S17A). Because of further heterogeneity within these clusters (e.g., *opsin*$^+$ presumptive photoreceptors were present together, but not as a separate cluster), data from these two cell sets (ciliated, not ciliated) were each taken in isolation for further subclustering. This yielded 37 nonciliated neuron subclusters (Fig. 4E and figs. S17B, S18, A and B, and S19 to S21) and 25 putatively ciliated neuron subclusters (Fig. 4F and figs. S17B, S22, A and B, and S23B). We assessed the localization of cells associated with 46 of 62 of these subclusters by FISH using subcluster-specific markers. The observed cell types had a wide range of patterns including rare cell types such as photoreceptor neurons (Fig. 4, E and F, and figs. S18 to S23). Many genes had enriched expression in multiple clusters; the distribution of neural cell types they represented was defined by a combinatorial set of markers (figs. S18 to S23). A number of identified cell types from different subclusters displayed similar localization patterns. However, FISH demonstrated no overlap in subcluster-specific markers, consistent with the SCS data (Fig. 4G). For several neural subtypes, we found *smedwi-1*$^+$ candidate precursor cells. Four nonciliated neuron subclusters (subclusters 1, 2, 4, and 12) and a single ciliated neuron subcluster (subcluster 1) were enriched in *smedwi-1* expression (fig. S24A). Noncilated neuron subcluster 4 also expressed *gata4/5/6*, as did six *smedwi-1*$^-$ clusters (clusters

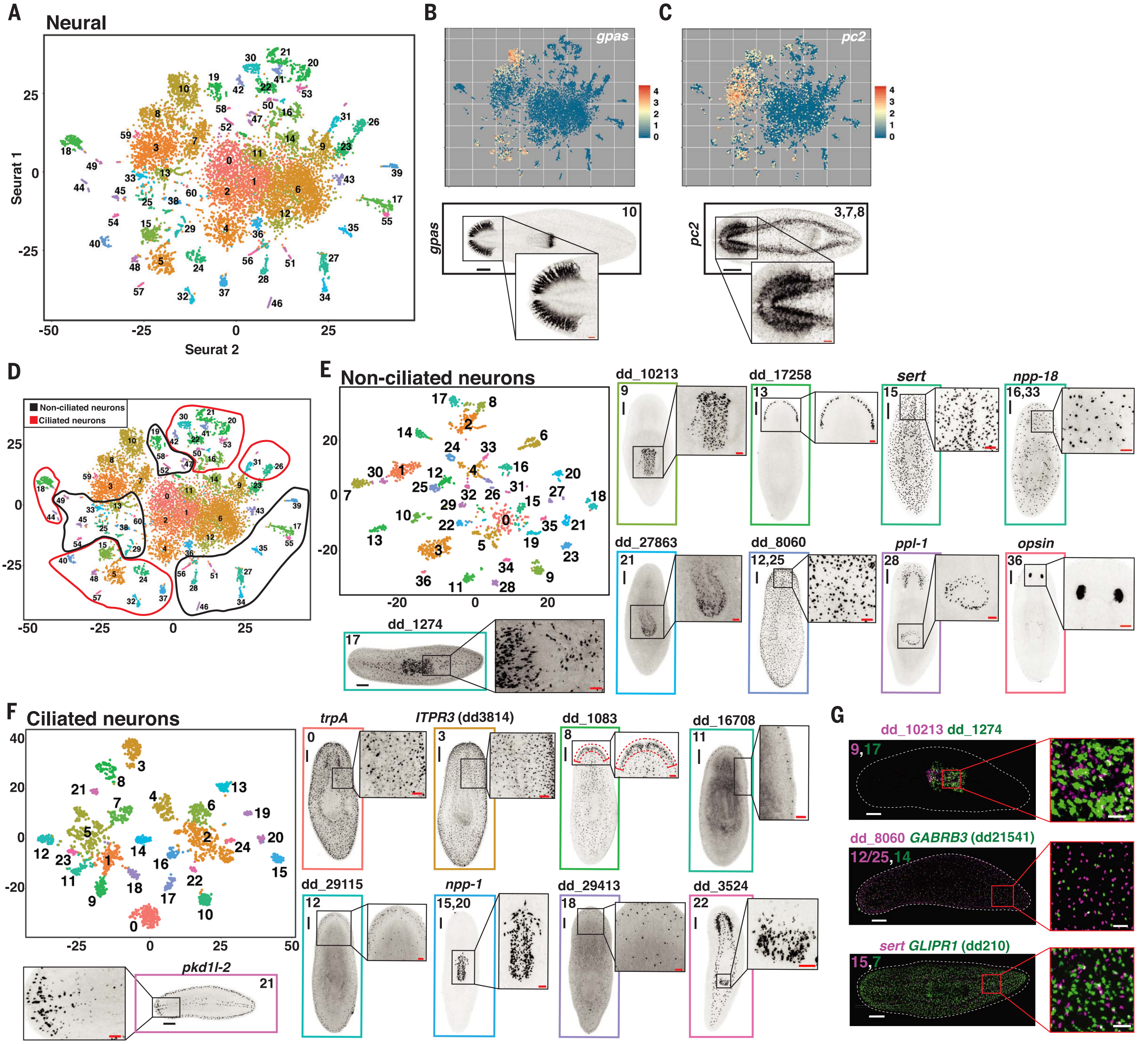


Fig. 4. Subclustering of neural cells reveals known and novel cell populations. (**A**) t-SNE representation of the neural subcluster. (**B** and **C**) Top: t-SNE plots colored by expression of *gpas* (B) and *pc-2* (C). Bottom: FISH for *gpas* (B) and *pc-2* (C) labeled with the associated neural subcluster. (**D**) t-SNE plot in (A) overlaid with outlines indicating the ascribed identity of each subcluster as ciliated or nonciliated. (**E** and **F**) t-SNE representation of subclustered cells identified in (D) as nonciliated (E) or ciliated (F). (**G**) Double FISH images of three sets of nonciliated neuron genes enriched in separate subclusters. Numbers indicate the associated nonciliated neuron subcluster(s) for each marker. Scale bars: whole-animal images, 200 μm; insets, 50 μm.

14,16/33, 24, 26, and 32) that radiated out from central *smedwi-1*$^+$ cells, raising the possibility that these *smedwi-1*$^+$ cells constitute precursors for these populations (fig. S24B).

The pharynx is a muscular tube used for feeding and defecation (*10*). It is contained within an epithelial cavity and connects to the intestine at its anterior end via an esophagus. Pharyngeal muscle cells and pharyngeal neurons clustered together with the other muscle cells and neurons of the body (figs. S25A and S26A). Other pharynx-associated cells, including cells from isolated pharynges and surrounding tissue, constituted the other major pharynx clusters. These non-neural, nonmuscle pharynx and pharynx-associated cells (Fig. 1C, n = 1083 cells) were subclustered, and FISH was performed on cluster-enriched markers (Fig. 5A and fig. S25, B and C). Subclusters included pharyngeal cavity epithelium cells (clusters 7 and 8), the epithelial pharynx lining (clusters 1 and 5), the mouth and esophagus (cluster 9), cells near the pharynx opening (cluster 6), and cells that constitute the connection to the planarian body (cluster 4). FISH confirmed nonoverlapping expression patterns for markers of tested separate cell populations (Fig. 5B).

Planarian muscle expresses *collagen* in addition to canonical muscle genes such as *troponin* and *tropomyosin* (*19*). Muscle exists in a subepidermal body wall layer, in the pharynx, surrounding the intestine, and in a DV domain (*46*). Subclustering 5014 muscle cells (Fig. 1C) revealed seven *smedwi-1*$^+$ candidate precursor subclusters (clusters 0, 1, 3, 4, 5, 10, and 11) (Fig. 2M), as well as subclusters containing body wall muscle (cluster 7), pharyngeal muscle (cluster 2, 8, 9, and 12) (fig. S26A), a population of muscle cells enriched around the intestine (cluster 6), and an unidentified population (cluster 13) (Fig. 5C and fig. S26, B and C). Markers for body wall muscle (cluster 7) and cluster 13 were expressed in nonoverlapping cells by FISH (Fig. 5D).

Whereas some molecular characterization existed for the seven broad planarian tissue classes previously mentioned, very little is known regarding the cellular composition of the two remaining classes. The parenchymal class (Fig. 1C) (*24*) was highly heterogeneous, with subclustering of 2120 cells identifying many distinct cell populations (Fig. 5E and figs. S27, A and B, and S28B). In addition to eight *smedwi-1*$^+$ putative precursor subclusters (clusters 0, 1, 2, 3, 4, 6, 8, and most of 9) (Fig. 2N), parenchymal cell subclustering revealed 13 well-separated differentiated cell subclusters. FISH showed that each of these differentiated cell populations were present as scattered cells, presumably within a mesenchymal tissue layer called the parenchyma that surrounds major planarian organs (*10*). Previous morphological studies determined that the parenchyma is composed of multiple gland cells, neoblasts, and "fixed parenchymal cells" characterized through histological and electron microscopy studies as a likely phagocytic cell with long cellular processes filling most of the parenchymal space (*10*, *47*, *48*). Some identified parenchymal subclusters appeared to be gland cells, displaying processes extending to the epidermis, defining transcriptomes for these cells. Candidate gland cell types included two that were exclusively dorsal (clusters 16 and 17), two exclusively lateral (clusters 10 and 13, including marginal adhesive gland cells and an unknown cell population), four present both dorsally and ventrally (clusters 7, 11, 14, and 15), and one present ventrally near the brain (cluster 19). Three subclusters (clusters 5, 12, and 18) contained cells with patterns similar to those of planarian neoblasts, but were not neoblasts. Finally, a single subcluster contained large cells surrounding the pharynx (small group of cluster 9 cells) and were enriched for expression of previously identified metalloprotease-encoding genes (*49*). Three pairs of parenchymal subclusters (six subclusters total) were confirmed to exist in nonoverlapping populations by FISH (Fig. 5F).

The transcription factor–encoding gene *nkx6-like* was expressed in a parenchymal cell population marked by dd_515. Inhibition of *nkx6-like* ablated dd_515 cells, while not affecting a distinct, non-enriched parenchymal cell population marked by dd_385 (Fig. 5G). These results further highlight the potential to use the data to ablate many specific cell types in the animal.

The final major class of cells, the *cathepsin*$^+$ group, contained 7034 cells (Fig. 1D). This group of clusters contained recently described glia and pigment cells (*12*, *13*, *50*). Subclustering of *cathepsin*$^+$ cells identified four subclusters expressing *smedwi-1* that represented putative precursor cells (clusters 0, 1, 3, and 6) (Fig. 2M), a glial subcluster (cluster 15), and two pigment cell populations (clusters 11 and 14), identifying transcriptomes for these cell types (Fig. 6A and figs. S29, A and B, and S30B). Eight *cathepsin*$^+$ subclusters represented previously unidentified cell populations. FISH revealed striking, elaborate morphologies for most of these cells, involving long processes and unique distributions (Fig. 6A and figs. S29B and S30B). Cells from subclusters 5 and 10 were spread throughout the planarian body, with long processes filling substantial parenchyma space. Subcluster 8 represented cells specific to the pharynx. Subcluster 9 cells were scattered throughout the animal. Subclusters 4 and 16 identified cells with dense aggregated foci of elaborate processes at scattered locations throughout the animal that lacked definitive positions—an unusual and unanticipated cell type distribution. FISH identified markers labeling cell bodies of these cells, revealing that the aggregates comprised many cells (Fig. 6B). Subclusters 12 and 13 also exhibited processes with visible cell bodies. Subcluster 12 cells were largely subepidermal. The most elaborate of these newly identified cells (subclusters 5 and 10) were excluded from the intestine and brain, but had processes around the branches of the intestine and protonephridia and interspersed within the cephalic ganglia (Fig. 6, C to E, and fig. S31, A and B). FISH confirmed nonoverlapping expression patterns for two tested subclusters (fig. S31C).

Subcluster 7 of the *cathepsin*$^+$ group of cells was enriched in expression of genes with expression spanning into clusters 5 and 10 (fig. S32A). Similarly, expression of cluster 2 marker genes spanned into clusters 4 and 16 (fig. S32B). These cells might reflect transition or variant states of cells for clusters 5/10 and 4/16, respectively. SMEDWI-1 protein perdures in neoblast progeny after loss of *smedwi-1* mRNA, allowing detection of newly produced neoblast progeny (*27*). *MAP3K5* (dd4849)$^+$ cells, which were predicted to be expressed in cells transitioning from the *smedwi-1*$^+$ state in the *cathepsin*$^+$ cell plot, were SMEDWI-1$^+$/*smedwi-1*$^-$, supporting the interpretation that these cells are progenitors in the *cathepsin*$^+$ cell lineage (fig. S32, C and D). The Monocle toolkit was also used to build single-cell trajectories for these clusters, with data closely recapitulating the transition states identified by Seurat (Fig. 6F, fig. S33, and table S3).

A near complete discovery of planarian cell type transcriptomes

The number of cell types identified in this study vastly exceeded prior planarian SCS data (*24*). Within the neuronal subclusters, a 17-cell subcluster represented photoreceptor neurons (Fig. 4E), which are present at ~100 total cells in a medium-sized (~2 to 3 mm) animal (*20*). Therefore, our data should have readily included unknown cell types as rare as photoreceptor neurons. Similarly, an average-sized planarian has ~60 *cintillo*$^+$ neurons, and our data included 10 *cintillo*$^+$ neurons (fig. S34A). These cells were grouped within a larger subcluster (cluster 3) of nonciliated neurons (fig. S34B), suggesting that even further subclustering of this "subcluster 3" could reveal additional distinct cell types. Indeed, *cintillo*$^+$ cells emerged as a unique cluster from such additional (fourth tier) subclustering of original data (fig. S34C and table S2). Esophagus cells, connecting pharynx to intestine, clustered with mouth cells in the pharynx subclustering data (fig. S34D). About 50 of these cells exist in an average-sized animal, and three such cells were present in the data (fig. S34, A and D). Several known rare cell types did not separate into individual clusters, although most could still be identified in the data, which suggests that the data are largely saturated for rare cell types. These include anterior pole cells, which function as an anterior organizer (*51*–*53*); *notum*$^+$ neurons in the brain (*54*); and posterior pole cells (*45*), each of which are among the rarest known cell types in the animal, with only ~10 each present in an average animal (fig. S34A). Five anterior pole cells, 10 *notum*$^+$ neurons, and one posterior pole cell were identified in the data (fig. S34, E to G). Similarly, ~25 *ovo*$^+$ eye progenitors (*42*) and ~90 *nanos*$^+$ germ cells (*55*) are present in an average animal (fig. S34A). Two eye progenitors and 19 germ cells were identified in the data (fig. S34, H and I). In addition to the asexual strain of *S. mediterranea* used in this study, a sexual strain of cross-fertilizing hermaphrodites exists. We sequenced 8455 cells from this strain, adding sexual strain cells to this resource as well, including seven yolk cells and seven testes cells in addition to the 19 germ cells described above (fig. S35, A to D).

Further sequencing of sexual cell types could be a target for future studies. Together, our data indicate that we have essentially reached saturation for determining the cell type transcriptomes of asexual planarians.

Discovery of novel patterning genes

Planarians constitutively express dozens of genes associated with patterning (PCGs) in complex spatial patterns across body axes (18). PCGs are almost exclusively expressed in muscle (19). AP-axis PCGs are well established, including with muscle SCS (56). Muscle cells did not subcluster according to their anatomical positions (Fig. 7A and fig. S36, A to C). However, we reasoned that expression of known PCGs could ascribe locations

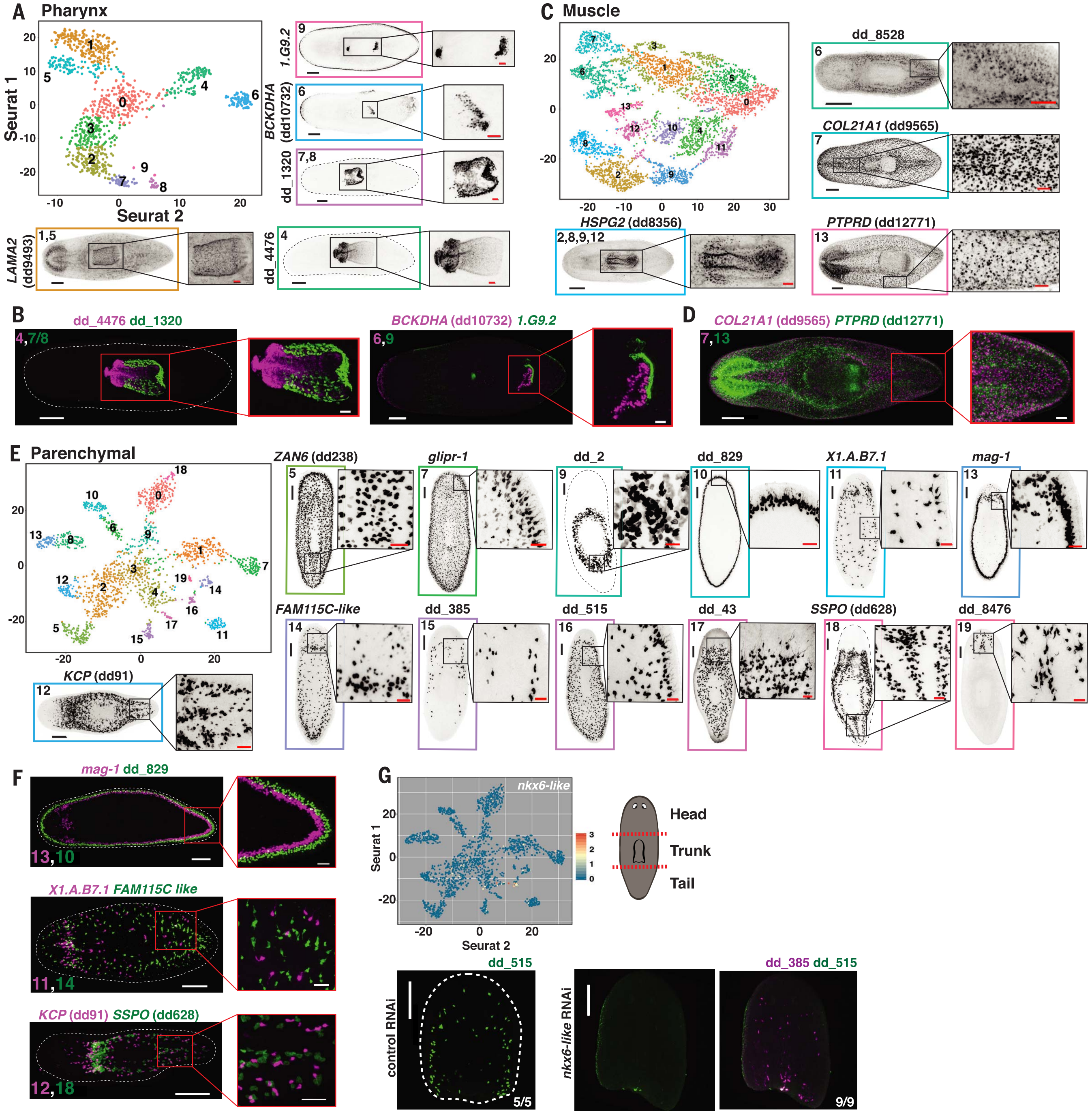


Fig. 5. Tissue subclustering identifies cell populations of poorly characterized tissues. (**A**) t-SNE representation of the pharynx subcluster. FISH images are labeled by their associated cluster(s). (**B**) Double FISH images of pharynx markers enriched in separate subclusters. Numbers indicate the associated pharynx subcluster(s) for each marker. (**C**) t-SNE representation of the muscle subcluster. (**D**) Double FISH images of two muscle markers enriched in separate subclusters. Numbers indicate the associated muscle subcluster for each marker. (**E**) t-SNE representation of the parenchymal subcluster. (**F**) Double FISH images of three sets of parenchymal markers enriched in separate subclusters. Numbers indicate the associated parenchymal subcluster for each marker. (**G**) Top left: Parenchymal t-SNE plot colored by expression of *nkx6-like*. Top right: Illustration of cutting scheme used to generate fragments. Bottom: dd_515 and dd_385 FISH of control and *nkx6-like* RNAi animals. Animal sections were cut and fixed 23 days after the start of dsRNA feedings. Scale bars: whole-animal/fragment images, 200 μm; insets, 50 μm.

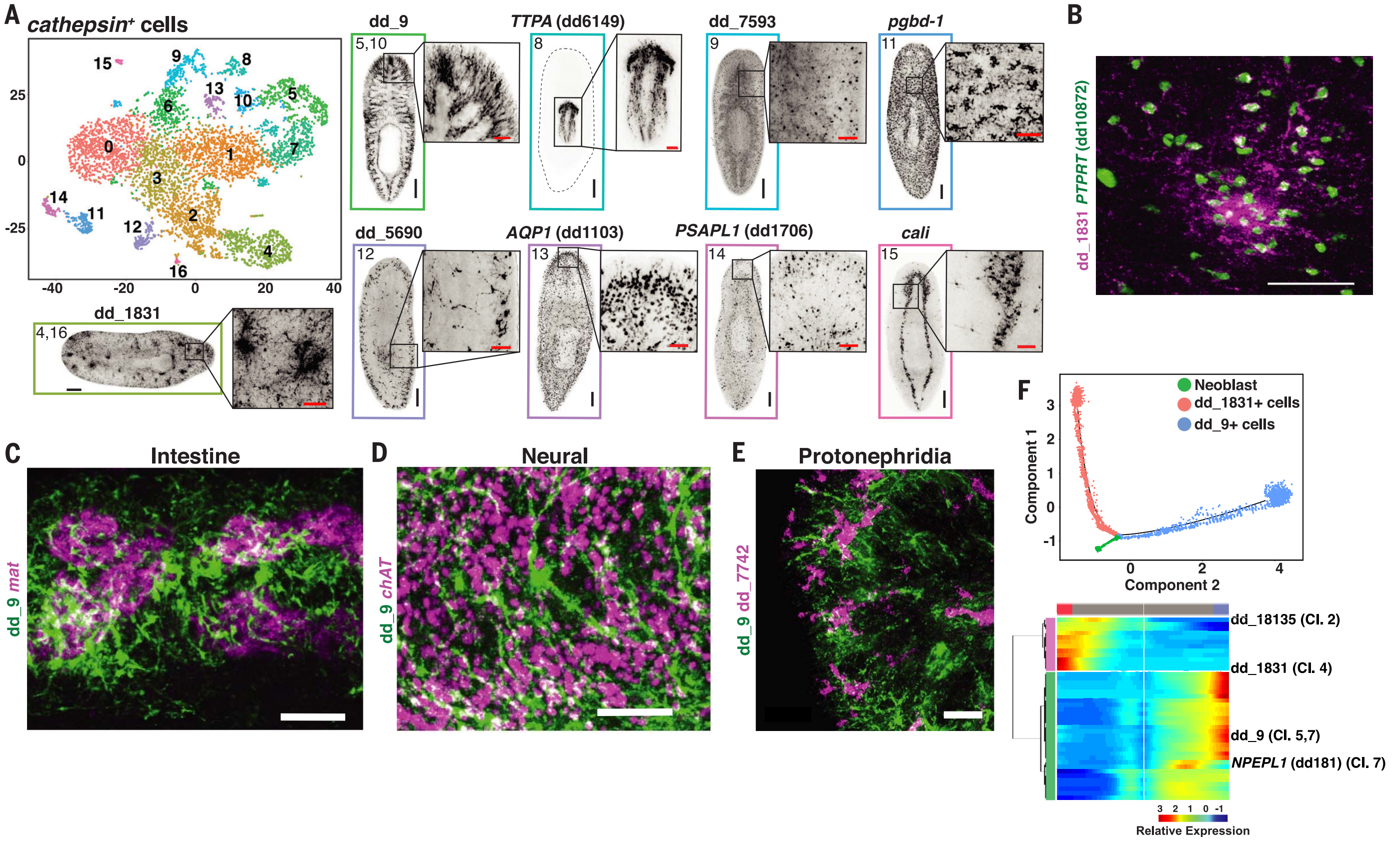
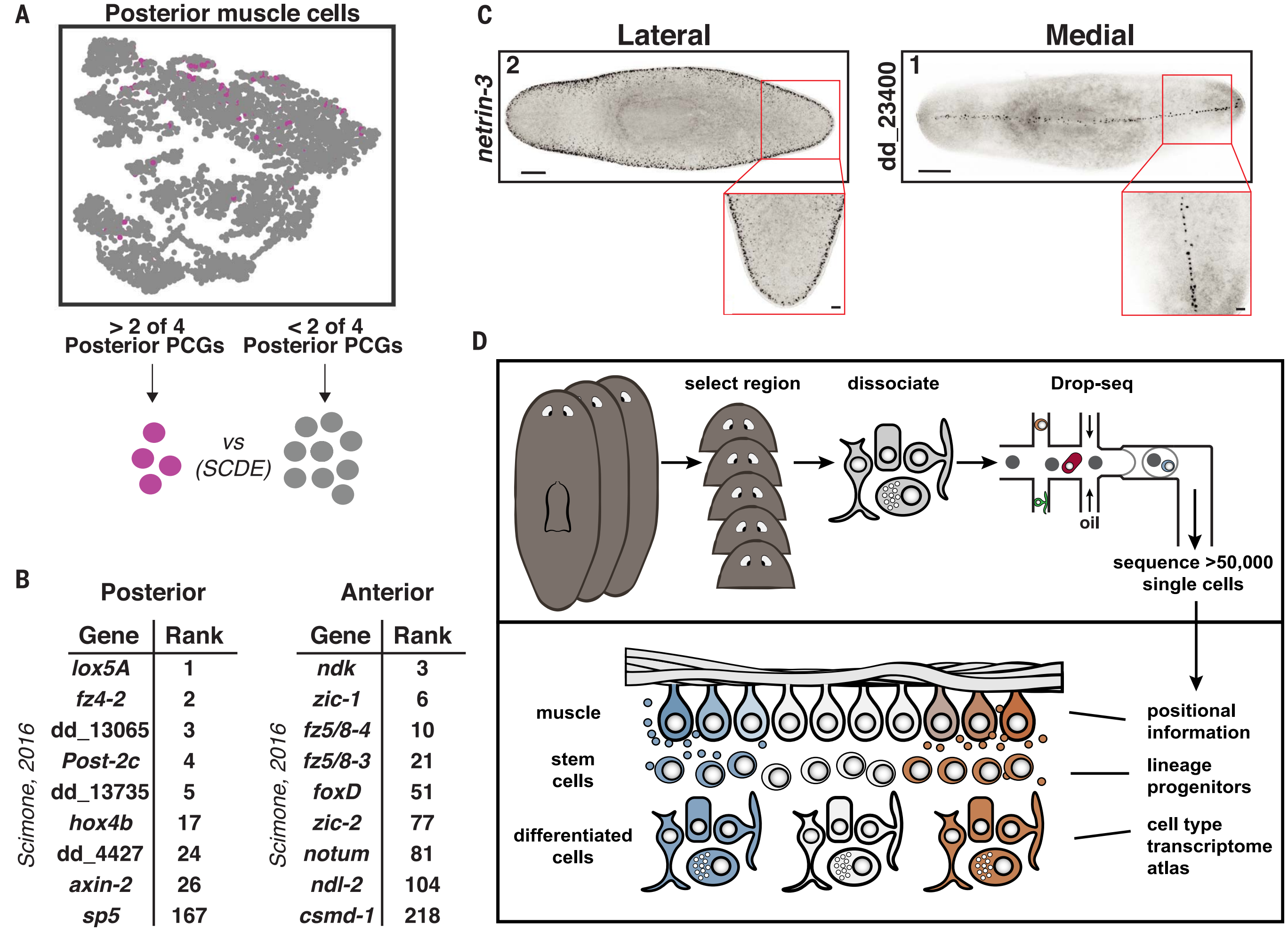


Fig. 6. Tissue subclustering reveals a previously unidentified class of cells. (**A**) t-SNE representation of the *cathepsin*$^+$ cell subcluster. FISH images are labeled by their associated cluster(s). Images associated with subclusters 5/10 and 8 are single slices in the animal. All other images are maximum intensity projections. (**B**) Double FISH for two *cathepsin*$^+$ cell markers enriched in the same subclusters, 4 and 16. (**C** to **E**) FISH for dd_9 and *mat* (C), *ChAT* (D), and dd_7742 (E). (**F**) Top: Cell trajectory of dd_1831$^+$ and dd_9$^+$ *cathepsin*$^+$ cell lineages produced by Monocle. Cells are colored by identity. Bottom: Heat map of branch dependent genes (q value $< 10^{-175}$) across cells plotted in pseudo-time (41). Cells, columns; genes, rows. Beginning of pseudo-time is at center of heat map. "Cl." annotation indicates a log-fold enrichment ≥ 1 of the gene in that *cathepsin*$^+$ cell Seurat cluster. Scale bars: whole-animal images, 200 μm; insets and (B) to (E), 50 μm.

to muscle cells in the data. Because of variability in the expression of any one PCG, muscle cell regional identity was determined on the basis of expression of at least two PCGs. For example, posterior muscle cells were identified by coexpression of at least two of the four posterior PCGs *wnt11-1*, *wnt11-2*, *fz4-1*, and *wntP-2*, yielding 163 cells (Fig. 7A and fig. S36D). Differential expression analysis using the algorithm SCDE (57) was performed on these 163 cells against the 4851 other muscle cells (Fig. 7A and table S4). Strikingly, nine of the differentially expressed genes were identified by Scimone *et al.* (56) as posterior-enriched; eight of these were within the top 26 genes identified by differential expression analysis (Fig. 7B, hypergeometric $P = 2.75 \times 10^{-9}$). A similar analysis on 837 anterior muscle cells was also performed (fig. S36, A and D, and table S4). Nine of the differentially expressed genes were identified by Scimone *et al.* (56); four of the genes were within the top 25 genes identified by SCDE (Fig. 7B, hypergeometric $P = 5.80 \times 10^{-5}$). We also applied this approach to the less well studied ML axis. We identified 62 lateral muscle cells, and FISH with 15 of the top genes identified seven with lateral muscle expression (Fig. 7C and fig. S36, B and D to G) (58, 59). We isolated 90 medial muscle cells, and the top-ranked gene displayed a striking thin stripe of expression down the dorsal midline (Fig. 7C and fig. S36, C, D, and H). Together, these results demonstrate the power of deep SCS for identifying regional gene expression, such as that involved in patterning, in adult animal tissues.

Discussion

RNA sequencing of >50,000 cells (in total, 66,783 cells were sequenced) of the planarian *S. mediterranea* allowed the identification of transcriptomes for most to all cell types of an adult animal. This includes transcriptomes for cell types present as rarely as 10 cells in an animal with 10^5 to 10^6 cells, which strongly suggests that we have reached near-saturation. Sequencing of different body regions and assessment of rare cell type coverage in an iterative process enabled us to reach this saturation level. Some cell types might escape detection by this technique if they are exceptionally rare or hard to dissociate from the animal. Our data did indicate that some cell types were preferentially recovered according to the abundance of that cell type by FISH, whereas others were less represented (fig. S37, A and B). In particular, *prog-1*$^+$ epidermal progenitor cells were highly overrepresented in the data relative to their prevalence in the animal, perhaps because their small size made their isolation easier (fig. S37A). Absent *prog-1*$^+$ cells, most other cell types analyzed were represented similarly to their relative abundance in the animal (fig. S37B). Regardless of differences in ease of dissociation between cell types, we recovered data from all known cell types assessed. Not every known rare cell type emerged as a separable cluster; that is, these cells were sometimes embedded within a larger cluster. In some instances, further rounds of subclustering based on such knowledge resulted in splitting of subclusters into additional subclusters. Therefore, further subclustering analyses and even deeper sequencing will likely continue to enhance the capacity to computationally isolate rare cell types from other clusters. Nonetheless, the transcriptomes for such rare cell types are present in our data and can be studied by



Scimone, 2016

Posterior Gene	Rank	Anterior Gene	Rank
lox5A	1	*ndk*	3
fz4-2	2	*zic-1*	6
dd_13065	3	*fz5/8-4*	10
Post-2c	4	*fz5/8-3*	21
dd_13735	5	*foxD*	51
hox4b	17	*zic-2*	77
dd_4427	24	*notum*	81
axin-2	26	*ndl-2*	104
sp5	167	*csmd-1*	218

Fig. 7. Identification of new regionally expressed genes in muscle. (**A**) Top: t-SNE plot colored by muscle cells positive for expression ≥ 0.5 [ln(UMI-per-10,000 + 1)] of two of the four posterior PCGs *wnt11-1*, *wnt11-2*, *fz-4*, and *wntP-2*. Pink, positive cells; gray, negative cells. Bottom: Transcriptomes for posterior muscle cells were compared to all other muscle cells by SCDE. (**B**) List of differentially expressed genes in posterior and anterior muscle cells that were identified in Scimone *et al.* (*56*). Rank indicates the rank of the gene in our analysis. (**C**) FISH images of one lateral and one medial expressed gene ranked highly in this analysis (*59*). Number indicates gene rank in the list generated by SCDE. Scale bars: whole-mount images, 200 μm; insets, 50 μm. (**D**) Illustration highlighting the capacity of the data set to identify almost all cell types in the planarian, as well as specialized neoblast progenitors and novel patterning information from the adult animal.

searching for the desired cells. Another challenge inherent in assessing saturation of cell type sequencing is ambiguity with the term cell type. Gene expression heterogeneity exists within well-defined clusters and could reflect differences attributable to technical sampling error, cell type state differences, or robust differences in biological function. Further in vivo morphological and functional studies with identified cell clusters, further computational analyses, and even more sequencing data can continue to refine the knowledge of biologically important cell type differences.

Cell types have been previously identified largely through morphological descriptions and perhaps a few marker genes. Determining cell type transcriptomes with large-scale SCS is a powerful new approach to defining the cell type constitution of a tissue, an organ, or even a complete animal. In our study, we identified a large number of previously uncharacterized planarian cell populations across multiple tissues. This included multiple cell populations (in the *cathepsin*$^+$ group) previously undescribed at the molecular level. One cell population, defined by dd_9 expression, had long processes filling parenchymal space and surrounding, but excluded from, other planarian tissues. This pattern is reminiscent of "fixed parenchymal cells," a largely uncharacterized cell population described by histology and electron microscopy (EM) (*48*). Previous EM work suggested that fixed parenchymal cells are likely phagocytic, with clearly observed lysosomes; dd_9$^+$ cells highly expressed genes encoding a variety of digestive enzymes and endocytosis proteins, providing further support for this hypothesis (table S2). The biology of these *cathepsin*$^+$ cells and all the other diverse cell types identified in this work can now be studied in depth using identified transcriptomes and the tools of planarian biology research. For instance, we show for two case studies above that RNAi of a gene encoding a transcription factor with enriched expression in a candidate cell lineage leads to ablation of the predicted differentiated cell.

Generating transcriptomes for most to all cell types in an animal will be invaluable for studying gene function and the biology and evolution of a large range of important cell types. Because of their phylogenetic position within the Spiralian superphylum (*60*), major cell types found across diverse bilaterians (e.g., shared between humans and *Drosophila*, *C. elegans*, molluscs, annelids, and/or other bilaterians) should have been present in the last common ancestor of planarians and humans. As such, studying the transcriptomes and associated genes with cell type–enriched expression in this data set can allow characterization of the gene function underlying the biology of these cells.

Planarian biology presents many features that made this organism attractive for comprehensive SCS. Planarians are a model for studying numerous important problems in regeneration, stem cell biology, patterning, and evolution. At a single time point—the adult—there exist progenitors for essentially all cell types and the patterning information for guiding new cell type production. We identified the transcriptomes of numerous candidate transition states in lineages from pluripotent stem cell to diverse differentiated cell types. Furthermore, we used the data to identify novel regionally expressed genes in planarian muscle (the site of patterning gene expression). Together, these results illustrate the capacity of our data set to define cell type transcriptomes, identify lineage transition states, and ascertain novel patterning information, all from a single time point (Fig. 7D). We propose that this atlas-like data set of cell type transcriptomes, much like the genome sequence of an animal, can serve as a resource fueling an immense amount of research, not only in planarians but in other bilaterians with similar cell types. To facilitate such study, we developed an online resource that generates cluster expression data, for any gene, across all clusters and subclusters (digiworm.wi.mit.edu). Case study model organisms have proved to be valuable testing grounds for developing approaches to complete genome sequencing; these planarian SCS data demonstrate an approach to near-to-complete cell type transcriptome identification that could be applied broadly to diverse organisms with varying degrees of information about cell type composition. The remarkable ability of single-cell RNA sequencing to reach nearly complete saturation of transcriptome identification for all the cell types of an animal represents a powerful approach for describing the anatomy of complete organisms at the molecular level.

REFERENCES

1. C. elegans Sequencing Consortium, Genome sequence of the nematode *C. elegans*: A platform for investigating biology. *Science* **282**, 2012–2018 (1998). doi: 10.1126/science.282.5396.2012; pmid: 9851916
2. E. S. Lander *et al.*, Initial sequencing and analysis of the human genome. *Nature* **409**, 860–921 (2001). doi: 10.1038/35057062; pmid: 11237011
3. J. C. Venter *et al.*, The sequence of the human genome. *Science* **291**, 1304–1351 (2001). doi: 10.1126/science.1058040; pmid: 11181995
4. D. A. Jaitin *et al.*, Massively parallel single-cell RNA-seq for marker-free decomposition of tissues into cell types. *Science* **343**, 776–779 (2014). doi: 10.1126/science.1247651; pmid: 24531970
5. A. K. Shalek *et al.*, Single-cell RNA-seq reveals dynamic paracrine control of cellular variation. *Nature* **510**, 363–369 (2014). doi: 10.1038/nature13437; pmid: 24919153
6. E. Z. Macosko *et al.*, Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. *Cell* **161**, 1202–1214 (2015). doi: 10.1016/j.cell.2015.05.002; pmid: 26000488
7. J. Cao *et al.*, Comprehensive single-cell transcriptional profiling of a multicellular organism. *Science* **357**, 661–667 (2017). doi: 10.1126/science.aam8940; pmid: 28818938
8. X. Han *et al.*, Mapping the mouse cell atlas by Microwell-seq. *Cell* **172**, 1091–1107.e17 (2018). doi: 10.1146/annurev.cellbio.20.010403.095114; pmid: 15473858
9. P. W. Reddien, A. Sánchez Alvarado, Fundamentals of planarian regeneration. *Annu. Rev. Cell Dev. Biol.* **20**, 725–757 (2004). doi: 10.1146/annurev.cellbio.20.010403.095114; pmid: 15473858
10. L. H. Hyman, *The Invertebrates: Platyhelminthes and Rhynchocoela, the Acoelomate Bilateria* (McGraw-Hill, 1951), vol. II.
11. J. Baguñà, R. Romero, Quantitative analysis of cell types during growth, degrowth and regeneration in the planarians *Dugesia mediterranea* and *Dugesia tigrina*. *Hydrobiologia* **84**, 181–194 (1981). doi: 10.1007/BF00026179
12. I. E. Wang, S. W. Lapan, M. L. Scimone, T. R. Clandinin, P. W. Reddien, Hedgehog signaling regulates gene expression in planarian glia. *eLife* **5**, e16996 (2016). doi: 10.7554/eLife.16996; pmid: 27612382
13. R. H. Roberts-Galbraith, J. L. Brubacher, P. A. Newmark, A functional genomics screen in planarians reveals regulators of whole-brain regeneration. *eLife* **5**, e17002 (2016). doi: 10.7554/eLife.17002; pmid: 27612384
14. D. E. Wagner, I. E. Wang, P. W. Reddien, Clonogenic neoblasts are pluripotent adult stem cells that underlie planarian regeneration. *Science* **332**, 811–816 (2011). doi: 10.1126/science.1203983; pmid: 21566185
15. M. L. Scimone, K. M. Kravarik, S. W. Lapan, P. W. Reddien, Neoblast specialization in regeneration of the planarian *Schmidtea mediterranea*. *Stem Cell Rep.* **3**, 339–352 (2014). doi: 10.1016/j.stemcr.2014.06.001; pmid: 25254346
16. P. W. Reddien, Specialized progenitors and regeneration. *Development* **140**, 951–957 (2013). doi: 10.1242/dev.080499; pmid: 23404104
17. P. A. Newmark, A. Sánchez Alvarado, Bromodeoxyuridine specifically labels the regenerative stem cells of planarians. *Dev. Biol.* **220**, 142–153 (2000). doi: 10.1006/dbio.2000.9645; pmid: 10753506
18. P. W. Reddien, Constitutive gene expression and the specification of tissue identity in adult planarian biology. *Trends Genet.* **27**, 277–285 (2011). doi: 10.1016/j.tig.2011.04.004; pmid: 21680047
19. J. N. Witchley, M. Mayer, D. E. Wagner, J. H. Owen, P. W. Reddien, Muscle cells provide instructions for planarian regeneration. *Cell Rep.* **4**, 633–641 (2013). doi: 10.1016/j.celrep.2013.07.022; pmid: 23954785
20. S. A. LoCascio, S. W. Lapan, P. W. Reddien, Eye absence does not regulate planarian stem cells during eye regeneration. *Dev. Cell* **40**, 381–391.e3 (2017). doi: 10.1146/annurev.cellbio.20.010403.095114; pmid: 15473858
21. S. Y. Liu *et al.*, Reactivating head regrowth in a regeneration-deficient planarian species. *Nature* **500**, 81–84 (2013). doi: 10.1038/nature12414; pmid: 23883932
22. R. Satija, J. A. Farrell, D. Gennert, A. F. Schier, A. Regev, Spatial reconstruction of single-cell gene expression data. *Nat. Biotechnol.* **33**, 495–502 (2015). doi: 10.1038/nbt.3192; pmid: 25867923
23. A. McDavid *et al.*, Data exploration, quality control and testing in single-cell qPCR-based gene expression experiments. *Bioinformatics* **29**, 461–467 (2013). doi: 10.1093/bioinformatics/bts714; pmid: 23267174
24. O. Wurtzel *et al.*, A generic and cell-type-specific wound response precedes regeneration in planarians. *Dev. Cell* **35**, 632–645 (2015). doi: 10.1016/j.devcel.2015.11.004; pmid: 26651295
25. P. W. Reddien, N. J. Oviedo, J. R. Jennings, J. C. Jenkin, A. Sánchez Alvarado, SMEDWI-2 is a PIWI-like protein that regulates planarian stem cells. *Science* **310**, 1327–1330 (2005). doi: 10.1126/science.1116110; pmid: 16311336
26. D. E. Wagner, J. J. Ho, P. W. Reddien, Genetic regulators of a pluripotent adult stem cell system in planarians identified by RNAi and clonal analysis. *Cell Stem Cell* **10**, 299–311 (2012). doi: 10.1016/j.stem.2012.01.016; pmid: 22385657
27. T. Guo, A. H. Peters, P. A. Newmark, A Bruno-like gene is required for stem cell maintenance in planarians. *Dev. Cell* **11**, 159–169 (2006). doi: 10.1016/j.devcel.2006.06.004; pmid: 16890156
28. J. C. van Wolfswinkel, D. E. Wagner, P. W. Reddien, Single-cell analysis reveals functionally distinct classes within the planarian stem cell compartment. *Cell Stem Cell* **15**, 326–339 (2014). doi: 10.1016/j.stem.2014.06.007; pmid: 25017721
29. M. L. Scimone, M. Srivastava, G. W. Bell, P. W. Reddien, A regulatory program for excretory system regeneration in planarians. *Development* **138**, 4387–4398 (2011). doi: 10.1242/dev.068098; pmid: 21937596
30. X. He *et al.*, FOX and ETS family transcription factors regulate the pigment cell lineage in planarians. *Development* **144**, 4540–4551 (2017). doi: 10.1242/dev.156349; pmid: 29158443
31. O. Wurtzel, I. M. Oderberg, P. W. Reddien, Planarian epidermal stem cells respond to positional cues to promote cell-type diversity. *Dev. Cell* **40**, 491–504.e5 (2017). doi: 10.1146/annurev.cellbio.20.010403.095114; pmid: 15473858
32. G. T. Eisenhoffer, H. Kang, A. Sánchez Alvarado, Molecular analysis of stem cells and their descendants during cell turnover and regeneration in the planarian *Schmidtea mediterranea*. *Cell Stem Cell* **3**, 327–339 (2008). doi: 10.1016/j.stem.2008.07.002; pmid: 18786419
33. K. C. Tu *et al.*, Egr-5 is a post-mitotic regulator of planarian epidermal differentiation. *eLife* **4**, e10501 (2015). doi: 10.7554/eLife.10501; pmid: 26457503
34. S. J. Zhu, S. E. Hallows, K. W. Currie, C. Xu, B. J. Pearson, A mex3 homolog is required for differentiation during planarian stem cell lineage development. *eLife* **4**, e07025 (2015). doi: 10.7554/eLife.07025; pmid: 26114597
35. H. Thi-Kim Vu *et al.*, Stem cells and fluid flow drive cyst formation in an invertebrate excretory organ. *eLife* **4**, e07405 (2015). doi: 10.7554/eLife.07405; pmid: 26057828
36. J. C. Rink, H. Thi-Kim Vu, A. Sánchez Alvarado, The maintenance and regeneration of the planarian excretory system are regulated by EGFR signaling. *Development* **138**, 3769–3780 (2011). doi: 10.1242/dev.066852; pmid: 21828097
37. B. H. Willier, L. H. Hyman, S. A. Rifenburgh, A histochemical study of intracellular digestion in triclad flatworms. *J. Morphol.* **40**, 299–340 (1925). doi: 10.1002/jmor.1050400205
38. S. Ishii, Electron microscopic observations on the planarian tissues II. The intestine. *Fukushima J. Med. Sci.* **12**, 67–87 (1965). pmid: 5863488
39. D. J. Forsthoefel *et al.*, An RNAi screen reveals intestinal regulators of branching morphogenesis, differentiation, and stem cell proliferation in planarians. *Dev. Cell* **23**, 691–704 (2012). doi: 10.1016/j.devcel.2012.09.008; pmid: 23079596
40. R. M. Zayas, F. Cebrià, T. Guo, J. Feng, P. A. Newmark, The use of lectins as markers for differentiated secretory cells in planarians. *Dev. Dyn.* **239**, 2888–2897 (2010). doi: 10.1002/dvdy.22427; pmid: 20865784
41. C. Trapnell *et al.*, The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nat. Biotechnol.* **32**, 381–386 (2014). doi: 10.1038/nbt.2859; pmid: 24658644
42. S. W. Lapan, P. W. Reddien, Transcriptome analysis of the planarian eye identifies *ovo* as a specific regulator of eye regeneration. *Cell Rep.* **2**, 294–307 (2012). doi: 10.1016/j.celrep.2012.06.018; pmid: 22884275
43. S. W. Lapan, P. W. Reddien, *dlx* and *sp6-9* Control optic cup regeneration in a prototypic eye. *PLOS Genet.* **7**, e1002226 (2011). doi: 10.1371/journal.pgen.1002226; pmid: 21852957
44. F. Cebrià *et al.*, Dissecting planarian central nervous system regeneration by the expression of neural-specific genes. *Dev. Growth Differ.* **44**, 135–146 (2002). doi: 10.1046/j.1440-169x.2002.00629.x; pmid: 11940100
45. C. P. Petersen, P. W. Reddien, *Smed-betacatenin-1* is required for anteroposterior blastema polarity in planarian regeneration. *Science* **319**, 327–330 (2008). doi: 10.1126/science.1149943; pmid: 18063755
46. F. Cebrià, Planarian body-wall muscle: Regeneration and function beyond a simple skeletal support. *Front. Cell Dev. Biol.* **4**, 8 (2016). doi: 10.3389/fcell.2016.00008; pmid: 26904543
47. K. J. Pedersen, Some features of the fine structure and histochemistry of planarian subepidermal gland cells. *Z. Zellf.* **50**, 121–142 (1959). doi: 10.1007/BF00350411
48. K. J. Pedersen, Studies on the nature of planarian connective tissue. *Z. Zellf.* **53**, 569–608 (1961). doi: 10.1007/BF00339508
49. P. A. Newmark, P. W. Reddien, F. Cebrià, A. Sánchez Alvarado, Ingestion of bacterially expressed double-stranded RNA inhibits gene expression in planarians. *Proc. Natl. Acad. Sci. U.S.A.* **100** (suppl. 1), 11861–11865 (2003). doi: 10.1073/pnas.1834205100; pmid: 12917490
50. B. M. Stubenhaus *et al.*, Light-induced depigmentation in planarians models the pathophysiology of acute porphyrias. *eLife* **5**, e14175 (2016). doi: 10.7554/eLife.14175; pmid: 27240733
51. M. L. Scimone, S. W. Lapan, P. W. Reddien, A forkhead transcription factor is wound-induced at the planarian midline and required for anterior pole regeneration. *PLOS Genet.* **10**, e1003999 (2014). doi: 10.1371/journal.pgen.1003999; pmid: 24415944
52. M. C. Vogg *et al.*, Stem cell-dependent formation of a functional anterior regeneration pole in planarians requires Zic and Forkhead transcription factors. *Dev. Biol.* **390**, 136–148 (2014). doi: 10.1016/j.ydbio.2014.03.016; pmid: 24704339
53. C. P. Petersen, P. W. Reddien, Polarized *notum* activation at wounds inhibits Wnt function to promote planarian head regeneration. *Science* **332**, 852–855 (2011). doi: 10.1126/science.1202143; pmid: 21566195

54. E. M. Hill, C. P. Petersen, Wnt/Notum spatial feedback inhibition controls neoblast differentiation to regulate reversible growth of the planarian brain. *Development* **142**, 4217–4229 (2015). doi: 10.1242/dev.123612; pmid: 26525673
55. Y. Wang, R. M. Zayas, T. Guo, P. A. Newmark, *nanos* function is essential for development and regeneration of planarian germ cells. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 5901–5906 (2007). doi: 10.1073/pnas.0609708104; pmid: 17376870
56. M. L. Scimone, L. E. Cote, T. Rogers, P. W. Reddien, Two FGFRL-Wnt circuits organize the planarian anteroposterior axis. *eLife* **5**, e12845 (2016). doi: 10.7554/eLife.12845; pmid: 27063937
57. P. V. Kharchenko, L. Silberstein, D. T. Scadden, Bayesian approach to single-cell differential expression analysis. *Nat. Methods* **11**, 740–742 (2014). doi: 10.1038/nmeth.2967; pmid: 24836921
58. S. Barberán, J. M. Martín-Durán, F. Cebrià, Evolution of the EGFR pathway in Metazoa and its diversification in the planarian *Schmidtea mediterranea*. *Sci. Rep.* **6**, 28071 (2016). doi: 10.1038/srep28071; pmid: 27325311
59. M. L. Scimone, L. E. Cote, P. W. Reddien, Orthogonal muscle fibres have different instructive roles in planarian regeneration. *Nature* **551**, 623–628 (2017). pmid: 29168507
60. C. E. Laumer *et al*., Spiralian phylogeny informs the evolution of microscopic lineages. *Curr. Biol.* **25**, 2000–2006 (2015). doi: 10.1016/j.cub.2015.06.068; pmid: 26212884

ACKNOWLEDGMENTS

We thank M. L. Scimone, C. McQuestion, and K. D. Atabay for their help in the targeted dissociation of tissue from the planarian brain; all Reddien Lab members for valuable comments and discussion; and E. Z. Macosko, M. Goldman, and S. A. McCarroll for making protocols available. **Funding:** We acknowledge NIH (R01GM080639) support. P.W.R. is an Investigator of the Howard Hughes Medical Institute and an associate member of the Broad Institute of Harvard and MIT. We thank the Eleanor Schwartz Charitable Foundation for support. **Author contributions:** P.W.R. supervised. C.T.F. and P.W.R. designed experiments and wrote the manuscript. C.T.F., P.W.R., and O.W. analyzed data. C.T.F. and T.H. built and optimized the Drop-seq setup. C.T.F. developed the data processing pipeline. C.T.F. and K.M.K. developed the pipeline for clustering analysis. C.T.F. and P.W.R. performed planarian tissue extractions. C.T.F. performed all other planarian experiments. O.W. generated the online resource. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** All raw and processed data files associated with this study have been deposited to Gene Expression Omnibus (GEO) under the accession number GSE111764.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/360/6391/eaaq1736/suppl/DC1
Materials and Methods
Figs. S1 to S37
Tables S1 to S5
References (*61*–*93*)

9 October 2017; accepted 5 April 2018
Published online 19 April 2018
10.1126/science.aaq1736

RESEARCH ARTICLE SUMMARY

SINGLE-CELL ANALYSIS

Cell type atlas and lineage tree of a whole complex animal by single-cell transcriptomics

Mireya Plass,* Jordi Solana,* F. Alexander Wolf, Salah Ayoub, Aristotelis Misios, Petar Glažar, Benedikt Obermayer, Fabian J. Theis, Christine Kocks, Nikolaus Rajewsky†

INTRODUCTION: Understanding the differentiation of stem cells into the vast amount of cell types that form the human body is a central problem of basic and medical science. The recent advances in single-cell sequencing techniques now make it possible to capture the transcriptomes of thousands of cells in a fast and cost-effective manner, opening a new way to study the cell composition of organs, tissues, and developmental stages. Yet, single-cell transcriptomics per se just provides a snapshot of cellular dynamics and transient cell populations. Computational algorithms have emerged that infer a pseudotemporal ordering of cells based on comparison of their transcriptomic profiles, allowing new insights into stem cell biology and tissue differentiation. However, these algorithms were designed for relatively simple scenarios, such as the differentiation of cells belonging to a specific lineage or the lineage relationships among cells from a particular tissue, and cannot evaluate all possible cellular differentiation trajectories in complex animals. To this end, we use single-cell transcriptomic approaches to improve our molecular understanding of how stem cells differentiate into the set of cell types that make an entire complex adult animal.

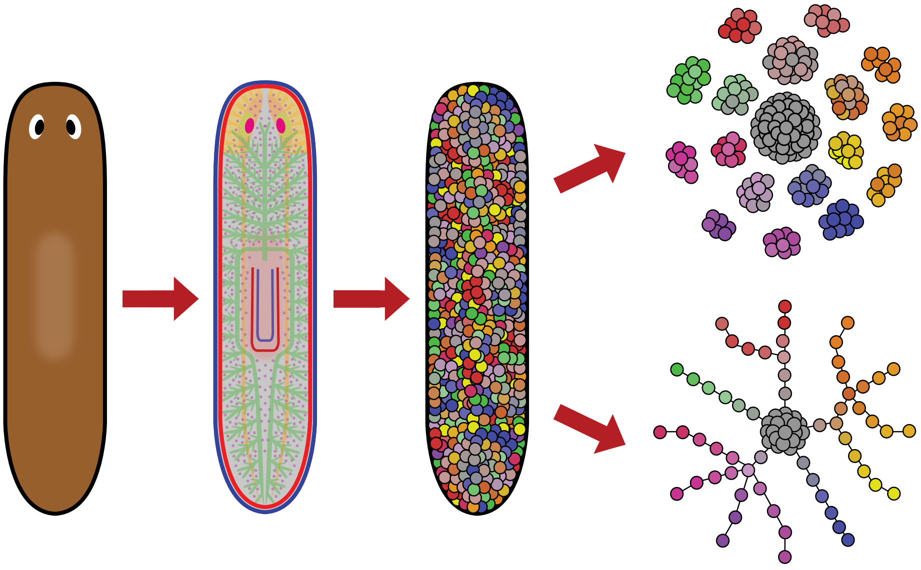
RATIONALE: Freshwater planarians such as *Schmidtea mediterranea* offer a unique opportunity to approach this question. These animals are immortal and constantly renew and regenerate all tissues owing to the presence of a large pool of pluripotent stem cells that continuously differentiate into all mature cell types. Therefore, we reasoned that an unbiased single-cell transcriptomic approach should allow us to capture not only terminally differentiated cell types but also intermediate cellular states, possibly enabling cell lineage reconstruction of the whole animal from transcriptomic data.

RESULTS: We performed massively parallel single-cell transcriptomics profiling of thousands of cells from adult planarians. At the molecular level, we identified and characterized dozens of cell types, including stem cells, progenitors, and terminally differentiated cells. We then applied a new computational algorithm, partition-based graph abstraction (PAGA), which can predict a lineage tree for the whole animal in an unbiased way. By combining the predictions from PAGA with several independent lines of evidence, including single-cell transcriptome data from purified stem cells and stem cell–depleted animals, analysis of gene expression dynamics, and a method called velocyto that predicts future gene expression from mRNA metabolism, we produced a consolidated lineage tree that included all identified cell types rooted to a single stem cell group. We used this information to identify gene sets co-regulated during the differentiation of many specific cell types. To show the power of our approach, we applied single-cell transcriptomics to regenerating planarians and characterized how each cell type in the adult planarian body dynamically responds to regenerative body remodeling at the transcriptomic and cellular levels. Our results highlight that some cell types that had been previously overlooked in molecular studies quickly decrease their abundance, indicating that they may serve as an energy reservoir that fuels the regeneration process.

ON OUR WEBSITE

Read the full article at http://dx.doi.org/10.1126/science.aaq1723

CONCLUSION: We have shown that it is possible to use single-cell transcriptomics to (i) build a cell atlas of an adult animal, (ii) reconstruct the lineage relationships of its cells in an unbiased way, and (iii) identify gene sets which likely contain genes that are involved in programming the lineage tree. Moreover, we demonstrated how single-cell transcriptomics can be used to study complex and dynamic cellular processes such as regeneration. Notably, our approach is applicable to other model and non–model organisms, assuming that their differentiation processes are sampled with sufficient time resolution. To foster future studies, we provide a detailed tutorial on the application of our approach, and we make our data available through an interactive web interface. This study opens the door to powerful approaches for understanding molecular mechanisms of development and regeneration in animals. ■



A lineage tree for complex animals from single-cell transcriptomics. Planarians are multicellular organisms. They contain adult pluripotent stem cells that continuously renew all tissues and differentiate into all adult cell types. Using single-cell transcriptomics, we characterized all major mature cell types and many intermediate cellular states. We then derived a lineage tree describing planarian stem cell differentiation into all mature cell types of the animal.

The list of author affiliations is available in the full article online.
*These authors contributed equally to this work.
†Corresponding author. Email: rajewsky@mdc-berlin.de

Cite this article as M. Plass *et al.*, *Science* 360, eaaq1723 (2018). DOI: 10.1126/science.aaq1723

Plass *et al.*, *Science* **360**, 875 (2018) 25 May 2018 **1 of 1**

RESEARCH ARTICLE

SINGLE-CELL ANALYSIS

Cell type atlas and lineage tree of a whole complex animal by single-cell transcriptomics

Mireya Plass,[1]* Jordi Solana,[1]*† F. Alexander Wolf,[2] Salah Ayoub,[1] Aristotelis Misios,[1] Petar Glažar,[1] Benedikt Obermayer,[1]‡ Fabian J. Theis,[2,3] Christine Kocks,[1] Nikolaus Rajewsky[1]§

Flatworms of the species *Schmidtea mediterranea* are immortal—adult animals contain a large pool of pluripotent stem cells that continuously differentiate into all adult cell types. Therefore, single-cell transcriptome profiling of adult animals should reveal mature and progenitor cells. By combining perturbation experiments, gene expression analysis, a computational method that predicts future cell states from transcriptional changes, and a lineage reconstruction method, we placed all major cell types onto a single lineage tree that connects all cells to a single stem cell compartment. We characterized gene expression changes during differentiation and discovered cell types important for regeneration. Our results demonstrate the importance of single-cell transcriptome analysis for mapping and reconstructing fundamental processes of developmental and regenerative biology at high resolution.

Understanding differentiation from stem cells into the different cell types that make up the human body is a central problem of basic and medical science. Although numerous mechanisms of cellular differentiation have been identified and many cell types have been characterized, it will require a huge coordinated undertaking to systematically map all human cell types and cellular differentiation states (*1*). Owing to the advances in single-cell transcriptomics, it has already been possible to study the cell type composition of mammalian organs and tissues (*2*–*6*), as well as development stages (*7*, *8*). However, single-cell transcriptomics provide just a snapshot of the dynamics of the cell populations unless cells can be traced or tagged experimentally (*9*–*12*). Thus, reconstructing cell lineages from stem cells to differentiated cells remains a challenge. Recently, algorithms to order developmental states and compute lineage trees based on comparing single-cell transcriptomes have made considerable progress (*13*–*15*) and have revealed insights into stem cell biology (*16*) and tissue differentiation (*17*–*20*). However, these algorithms have been developed for the study of differentiation in specific cell lineages or tissues and are not suitable to reconstruct all the cell differentiation trajectories present in complex animals.

Given these problems, can single-cell transcriptomic approaches improve our molecular understanding of how stem cells differentiate into all the cell types of an entire complex adult animal? Freshwater planarians such as *Schmidtea mediterranea* offer a unique opportunity to answer this question. Planarians are immortal, as they contain as adults a large pool of pluripotent stem cells (neoblasts) that continuously differentiate to all mature cell types to turn over all tissues (*21*). Hence, all cell differentiation pathways are constantly active in adult individuals. We therefore reasoned that an unbiased single-cell transcriptomics approach should yield terminally differentiated cell types as well as many intermediate cellular states, making planarians an ideal model system to attempt the lineage reconstruction of a whole animal.

Here, we performed highly parallel droplet-based single-cell transcriptomics, Drop-seq (*3*), to characterize planarian cell types. We molecularly characterized dozens of cell types and uncovered many new ones. By applying a newly developed algorithm, partition-based graph abstraction (PAGA), which reconciles the principles of clustering and pseudotemporal ordering (*22*), and combining it with independent computational and experimental approaches, we derive a consolidated lineage tree that includes all identified cell types rooted to a single stem cell cluster. Along this tree, we identify 48 gene sets that are co-regulated during the differentiation of specific cell types. Finally, we used single-cell transcriptomics to characterize cellular processes that happen during regeneration. Our results reveal a strong depletion of newly characterized cell types, suggesting that these cells are used as an energy source for regeneration.

A high-resolution cell type atlas for planaria

To comprehensively characterize different cell types and progenitor stages present in adult planarians, we performed genome-wide expression profiling in individual cells using nanoliter droplets (Drop-seq) (*3*) of cells isolated from whole adult animals. These cells were obtained, after dissociation, by fluorescence-activated cell sorting (FACS), which separated intact live cells from dead cells and enucleated cellular debris (Fig. 1A). From 11 independent experiments, we captured a total of 21,612 cells. We detected on average 494 genes and ~970 transcripts [identified by using "unique molecular identifiers" (UMIs)] per cell. The individual data sets correspond to five wild-type samples (10,866 cells), two RNA interference (RNAi) samples (3314 cells), a high-DNA content G_2/M population corresponding to cycling planarian stem cells (typically defined as x-ray–sensitive "X1 cells"; 981 cells) (*23*, *24*), and three wild-type regeneration samples (6451 cells; table S1). Sequencing depth was comparable across samples (fig. S1A). Biological replicates showed highly correlated gene expression profiles (fig. S1B). In addition, all samples showed high correlation with published RNA-sequencing (RNA-seq) data from equivalent bulk cell populations (*24*–*27*) (fig. S1C). We pooled and analyzed all single-cell data sets together using Seurat (*3*). Eight of 11 samples were fixed with methanol (*28*) or frozen with dimethyl sulfoxide (DMSO) (*29*) to facilitate sample handling. To assess batch effects, we compared the overall quality across wild-type samples. Cells from each batch were distributed similarly on the t-distributed stochastic neighbor embedding (tSNE) (fig. S1D), which resulted in comparable proportions of cells per cluster (fig. S1E and table S2). Although we observed a mild bias in gene expression due to the preservation procedure of the samples, clustering was not affected (fig. S1F). However, we observed differences in the number of UMIs per cell across clusters (fig. S1G). Together, these analyses confirmed that sample preparation did not compromise data quality or introduce bias. Therefore, we clustered the expression profiles of the individual cells from all samples together using Seurat (*3*). In total, we identified 51 cell clusters (Fig. 1B).

We elucidated the cell type identity of clusters by examining marker genes and comparing them to those in previous studies (fig. S2, A and B) (supplementary note 1). The largest cluster and 14 smaller clusters located in the center of the tSNE plot express combinations of well-known stem cell markers (fig. S3A), such as *Smedwi-1*, *Smedtud-1*, and *bruli* (fig. S3B). The remaining clusters corresponded to the previously described neural, epidermal, secretory, muscle, gut, and

---

[1]Laboratory for Systems Biology of Gene Regulatory Elements, Berlin Institute for Medical Systems Biology, Max-Delbrück Center for Molecular Medicine, Berlin, Germany. [2]Helmholtz Zentrum München, German Research Center for Environmental Health, Institute of Computational Biology, Neuherberg, Germany. [3]Department of Mathematics, Technische Universität München, München, Germany.
*These authors contributed equally to this work. †Present address: Department of Biological and Medical Sciences, Oxford Brookes University, Oxford, UK. ‡Present address: Core Unit Bioinformatics, Berlin Institute of Health, Berlin, Germany.
§Corresponding author. Email: rajewsky@mdc-berlin.de

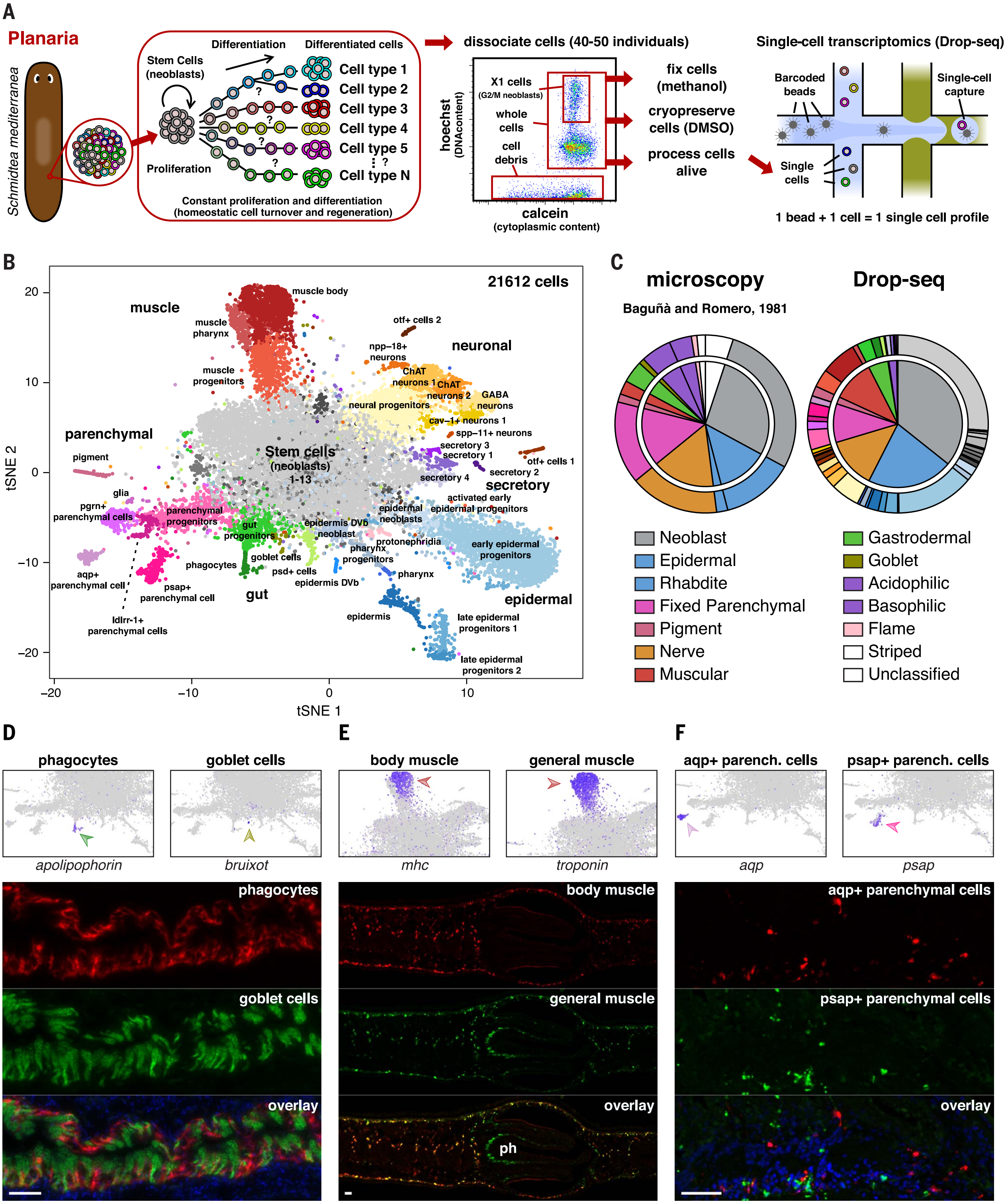


Fig. 1. Cell type atlas by single-cell transcriptomics. (**A**) Experimental workflow. (**B**) tSNE representation of the single-cell transcriptomics data with clusters colored according to the expression of previously published marker genes as follows: gray, neoblasts; orange, neuronal lineage; red, muscle; purple, secretory; blue, epidermal lineages; pink, protonephridia; green, gut; magenta, parenchymal lineages. (**C**) Proportions of cell types identified by Baguñà and Romero by microscopy (left) and as identified by tallying up our annotated Drop-seq clusters (right). The outer ring shows the proportion of each individual cluster, which includes neoblasts, epidermal (epidermal and rhabdite), parenchymal (fixed parenchymal), pigment, neuronal (nerve), muscular, gut (gastrodermal and goblet), secretory (acidophilic and basophilic), and protonephridia (flame) cells. We did not find "striped" cells in our data set. Overall, we find many subtypes for each of the original cell types. (**D** to **F**) tSNE plots (upper panels) showing the expression of marker genes and their expression patterns in adult animals with double in situ hybridizations on tissue sections (lower panels). Nuclei in (D) and (F) were stained with Hoechst and are shown in blue in the overlay. Scale bars: 100 μm. The color scale for tSNE plots ranges from light gray (no expression) to blue (high expression).

protonephridia cell types (fig. S2, A and B). However, in each of these categories, we found several distinct clusters (Fig. 1B) that express different combinations of marker genes (table S3 and fig. S4). This result suggests that our approach can distinguish more cell types than previous studies.

Single-cell transcriptomics unveils previously uncharacterized cell types

In the 1980s, Baguñà and Romero used microscopy to morphologically characterize and count all major cell types in *S. mediterranea* (*30*). We used this resource as a reference to validate the cell types identified by our Drop-seq data and cluster annotation. Even though the microscopy data are of a qualitative nature, we observed a strong correlation between it and our molecular, unbiased Drop-seq annotation (Fig. 1C), suggesting that FACS sorting, cryopreservation or fixation, and cell capture in nanodroplets did not influence cell type proportions. We validated the identity of several clusters by designing RNA probes targeting marker genes and performing in situ hybridizations, both whole mount and in histological sections (fig. S5). We could confirm major known cell types such as different types of neurons, muscle, protonephridia, epidermis, and secretory cells. We identified the two main cell types of the planarian gut: phagocytes (Fig. 1D, red) and goblet cells (Fig. 1D, green), and discovered markers of planarian goblet cells, including a gene without apparent homologs in other phyla. We named this gene *bruixot*. We also distinguished body and pharynx muscle (Fig. 1E). General muscle markers colocalized with body muscle markers throughout the entire body except in the pharynx (Fig. 1E). Pharynx muscle was characterized by the expression of *laminin* (*31*) (fig. S5). The protonephridia cluster (0.3% of our wild-type cells) contained the two main cell types of these organs, flame and tubular cells (*32*) (figs. S4 and S5). In some cases, cell clusters contain several similar subtypes that we cannot distinguish at this resolution. For instance, previously described markers of eye pigment cup cells and photoreceptor neurons (*33*) are expressed in pigment and ChAT neurons 2 clusters, respectively, indicating that the former are subtypes of the latter (fig. S6).

We also validated a recently discovered epidermis cell type, which marks the boundary between the dorsal and ventral parts of planarians (fig. S5) (*34*). Additionally, we identified an epidermal-related pharynx cell type (fig. S5) and several parenchymal cell types previously undescribed molecularly (Fig. 1B and fig. S5). Among parenchymal clusters, we found a diversity of nonoverlapping cells types, including aqp+ and the psap+ parenchymal cells (Fig. 1F), which probably collectively correspond to the previously described fixed parenchymal cells (*30*, *35*), pigment cells (cluster 44) (*36*, *37*), and glial cells (*38*, *39*) (cluster 47) (figs. S4 and S5). Altogether, these results show that we can identify known as well as unknown cell types using single-cell transcriptomics and measure their abundances in a reproducible way.

To investigate the function of newly identified cell types, we used pathway and gene set overdispersion analysis (PAGODA) (*40*) to identify variable gene sets with particular Gene Ontology (GO) terms annotated (fig. S7 and supplementary note 2). The clustering that emerges with these gene sets roughly recapitulates the one obtained with Seurat, showing the robustness of our clustering approach (fig. S7A). This analysis revealed that neoblasts and progenitors are functionally similar, both expressing gene signatures enriched for GO terms related to RNA processing. Additionally, parenchymal clusters showed enrichment for GO terms related to "lysosome," "extracellular region," and hydrolytic enzymes and appear to share metabolic functions with gut cells (fig. S7B).

Single-cell transcriptomics of purified stem cells and stem cell–depleted animals reveals stem, progenitor, and differentiated cell populations

The great diversity of cell types identified, which included stem cells, differentiated cells, and presumably many progenitor cells, offered a unique opportunity for exploring stem cell differentiation and lineage relationships between all cell clusters. We focused on the X1 cell sample, which is enriched in G_2/M neoblasts (*23*, *24*), and the *histone 2b* (*h2b*) RNAi-treated whole-planaria sample, in which stem and progenitor cell populations are depleted (*41*). Cells from these data sets showed a clear distribution pattern: X1 cells were located in the middle of the tSNE plot (Fig. 2A, red dots), whereas *h2b*(*RNAi*)-resistant cells were clearly enriched in the periphery (Fig. 2A, blue dots). This distribution was specific and not the result of batch effects, as evident from the respective control samples [Fig. 2A, X1 control and *gfp*(*RNAi*) samples]. Given that each data set

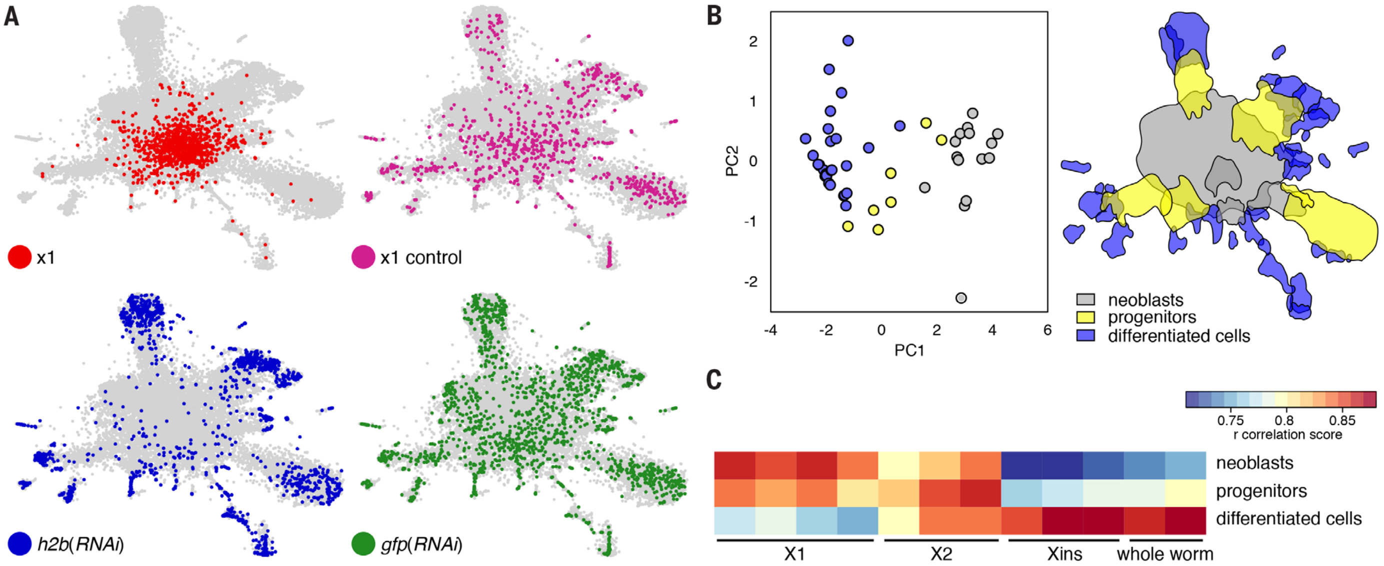


Fig. 2. Neoblast ablation and enrichment experiments show stem and progenitor clusters. (**A**) tSNE plots showing the distribution of the cells of an X1 FACS-sorted sample (red) and its whole-cell population control (x1 control, magenta), and a *h2b*(*RNAi*) sample with its negative control [*gfp*(*RNAi*), green]. X1 cells are enriched in the center of the plot, whereas *h2b*(*RNAi*) cells are enriched in the periphery. (**B**) PCA analysis considering the expression level of neoblast marker genes and the log odds ratio of the amount of cells per cluster from *h2b*(*RNAi*) and X1 experiments compared to wild-type and control samples separates neoblasts (gray), progenitor clusters (yellow), and differentiated cell clusters (blue). The location of these clusters is shown on the tSNE plot on the right. (**C**) Gene expression correlation between bulk RNA-seq data from FACS-sorted X1, X2, Xins populations, and whole worms and the pooled clusters as defined in (B). Neoblasts show a stronger correlation with X1, progenitors with X2, and differentiated cells with Xins and whole worms.

is enriched in particular cell populations, we reasoned that they could be used to distinguish cells in varying differentiation states. We quantified the fraction of cells per cluster from the X1 and *h2b*(*RNAi*) samples and compared them to wild-type and control samples. We performed a principal component analysis (PCA) using these cellular proportions as well as the mean expression of the three top neoblast markers (*Smedwi-1*, *tub-α1*, and *h2b*) in each cluster. The first two principal components resulting from this analysis separated clusters according to their gene expression profiles as neoblasts, progenitors, and differentiated cell clusters (Fig. 2B). Mapping onto the tSNE revealed that progenitor cell clusters were located between differentiated cells and neoblast clusters (Fig. 2B). To corroborate the differentiation state of the cells in the different clusters, we pooled the cells in each group and correlated their gene expression profiles to previously described FACS populations. Neoblast clusters best correlated with X1 populations, corresponding to high-content DNA G_2/M neoblasts; progenitor clusters correlated with X2 populations, a mixture of G_1/S neoblasts and early progenitors; and differentiated cell clusters correlated with Xins samples, a pool of all differentiated cells (Fig. 2C). Altogether, our functional experiments reveal the stem, progenitor, or differentiated status of each cell cluster.

Computational lineage reconstruction predicts a single tree for all major planarian cell differentiation trajectories

Existing methods to investigate cell differentiation using single-cell transcriptomics data were designed to study individual lineages or organs, allow few branching trajectories (*13*, *15*, *18*), and often require high sequencing depth and associated costs (*16*). To overcome these limitations, we developed the general framework of PAGA, which reconciles clustering and pseudotemporal ordering algorithms and allows the inference of complex cell trajectories and differentiation trees (*22*). Starting from the neighborhood graph of single cells, in which cells are represented as nodes, the algorithm quantifies the connectivity of cell clusters and generates a much simpler abstracted graph in which nodes correspond to the clusters identified by Seurat and edges represent putative transitions between clusters. The differentiation tree is then computed as the treelike subgraph in the abstracted graph that best explains all continuous progressions along the original single-cell graph (supplementary note 3).

When running this algorithm, without any assumptions about the tree structure, we obtained an abstracted graph that shows high confidence of the branching events (Fig. 3A) from which we can derive a single differentiation tree that included all the cell types and linked them to a single root, the neoblast 1 cluster. This tree defines independent differentiation branches for all the major tissues such as neurons, muscle, parenchyma, and gut (Fig. 3A). Additionally, the tree reflects the relation between different groups of cells. For example, it predicts the existence of independent progenitor cells for the epidermis dorsoventral boundary and the pharynx cell type lineages, although both lineages are related to the epidermal lineage. By contrast, it shows the presence of a shared progenitor for all parenchymal lineages despite containing cell types as different as glia and pigment cells. The connections in the tree are highly consistent with the continuity of gene expression patterns along the various lineages (fig. S8A) except for two cases: The epidermis cluster itself is disconnected from epidermal lineage, and muscle pharynx is connected to muscle body instead of muscle progenitors (fig. S8A). Together, from 51 clusters (with 1275 possible transitions between them), PAGA predicts 53 transitions that are mainly consistent with our marker-based analysis.

Furthermore, PAGA yields a pseudotemporal ordering of individual cells within each cluster consistent with our stem cell ablation and purification experiments and therefore confidently predicts their differentiation status, even for cell types for which separate progenitor clusters could not be identified (fig. S8B). For instance, when we sorted the goblet cells by pseudotime, we observed a higher percentage of X1 cells in early pseudotime and *h2b*(*RNAi*) cells in the late pseudotime (fig. S8B). To validate this observation, we performed double FISH (fluorescence in situ hybridization) of *bruixot*, our newly identified goblet cell marker, and *adb* (*aprenent de bruixot*), a gene expressed earlier in the goblet cluster pseudotime (fig. S8B). Consistently, *adb* was expressed in the gut (fig. S8C) overlapping with *bruixot*, but staining more cells located in the periphery of the gut that clearly lacked goblet cell morphology (fig. S8D). This indicates that *adb* is a marker of immature goblet cells and that computationally estimated pseudotime correctly orders cells according to their differentiation status.

Although the tree predicts the connectivity of cell clusters, it does not give any information about the direction of the trajectories. Thus, we used the tree topology to estimate the developmental potency of each cluster, i.e., their ability to give rise to other cells. We developed a potency score that is conceptually similar to the stemID score previously proposed to identify stem cells (*16*) but additionally estimates pluripotency versus multi- or unipotency of cell populations. It is computed as the normalized degree of each cluster in the abstracted graph (supplementary note 4). This analysis showed that neoblast 1, the largest stem cell cluster, had a score of 1 (Fig. 3B), correctly assigning pluripotency to neoblasts as expected from earlier studies (*21*). We note that the potency score is independent of prior information and therefore can be used to identify stem cells from single-cell transcriptomics data alone, a feature that is particularly useful in less well-studied non–model organisms. Progenitor clusters showed lower potency than neoblasts and higher potency scores than differentiated cells (Fig. 3B), in agreement with a gradual potency loss. To assess the stem cell and progenitor status of the clusters connected in the center of the PAGA topology, we mapped X1 and *h2b*(*RNAi*) data onto the tree. Most X1 cells were located in the neoblast 1 cluster (Fig. 3C), whereas *h2b*(*RNAi*)-resistant cells were more enriched in the leaves of the tree (Fig. 3D). Thus, both PAGA and stem cell ablation and purification independently support the stem and progenitor status of these clusters.

The remaining neoblast clusters had lower potency scores than the neoblast cluster 1 and were connected to it. These clusters share the majority of marker genes with the neoblast 1 cluster (table S3) and do not correspond to previously identified specialized neoblasts of the sigma, gamma, and zeta class (*26*, *42*, *43*) (fig. S9). Although some of these neoblast clusters are connected to differentiated cell types (Fig. 3A), most do not give rise to differentiated cell types, raising the possibility that they represent neoblasts in different metabolic, cell cycle, or activation states (supplementary note 5).

We detect expression of specialized neoblast markers among both neoblast and progenitor clusters (figs. S10 and S11). Although present in neoblasts, sigma markers were most highly expressed in neural and muscle progenitors, gamma markers in gut, and parenchymal progenitors and zeta markers in epidermal progenitors (fig. S9D). These clusters are mostly devoid of X1 cells (Fig. 3C) and therefore correspond mainly to postmitotic progenitors.

RNA velocity confirms lineage relationships predicted by PAGA

To independently validate the differentiation trajectories predicted by PAGA, we used velocyto (*44*). This method computes RNA velocity, defined as the rate of change of mRNA levels for a gene in time, in every single cell. In differentiating cells in which changes in gene expression are dominated by changes in transcription rates, the ratio of unspliced to spliced reads for a given gene within a cell will be proportional to the temporal change of the logarithm of spliced reads (or mature mRNAs) (*44*). Thus, one can estimate the future mRNA level of a gene by computing its velocity and a linear fit. By aggregating over many genes in a cell, one can estimate the cellular expression state to which the cell is apparently moving in time. We estimated mRNA velocities for each cell and projected the estimated future cell states onto the tSNE, which describe the paths predicted by the mRNA velocity model (Fig. 3E and fig. S12A). These paths show a highly homogeneous stem cell population that moves slowly to progenitors, which will differentiate to mature cell types. The long arrows at the edges of the clusters likely are due to the averaging on the force field, as they do not appear when individual arrows are plotted (fig. S12A). These paths largely agree with the trajectories predicted by PAGA and also confirmed the connection between muscle progenitors and pharynx muscle predicted from gene expression changes (fig. S12A). Additionally, velocyto can also model longer cell trajectories to identify their root (Fig. 3F) and terminal end points (Fig. 3G), which corresponded to the tSNE regions containing stem cells and terminally differentiated cells, respectively. Velocyto cannot provide information from disconnected clusters. As a result, all disconnected clusters contain differentiation trajectories with independent start and end points (Fig. 3, F and G).

The estimates of RNA dynamics obtained with velocyto also identified regions where genes are mainly induced or repressed compared to the steady-state level. This information can be helpful to investigate relations between clusters that appear disconnected on the tSNE. We used these estimates to study the expression of marker genes from the epidermis cluster. These genes are clearly induced in epidermal progenitors and become repressed in mature epidermis, where

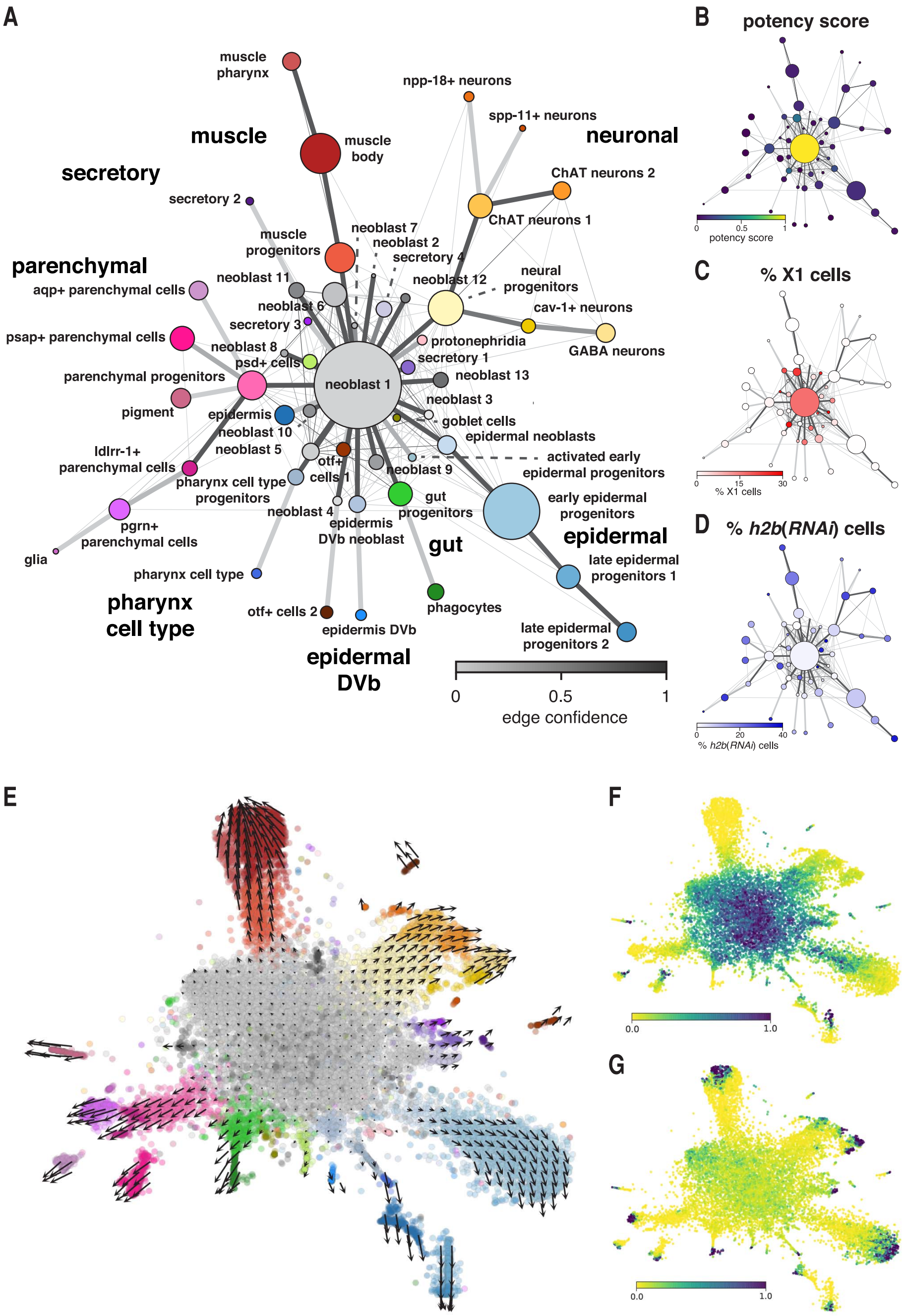


Fig. 3. Lineage tree reconstruction by PAGA and velocyto. (**A**) Abstracted graph showing all the possible edges with a probability higher than 10^{-6} connecting two clusters and their confidence. Each node corresponds to each of the clusters identified with Seurat. The size of nodes is proportional to the amount of cells in the cluster. The most probable path connecting the clusters is plotted on top, with thicker edges. (**B**) Lineage tree colored according to potency score, which ranges from blue (0) to yellow (1). (**C** and **D**) Lineage trees colored according to the percentage of X1 (C) or *h2b*(*RNAi*)-resistant (D) cells in each cluster. X1 cells are most abundant in the neoblast 1 cluster, whereas *h2b*(*RNAi*)-resistant cells are mostly located in the leaves of the tree. (**E**) Velocyto force field showing the average differentiation trajectories (velocity) for cells located in different parts of the tSNE plot. (**F** and **G**) Root (F) and terminal end points (G) obtained after modeling the transition probabilities derived from the RNA velocity by using a Markov process. The color scale represents the density of the end points of the Markov process and ranges from yellow (low) to blue (high).

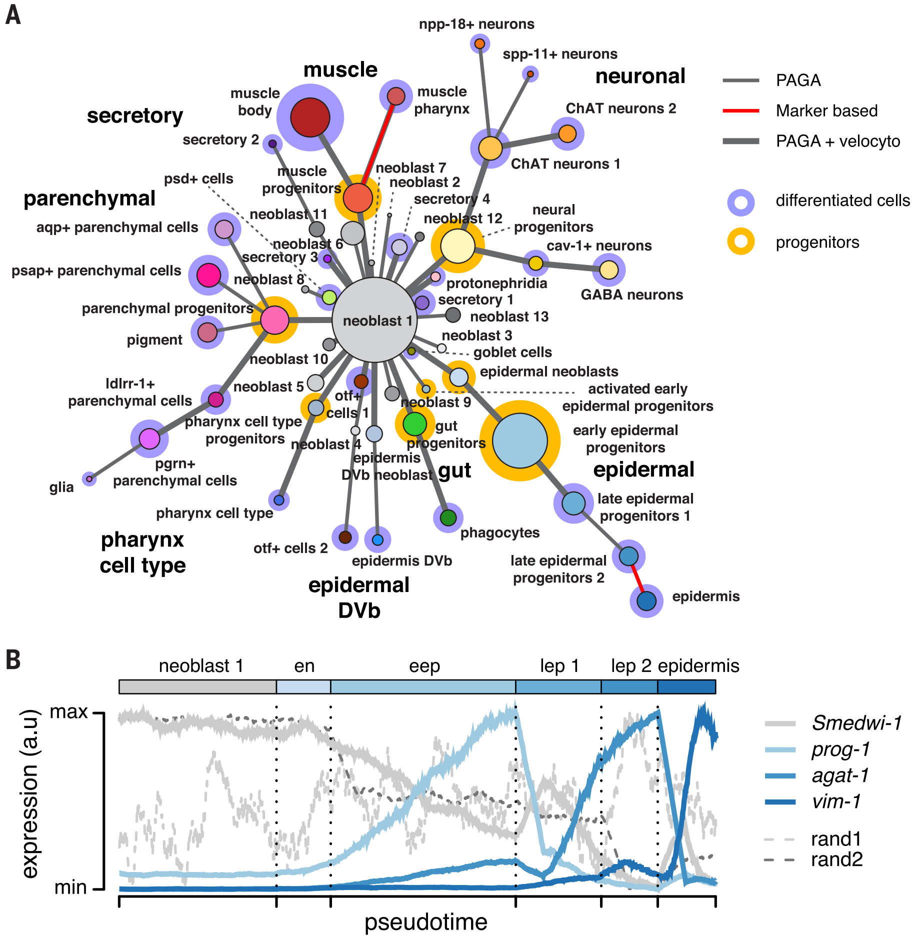


Fig. 4. Consolidated lineage tree of planarian stem cell differentiation into all major types. (**A**) Consolidated lineage tree including four independent sources of evidence. The topology of the tree is shown according to PAGA, and marker-based connections are shown with red edges. Velocyto-supported connections are shown with thick edges. Progenitor and differentiated cell clusters according to neoblast ablation and enrichment experiments are shown with yellow and blue halos, respectively. (**B**) Gene expression changes of marker genes for the individual stages during epidermal differentiation (in pseudotime). Relative expression of marker genes from neoblast (*Smedwi-1*), early (*prog-1*) and late (*agat-1*) progenitors, as well as from the epidermis (*vim-1*). A maximum of 1000 cells from neoblast 1, epidermal neoblasts (en), early epidermal progenitors (eep), late epidermal progenitors 1 (lep 1) and 2 (lep 2), and epidermis were sampled. Gray thin dashed lines show the expression of *Smedwi-1* after randomly permuting cells (rand1) or after randomly sorting cells within each cluster (rand2).

they are highly expressed (fig. S12B). Thus, mRNA metabolism patterns provide additional support to the differentiation trajectory connecting late epidermal progenitors to epidermis that we predicted on the basis of gene expression changes (fig. S8A).

A consolidated lineage tree of planarian stem cell differentiation into all major cell types

Taken together, our results show that both computational and experimental methods agree in the identification of stem cells, progenitors, and differentiated cells. By combining all four independent lines of evidence (PAGA, gene expression changes, stem cell ablation and enrichment experiments, and velocyto), we provide a single consolidated tree that models stem cell differentiation trajectories into all identified cell types of adult planarians (Fig. 4A). The resulting cell lineage tree correctly recapitulates the known expression changes described during epidermal differentiation (*26*, *34*, *45*) (Fig. 4B). We observed a continuous decrease of the expression of *Smedwi-1*, a well-characterized neoblast marker (fig. S3), with pseudotime progression, whereas early (*prog-1*) and late (*agat-1*) epidermal progenitor, as well as mature epidermis (*vim-1*) markers, increased their expression at consecutive time points (Fig. 4B).

According to the consolidated lineage tree, neoblasts (35% of our wild-type cells) differentiate into at least 23 independent cell lineages. There are six major differentiation fates (57% of cells) (table S2), each representing more than 1% of total cells: epidermal, parenchymal, neural, muscle, gut, and a pharynx cell type. For these major fates, we identified progenitor and differentiated states. Additionally, we identified 10 minor lineages (6% of cells; each less abundant than 1% of total cells) that differentiate from the neoblasts, but for which we were unable to identify progenitors.

Self-organizing maps identify gene programs underlying cell differentiation

We used our data to identify gene sets that coordinately change their expression during differentiation. For this analysis, we discarded all cells from neoblast clusters that did not give rise to differentiated cell types in our consolidated cell lineage tree. The remaining cells were ordered following the tree for each lineage and sorted within each cluster according to their pseudotemporal ordering (Fig. 5A). Subsequently, we used self-organizing maps (SOMs) (*46*) to identify 48 sets of highly variable genes that coordinately change their expression during differentiation (*47*) (Fig. 5B, fig. S13, and table S4). Many of these sets contain some genes previously known to be expressed in the respective lineages and in some cases involved in their differentiation (table S5). For instance, gene sets 10 and 11 contain genes that are highly expressed in neoblast and progenitor clusters, such as *Smedwi-1* and *tub-α1*, whose expression drops during differentiation (Fig. 5B and fig. S13). Similarly, we found gene sets that are regulated along muscle, neuronal, parenchymal, gut, and epidermal differentiation (Fig. 5B, top row). They contain genes expressed in these lineages, such as *mhc* for the muscle and *chat* for the neuronal lineage, but also include well-known regulators of their differentiation, such as *myoD* (*48*) and *coe* (*49*). As a consequence of analyzing all detected planarian cell lineages simultaneously, we not only identified gene sets involved in lineage specific programs but also gene sets co-regulated during the differentiation of several fates (Fig. 5B, middle and bottom row). Taken together, these results show that single-cell transcriptomics of a whole organism allows the reconstruction of specific differentiation events for many differentiation fates in parallel, enabling the identification of previously undetected combinations of co-regulated genes.

Molecular profiling of planarian regeneration by single-cell transcriptomics

Freshwater planarians are well known for their remarkable regenerative capacities. Planarians can be cut into small pieces, and each piece (except for the pharynx and the most anterior tip of the head, which are devoid of neoblasts) can regenerate a complete, albeit much smaller, organism in a matter of days. This process is dynamically complex and involves the orchestration of all cellular differentiation pathways. The animal does not grow (as it cannot eat) during the process. Thus, the truncated body fragments need to reshape their body proportions to adjust to their new size by a process termed morphallaxis ([50](#)). It is still largely unknown how each individual cell type behaves in this process.

Given the detected cell type abundances and the cell differentiation tree of steady-state adult animals, we asked if we could use Drop-seq to profile the cellular and transcriptomic changes that occur during regeneration. We cut planarians into five to seven pieces, discarded the head piece, and prepared the remaining body pieces for single-cell transcriptomics immediately after cutting (day 0), and 2 and 4 days after cutting (Fig. 6A and fig. S14). We compared regenerating samples to those at day 0 using Seurat and detected hundreds of differentially expressed genes in both samples (tables S6 and S7). By pooling all cells, we could detect up-regulation after 2 days of regeneration of 16 of the 128 wound-induced genes described in a previous study (*42*, *51*) (table S8 and fig. S15A). The shallowness of Drop-seq data makes it difficult to assess differences in lowly expressed genes. However, Drop-seq allows the cell types that undergo these changes to be distinguished, showing that *runt-1* and *egr-2* are up-regulated in the neoblast 1 cluster (fig. S15B) and *jun-1* in the muscle body cluster (fig. S15C and tables S6 and S7).

All cells from the regenerating samples fall into clusters that are present in wild-type samples, indicating an absence of regeneration-specific types or trajectories (table S2). However, our analysis revealed notable changes in cell composition during regeneration (Fig. 6B and fig. S16). On one hand, we observed a large increase in the number of neoblasts, consistent with an increase in mitotic activity, and of neural progenitors, reflecting active neurogenesis to replace missing brain structures after head removal. On the other hand, we detected that both parenchymal progenitor cells and differentiated parenchymal cell types were depleted (Fig. 6B), indicating that these cells are cleared in the process of reshaping the planarian tissue. The cell proportion changes at days 2 and 4 were clearly correlated (Fig. 6C), indicating that *aqp*-positive parenchymal cells are the most depleted cell type. We experimentally confirmed this observation by in situ hybridization on planarian tissue sections (Fig. 6D) and counting *aqp*-positive parenchymal cells (Fig. 6E) (Mann Whitney U test p-value < 10^{-7}). Our results indicate that parenchymal cells are highly depleted upon regeneration, implying that they may be used to metabolically fuel the regeneration process (*52*).

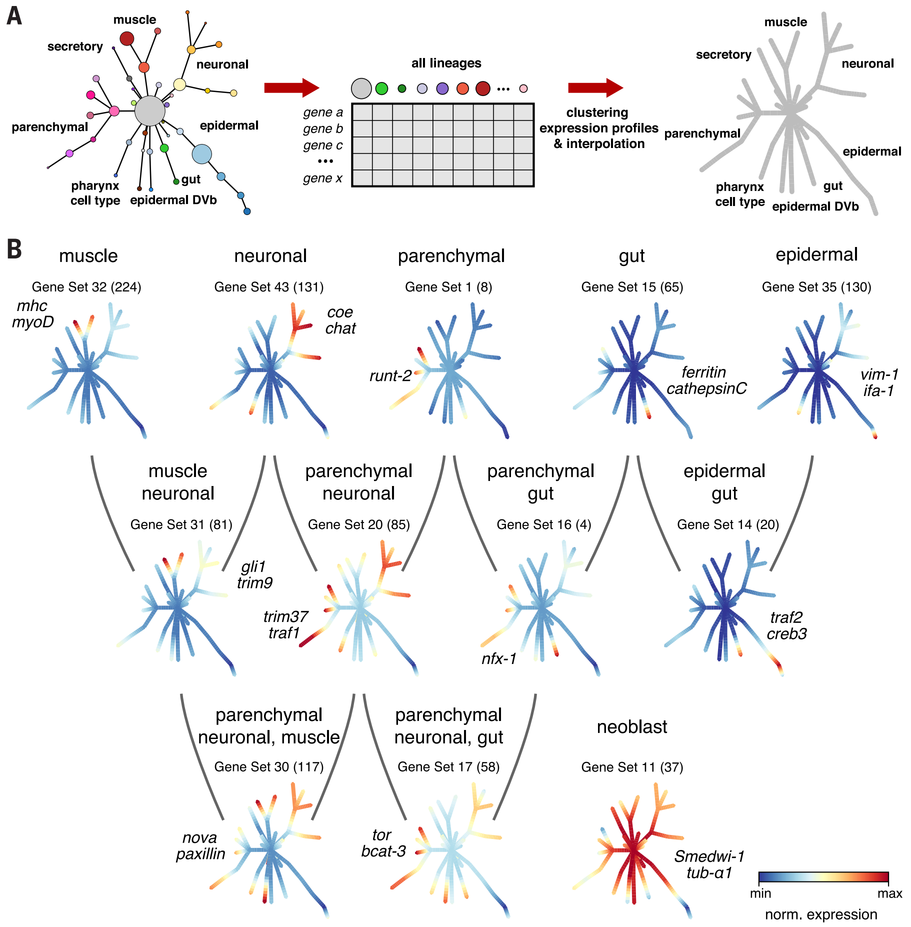


Fig. 5. Identification of gene sets regulated and co-regulated in cell differentiation. (**A**) Schematic workflow of the analysis performed to identify gene sets involved in the differentiation processes. Pseudotemporal ordering of the cells from all lineages and clustering of variable genes using SOMs allowed the identification of 48 gene sets. (**B**) Graphical representation of gene expression changes during cell differentiation of 12 gene sets. For each gene set, the normalized expression of the genes is shown on the edges of the tree and ranges from blue (low expression) to red (high expression). Next to each tree, representative genes from each gene set are highlighted.

Discussion

In this study, we made use of the stem cell population and the extreme regeneration capabilities of adult flatworms to generate an atlas of cell types at high resolution. We identified, quantified, and molecularly characterized 37 cell types, including 23 terminally differentiated cell types, and numerous progenitor and stem cell clusters. Although our sequencing data are relatively shallow, molecular characterization of cell types with computational methods was robust, agreed well with previously published microscopy data, and revealed progenitor and differentiated cells. This implies that the grouping of incomplete transcriptomes of thousands of cells into clusters did not suffer appreciably from capture rates or other confounding factors.

The resolution of our data depends on both the number of cells sequenced and the number of genes detected per cell. Considering only wild-type and control samples (~11,000 cells), we can identify differentiated cell clusters containing about 10 cells. Therefore, we estimate that cells present at a frequency of <1:1000, such as *cintillo*+ cells (*53*) and photoreceptor neurons (*33*), will be missed by our approach. Furthermore, we failed to identify certain neoblast subpopulations previously described in the literature (*26*, *42*). This result could be due to the low sensitivity of Drop-seq, which captures only a fraction of mRNAs in a cell. However, we do detect the expression of the proposed marker genes of these subpopulations. They appear to be spread among neoblasts and progenitor clusters (fig. S9), which still express neoblast markers such as *Smedwi-1* at low levels (figs. S10 and S11). This result indicates that the boundary between stem cells and lineage-committed progenitors is probably not sharp. Further studies will help in describing and delimiting these boundaries.

Projecting high-dimensional gene expression data of thousands of transcriptomes onto a two-dimensional plot [for example, by the widely used tSNE method (*54*)] visually reveals clusters. However, it is impossible to infer the relationships among them, as the distances between clusters cannot be interpreted as differentiation trajectories. To solve this problem, we used computational and experimental methods to reconstruct a lineage tree. PAGA and velocyto provide two complementary approaches to study cell differentiation by using single-cell transcriptomics. Whereas velocyto allows the differentiation trajectories (including their directionality) of individual cells to be located within a cell continuum on the basis of mRNA metabolism, PAGA allows the average differentiation paths of a group of

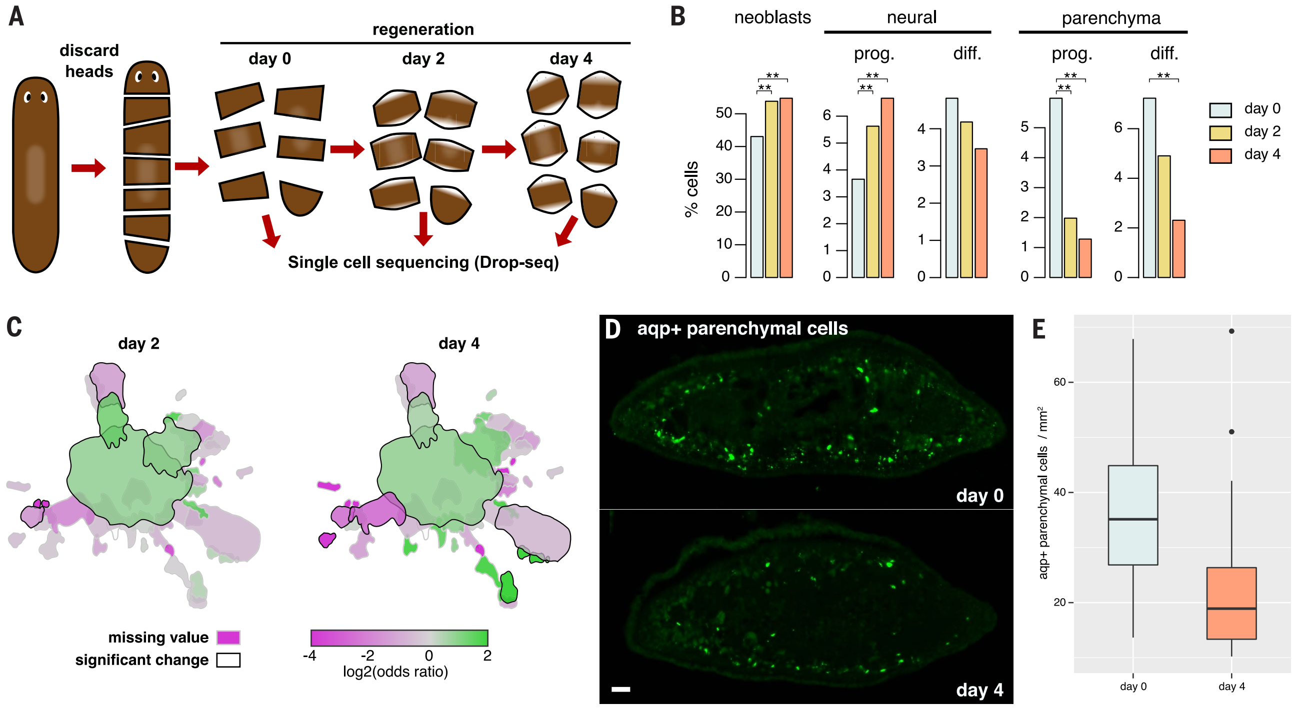


Fig. 6. Molecular profiling of regeneration by single-cell transcriptomics. (**A**) Experimental workflow: Planarians were cut into small pieces; head pieces were discarded and the remaining pieces were processed for single-cell RNA sequencing 0, 2, and 4 days after cutting. (**B**) Quantification of neoblasts, neural progenitors, and differentiated clusters and parenchymal progenitors and differentiated clusters. Significant differences calculated using Fisher's exact test with an adjusted *p*-value < 0.001 are marked with **. (**C**) Cluster outlines colored according to the log2 (odds ratio) of changes in regeneration at day 2 (left) and day 4 (right) versus day 0, showing enriched clusters in green colors and depleted clusters in magenta colors. Significant changes are indicated by black solid outlines. (**D** and **E**) In situ hybridization on sections (D) and quantification (E) of aqp+ parenchymal cells in regenerating planarians after 0 and 4 days of regeneration. Mann Whitney U test *p*-value < 10^{-7}. Scale bar: 100 μm.

cells to be inferred, even when they are disconnected. Thus, combining these two computational methods results in a robust lineage prediction. This prediction is supported by the continuity of expression of marker genes and the mapping of stem cells and differentiated cells on the tree (Fig. 4A), and validates known differentiation trajectories such as that of the epidermal lineage (Fig. 4B) (*34*).

We used PAGA (*22*) to reconstruct in an unbiased way the lineage tree of all major planarian cells. This method, although indirect, allows the lineage information to be reconstructed from the transcriptomic snapshot of individual cells. Additionally, in contrast to high-throughput lineage-tracing methods (*9*–*12*), which rely on the use of transgenic or CRISPR-Cas tools, it can be applied to every species provided that single cells can be isolated and sequenced. Using this method, we identified de novo planarian stem cells and predict their differentiation paths to at least 23 different lineages, including several multipotent progenitor populations. Notably, these tools can readily be applied to other organisms to identify de novo stem cells, identify their differentiation trajectories, and estimate the developmental potency of the resulting cell populations.

Pseudotemporal ordering of cells along these lineages allowed us to discover gene sets that are putatively involved in differentiation programs, highlighting the similarities and differences that exist across tissues and identifying several genes known to be involved in cell differentiation not only in planarians but also in other species, such as *myoD*, *nkx6*, and *pax6*. Further characterization of these gene sets should be the subject of future studies. As we show for the *h2b* RNAi phenotype, Drop-seq also allows profiling of perturbation studies at both the transcriptomic and cellular levels. In general, and beyond planaria, we can foresee that future studies will use single-cell transcriptomics coupled to loss-of-function experiments to unravel the specific developmental functions of genes or regulatory networks.

Furthermore, we used single-cell transcriptomics to profile cellular abundance changes upon regeneration. Our experiments revealed that several of our newly described parenchymal cell types are depleted in regeneration. These cell types had been largely overlooked in molecular studies but had been described, on the basis of microscopy, in the literature decades ago. This is in part due to the unbiased nature of both microscopic and single-cell transcriptomic data. These parenchymal cells are highly enriched in lysosomes and other vacuoles and might be an energy reservoir that regenerating planarians mobilize to fuel regeneration.

To make our data easily accessible, we built an interactive app that allows users to query and interpret all sequencing data (https://shiny.mdc-berlin.de/psca). We also provide a detailed tutorial for the lineage reconstruction algorithm PAGA that we hope will serve as a reference for future studies (https://github.com/rajewsky-lab/planarian_lineages). Together, our results show that single-cell expression profiling can be used to systematically annotate cell types of entire animals, to reconstruct stem cell differentiation lineages and identify gene sets that likely contain genes involved in programming them, and to study complex processes such as regeneration (and their relation to lineages used in normal development) at single-cell resolution. Our results and methods demonstrate that single-cell approaches will become an indispensable method to study developmental and regeneration biology.

Methods summary

Single-cell transcriptomic profiling of asexual adult planarians from the species *S. mediterranea* was performed with Drop-seq. Single-cell suspensions were prepared by dissociating cells from adult planarians with trypsin. We used FACS to discard broken and dead cells. Cells were either directly processed for Drop-seq or preserved in methanol or DMSO for later processing. For RNAi experiments, animals were injected with double-stranded RNA against the coding region of *h2b* or *gfp* for three consecutive days and kept at 20°C; their cells were prepared for FACS and single-cell transcriptomics 5 days after the third injection. For regeneration experiments, animals ranging from 4 to 10 mm in size were cut into five to seven pieces; the head pieces were discarded, and the remaining pieces were processed for Drop-seq immediately, or 2 or 4 days after cutting.

Computational analysis of the sequenced samples was done with Drop-seq tools and the Seurat package (*3*). Briefly, reads were mapped to the *S. mediterranea* dd_Smed_v6 transcriptome and processed with Drop-seq tools and custom perl scripts to generate Digital Gene Expression (DGE) matrices for each sample. Finally, all DGE matrices were joined. Variable genes across all clusters were used to perform a PCA. The first 50 principal components obtained were then tested for statistical significance, and those with a p-value $< 10^{-5}$ were used to perform clustering. The robustness of the obtained clusters was assessed, and spurious clusters were merged to obtain a final set of 51 clusters.

Cell type identification was performed by calculating marker genes for each cluster. Manual inspection, comparison to previously published single-cell data, and experimental validation using in situ hybridizations of the marker genes reported allowed the identification of the different cell populations. Additional characterization of the identified cell types was performed by characterizing GO-term–based gene sets with PAGODA (*40*). Experimental validation of cell types was done by whole-mount in situ hybridization and in situ hybridization on histological sections as previously described and using probes complementary to marker genes (table S9).

Lineage reconstruction was done by combining the unsupervised graph obtained with the PAGA algorithm (*22*) with velocyto (*44*), gene expression analysis, and experimental data from *h2b*(*RNAi*) and X1 FACS-sorted cells. To calculate RNA velocity with velocyto, we mapped the reads from all data sets to the planarian genome to extract spliced and unspliced reads. These analyses allowed pseudotemporal ordering of cells that was used to identify gene sets that change during stem cell differentiation by using self-organizing maps.

To perform cell counting of regenerating planarians, positive cells were automatically counted by using a custom script for ImageJ (https://imagej.net).

REFERENCES AND NOTES

1. A. Regev *et al*., The Human Cell Atlas. *eLife* **6**, e27041 (2017). doi: 10.7554/eLife.27041; pmid: 29206104
2. D. A. Jaitin *et al*., Massively parallel single-cell RNA-seq for marker-free decomposition of tissues into cell types. *Science* **343**, 776–779 (2014). doi: 10.1126/science.1247651; pmid: 24531970
3. E. Z. Macosko *et al*., Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. *Cell* **161**, 1202–1214 (2015). doi: 10.1016/j.cell.2015.05.002; pmid: 26000488
4. K. Shekhar *et al*., Comprehensive classification of retinal bipolar neurons by single-cell transcriptomics. *Cell* **166**, 1308–1323.e30 (2016). doi: 10.1016/j.cell.2016.07.054; pmid: 27565351
5. A. C. Villani *et al*., Single-cell RNA-seq reveals new types of human blood dendritic cells, monocytes, and progenitors. *Science* **356**, eaah4573 (2017). doi: 10.1126/science.aah4573; pmid: 28428369
6. D. Grün *et al*., Single-cell messenger RNA sequencing reveals rare intestinal cell types. *Nature* **525**, 251–255 (2015). doi: 10.1038/nature14966; pmid: 26287467
7. N. Karaiskos *et al*., The *Drosophila* embryo at single-cell transcriptome resolution. *Science* **358**, 194–199 (2017). doi: 10.1126/science.aan3235; pmid: 28860209
8. J. Cao *et al*., Comprehensive single-cell transcriptional profiling of a multicellular organism. *Science* **357**, 661–667 (2017). doi: 10.1126/science.aam8940; pmid: 28818938
9. A. McKenna *et al*., Whole-organism lineage tracing by combinatorial and cumulative genome editing. *Science* **353**, aaf7907 (2016). doi: 10.1126/science.aaf7907; pmid: 27229144
10. K. L. Frieda *et al*., Synthetic recording and in situ readout of lineage information in single cells. *Nature* **541**, 107–111 (2017). doi: 10.1038/nature20777; pmid: 27869821
11. B. Spanjaard *et al*., Simultaneous lineage tracing and cell-type identification using CRISPR–Cas9-induced genetic scars. *Nat. Biotechnol.* **2018**, 24 (2018). doi: 10.1038/nbt.4124; pmid: 29644996
12. A. Alemany, M. Florescu, C. S. Baron, J. Peterson-Maduro, A. van Oudenaarden, Whole-organism clone tracing using single-cell sequencing. *Nature* **556**, 108–112 (2018). doi: 10.1038/nature25969; pmid: 29590089
13. L. Haghverdi, M. Büttner, F. A. Wolf, F. Buettner, F. J. Theis, Diffusion pseudotime robustly reconstructs lineage branching. *Nat. Methods* **13**, 845–848 (2016). doi: 10.1038/nmeth.3971; pmid: 27571553
14. M. Setty *et al*., Wishbone identifies bifurcating developmental trajectories from single-cell data. *Nat. Biotechnol.* **34**, 637–645 (2016). doi: 10.1038/nbt.3569; pmid: 27136076
15. X. Qiu *et al*., Reversed graph embedding resolves complex single-cell trajectories. *Nat. Methods* **14**, 979–982 (2017). doi: 10.1038/nmeth.4402; pmid: 28825705
16. D. Grün *et al*., De novo prediction of stem cell identity using single-cell transcriptome data. *Cell Stem Cell* **19**, 266–277 (2016). doi: 10.1016/j.stem.2016.05.010; pmid: 27345837
17. F. Notta *et al*., Distinct routes of lineage development reshape the human blood hierarchy across ontogeny. *Science* **351**, aab2116 (2016). doi: 10.1126/science.aab2116; pmid: 26541609
18. J. Shin *et al*., Single-cell RNA-Seq with Waterfall reveals molecular cascades underlying adult neurogenesis. *Cell Stem Cell* **17**, 360–372 (2015). doi: 10.1016/j.stem.2015.07.013; pmid: 26299571
19. B. Treutlein *et al*., Reconstructing lineage hierarchies of the distal lung epithelium using single-cell RNA-seq. *Nature* **509**, 371–375 (2014). doi: 10.1038/nature13173; pmid: 24739965
20. L. Velten *et al*., Human haematopoietic stem cell lineage commitment is a continuous process. *Nat. Cell Biol.* **19**, 271–281 (2017). doi: 10.1038/ncb3493; pmid: 28319093
21. D. E. Wagner, I. E. Wang, P. W. Reddien, Clonogenic neoblasts are pluripotent adult stem cells that underlie planarian regeneration. *Science* **332**, 811–816 (2011). doi: 10.1126/science.1203983; pmid: 21566185
22. F. A. Wolf *et al*., Graph abstraction reconciles clustering with trajectory inference through a topology preserving map of single cells. bioRxiv 208819 [Preprint]. 25 October 2017.
23. T. Hayashi, M. Asami, S. Higuchi, N. Shibata, K. Agata, Isolation of planarian X-ray-sensitive stem cells by fluorescence-activated cell sorting. *Dev. Growth Differ.* **48**, 371–380 (2006). doi: 10.1111/j.1440-169X.2006.00876.x; pmid: 16872450
24. P. Önal *et al*., Gene expression of pluripotency determinants is conserved between mammalian and planarian stem cells. *EMBO J.* **31**, 2755–2769 (2012). doi: 10.1038/emboj.2012.110; pmid: 22543868
25. R. M. Labbé *et al*., A comparative transcriptomic analysis reveals conserved features of stem cell pluripotency in planarians and mammals. *Stem Cells* **30**, 1734–1745 (2012). doi: 10.1002/stem.1144; pmid: 22696458
26. J. C. van Wolfswinkel, D. E. Wagner, P. W. Reddien, Single-cell analysis reveals functionally distinct classes within the planarian stem cell compartment. *Cell Stem Cell* **15**, 326–339 (2014). doi: 10.1016/j.stem.2014.06.007; pmid: 25017721
27. J. Solana *et al*., Conserved functional antagonism of CELF and MBNL proteins controls stem cell-specific alternative splicing in planarians. *eLife* **5**, e16797 (2016). doi: 10.7554/eLife.16797; pmid: 27502555
28. J. Alles *et al*., Cell fixation and preservation for droplet-based single-cell transcriptomics. *BMC Biol.* **15**, 44 (2017). doi: 10.1186/s12915-017-0383-5; pmid: 28526029
29. A. Guillaumet-Adkins *et al*., Single-cell transcriptome conservation in cryopreserved cells and tissues. *Genome Biol.* **18**, 45 (2017). doi: 10.1186/s13059-017-1171-9; pmid: 28249587
30. J. Baguñà, R. Romero, Quantitative-analysis of cell-types during growth, Degrowth and regeneration in the planarians Dugesia Mediterranea and Dugesia Tigrina. *Hydrobiologia* **84**, 181–194 (1981). doi: 10.1007/BF00026179
31. F. Cebrià, P. A. Newmark, Morphogenesis defects are associated with abnormal nervous system regeneration following roboA RNAi in planarians. *Development* **134**, 833–837 (2007). doi: 10.1242/dev.02794; pmid: 17251262
32. J. C. Rink, H. T.-K. Vu, A. Sánchez Alvarado, The maintenance and regeneration of the planarian excretory system are regulated by EGFR signaling. *Development* **138**, 3769–3780 (2011). doi: 10.1242/dev.066852; pmid: 21828097
33. S. W. Lapan, P. W. Reddien, Transcriptome analysis of the planarian eye identifies ovo as a specific regulator of eye regeneration. *Cell Reports* **2**, 294–307 (2012). doi: 10.1016/j.celrep.2012.06.018; pmid: 22884275
34. O. Wurtzel, I. M. Oderberg, P. W. Reddien, Planarian epidermal stem cells respond to positional cues to promote cell-type diversity. *Dev. Cell* **40**, 491–504.e5 (2017). doi: 10.1016/j.devcel.2017.02.008; pmid: 28292427
35. K. J. Pedersen, Studies on the nature of planarian connective tissue. *Z. Zellf. Mikrosk. Anat.* **53**, 569–608 (1961). doi: 10.1007/BF00339508
36. B. M. Stubenhaus *et al*., Light-induced depigmentation in planarians models the pathophysiology of acute porphyrias. *eLife* **5**, e14175 (2016). doi: 10.7554/eLife.14175; pmid: 27240733
37. C. Wang *et al*., Forkhead containing transcription factor Albino controls tetrapyrrole-based body pigmentation in planarian. *Cell Discov.* **2**, 16029 (2016). doi: 10.1038/celldisc.2016.29; pmid: 27551436
38. I. E. Wang, S. W. Lapan, M. L. Scimone, T. R. Clandinin, P. W. Reddien, Hedgehog signaling regulates gene expression in planarian glia. *eLife* **5**, e16996 (2016). doi: 10.7554/eLife.16996; pmid: 27612382
39. R. H. Roberts-Galbraith, J. L. Brubacher, P. A. Newmark, A functional genomics screen in planarians reveals regulators of whole-brain regeneration. *eLife* **5**, e17002 (2016). doi: 10.7554/eLife.17002; pmid: 27612384
40. J. Fan *et al*., Characterizing transcriptional heterogeneity through pathway and gene set overdispersion analysis. *Nat. Methods* **13**, 241–244 (2016). doi: 10.1038/nmeth.3734; pmid: 26780092
41. J. Solana *et al*., Defining the molecular profile of planarian pluripotent stem cells using a combinatorial RNAseq, RNA interference and irradiation approach. *Genome Biol.* **13**, R19 (2012). doi: 10.1186/gb-2012-13-3-r19; pmid: 22439894
42. O. Wurtzel *et al*., A generic and cell-type-specific wound response precedes regeneration in planarians. *Dev. Cell* **35**, 632–645 (2015). doi: 10.1016/j.devcel.2015.11.004; pmid: 26651295
43. A. M. Molinaro, B. J. Pearson, In silico lineage tracing through single cell transcriptomics identifies a neural stem cell population in planarians. *Genome Biol.* **17**, 87 (2016). doi: 10.1186/s13059-016-0937-9; pmid: 27150006
44. G. La Manno *et al*., RNA velocity in single cells. bioRxiv 206052 [Preprint]. 19 October 2017.
45. G. T. Eisenhoffer, H. Kang, A. Sánchez Alvarado, Molecular analysis of stem cells and their descendants during cell turnover and regeneration in the planarian *Schmidtea mediterranea*. *Cell Stem Cell* **3**, 327–339 (2008). doi: 10.1016/j.stem.2008.07.002; pmid: 18786419
46. T. Kohonen, The self-organizing map. *Proc. IEEE* **78**, 1464–1480 (1990). doi: 10.1109/5.58325
47. P. Tamayo *et al*., Interpreting patterns of gene expression with self-organizing maps: Methods and application to hematopoietic differentiation. *Proc. Natl. Acad. Sci. U.S.A.* **96**, 2907–2912 (1999). doi: 10.1073/pnas.96.6.2907; pmid: 10077610
48. M. L. Scimone, L. E. Cote, P. W. Reddien, Orthogonal muscle fibres have different instructive roles in planarian regeneration. *Nature* **551**, 623–628 (2017). doi: 10.1038/nature24660; pmid: 29168507
49. M. W. Cowles, K. C. Omuro, B. N. Stanley, C. G. Quintanilla, R. M. Zayas, COE loss-of-function analysis reveals a genetic program underlying maintenance and regeneration of the nervous system in planarians. *PLOS Genet.* **10**, e1004746 (2014). doi: 10.1371/journal.pgen.1004746; pmid: 25356635
50. K. Agata, Y. Saito, E. Nakajima, Unifying principles of regeneration I: Epimorphosis versus morphallaxis. *Dev. Growth Differ.* **49**, 73–78 (2007). doi: 10.1111/j.1440-169X.2007.00919.x; pmid: 17335428

51. D. Wenemoser, P. W. Reddien, Planarian regeneration involves distinct stem cell responses to wounds and tissue absence. *Dev. Biol.* **344**, 979–991 (2010). doi: 10.1016/j.ydbio.2010.06.017; pmid: 20599901
52. C. González-Estévez, D. A. Felix, A. A. Aboobaker, E. Saló, Gtdap-1 promotes autophagy and is required for planarian remodeling during regeneration and starvation. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 13373–13378 (2007). doi: 10.1073/pnas.0703588104; pmid: 17686979
53. N. J. Oviedo, P. A. Newmark, A. Sánchez Alvarado, Allometric scaling and proportion regulation in the freshwater planarian *Schmidtea mediterranea*. *Dev. Dyn.* **226**, 326–333 (2003). doi: 10.1002/dvdy.10228; pmid: 12557210
54. L. Van Der Maaten, G. Hinton, Visualizing Data using t-SNE. *J. Mach. Learn. Res.* **9**, 2579–2605 (2008).

ACKNOWLEDGMENTS

We thank J. Alles and A. Boltengagen for help with single-cell sequencing and members of the Rajewsky lab for discussions. We thank S. Linnarsson and G. La Manno for help running velocyto. We also thank N. Friedman and A. García-Pérez for helpful comments. **Funding:** Work on this project was funded by the German Center for Cardiovascular Research (DZHK BER 1.2 VD), the DFG (grant RA 838/5-1), the BMBF (grant 01ZX1711A), and the Helmholtz Association (Incubator grant sparse2big, grant ZT-I-0007). F.A.W. acknowledges the support of the Helmholtz Postdoc Program, Initiative and Networking Fund of the Helmholtz Association. **Author contributions:** J.S., M.P., C.K., and N.R. designed the project. J.S., C.K., and N.R. designed all experiments. S.A., J.S., and C.K. performed all experiments. M.P. performed and coordinated computational analyses. F.A.W. performed the PAGA analysis with the supervision of F.J.T. A.M. performed velocyto analyses. P.G. developed the Shiny app and performed image analysis. B.O. developed the visualization of gene expression along the lineage tree. All authors discussed and interpreted the data. M.P., J.S., and N.R. wrote the manuscript with input from all other authors. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** The sequencing data generated are available in Gene Expression Omnibus under the accession GSE103633. The single-cell data generated can be interactively accessed at https://shiny.mdc-berlin.de/psca. Manuals to run and reproduce velocyto and PAGA are available on Github (https://github.com/rajewsky-lab/planarian_lineages).

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/360/6391/eaaq1723/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S22
Tables S1 to S10
References (*55*–*74*)

9 October 2017; resubmitted 14 February 2018
Accepted 12 April 2018
Published online 19 April 2018
10.1126/science.aaq1723

RESEARCH ARTICLE SUMMARY

MICROBIOME

Gut microbiome–mediated bile acid metabolism regulates liver cancer via NKT cells

Chi Ma, Miaojun Han, Bernd Heinrich, Qiong Fu, Qianfei Zhang, Milan Sandhu, David Agdashian, Masaki Terabe, Jay A. Berzofsky, Valerie Fako, Thomas Ritz, Thomas Longerich, Casey M. Theriot, John A. McCulloch, Soumen Roy, Wuxing Yuan, Vishal Thovarai, Shurjo K. Sen, Mathuros Ruchirawat, Firouzeh Korangy, Xin Wei Wang, Giorgio Trinchieri, Tim F. Greten*

INTRODUCTION: Primary liver tumors and liver metastasis currently represent the leading cause of cancer-related deaths. The liver intimately cross-talks with the gut and performs many essential functions related to digestion, metabolism of nutrients, and clearance of bacterial metabolites. Diseased livers are often associated with altered gut bacterial composition, or dysbiosis, and it has been suggested that gut bacterial products contribute to malignant transformation of hepatocytes. The liver is exposed to the gut microbiome through the portal vein and is an immunological organ that is heavily populated by immune cells. Emerging studies have shown that gut commensal bacteria are important regulators of antitumor immunity. Although it has been established that the gut microbiome influences the efficacy of cancer immunotherapy, the role of gut bacteria in antitumor surveillance in the liver is poorly understood.

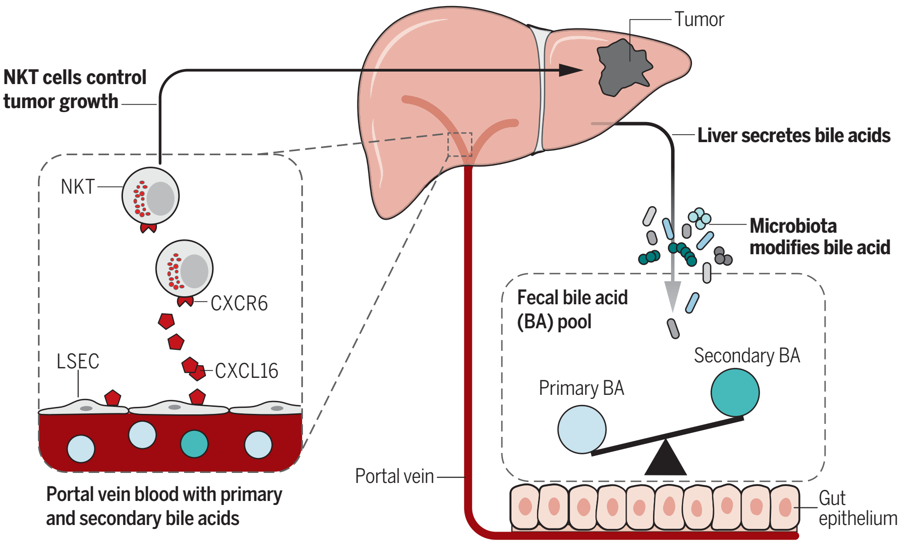
RATIONALE: The liver is exposed to gut bacterial metabolites and products by way of blood from the intestine, which comprises 70% of the whole liver blood supply. Changes in the gut microbiome may affect immune cell function in the liver, and commensal bacteria can mediate the metabolism of primary into secondary bile acids, which recirculate back into the liver through the enterohepatic circulation. Given that bile acids are known to be involved in liver cancer development, we focused on the role of bile acids in immunosurveillance of tumors growing in the liver. We altered gut bacteria and examined changes of hepatic immune cells and antitumor immunity directed against liver tumors. Uncovering how the gut microbiome uses bile acids to shape immunity to liver cancer may have future therapeutic applications.

RESULTS: Using one primary liver model and three liver metastasis models, we found that altering commensal gut bacteria induced a liver-selective antitumor effect. A selective increase of hepatic CXCR6$^+$ natural killer T (NKT) cells was observed, independent of mouse strain, gender, or presence of liver tumors. The accumulated hepatic NKT cells showed an activated phenotype and produced more interferon-γ upon antigen stimulation. In vivo studies using both antibody-mediated cell depletion and NKT-deficient mice confirmed that NKT cells mediated the inhibition of tumor growth in the liver. Further investigation showed that NKT cell accumulation was regulated by the expression of CXCL16, the solo ligand for CXCR6, on liver sinusoidal endothelial cells, which form the lining of liver capillaries and the first barrier for the blood coming from the gut entering the liver. Primary bile acids increased *CXCL16* expression, whereas secondary bile acids showed the opposite effect. Removing gram-positive bacteria by antibiotic treatment with vancomycin, which contains the bacteria mediating primary-to-secondary bile acid conversion, was sufficient to induce hepatic NKT cell accumulation and decrease liver tumor growth. Feeding secondary bile acids or colonization of bile acid–metabolizing bacteria, reversed both NKT cell accumulation and inhibition of liver tumor growth in mice with altered gut commensal bacteria. In nontumor liver tissue from human patients with primary liver cancer, primary bile acid chenodeoxycholic acid (CDCA) levels correlated with *CXCL16* expression, whereas an inverse correlation was observed with secondary bile acid glycolithocholate (GLCA), suggesting that the finding may apply to humans.

ON OUR WEBSITE

Read the full article at http://dx.doi.org/10.1126/science.aan5931

CONCLUSION: We describe a mechanism by which the gut microbiome uses bile acids as messengers to control a chemokine-dependent accumulation of hepatic NKT cells and antitumor immunity in the liver, against both primary and metastatic liver tumors. These findings not only have possible implications for future cancer therapeutic studies but also provide a link between the gut microbiome, its metabolites, and immune responses in the liver. ■



Gut microbiome modulates liver cancer through bile acid–regulated NKT cells. Gut microbiome uses bile acids as a messenger to regulate chemokine CXCL16 level on liver sinusoidal endothelial cells (LSEC) and thus controls the accumulation of CXCR6$^+$ hepatic NKT cells. The accumulated NKT cells have an activated phenotype and inhibit liver tumor growth.

The list of author affiliations is available in the full article online.
*Corresponding author. Email: tim.greten@nih.gov
Cite this article as C. Ma *et al.*, *Science* **360**, eaan5931 (2018). DOI: 10.1126/science.aan5931

Ma *et al.*, *Science* **360**, 876 (2018) 25 May 2018 **1 of 1**

RESEARCH ARTICLE

MICROBIOME

Gut microbiome–mediated bile acid metabolism regulates liver cancer via NKT cells

Chi Ma,[1] Miaojun Han,[1] Bernd Heinrich,[1] Qiong Fu,[1] Qianfei Zhang,[1] Milan Sandhu,[1] David Agdashian,[1] Masaki Terabe,[2] Jay A. Berzofsky,[2] Valerie Fako,[3] Thomas Ritz,[4] Thomas Longerich,[4,5] Casey M. Theriot,[6] John A. McCulloch,[7] Soumen Roy,[7] Wuxing Yuan,[7,8] Vishal Thovarai,[7,8] Shurjo K. Sen,[7,8] Mathuros Ruchirawat,[9] Firouzeh Korangy,[1] Xin Wei Wang,[3,10] Giorgio Trinchieri,[7] Tim F. Greten[1,10]*

Primary liver tumors and liver metastasis currently represent the leading cause of cancer-related death. Commensal bacteria are important regulators of antitumor immunity, and although the liver is exposed to gut bacteria, their role in antitumor surveillance of liver tumors is poorly understood. We found that altering commensal gut bacteria in mice induced a liver-selective antitumor effect, with an increase of hepatic CXCR6$^+$ natural killer T (NKT) cells and heightened interferon-γ production upon antigen stimulation. In vivo functional studies showed that NKT cells mediated liver-selective tumor inhibition. NKT cell accumulation was regulated by CXCL16 expression of liver sinusoidal endothelial cells, which was controlled by gut microbiome-mediated primary-to-secondary bile acid conversion. Our study suggests a link between gut bacteria–controlled bile acid metabolism and liver antitumor immunosurveillance.

The gut microbiome has emerged as a critical factor regulating antitumor immunity controlling the efficacy of chemo- and immunotherapies (*1*–*6*). It is noteworthy that the liver is exposed to bacterial components and metabolites through the portal vein, and profound effects of the gut microbiome on hepatocellular carcinoma (HCC) have been described (*7*, *8*). Secondary hepatic malignancies (liver metastases) account for 95% of all hepatic cancers, and the liver is the most common site for organ metastasis in the body (*9*). To evaluate how the gut microbiome shapes antitumor immunity in the liver, we investigated the effects of gut commensal bacteria on both primary HCC and liver metastasis in mouse models.

Alterations in the gut microbiome suppress liver tumors in multiple mouse models

Spontaneous HCC was induced using MYC transgenic mice as described before (*10*). An antibiotic cocktail (ABX, consisting of vancomycin, neomycin, and primaxin) was added to drinking water to deplete gut commensal bacteria (*3*). The antibacterial efficacy of ABX was confirmed, and the cocktail was not toxic to the liver (fig. S1, A to C). Consistent with previous findings, fewer and smaller HCC were found in ABX-treated MYC mice (Fig. 1A and fig. S1, D and E). We extended our studies to a subcutaneous implantation model (EL4 thymoma) to study potential systemic effects. ABX treatment did not affect the growth of subcutaneous EL4 tumors in syngeneic C57BL/6 mice (fig. S1F). In contrast, fewer spontaneous liver metastases were seen in mice with large subcutaneous EL-4 tumors upon ABX treatment (Fig. 1B). To confirm this liver-selective antitumor effect, we used an intrasplenic tumor injection model (*11*) and found a robust decrease of B16 liver metastasis (Fig. 1C). Unlike in the liver, lung metastases were increased by ABX when the same B16 tumor cells were injected into the tail vein (fig. S1, G and H). Similar results were observed in BALB/c mice using A20 tumors (*12*) (Fig. 1D and fig. S1, I and J). These findings suggest that modulating gut commensal bacteria can specifically modify growth kinetics of intrahepatic tumors.

Hepatic NKT cell accumulation precedes tumor inhibition

To explore the mechanism behind tumor suppression, we studied immune cell subsets in EL4-tumor–bearing mice on ABX treatment. A prominent expansion of hepatic natural killer T (NKT) cells and CD8$^+$ T cells was found (Fig. 2A and fig. S2A), whereas no changes were found in other immune subsets (B cells, CD4$^+$ T, NK, γ/δ T cells and G-MDSC). The accumulation of hepatic NKT cells, but not CD8$^+$ T cells, was also observed in ABX-treated MYC mice bearing HCC (Fig. 2B and fig. S2B), and splenic NKT cells remained unchanged, suggesting a liver-specific effect (fig. S2C).

To understand how the gut microbiome mediates hepatic NKT cell accumulation, tumor-free mice were used. ABX-treated C57BL/6 or BALB/c mice had more absolute and relative hepatic NKT cells than untreated mice, which was not observed in the spleen and was independent of gender (Fig. 2, C to E, and fig. S2, D, E, G, and H).

The chemokine receptor CXCR6 mediates NKT cell survival and accumulation in the liver (*13*), and we found that all hepatic NKT cells expressed CXCR6 (Fig. 2F) as expected. NKT cells make up the majority of hepatic CXCR6$^+$ cells (~75%) (Fig. 2G), which did not change after ABX treatment (Fig. 2H). ABX treatment caused a ~2-fold increase of CXCR6$^+$ cells in the liver of C57BL/6 and BALB/c mice (Fig. 2I). CXCR6 is also expressed on T cells (*14*), and we observed that both hepatic CXCR6$^+$CD62L^{low}CD44^{hi} effector memory CD4$^+$ and CD8$^+$ T cells increased after ABX treatment (fig. S2, J to L). The observation that most NKT cells were effector memory CD44^{hi}CD62L^{low} cells (fig. S3A) prompted us to check activation status. Almost all hepatic NKT cells were the activated CD69^{hi} subtype in both tumor-bearing (fig. S3, B and C) and tumor-free mice (fig. S3D). CD25 and 4-1BB, two additional activation markers, did not change in hepatic NKT cells after ABX treatment (fig. S3, E and F). The observed higher CD69 expression but no change in 4-1BB activation level suggests that the NKT cells in ABX-treated mice are more active (Fig. 2K and fig. S3, J to K). NKT cells can exert diverse functions by rapidly releasing cytokines after activation. Thus, cytokine expression was measured from NKT cells after in vivo antigen-specific stimulation by injecting mice with α-galactosylceramide (aGalCer)–loaded tumor cells. Higher interferon-γ (IFN-γ) was detected in hepatic NKT cells from mice that received ABX treatment, whereas tumor necrosis factor (TNF) levels did not change (Fig. 2J and fig. S3, G to I). IFN-γ production by NKT cells has been shown to be instrumental for NKT-initiated tumor immunity. These data suggest that depleting gut commensal bacteria endows hepatic NKT cells with a stronger antitumor function. No change was observed for in vivo cytotoxicity of NKT cells after ABX treatment (fig. S3, L and M).

---

[1]Gastrointestinal Malignancy Section, Thoracic and Gastrointestinal Oncology Branch, Center for Cancer Research, National Cancer Institute, National Institutes of Health, Bethesda, MD 20892, USA. [2]Vaccine Branch, Center for Cancer Research, National Cancer Institute, National Institutes of Health, Bethesda, MD 20892, USA. [3]Laboratory of Human Carcinogenesis, Center for Cancer Research, National Cancer Institute, National Institutes of Health, Bethesda, MD, USA. [4]Institute of Pathology, University Hospital RWTH Aachen, Aachen 52074, Germany. [5]Institute of Pathology, University Hospital Heidelberg, Heidelberg 69120, Germany. [6]Department of Population Health and Pathobiology, College of Veterinary Medicine, North Carolina State University, Raleigh, NC 27607, USA. [7]Cancer and Inflammation Program, Center for Cancer Research, National Cancer Institute, National Institutes of Health, Bethesda, MD 20892, USA. [8]Leidos Biomedical Research, Inc, Microbiome Sequencing Core, National Cancer Institute, National Institutes of Health, Bethesda, MD 20892, USA. [9]Chulabhorn Research Institute, Bangkok, Thailand. [10]NCI CCR Liver Cancer Program, Bethesda, MD, USA.
*Corresponding author. Email: tim.greten@nih.gov

Finally, we studied NKT subpopulations based on the expression of transcriptional factors (*15*). The majority of hepatic NKT cells were NKT1, and the levels of NKT subsets did not change (fig. S3N). Interestingly, PLZF, which is required for NKT development (*16*), significantly decreased in the liver after ABX treatment (fig. S3O), which was not apparent in the spleen (fig. S3, P and Q).

Regulatory T cells (T$_{regs}$) are important modulators in tumor progression. The gut microbiome has been reported to affect T$_{regs}$ (*17*); however, we found no change in Foxp3$^+$CD4$^+$ subsets in either the liver or spleen of ABX-treated C57BL/6 mice (fig. S3, R and S). Taken together, our results showed that altering the gut microbiome caused accumulation of hepatic NKT and effector memory CD4$^+$ or CD8$^+$ T cells. The activated status of NKT cells leads to higher levels of IFN-γ upon encounter with antigen-loaded tumor cells, thereby fostering a tumor-rejecting environment.

Hepatic NKT cells mediate tumor inhibition

Antibody-mediated cell depletion was performed to investigate the specific function of individual immune cell populations controlling liver tumor immunity in ABX-treated mice. ABX-pretreated C57BL/6 mice were given intrasplenic injection of B16 tumor cells. T cell depletion was performed 1 day before tumor injection (Fig. 3A). Removing all the three major hepatic T cell subsets (CD4$^+$ T, CD8$^+$ T, and CD4$^+$ NKT cells) completely reversed the inhibition of liver metastasis caused by elimination of gut commensal bacteria (Fig. 3, B to D, and fig. S4, A to C), whereas depleting CD8$^+$ T cells alone only had minor effects.

The antitumor activity of NKT cells can cause direct killing of CD1d-expressing tumors (*18*–*20*), and we noted that all three (B16, EL4, and A20) of our tumor models tested expressed CD1d (fig. S4D). In addition, NKT cells can recognize tumor antigen from CD1d^{neg} tumors through cross-presentation by professional antigen-presenting cells (*21*). To investigate the role of NKT cells in CD1d-expressing tumors, we used CD1d-knockout mice (which completely lack NKT cells) (*22*) and CXCR6-knockout mice (which have a selective NKT deficiency in the liver) (*13*). We confirmed loss of hepatic NKT cells in these mice (fig. S4E) and then induced liver tumors by intravenous injection of EL4 tumor cells (*23*). Depleting gut microbiome reduced EL4 liver tumor burden in wild-type mice, but no reduction in liver tumor size was found in either CD1d-knockout or CXCR6-knockout mice after ABX treatment (Fig. 3, E and F). Intrasplenic injection of B16 tumor cells was repeated in CD1d-knockout or CXCR6-knockout mice, and similar results were observed (fig. S4, F to H). These findings suggest that hepatic NKT cells are necessary for effects on tumor growth in the liver induced by changes in the gut microbiome.

A bile acid/CXCL16/CXCR6 axis controls hepatic NKT accumulation

Virtually all hepatic NKT cells express CXCR6 (Fig. 2F), and ABX treatment increased the accumulation of CXCR6$^+$ cells into the liver (Fig. 2I), whereas CXCR6 mean fluorescence intensity did not change on NKT cells (fig. S5A). Therefore, CXCL16, the only ligand for CXCR6 (*24*), was further studied. As expected, higher *Cxcl16* mRNA levels were found in the liver of ABX-treated mice (Fig. f4A). This increase of *Cxcl16* mRNA was not detected in the lung (fig. S5B). Liver sinusoidal endothelial cells (LSECs) have been reported to be the major source of CXCL16 production in the liver (*13*). To identify the source of CXCL16, we isolated LSECs from ABX-treated mice. Figure 4B shows that there was an almost 2-fold increase of *Cxcl16* mRNA in LSECs from ABX-treated mice. Consistent with the previous report, LSECs have a much higher basal level of *Cxcl16* mRNA, and ABX treatment did not affect *Cxcl16* mRNA expression in non-LSECs. The increase of CXCL16 protein in LSECs was confirmed by immunohistochemistry staining (Fig. 4C). In addition, forced CXCL16 expression in the liver increased hepatic NKT levels (fig. S5C). Together,

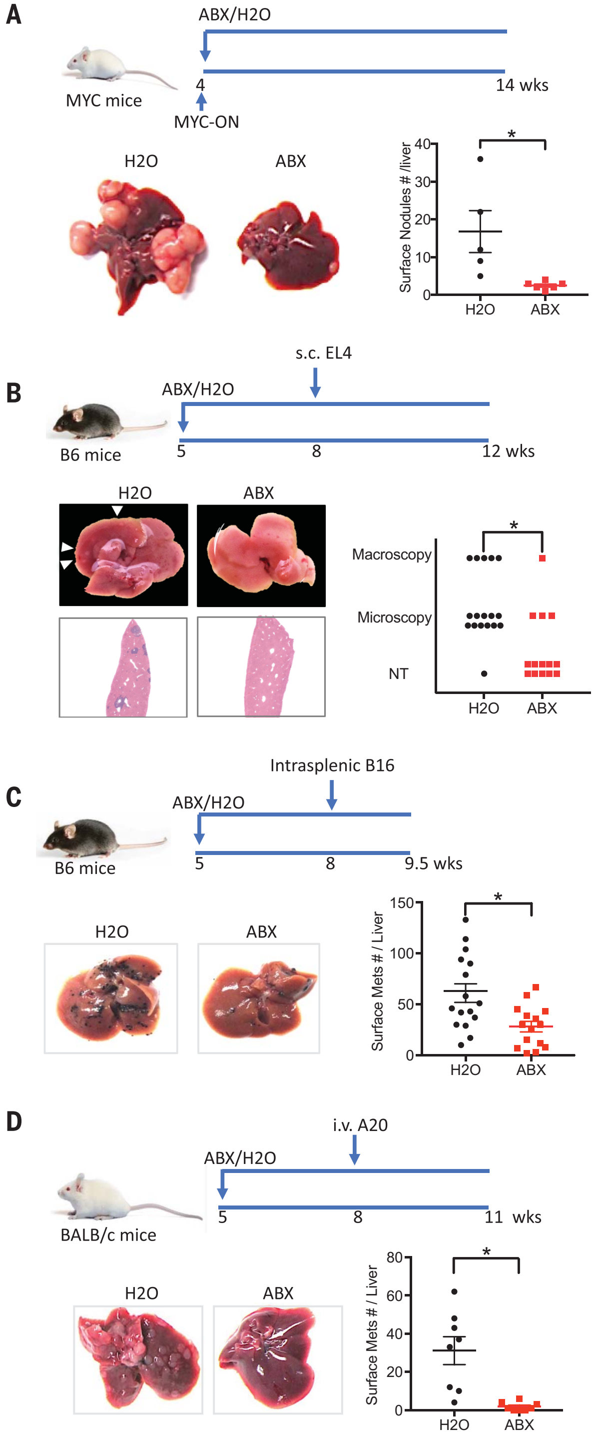


Fig. 1. Depleting gut commensal bacteria inhibits liver tumor growth in multiple mouse models. (**A**) MYC transgene was turned on at the age of 4 weeks. MYC-ON mice were fed with ABX or regular water. Ten weeks later, mice were killed and liver surface tumor nodules were counted. Representative liver images are shown. Data represent mean ± SEM of two pooled experiments. n = 5 for H_2O, 6 for ABX. $P < 0.05$, Student's t test. (**B**) C57BL/6 mice were treated with ABX or H_2O for 3 weeks before receiving subcutaneous EL4 tumor cell injection. Four weeks later, liver metastases were determined. Representative images of five pooled experiments are shown. n = 17 for H_2O, 12 for ABX. $P < 0.05$, Chi-square test. (**C**) ABX- or H_2O-pretreated C57BL/6 mice were given intrasplenic B16 tumor cell injection. One and a half weeks later, liver metastases were measured. Representative images are shown. Data represent mean ± SEM of five pooled experiments. n = 18 for H_2O, 15 for ABX. $P < 0.05$, Student's t test. (**D**) BALB/c mice were treated with ABX or H_2O for 3 weeks. Then mice received A20 tumor cell tail vein injection. Three weeks later, liver metastases were counted. Representative images are shown. Data represent mean ± SEM of two pooled experiments. $n = 8$, $P < 0.05$, Student's t test.

these results suggest that ABX treatment causes LSECs to produce more CXCL16 and recruits NKT cells to the liver.

CXC16 has both cell-surface and secreted forms; the cell-surface form has been identified as a scavenger receptor for phosphatidylserine and lipoprotein and is involved in lipid metabolism (25). Gut commensal bacteria are well known to mediate bile acid metabolism in the intestine, and the gut microbiome reportedly regulates bile acid composition (26). Therefore, we searched for a possible link between bile acids and CXCL16. First, we used cholestyramine, a bile acid sequestrant, to block the enterohepatic circulation, thus reducing bile acid levels in the liver (27).

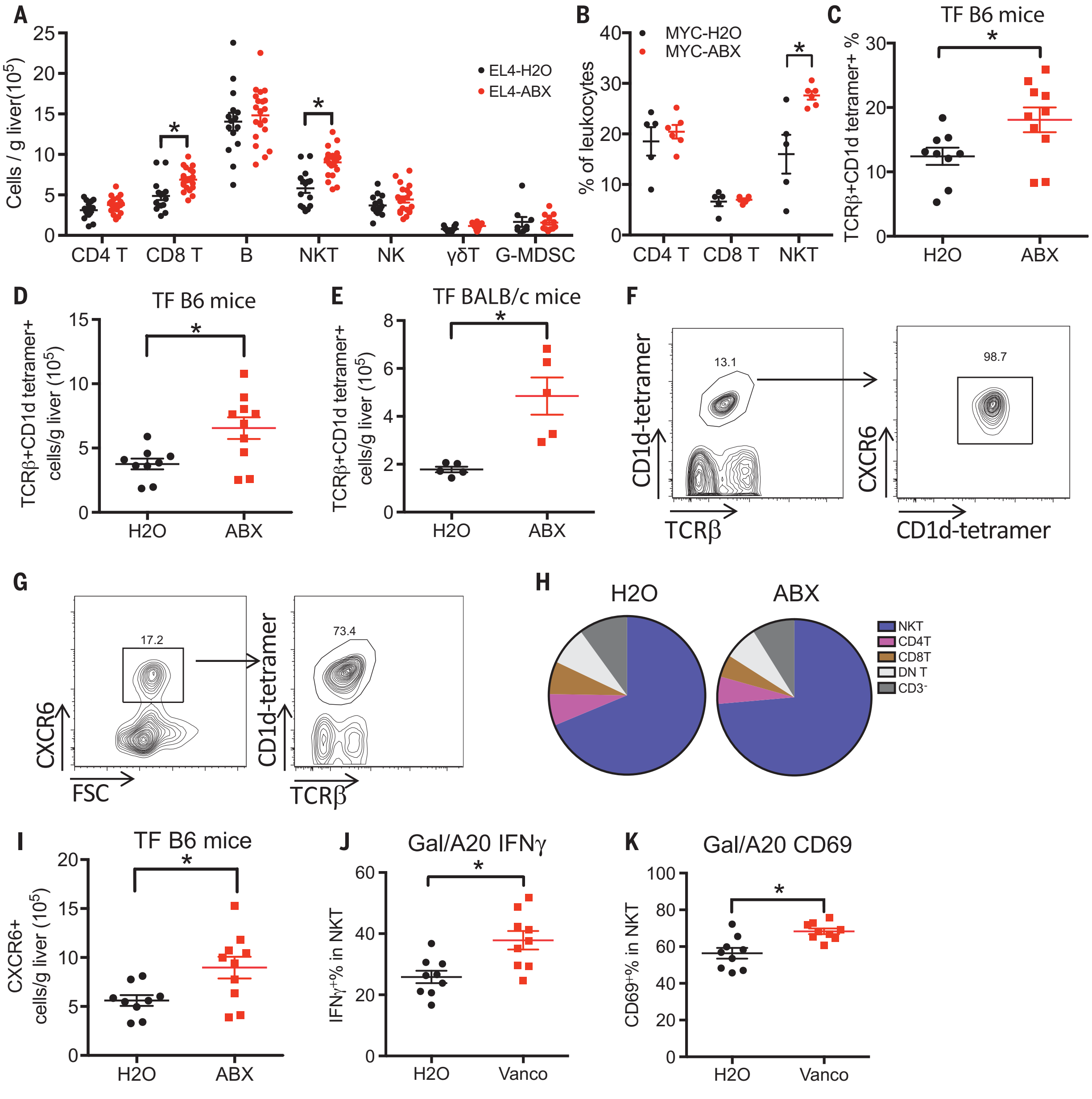


Fig. 2. Hepatic NKT cells increase after depleting gut commensal bacteria. (**A**) ABX- or H_2O-pretreated C57BL/6 mice were given subcutaneous EL4 tumor injection. Two and a half weeks later, liver infiltrating immune cells were measured. Data represent mean ± SEM of three pooled experiments. $n = 15$ for EL4-H_2O, 20 for EL4-ABX. $P < 0.05$, two-way analysis of variance (ANOVA). (**B**) Hepatic NKT, CD4, and CD8 T cell levels of MYC mice described in Fig. 1A. Data represent mean ± SEM of two pooled experiments. $P < 0.05$, two-way ANOVA. (**C** and **D**) Hepatic NKT cells in tumor-free C57BL/6 mice were fed with ABX or H_2O. Data represent mean ± SEM of two pooled experiments. $n = 9$ for H_2O, 10 for ABX. $P < 0.05$, Student's *t* test. (**E**) Hepatic NKT cells in tumor-free BALB/c mice fed with ABX or H_2O. Data represent mean ± SEM of two pooled experiments. $n = 5$, $P < 0.05$, Student's *t* test. (**F**) Representative CXCR6 staining in hepatic NKT cells from three independent experiments. (**G**) Representative NKT cell staining in CXCR6$^+$ liver infiltrating mononuclear cells from three independent experiments. (**H**) Composition of CXCR6$^+$ liver infiltrating mononuclear cells in tumor-free C57BL/6 mice fed with ABX or H_2O. DN T: double negative T cells. Data represent pooled results from three experiments. (**I**) Levels of CXCR6$^+$ liver-infiltrating cells in tumor-free C57BL/6 mice fed with ABX or H_2O. Data represent mean ± SEM of two pooled experiments. $n = 9$ for H_2O, 10 for ABX. $P < 0.05$, Student's *t* test. (**J** and **K**) IFN-γ and CD69 levels of hepatic NKT cells after in vivo stimulation by injecting aGalCer-loaded A20 tumor cells (Gal/A20) into vancomycin (Vanco) or H_2O-fed BABL/c mice. Data represent mean ± SEM of two pooled experiments. $n = 9$, $P < 0.05$, Student's *t* test.

Cholestyramine treatment increased hepatic NKT and CXCR6$^+$ cells, but not CD4$^+$ T or CD8$^+$ T cells (Fig. 4D and fig. S5D). In parallel, *Cxcl16* mRNA was up-regulated in the liver (fig. S5E). This result suggests that bile acids are involved in the accumulation of NKT cells in the liver through CXCL16 regulation. To identify the bile acids involved in NKT cell regulation, the liver bile acid profile was determined. As reported (*26*), control mouse liver contains the highest amount of primary bile acid taurocholic acid (TCA), followed by primary bile acid tauro-β-muricholic acid (T-β-MCA), and secondary bile acid tauro-ω-muricholic acid (T-ω-MCA) (Fig. 4E). ABX treatment did not affect liver TCA but significantly increased primary bile acids T-β-MCA and β-MCA. Gut commensal bacteria convert primary bile acids into secondary bile acids (*28*). As expected, secondary bile acids T-ω-MCA, taurodeoxycholic acid (TDCA), ω-MCA, taurolithocholic acid (TLCA), and tauroursodeoxycholic acid (TUDCA) were reduced in ABX-treated mice (Fig. 4E).

To connect these findings, we studied the effect of bile acids on CXCL16 expression in primary murine LSECs in vitro. Secondary bile acid ω-MCA decreased *cxcl16* mRNA expression, whereas the primary bile acid T-β-MCA induced *Cxcl16* mRNA (Fig. 4F). Because cholestyramine treatment reduces both primary and secondary bile acids and caused NKT accumulation, we tested the possibility that primary and secondary bile acids have opposing effects on CXCL16 regulation. Indeed, the secondary bile acid compromised the primary acid-induced *Cxcl16* mRNA up-regulation, and even impaired *cxcl16* mRNA expression below the untreated baseline levels. (Fig. 4G). When mice were fed ω-MCA, a secondary bile acid, hepatic NKT accumulation was reversed, whereas feeding chenodeoxycholic acid (CDCA), a primary bile acid, enhanced NKT cell accumulation (Fig. 4H). These data demonstrate that primary and secondary bile acids influence hepatic NKT cell recruitment and have opposing effects in vivo.

Next, bile acids were fed to mice to investigate the effect on liver tumor growth. As expected, secondary bile acids LCA or ω-MCA reversed the inhibition of liver tumor growth caused by antibiotic treatment, and primary bile acid CDCA further enhanced tumor inhibition (Fig. 4I and fig. S5, F to H). Enforced CXCL16 expression in the liver (Fig. 4I and fig. S5F) or depletion of CD4$^+$ NKT cells (fig. S5G) eliminated the effect of secondary bile acids on liver tumor growth (Fig. 4I and fig. S5, F and G), suggesting that this effect was NKT cell mediated. Similarly, secondary bile acids failed to affect liver tumor development in liver-NKT deficient (CXCR6$^{-/-}$) mice (Fig. 4J) but still preserved the function to block the accumulation of transferred wild-type NKT cells caused by antibiotic treatment (fig. S5I). Therefore, depleting commensal bacteria with antibiotics leads to a model where primary bile acids that induce CXCL16 are preserved, whereas secondary bile acids that inhibit CXCL16 are reduced. This leads to up-regulation of CXCL16 in LSECs and accumulation of NKT cells in the liver.

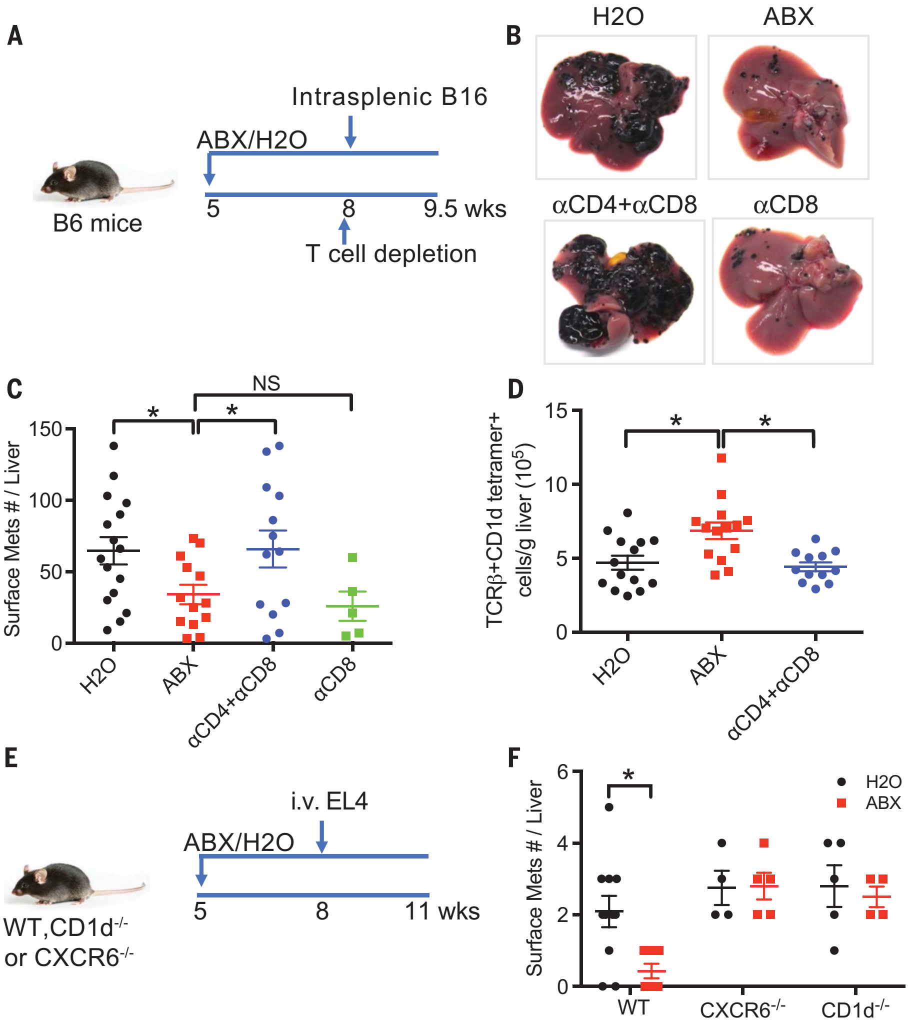


Fig. 3. Hepatic NKT cells mediate the inhibition of liver metastasis in mice. (**A** to **D**) B16 tumor cells were intrasplenically injected into mice pretreated with ABX or H_2O. One day before tumor injection, mice were given intraperitoneal injection of a combination of antibodies to CD4 (anti-CD4) (500 μg per mouse) and anti-CD8 (200 μg per mouse) or anti-CD8 alone (200 μg per mouse). Liver surface metastatic nodules were counted. Representative images are shown. Data represent mean ± SEM of two pooled experiments. n = 16 for H_2O, 14 for ABX, 13 for anti-CD4 + anti-CD8, 5 for anti-CD8. $P <$ 0.05, one-way ANOVA. (**E** and **F**) Loss of hepatic NKT abrogated the inhibition of liver metastasis caused by ABX. CXCR6$^{-/-}$, CD1d$^{-/-}$, or wild-type mice pretreated with ABX or H_2O were given EL4 tumor cell tail vein injections. Liver surface metastatic nodules were counted. Data represent mean ± SEM of two pooled experiments. n = 10 for Wt-H_2O, 7 for Wt-ABX, 4 for CXCR6$^{-/-}$-H_2O, 5 for CXCR6$^{-/-}$-ABX, 5 for CD1d$^{-/-}$-H_2O, 4 for CD1d$^{-/-}$-ABX. $P < 0.05$, two-way ANOVA.

Clostridium species regulates liver NKT accumulation in mice

Although ABX treatment efficiently reduced gut commensal bacteria load, it did not result in complete elimination of gut bacteria (fig. S1A). To rule out the possibility that the remaining bacteria mediate NKT accumulation, we repeated the experiment in germ-free mice. Whereas more hepatic NKTs were found in germ-free mice when compared with the matched SPF control mice (Fig. 5A and fig. S5, J and K) (*29*), no change was seen in TLR4-knockout mice (fig. S5L). Similarly, *Cxcl16* mRNA levels were higher in the liver of germ-free mice (Fig. 5B).

Next, we tried to identify the commensal bacteria responsible for the observed liver NKT accumulation. Because the ABX antibiotic cocktail used here contains three antibiotics with a different activity spectrum, individual antibiotic treatment was performed to narrow down the targeting bacteria. Vancomycin alone was sufficient to increase hepatic NKT cells, whereas neomycin had a marginal effect (Fig. 5C and fig. S5M). A clear increase of liver NKT cells was seen in mice fed with cefoperazone, but no significant changes in hepatic CD4$^+$ or CD8 T$^+$ cells were observed (fig. S5, N and O). Interestingly,

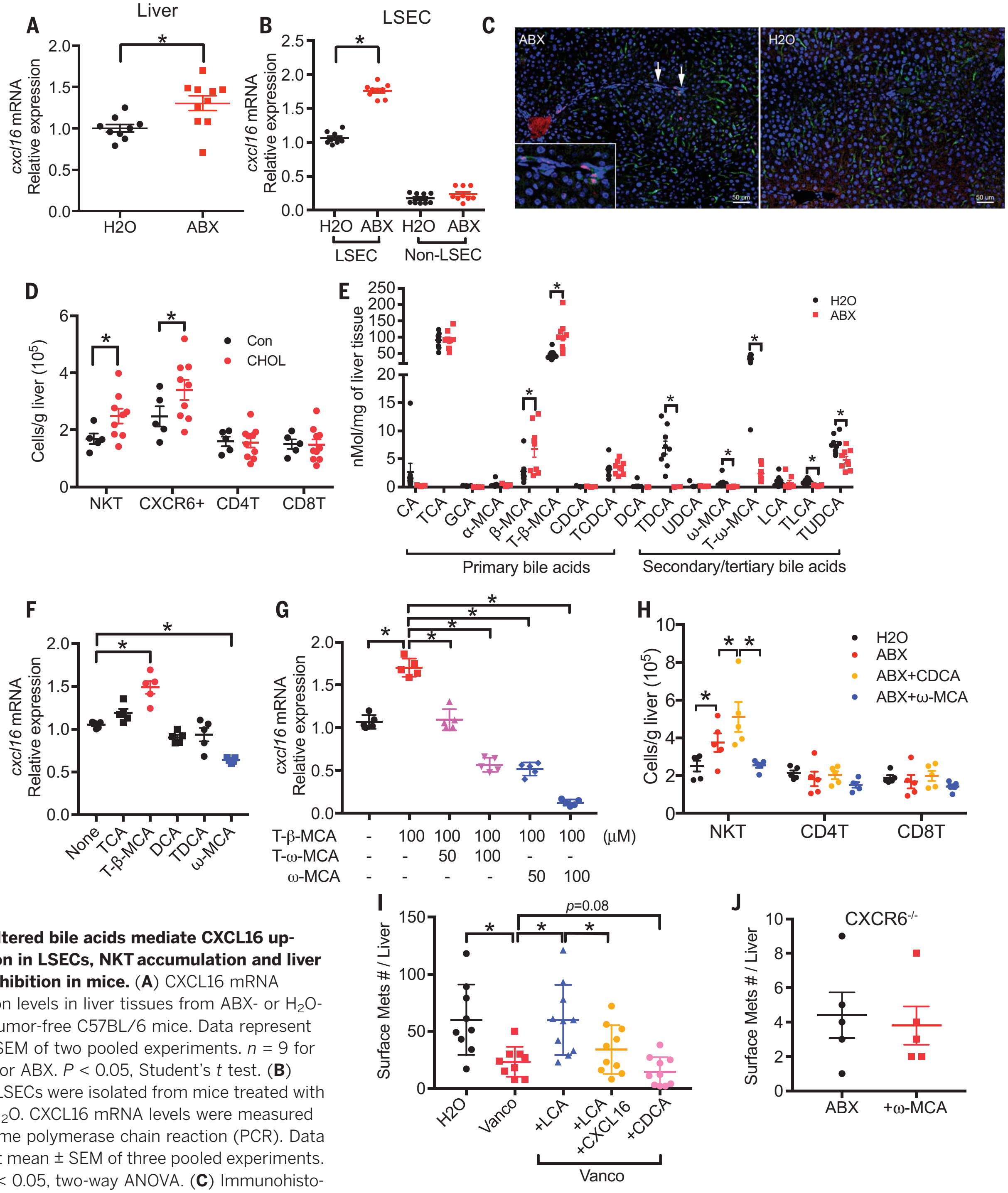


Fig. 4. Altered bile acids mediate CXCL16 upregulation in LSECs, NKT accumulation and liver tumor inhibition in mice. (**A**) CXCL16 mRNA expression levels in liver tissues from ABX- or H$_2$O-treated tumor-free C57BL/6 mice. Data represent mean ± SEM of two pooled experiments. $n = 9$ for H$_2$O, 10 for ABX. $P < 0.05$, Student's *t* test. (**B**) Primary LSECs were isolated from mice treated with ABX or H$_2$O. CXCL16 mRNA levels were measured by real-time polymerase chain reaction (PCR). Data represent mean ± SEM of three pooled experiments. $n = 9$, $P < 0.05$, two-way ANOVA. (**C**) Immunohistochemistry staining of CXCL16 and LYVE-1 in liver sections from ABX- or H$_2$O-treated mice. CXCL16$^+$ (red)/LYVE$^+$ (green) LSECs are highlighted by arrows and are shown in more detail in the insert. Scale bar, 50 μm. (**D**) Mice were fed with a 2% cholestyramine (CHOL) or control diet (Con). Hepatic NKT, CXCR6$^+$, CD4 T, and CD8 T cells were measured. Data represent mean ± SEM of two pooled experiments. $n = 5$ for control, 9 for cholestyramine diet. $P < 0.05$, two-way ANOVA. (**E**) Liver bile acids profile of ABX- or H$_2$O-treated mice. Data represent mean ± SEM of two pooled experiments. $n = 9$, $P < 0.05$, Student's *t* test. (**F** and **G**) Isolated LSECs were treated with different bile acids (F) or a combination of T-β-MCA with T-ω-MCA or ω-MCA (G). Data represent mean ± SEM of two pooled experiments. $n = 5$, $P < 0.05$, one-way ANOVA. (**H**) ABX-treated mice were given three times oral gavage of CDCA or ω-MCA three times at the dose of 6 mg per 15 g body weight. Hepatic NKT cell levels from different treatments were measured. Data represent mean ± SEM of two pooled experiments. $n = 5$, $P < 0.05$, two-way ANOVA. (**I**) A20 liver metastasis in vancomycin (Vanco) or H$_2$O-treated mice receiving LCA or CDCA. As indicated, one group of LCA-treated mice also received hydrodynamic injection to force CXCL16 expression in the liver. Data represent mean ± SEM of two pooled experiments. $n = 15$ for Vanco + LCA, others $n = 10$. $P < 0.05$, one-way ANOVA. (**J**) EL4 liver metastasis in ABX-treated CXCR6$^{-/-}$ mice with or without ω-MCA feeding. Data represent mean ± SEM of two pooled experiments. $n = 5$.

both vancomycin and cefoperazone target gram-positive bacteria and have been reported to increase primary bile acids and deplete secondary bile acids in the gut (*30*), which is consistent with our finding that bile acids also change in the liver and regulate NKT cell accumulation.

The 7α-dehydroxylation reaction is the most quantitatively important process performed by the gut bacteria in the production of secondary bile acids (*28*) and is restricted to gram-positive bacteria of the *Clostridium* cluster XIV (*31*). We found that vancomycin, which increased liver NKT cells, depleted *Clostridium*, whereas neomycin, which had little effect on liver NKT, had little effect on *Clostridium* (fig. S5P). To explore a role for *Clostridium* species on liver NKT cell accumulation, we used *C. scindens*, which is commonly found in both mice and humans (*32*) and has a conserved *bai* (bile acid inducible) gene operon for the 7α-dehydroxylation reaction (*28*, *33*). Mice were fed with vancomycin for 1 week to induce hepatic NKT cells, then vancomycin was stopped, and the mice were given *C. scindens* or vehicle (fig. S6A). The successful colonization with *C. scindens* was confirmed (fig. S6B). One day after oral gavage, fecal bacterial were analyzed by 16*S* ribosomal (rRNA) sequencing (Fig. 5D and fig. S6, C and D). Continuous vancomycin treatment caused a reduction of *Clostridiales* and *Bacteroidales* but an expansion of *Verrucomicrobiales* compared with the water control group. An early recovery of *Bacteroidales* was observed after vancomycin cessation. Gavage of *C. scindens* increased the *Clostridiales* population. A time course study showed that hepatic NKT levels started to drop between day 2 and day 4 after vancomycin withdrawal (Fig. 5E), suggesting recovery of gut commensal bacteria. *Clostridium* cluster XIV recovered in parallel

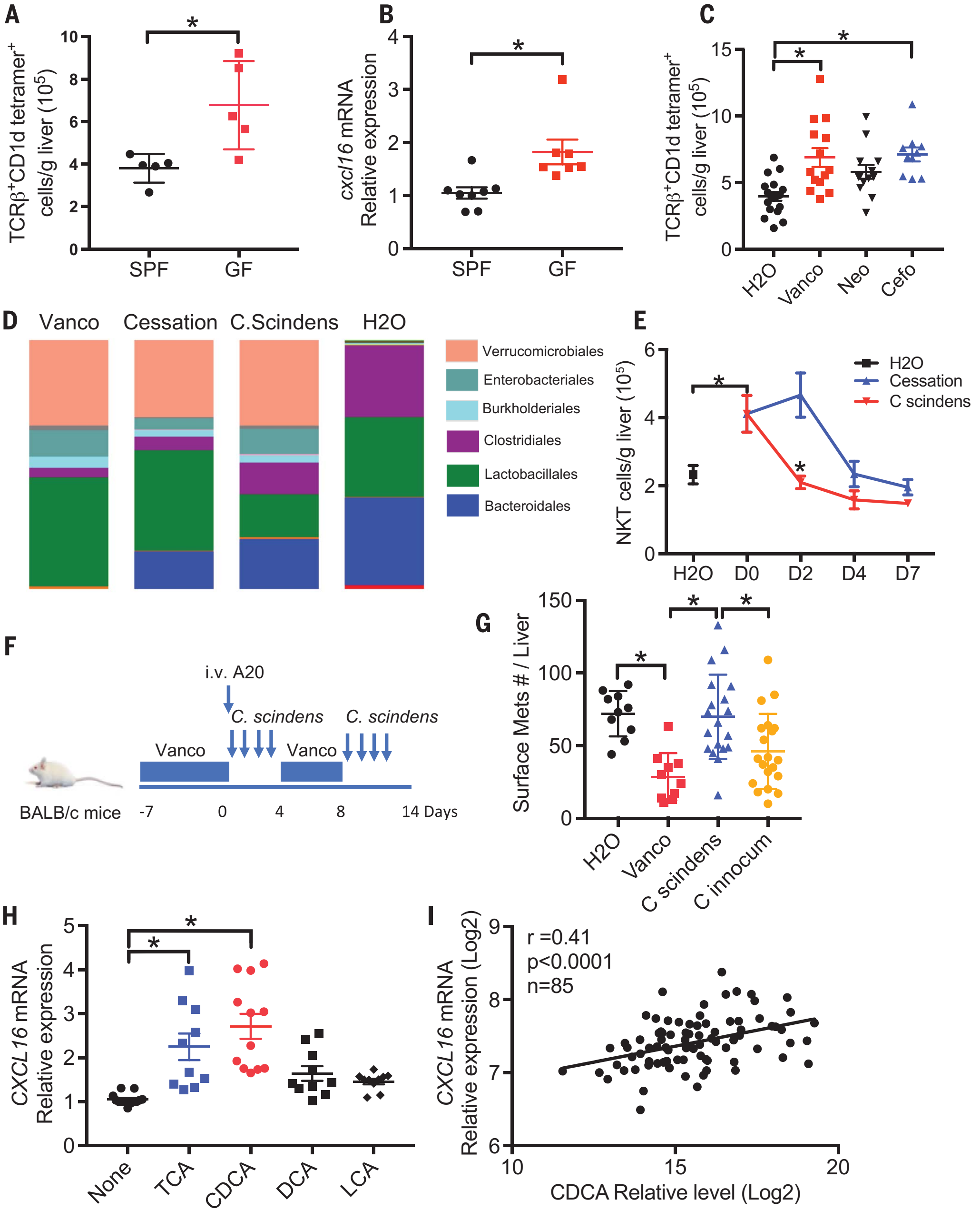


Fig. 5. *Clostridium* species influence bile acid–controlled NKT cell accumulation. (**A**) Hepatic NKT cell levels from germ-free mice or matched SPF mice were measured. Data represent mean ± SEM of two pooled experiments. $n = 5$, $P < 0.05$, Student's *t*-test. (**B**) *cxcl16* mRNA expression in liver tissue from germ-free or SPF mice. Data represent mean ± SEM of two pooled experiments. $n = 8$ for SPF, 7 for GF. $P < 0.05$, Student's *t*-test. (**C**) Naïve C57BL/6 mice were fed with vancomycin (Vanco), neomycin (Neo) or cefoperazone (Cefo). Hepatic NKT levels were determined. Data represent mean ± SEM of three pooled experiments. $n = 18$ for H_2O, 14 for vancomycin, 14 for neomycin, 10 for cefoperazone. $P < 0.05$, one-way ANOVA. (**D** and **E**) Mice were treated with vancomycin for 1 week and then gavaged with *C. scindens* or vehicle (cessation). Twenty-four hours after *C. scindens* gavage, 16*S* rRNA sequencing analysis of stool samples was performed. The relative abundance of OTUs in the fecal bacterial are shown (D). Time-course study of hepatic NKT levels was performed (E). Data represent mean ± SEM of two pooled experiments. $n = 10$ for H_2O, D0, *C. scindens* D4, and Cessation D4; 5 for *C. scindens* D2, Cessation D2, *C. scindens* D7, and Cessation D7. $P < 0.05$, two-way ANOVA. (**F** and **G**) A20 liver tumors were induced in mice treated with vancomycin or H_2O. Mice were colonized with *C. scindens* or control *C. innocuum* as illustrated in (F). Cumulative liver tumor counts are shown in (G). Data represent mean ± SEM of two pooled experiments. $n = 10$ for H_2O and Vanco, $n = 20$ for *C. scindens* and *C. innocuum*. $P < 0.05$, one-way ANOVA. (**H**) SK-HEP1 cells were treated with different bile acids. *CXCL*16 mRNA levels were measured by real-time PCR. Data represent mean ± SEM of three pooled experiments. $n > 10$, $P < 0.05$, one-way ANOVA. (**I**) Correlation between primary bile acid CDCA and *CXCL16* mRNA expression in nontumor liver tissues from hepatocellular carcinoma or cholangiocarcinoma patients of the TIGER cohort. Pearson correlation coefficient test was performed.

(fig. S6E). Colonization of *C. scindens* induced a rapid reduction of hepatic NKT cells (Fig. 5E) but did not affect other immune cells (fig. S6, F and G). As expected, colonization of mice with *C. scindens*, but not cessation of ABX treatment, resulted in a reduction of primary bile acids and recovery of secondary bile acids on day 2 (fig. S6H). Thus, our results suggest that bile acid–converting *Clostridium* species such as *C. scindens* are involved in the regulation of hepatic NKT cell accumulation. We next tested the effect of *C. scindens* on liver tumor growth. As expected, more liver tumors were found in *C. scindens*–colonized mice than mice kept on vancomycin or *C. innocum*–colonized mice (Fig. 5, F and G), suggesting that gut bacteria capable of metabolizing bile acids can regulate growth of liver tumors.

Bile acids control liver CXCL16 expression in humans

We extended our study to human samples to determine the effect of bile acids on *CXCL16* mRNA expression of human liver sinusoidal endothelial cells. Similar to the mouse study, primary bile acids CDCA and TCA induced *CXCL16* mRNA expression in SK-HEP1 cells (Fig. 5H). Next, the correlation between bile acids and *CXCL16* expression was tested in nontumor liver tissues from HCC or cholangiocarcinoma patients of the Thailand Initiative for Genomics and Expression Research in Liver Cancer (TIGER-LC) cohort (*34*). Primary bile acid CDCA levels correlated with *CXCL16* expression (Fig. 5I), whereas inverse correlation was seen for the secondary bile acid glycolithocholate (GLCA) (fig. S6I). The primary/secondary ratio was associated with *CXCL16* increase (fig. S6J), indicating that the opposing effect of bile acids on CXCL16 expression also exists in humans.

Mucosal-associated invariant (MAIT) cells are enriched in the liver and comprise 20 to 50% of hepatic lymphocytes in humans (*35*). Interestingly, CXCR6 is expressed on MAIT cells (*35*), which raises the question whether liver MAIT cells can also be controlled by gut bacteria through CXCL16 regulation. MAIT cells recognize bacterial derivatives and are involved in inflammatory liver diseases such as nonalcoholic steatohepatitis (NASH). Besides antitumor function, NKT cells have been reported as important regulators of autoimmune responses (*36*). Our study suggests that gut commensal bacteria and bile acids could be potential targets for controlling liver autoimmune diseases. In summary, we have describe a mechanism by which the gut microbiome uses bile acids as messengers to control the accumulation of hepatic NKT cells and thereby antitumor immunity in the liver of mice. These findings not only have possible implications for future cancer therapeutic studies in humans but also provide a link between the gut microbiome, its metabolites, and immune responses in the liver.

Materials and methods

Murine studies

SPF C57BL/6 and BALB/c mice were purchased from Charles River. CXCR6-knockout mice were purchased from Jackson laboratory. CD1d-knockout mice, LAP-tT and TRE-MYC mice have been previously described (*10*). Germ-free mice were provided by R. Goldszmid (Cancer and inflammation program, NIH). Newly purchased four-week old C57BL/6 or BALB/c mice were randomized into 5 mice/cage and housed for one week to normalize gut microbiome. Then mice were assigned into H_2O or ABX treatment groups. Mice in the ABX group received three-antibiotic cocktail in the drinking water containing vancomycin (Hospira, 0.5 g/L), neomycin (VETone, 0.5 g/L) and primaxin (Merck & CO, 0.5 g/L) as previously reported (*3*). In some experiments mice were given single antibiotic water, and cefoperazone (MP Biomedicals) was given at the concentration of 0.5 g/L. Fresh antibiotic water was replaced every other day. After 3 weeks of ABX pretreatment, mice were challenged with different tumor cell lines. B16-F1 and A20 cells were purchased from ATCC. EL4 cells were used as described (*37*). 1 × 10^6 EL4 tumor were given by subcutaneous or tail vein injection, 3 × 10^5 B16-F1 tumor cells were given by intrasplenic injection as described before (*11*), and 1 × 10^6 A20 tumor cell were given by tail vein injection. In some experiments mice were fed with a 2% cholestyramine diet made by Research Diets Inc (New Brunswick, NJ). Mice were treated with 500 μg anti-CD4 (clone GK1.5, BioxCell) or 200 μg anti-CD8 (clone 2.43, BioxCell) 24 hours before receiving tumor injection for depletion studies. For in vivo NKT cell stimulation, 1 × 10^6 aGalCer-loaded A20 tumor cells in the combination of brefeldi A (500 ug/mouse) were given by tail vein injection, and mice were sacrificed 3 hours after injection. aGalCer-loading was performed by incubate A20 cells with 1 μg/ml aGalCer overnight followed by 3 times of washing. At the experimental end points, mice were sacrificed for organ harvest. All experiments were conducted according local institution guidelines and approved by the Animal Care and Use Committee of the National Institutes of Health, Bethesda, USA.

Flow cytometry

Cells were surface-labeled with the indicated antibodies for 15 min at 4°C. Intracellular staining using a Foxp3/transcription factor staining buffer set (eBioscience) was used according to the manufacturer's instructions. Flow cytometry was performed on BD LSRFortessa platform and results were analyzed using FlowJo software version 9.3.1.2 (TreeStar). Dead cells were excluded by using live/dead fixable near-IR dead cell staining kit (ThermoFisher scientific). The following antibodies were used for flow cytometry analysis: anti-TCRb-BV510 (clone H57-587, Biolegend), PBS57/CD1d- tetramer-APC (NIH core facility), anti-CXCR6-FITC (clone SA051D1, Biolegend), anti CD3-PE (clone 17A2, Biolegend), anti-CD4-PE (clone RM4-5, Biolegend), anti-CD4- Alexa Fluor 700 (clone GK1.5, Biolegend), anti-CD8-BV210 (clone 53-6.7 Biolegend), anti-CD19-PerCP/Cy5.5 (clone eBio1D3, eBioscience), anti-CD49b (clone DX5, eBioscience), anti-TCRg/d-PE, (clone GL3, BD pharmigen), anti-CD11b-BV421 (clone M1/70, Biolegend), anti-Ly6G-Alexa Fluor 700 (clone 1A8, Biolegend), anti-Ly6C-AP (clone HK1.4, Biolegend), anti-CD44-PE/Cy7 (clone IM7, eBioscience), anti-CD62L- PerCP/Cy5.5 (MEL-14, Biolegend), anti-CD69-Pacific blue (clone H1.2F3, Biolegend), anti-CD25-FITC (clone 7D4, BD pharmigen), anti-4-1BB-PE (clone 17B5, Biolegend), anti-Foxp3-Alexa Fluor 488 (clone 22F6, Biolegend), anti-Tbet-Pacific Blue (clone 4B10, Biolegend), anti-RORg-PE (clone B2D, eBioscience), anti-PLZF-PerCP/Cy5.5 (clone 9E12, Biolegend), and anti-CD1d-PE (clone 1B1, eBioscience), anti-IFNg-PE (clone XMG1.2, BD Biosciences), anti-TNFa-PerCP/Cy5.5 (clone MP6-XT22, Biolegend). The following markers were used for identifying different immune cell subsets: $TCRb^+CD1d$-Teteramer$^+$ for NKT cells, $CD3^+CD4^{hi}$ for hepatic $CD4^+$ T cells, $CD3^+CD8^+$ for $CD8^+$ T cells, $CD3^-CD19^+$ for B cells, $CD3^-CD49b^+$ for NK cells, $CD3^+TCRg/d^+$ for g/d T cells, $CD11b^+Ly6G^+Ly6C^{lo}$ for G-MDSC. Absolute numbers were calculated by multiplying frequencies obtained from flow cytometry by the total live mononuclear cell count, and then divided by liver weight.

In vivo cytotoxcicity assay

Splenocytes isolated from naïve C57BL/6 mice were loaded with aGalCer (1 μg/ml) then labeled with high dose of CFSE as target cells. Unloaded cells were labeled with low dose of CFSE as control cells. Then $CFSE^{hi}$ target cell and $CFSE^{lo}$ controls cells were mixed at about 1:1 ratio. 10^7 mixed cells were injected intravenously into ABX or H_2O-treated C57BL/6 mice. Sixteen hours later, mice were killed and cytotoxicity was analyzed by flow cytometry. r = (%CFSElo / %CFSEhi); ro is the ratio of mixed cells without injection; % cytotoxicity = [1- (r_0/r)]x100.

Immunohistochemistry and quantification

Immunohistochemistry was performed on 3 μm sections obtained from formalin- fixed paraffin-embedded liver tissues of H_2O-treated (n = 5) or ABX-treated (n = 5) mice using the Opal 5-color IHC Kit (PerkinElmer, Waltham, MA, USA) according to the manufacturer´s instructions. The following primary antibodies were used: anti-CXCL16 (Bioss Antibodies, Woburn, MA, USA; bs-1441R, rabbit polyclonal, 1/4.000, Opal 620), anti-LYVE1 (Abcam, Cambridge, UK; rabbit polyclonal, 1/15.000, Opal 520). Slides were evaluated using the Vectra® 3 automated, high-throughput quantitative pathology imaging system (PerkinElmer) and the inForm software (PerkinElmer) for segmentation and quantification of $CXCL16^+$/$LYVE1^+$ cells.

Hepatic bile acid profiling

Fresh mice liver tissue was snap frozen in liquid nitrogen and then kept at –80°C. Hepatic bile acid composition was measured at West Coast Metabolomics Center at UC Davis using the targeted metabolite analysis service.

Liver sinusoidal endothelial cell preparation and bile acids treatment

Primary mice liver sinusoidal endothelia cells were isolated as previously described (*38*). Briefly, mice were CO2 euthanized, and then the portal vein was cannulated and the liver was perfused with 0.05% collagenase in Ca^{2+} deprived medium. Liver cells were dissociated and parenchymal cells were killed by incubation in 0.04% collagenase in Gey's balanced salt solution at 37°C for 15 min. Then density gradient centrifugation was performed using Nycodenz solution at the final solution of 1.089 g/cm^3. LSECs were isolated using anti-LSEC microbeads (Miltenyi) according to the manufacturer's instructions. LSECs or the human SK-HEP1 cell line (ATCC, HTB-52) were treated with different bile acids for 24 hours. Gene expression was analyzed by real-time PCR. TCA, CDCA, DCA, and TDCA were purchased from Sigma. T-b-MCA, w-MCA, and T- w-MCA were purchased from Steraloids Inc.

In vivo bile acids feeding

Mice were kept on ABX cocktail and fresh ABX was replaced every other day. Mice were fed with bile acids by oral gavage 48, 24 and 16 hours prior to sacrifice. For A20 tumor bearing mice, w-MCA, LCA or CDCA were given 3 times/week. Bile acids were dissolved in corn oil. w-MCA and CDCA were given at the dose of 6 mg/15 g body weight, LCA was given at the dose of 1 mg/15 g body weight.

Adoptive transfer of NKT cells

Donor NKT-rich cells were isolated from livers of wild-type C57BL/6 mice by autoMACs sorting of NK1.1$^+$ cells. Half million of NKT-rich cells were transferred into CXCR6$^{-/-}$ mice by tail vein injection. Two days later mice were scarified and the liver accumulation of transferred NKT cells were measured by flow cytometry as TCRb$^+$CD1d-Teratmer$^+$CXCR6$^+$ population.

Gut colonization with *Clostridium scindens*

Mice were fed with vancomycin in drinking water (Hospira, 0.5 g/L) for one week. Fresh antibiotic water was replaced every other day. One week later vancomycin was stopped, and the mice were given oral gavage of 10^9 *C. scindens* or vehicle (anaerobic glycerol) every day for 5 days. *C. scindens* was purchased from ATCC (35704), and grown under anaerobic conditions. One day after gavage, the colonization of *C. scindens* were confirmed by real-time PCR using primers specific for C. scindens. For A20 tumor study, BALB/c mice were fed with vancomycin for one week. Then vancomycin was stopped, and A20 tumor (1 × 10^6 cells) were injected intravenously. Mice were given oral gavage of 10^9 *C. scindens* or *C. innocuum* (ATCC 14501) every day for 4 days. Then mice were given a second round of vancomycin treatment for 4 days, followed by additional four days of oral gavage of *C. scindens* or *C. innocuum*. Fourteen days after tumor injection, mice were sacrificed and liver tumor burden was measured.

16S rRNA sequencing and analysis

Mouse stool DNA extraction and 16*S* V4 region amplification were performed on the liquid handling robots (Eppendorf, epMotion5075 and epMotion 5073). The V4 region of the 16*S* rDNA gene (515F-806R) was sequenced; generating partially overlapping, paired-end reads on the Illumina MiSeq platform. After quality control filtering; a total of 3,979,728 reads were processed with an average of 132,657 reads per sample. The demultiplexed FASTQ files containing the 16*S* rRNA gene sequences were filtered for chimeric sequences using the USEARCH (version 8.1.1831) utility's UCHIME implementation and the 'gold' database (version microbiomeutil-r20110519). The reads were then binned into Operational Taxonomic Units (OTUs) at 97% similarity using USEARCH's cluster_otus command. The OTUs thus obtained were classified and aligned using QIIME (1.9.1) scripts. The assign_taxonomy.py script was used to assign taxonomy using the default RDP method and the default GreenGenes database. This provided insight into the larger trends at higher taxonomic levels (such as order *Clostridiales*). The 16*S* rRNA sequencing data was deposited into Sequence Read Archive (SRA) public database with the accession number SRP136953.

Human studies

Nontumor specimens derived from a set of 142 patients of the TIGER-LC Consortium were used in this study (*34*). Transcript expression was measured using the Affymetrix Human Transcriptome Array 2.0. Data has been deposited into the Gene Expression Omnibus (GEO) public database at NCBI (GEO Series GSE76297). A total of 718 biochemical metabolite species were measured by Metabolon's Discover HD4 Platform. All expression and metabolite data were log2 transformed. Pearson correlation analysis was performed using GraphPad Prism 7 to determine correlation between CXCL16 gene expression and selected metabolites. Due to the missing information caused by detection limitations, 85 valid patient data were used to correlate CDCA and CXCL16 expression.

Statistical analysis

The sample sizes for animal studies were guided by previous murine studies in our laboratory. Statistical analysis was performed with GraphPad Prism 7 (GraphPad Software). The significant differences between groups were calculated by Student's unpaired *t* test, one-way, or two-way ANOVA (Tukey's and Bonferroni's multiple comparison test). Welch's corrections were used when variances between the groups were unequal. $P <$ 0.05 was considered as statistically significant.

REFERENCES AND NOTES

1. S. Viaud *et al*., The intestinal microbiota modulates the anticancer immune effects of cyclophosphamide. *Science* **342**, 971–976 (2013). doi: 10.1126/science.1240537; pmid: 24264990
2. M. Vétizou *et al*., Anticancer immunotherapy by CTLA-4 blockade relies on the gut microbiota. *Science* **350**, 1079–1084 (2015). doi: 10.1126/science.aad1329; pmid: 26541610
3. N. Iida *et al*., Commensal bacteria control cancer response to therapy by modulating the tumor microenvironment. *Science* **342**, 967–970 (2013). doi: 10.1126/science.1240527; pmid: 24264989
4. B. Routy *et al*., Gut microbiome influences efficacy of PD-1-based immunotherapy against epithelial tumors. *Science* **359**, 91–97 (2018). doi: 10.1126/science.aan3706; pmid: 29097494
5. V. Matson *et al*., The commensal microbiome is associated with anti-PD-1 efficacy in metastatic melanoma patients. *Science* **359**, 104–108 (2018). doi: 10.1126/science.aao3290; pmid: 29302014
6. V. Gopalakrishnan *et al*., Gut microbiome modulates response to anti-PD-1 immunotherapy in melanoma patients. *Science* **359**, 97–103 (2018). doi: 10.1126/science.aan4236; pmid: 29097493
7. S. Yoshimoto *et al*., Obesity-induced gut microbial metabolite promotes liver cancer through senescence secretome. *Nature* **499**, 97–101 (2013). doi: 10.1038/nature12347; pmid: 23803760
8. D. H. Dapito *et al*., Promotion of hepatocellular carcinoma by the intestinal microbiota and TLR4. *Cancer Cell* **21**, 504–516 (2012). doi: 10.1016/j.ccr.2012.02.007; pmid: 22516259
9. G. Disibio, S. W. French, Metastatic patterns of cancers: Results from a large autopsy study. *Arch. Pathol. Lab. Med.* **132**, 931–939 (2008). pmid: 18517275
10. C. Ma *et al*., NAFLD causes selective CD4(+) T lymphocyte loss and promotes hepatocarcinogenesis. *Nature* **531**, 253–257 (2016). doi: 10.1038/nature16969; pmid: 26934227
11. T. Eggert *et al*., Distinct Functions of Senescence-Associated Immune Responses in Liver Tumor Surveillance and Tumor Progression. *Cancer Cell* **30**, 533–547 (2016). doi: 10.1016/j.ccell.2016.09.003; pmid: 27728804
12. A. Marabelle *et al*., Depleting tumor-specific Tregs at a single site eradicates disseminated tumors. *J. Clin. Invest.* **123**, 2447–2463 (2013). doi: 10.1172/JCI64859; pmid: 23728179
13. F. Geissmann *et al*., Intravascular immune surveillance by CXCR6+ NKT cells patrolling liver sinusoids. *PLOS Biol.* **3**, e113 (2005). doi: 10.1371/journal.pbio.0030113; pmid: 15799695
14. C. H. Kim *et al*., Bonzo/CXCR6 expression defines type 1-polarized T-cell subsets with extralymphoid tissue homing potential. *J. Clin. Invest.* **107**, 595–601 (2001). doi: 10.1172/JCI11902; pmid: 11238560
15. Y. J. Lee, K. L. Holzapfel, J. Zhu, S. C. Jameson, K. A. Hogquist, Steady-state production of IL-4 modulates immunity in mouse strains and is determined by lineage diversity of iNKT cells. *Nat. Immunol.* **14**, 1146–1154 (2013). doi: 10.1038/ni.2731; pmid: 24097110
16. L. A. Pobezinsky *et al*., Let-7 microRNAs target the lineage-specific transcription factor PLZF to regulate terminal NKT cell differentiation and effector function. *Nat. Immunol.* **16**, 517–524 (2015). doi: 10.1038/ni.3146; pmid: 25848867
17. K. Atarashi *et al*., Induction of colonic regulatory T cells by indigenous Clostridium species. *Science* **331**, 337–341 (2011). doi: 10.1126/science.1198469; pmid: 21205640
18. E. Vivier, S. Ugolini, D. Blaise, C. Chabannon, L. Brossay, Targeting natural killer cells and natural killer T cells in cancer. *Nat. Rev. Immunol.* **12**, 239–252 (2012). doi: 10.1038/nri3174; pmid: 22437937
19. M. Terabe, J. A. Berzofsky, The role of NKT cells in tumor immunity. *Adv. Cancer Res.* **101**, 277–348 (2008). doi: 10.1016/S0065-230X(08)00408-9; pmid: 19055947
20. M. Terabe, J. A. Berzofsky, in *Encyclopedia of Immunobiology*, M. J. H. Ratcliffe, Ed. (Elsevier, 2016), vol. 4, pp. 460-469.
21. D. Y. Wu, N. H. Segal, S. Sidobre, M. Kronenberg, P. B. Chapman, Cross-presentation of disialoganglioside GD3 to natural killer T cells. *J. Exp. Med.* **198**, 173–181 (2003). doi: 10.1084/jem.20030446; pmid: 12847141
22. K. H. Sonoda, M. Exley, S. Snapper, S. P. Balk, J. Stein-Streilein, CD1-reactive natural killer T cells are required for development of systemic tolerance through an immune-privileged site. *J. Exp. Med.* **190**, 1215–1226 (1999). doi: 10.1084/jem.190.9.1215; pmid: 10544194
23. M. J. Smyth *et al*., Differential tumor surveillance by natural killer (NK) and NKT cells. *J. Exp. Med.* **191**, 661–668 (2000). doi: 10.1084/jem.191.4.661; pmid: 10684858
24. F. Balkwill, Cancer and the chemokine network. *Nat. Rev. Cancer* **4**, 540–550 (2004). doi: 10.1038/nrc1388; pmid: 15229479
25. T. Shimaoka *et al*., Molecular cloning of a novel scavenger receptor for oxidized low density lipoprotein, SR-PSOX, on macrophages. *J. Biol. Chem.* **275**, 40663–40666 (2000). doi: 10.1074/jbc.C000761200; pmid: 11060282
26. S. I. Sayin *et al*., Gut microbiota regulates bile acid metabolism by reducing the levels of tauro-beta-muricholic acid, a naturally occurring FXR antagonist. *Cell Metab.* **17**, 225–235 (2013). doi: 10.1016/j.cmet.2013.01.003; pmid: 23395169

27. Y. Zhang, C. D. Klaassen, Effects of feeding bile acids and a bile acid sequestrant on hepatic bile acid composition in mice. *J. Lipid Res.* **51**, 3230–3242 (2010). doi: 10.1194/jlr.M007641; pmid: 20671298
28. C. Staley, A. R. Weingarden, A. Khoruts, M. J. Sadowsky, Interaction of gut microbiota with bile acid metabolism and its influence on disease states. *Appl. Microbiol. Biotechnol.* **101**, 47–64 (2016). pmid: 27888332
29. T. Olszak *et al*., Microbial exposure during early life has persistent effects on natural killer T cell function. *Science* **336**, 489–493 (2012). doi: 10.1126/science.1219328; pmid: 22442383
30. C. M. Theriot, A. A. Bowman, V. B. Young, Antibiotic-Induced Alterations of the Gut Microbiota Alter Secondary Bile Acid Production and Allow for Clostridium difficile Spore Germination and Outgrowth in the Large Intestine. *MSphere* **1**, e00045-15 (2016). doi: 10.1128/mSphere.00045-15; pmid: 27239562
31. J. M. Ridlon, D. J. Kang, P. B. Hylemon, Bile salt biotransformations by human intestinal bacteria. *J. Lipid Res.* **47**, 241–259 (2006). doi: 10.1194/jlr.R500013-JLR200; pmid: 16299351
32. C. G. Buffie *et al*., Precision microbiome reconstitution restores bile acid mediated resistance to Clostridium difficile. *Nature* **517**, 205–208 (2015). doi: 10.1038/nature13828; pmid: 25337874
33. S. Bhowmik *et al*., Structure and functional characterization of a bile acid 7α dehydratase BaiE in secondary bile acid synthesis. *Proteins* **84**, 316–331 (2016). doi: 10.1002/prot.24971; pmid: 26650892
34. J. Chaisaingmongkol *et al*., Common Molecular Subtypes Among Asian Hepatocellular Carcinoma and Cholangiocarcinoma. *Cancer Cell* **32**, 57–70.e3 (2017). doi: 10.1016/j.ccell.2017.05.009; pmid: 28648284
35. A. Kurioka, L. J. Walker, P. Klenerman, C. B. Willberg, MAIT cells: New guardians of the liver. *Clin. Transl. Immunology* **5**, e98 (2016). doi: 10.1038/cti.2016.51; pmid: 27588203
36. L. Van Kaer, alpha-Galactosylceramide therapy for autoimmune diseases: Prospects and obstacles. *Nat. Rev. Immunol.* **5**, 31–42 (2005). doi: 10.1038/nri1531; pmid: 15630427
37. C. Ma *et al*., Anti-Gr-1 antibody depletion fails to eliminate hepatic myeloid-derived suppressor cells in tumor-bearing mice. *J. Leukoc. Biol.* **92**, 1199–1206 (2012). doi: 10.1189/jlb.0212059; pmid: 23077247
38. A. Limmer *et al*., Efficient presentation of exogenous antigen by liver endothelial cells to CD8+ T cells results in antigen-specific T-cell tolerance. *Nat. Med.* **6**, 1348–1354 (2000). doi: 10.1038/82161; pmid: 11100119

ACKNOWLEDGMENTS

We thank the NIH Tetramer Core Facility at Emory University for the CD1d-tetramer. **Funding:** T.F.G., M.T., J.A.B., X.W.W., and G.T. were supported by the Intramural Research Program of the NIH, NCI (ZIA BC 011345, ZO1 SC 004020, and Z01 BC 010876). C.M.T. was supported by R35GM119438. T.L. was supported by the Deutsche Forschungsgemeinschaft DFG (SFB/TRR57, project Q1). **Author contributions**: T.F.G. conceived the study. C.M. and M.H. designed and performed most of the experiments and analyzed data. B.H., Q.F., Q.Z., M.S., D.A., T.R., and W.Y. contributed to experiments and data analysis. M.T., J.B., C.M.T., T.L., F.K., V.T., S.K.S., and G.T. contributed to data analysis. J.A.M. and S.R. provided key material. V.F., M.R., and X.W.W. provided and analyzed human data. C.M. and T.F.G. prepared illustrations, and wrote the manuscript. All authors contributed to editing of the manuscript. **Competing interests**: C.M and T.F.G are inventors on a patent application (U.S. Patent Application no. 62/578,176) submitted by the National Cancer Institute that covers the use of bile acids to treat and/or prevent adverse liver conditions The authors declare no other competing interests. **Data and materials availability**: All data presented in this manuscript are available in the main text and the supplementary materials. **Accession numbers:** The transcript expression data of nontumor specimens of the TIGER-LC Consortium were deposited into the Gene Expression Omnibus (GEO) public data set with accession number GSE76297. The 16*S* rRNA sequencing data was deposited into the Sequence Read Archive (SRA) public database with accession number SRP136953.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/360/6391/eaan5931/suppl/DC1
Materials and Methods
Figs. S1 to S6
References

5 May 2017; resubmitted 27 February 2018
Accepted 12 April 2018
10.1126/science.aan5931

RESEARCH ARTICLE

PALEOCLIMATE

A 550,000-year record of East Asian monsoon rainfall from ^{10}Be in loess

J. Warren Beck,[1,2,3]* Weijian Zhou,[4,5]* Cheng Li,[3] Zhenkun Wu,[3,4] Lara White,[1,3] Feng Xian,[4] Xianghui Kong,[3,4] Zhisheng An[4,6]

Cosmogenic ^{10}Be flux from the atmosphere is a proxy for rainfall. Using this proxy, we derived a 550,000-year-long record of East Asian summer monsoon (EASM) rainfall from Chinese loess. This record is forced at orbital precession frequencies, with higher rainfall observed during Northern Hemisphere summer insolation maxima, although this response is damped during cold interstadials. The ^{10}Be monsoon rainfall proxy is also highly correlated with global ice-volume variations, which differs from Chinese cave δ^{18}O, which is only weakly correlated. We argue that both EASM intensity and Chinese cave δ^{18}O are not governed by high-northern-latitude insolation, as suggested by others, but rather by low-latitude interhemispheric insolation gradients, which may also strongly influence global ice volume via monsoon dynamics.

Chinese speleothem ^{18}O/^{16}O (d^{18}O) isotope ratio records (*1*, *2*), such as the Sanbao Cave record, are clearly linked to the Asian monsoon and provide some of the most high-resolution and best dated records of tropical paleoclimate available. Their extraordinary match with high-latitude (65°N) summer insolation variations suggests a strong coupling with high-latitude climate (*1*); however, the origins of this coupling and of the δ^{18}O variations themselves remains controversial. Wang *et al.* (*2*) originally interpreted Chinese speleothem δ^{18}O as reflecting changes in the mixing ratio of summer versus winter monsoon moisture, each differing in isotopic composition. Subsequent authors have instead suggested that cave δ^{18}O variations were a response to changing upstream Rayleigh fractionation (*3*) or changes in the seasonality of precipitation (*4*), or were a proxy for monsoon intensity (*1*). Still others (*5*) assert that cave δ^{18}O cannot be strictly a proxy for Asian monsoon intensity because according to them, intensity should be phase-lagged from 65°N solar insolation. Thus, it seems the question of what controls cave δ^{18}O remains unclear. Here, we reinterpret these cave records in the context of our ^{10}Be-based East Asian summer monsoon (EASM) rainfall record.

Meteoric ^{10}Be preserved in Pleistocene Chinese loess can be used as a proxy for monsoon paleorainfall. Our sampling site (N34.43°, E107.12°) was chosen in order to study the paleo-EASM, which dominates the regional climate there, providing ~85% of annual rainfall. ^{10}Be is a long-lived cosmogenic radionuclide produced in the atmosphere and does not form gaseous species but instead attaches to atmospheric dust, with deposition mainly mediated by wet precipitation events (*6*, *7*). Its flux can thus be quantitatively linked with rainfall amount, after correcting for geomagnetic field and recycling effects (supplementary materials). Once on the ground, ^{10}Be becomes an immobile oxidized component of the sediments, permitting extraction of a rainfall record with minimal signal loss from diagenesis. Direct comparison with the nearby U/Th–dated Sanbao Cave δ^{18}O record (*1*), which is also reportedly a record of monsoon strength, was made possible by establishing an age model for this record by correlating the large number of similar features observed in both records.

Main findings

Shown in Fig. 1A is our ^{10}Be-proxy rainfall record (Fig. 1A, red curve) compared with the Sanbao δ^{18}O record (Fig. 1A, yellow curve) (*1*). ^{10}Be reveals that EASM rainfall increases during Northern Hemisphere summer insolation maxima and decreases during insolation minima (Fig. 1A, gray dashed curve) (*8*). Fig. 1A shows that EASM rainfall intensity is inversely correlated with speleothem δ^{18}O; however, the scaling deviates during

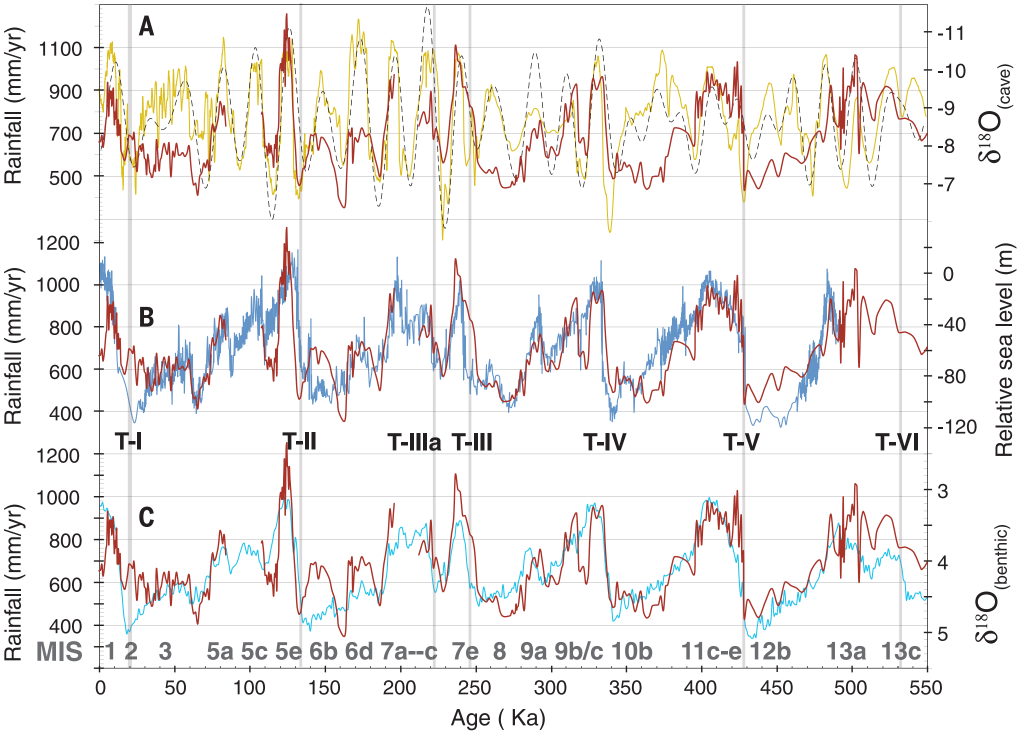


Fig. 1. ^{10}Be-proxy EASM rainfall versus Sanbao Cave d^{18}O, sea level, and benthic foraminiferal d^{18}O records. (**A**) Plot of ^{10}Be-based rainfall (red) versus Sanbao δ^{18}O record (yellow). The δ^{18}O scale is reversed. 65°N June insolation (dashed curve) is also shown (*8*) for reference. Locations of some MIS stages are labeled below (C). MIS stages 5b,c and 7b are missing from our profile (supplementary materials). (**B**) Plot of ^{10}Be-based rainfall (red) versus Red Sea glacio-eustatic sea level curve (blue) relative to modern sea level (*8*), both placed on Sanbao Cave age models. Main glacial terminations (vertical gray bars) are labeled below this figure. (**C**) Plot of ^{10}Be-based rainfall (red) versus LR04 stacked benthic δ^{18}O curve (blue) (*9*).

the cold interstadials of marine isotope (MIS)–3, -6, -8, -10, and -12, where our ^{10}Be-rainfall proxy has lower amplitude than that of the cave isotopes. As a result, we found that only 16.3% of the variance in the ^{10}Be-proxy rainfall record is correlated with the Sanbao record (fig. S1A). Nevertheless, both exhibit strong forcing at precessional periods.

On the other hand, our ^{10}Be-rainfall record is strongly correlated with global ice volume proxies. Shown in Fig. 1B is our record plotted against the Red Sea (RSL) paleosea level reconstruction (*8*), whereas in Fig. 1C we compare it with the LR04 benthic foraminifera δ^{18}O record (*9*). Our rainfall record exhibits very similar structure, with respectively 57 and 51.7% of its variance reflected in each of these two global ice volume proxies (fig. S1, B and C). This suggests that EASM intensity is closely coupled to high-latitude ice volume variations by some mechanism.

Surprisingly, this is not true of the Sanbao Cave δ^{18}O record, which is highly correlated with 65°N summer insolation, but only 10.6% of its variance is correlated with sea-level or global ice volume amplitude (Fig. 2). This creates a conundrum: Although cave δ^{18}O and ^{10}Be monsoon proxies are modestly correlated with each other, the former is highly correlated with insolation but not ice volume, whereas the latter is highly correlated with ice volume but not insolation. The Chinese cave community asserts that its isotope records imply that Asian monsoon intensity is controlled by orbital forcing of high-northern-latitude insolation (*1*, *10*); but if so, how can their proxy for monsoon intensity also be essentially uncorrelated with global ice volume? This seemingly contradicts a principal tenet of the Milankovitch theory, that high-northern-latitude solar insolation controls global ice volume and, through ice albedo feedback, also controls global temperature and, by corollary, tropical monsoon strength. How do we reconcile these two monsoon records, and can we do so without violating the Milankovitch theory?

Interpretation

We suggest that these conflicts may in part be resolved by reinterpreting the cave isotopes as a two-component moisture-mixing proxy instead of monsoon intensity and by recognition that monsoon intensity is determined mainly by low-latitude, not high-latitude, insolation variations. Why, then, are Chinese cave δ^{18}O records so strongly correlated with 65°N summer insolation (Fig. 3A)? The reason is that 65°N insolation exhibits nearly identical phase, net-range, and pattern of variations as those of the low-latitude (30°N to 30°S) June solar insolation gradient (Fig. 3B). This finding partly resolves the conflict between the cave records and Milankovitch theory because a connection with ice volume (and ice albedo) is not required if monsoons are forced from low latitudes. Still, this observation weakens the concept that global climate control emanates principally from high-northern-latitude insolation.

The idea that global climate control should chiefly emanate from low latitudes seems a more natural perspective because the amount of solar energy falling between 30°N and 30°S (that is, the region controlling the monsoons) is more than an order of magnitude larger than that falling above 65°N. Furthermore, the monsoons are the primary agent by which this incident energy is transported aloft or to high latitude, where it can be reradiated to space. When coupled with our observation that the ^{10}Be monsoon intensity proxy is highly correlated with global ice volume, this model leads us to suggest that the monsoons may in fact be an important driver of high-latitude climate, rather than the reverse. This model thus deviates from classical Milankovitch theory, although it retains the idea that orbital forcing of insolation is important.

We argue that it is the low-latitude interhemispheric insolation gradient that controls both EASM intensity and Chinese cave δ^{18}O. We assert that this gradient modulates the pattern of upper-tropospheric outflow from the Asian Monsoons to the North Pacific versus Southern Indian Ocean subtropical highs and in so doing regulates the relative strength of Hadley and Walker circulations in the Indo-Pacific sectors. This in turn influences the relative strength of the Indian summer monsoon (ISM) and Western North Pacific summer monsoon (WNPSM). In this model, Chinese cave isotopes are explained by mixing between ^{18}O-depleted moisture derived from the ISM and relatively ^{18}O-undepleted moisture nominally derived from the WNPSM, both of which contribute to EASM rainfall. ISM moisture is generally more δ^{18}O-depleted than WNPSM-sourced moisture because of rainout during transport across India or from greater interaction with topography of the Tibetan Plateau (TP) before arrival in central China (*11*–*13*). Orbital forcing of the interhemispheric insolation gradient, we argue, results in a higher ISM (δ^{18}O-depleted) moisture fraction during boreal summer precession maxima, and vice versa. When this gradient is large, EASM rainfall amount also increases, although the effect is damped during cold interstadials because of persistent summer snow cover on the TP or orbital modulation of Hadley circulation strength (*14*).

Modern evidence

Modern observations (*15*) do in fact show that it is differential heating between the hemispheres at low latitudes—not high-latitude insolation—that controls tropical monsoon development. Modern observations also show that the Asian monsoon is linked to both meridional Hadley and zonal Walker circulation pathways via monsoon outflow to the southern Indian ocean subtropical high (SISH) and north Pacific subtropical high (NPSH) (*16*–*19*). These two subtropical highs are key elements of Hadley and Walker circulation, in which rising air generated by the tropical monsoons descends in the subtropics. We assert that

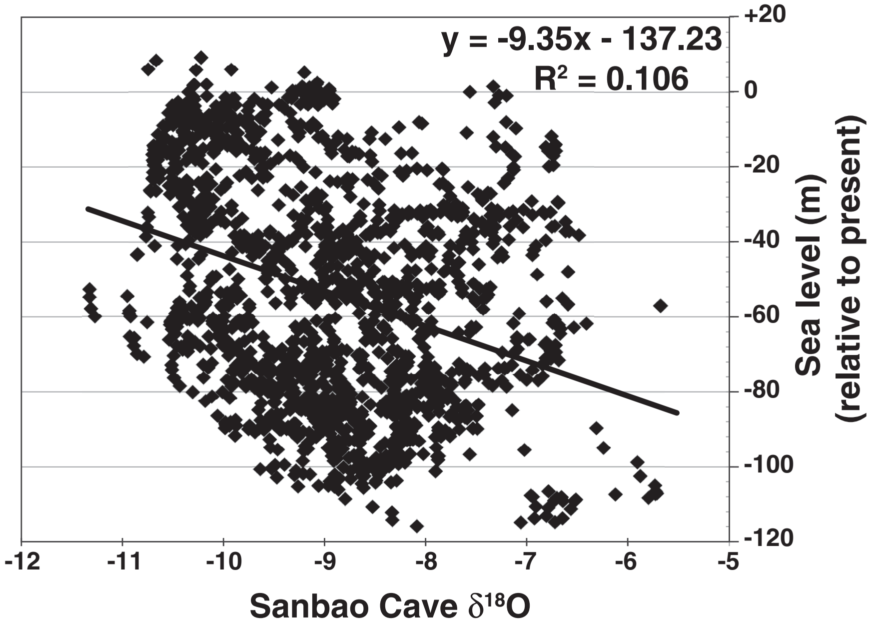


Fig. 2. Correlation between Red Sea glacio-eustatic sea-level reconstruction and Sanbao Cave δ^{18}O, on the same time scale. Only 10.6% of the variance in the cave isotope record is correlated with sea-level changes (*1*, *8*). Because of different sampling intervals, 3000-year smoothing filters have been applied to each record.

---

[1]Department of Geosciences, University of Arizona, Tucson, AZ 85721, USA. [2]Department of Physics, University of Arizona, Tucson, AZ 85721, USA. [3]NSF–Arizona Accelerator Mass Spectrometry Laboratory, University of Arizona, Tucson, AZ 85721, USA. [4]State Key Laboratory of Loess and Quaternary Geology, Institute of Earth Environment, Chinese Academy of Sciences, Xi'an 710075, China. [5]Qingdao National Laboratory for Marine Science and Technology, Qingdao, China. [6]Institute of Global Environmental Change, Jiaotong University, Xi'an 710061, China.
***Corresponding author. Email: wbeck@physics.arizona.edu (J.W.B.); weijian@loess.llqg.ac.cn (W.Z.)**

orbital forcing of solar insolation can change the ratio of monsoon outflow to these two regions differentially, intensifying the subtropical anticyclone in one sector while weakening it in the other. This in turn modulates the relative strength of the westerlies and trade winds in each sector, with both intensifying in one sector at the expense of the other owing to conservation of angular momentum. But Indian Ocean trade winds couple to the Somali jet (*15*, *19*), which influences ISM strength. North Pacific trade winds, on the other hand, influence the transport of moisture into central China from the WNPSM (*20*). Hence, differentially modulating the trade wind strength in these two sectors affects the mixing ratio of moisture arriving into central China derived from these two monsoons, and this, we believe, is what regulates both cave δ^{18}O and to a lesser extent EASM intensity, as seen in our ^{10}Be proxy.

That Hadley and Walker circulation in the Indian and North Pacific Ocean sectors are inversely coupled to each other via the pattern of upper-tropospheric monsoon outflow (*16*, *17*) is well documented in the seasonal cycle of atmospheric circulation (*18*, *21*). Observations show that during Boreal summer, the southern Hadley cell strengthens while the northern Hadley cell weakens, mainly in response to intensification of the Asian Monsoons (figs. S2 to S4). These two Hadley cells join at the intertropical convergence zone (ITCZ), which is the rising limb of both cells. During boreal summer, the ITCZ migrates into the Northern Hemisphere (figs. S2 and S3) in response to warmer northern and cooler Southern Hemisphere tropics and subtropics (*22*). Over Asia, this migration is accompanied by northward migration of the upper-tropospheric subtropical zonal westerly jet (*23*, *24*), which moves north of the TP during mid-summer. When it does so, it enhances the Tibetan high circulation pattern, which helps to redirect Asian monsoon outflow to the southern Indian Ocean subtropical high. At the same time, boreal summer heating over the TP (~30°N) drives onshore air mass transport, in which sensible heating from latent heat release in the mid-troposphere strongly drives convective and monsoon general circulation (*15*), producing the world's strongest center of atmospheric circulation. During austral summer, the ITCZ moves back into the Southern Hemisphere and the subtropical jet returns to the south of the TP as the heating centers reverse hemisphere, strengthening the northern and weakening the southern Hadley cells.

Thus, during boreal summer, northward movement of the ITCZ accompanied by heating over the TP strongly enhances northward cross-equatorial atmospheric mass transport in the lower troposphere associated with Hadley circulation. This requires an upper-tropospheric return flow to the Southern Hemisphere for mass balance. When coupled with migration of the subtropical jet, these factors strengthen the overturning circulation in the southern Hadley cell and strengthen both the ISM and WNPSM (*15*, *19*, *21*, *23*). Much of this lower tropospheric cross-equatorial mass flow occurs in the northwest Indian Ocean Somali jet, although a lesser part [~1/3 today (*25*)] occurs in several broad low-level flows north-northeast of Australia and the Bay of Bengal (Fig. 4). These two regions of low-level cross-equatorial flow are part of the reason that the southern Hadley cell completely dominates the global meridional, overturning circulation during boreal summer (*18*, *21*). The other reason is that the southern subtropical highs (~30°S) intensify during boreal summer in response to colder austral winter (June–July–August) temperatures. These southern highs (surface anticyclones) form the descending limb of the southern Hadley cell and are further intensified by strong upper-tropospheric return flows of cold dry air from the Northern Hemisphere that balance the low-level northward monsoonal flows during boreal summer (Fig. 4) (*18*, *19*, *26*). Cold sea surface temperature (SST) in the southern subtropics further enhances these anticyclones by inhibiting evaporation from the sea surface, whereas warmer SST in the northern subtropics allows increased evaporation into the lower to middle troposphere there, weakening those surface highs (*27*).

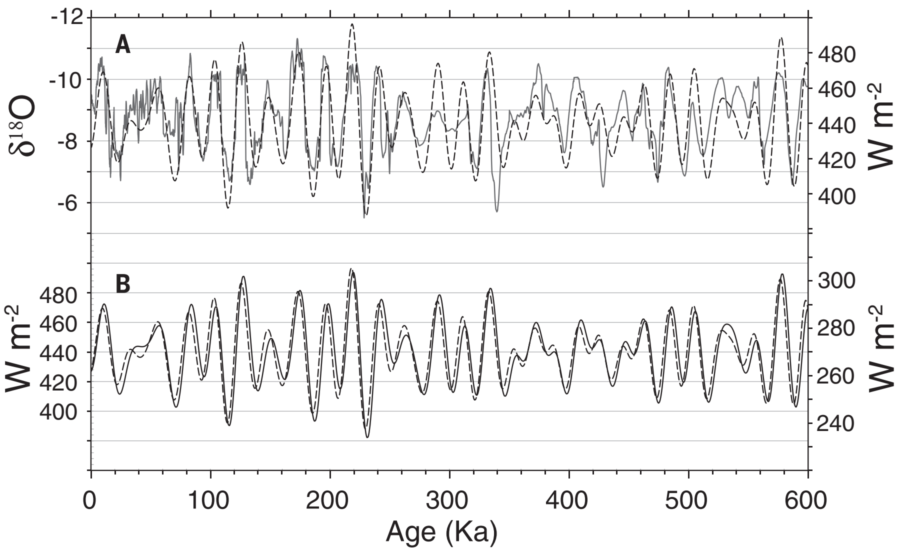


Fig. 3. High-latitude versus low-latitude solar insolation gradient variations compared to Sanbao Cave d^{18}O. (**A**) Plot of 65°N June solar insolation (dashed curve) (*39*) versus Sanbao Cave δ^{18}O record (solid curve) (*1*). The δ^{18}O scale is reversed. (**B**) Plot of 65°N June solar insolation (dashed curve) versus (30°N to 30°S) June solar insolation difference (solid curve). Visible are the nearly identical pattern and phase and the similar magnitude of changes.

Thus, the seasonal cycle of differential interhemispheric heating affects where the monsoon outflow goes. Today, although some Asian monsoon outflow falls into the Sahara and Asian interior, the remainder is mainly split between the NPSH and the SISH. Numerous modeling studies (*16*, *17*), reanalysis of modern climate data (*18*, *19*, *21*), and Lagrangian tracers studies of ozone (*28*) all show that this upper-tropospheric monsoon outflow to the Southern Indian Ocean is very strong during boreal summer and weak to the NPSH but reverses during austral summer (fig. S4). These models also show that these strong outflows lead to intensified trade winds in the colder hemisphere (*16*, *29*), where today stronger Indian Ocean trade winds couple directly to the Somali jet, intensifying the Indian Ocean tropical westerlies and the ISM (Fig. 4).

Variations in this cyclic seasonal pattern of monsoon outflow also occur, and reanalysis of modern climate data reveals that these variations are linked to changes in SST. In particular, there is an antiphased SST coupling between the Indian and west central Pacific oceans so that when SST in the Indian ocean is anomalously warm, SST in the west central Pacific is anomalously cool (*30*). This pattern is associated with stronger WNPSM and stronger easterly trade winds in the west Pacific, but weaker tropical westerlies over the Indian ocean and weaker ISM, and vice versa.

More generally, modern climate data show that strong Asian monsoon years are indeed associated with both anomalously strong northward movement of the ITCZ as well as anomalous cooling in the Southern Hemisphere subtropics (*25*, *31*). At the same time, stronger 20th-century ISM correlates with weakening and eastward retreat of the NPSH (*20*), which we argue is linked to weaker monsoonal outflow to the NPSH and stronger outflow to the SISH, as required by mass balance.

Putting this all together, reanalysis of modern climate data shows that strong Asian summer monsoons produce strong upper-tropospheric outflow from the Asian monsoons to the SISH but weak outflow to the NPSH and vise versa. It substantiates the idea that there is an interhemispheric oscillation linking Indian and Pacific ocean SST to trade wind velocity over the Northwest Pacific and Indian ocean sectors, and that the relative trade wind strength regulates the relative flux of moisture in the EASM from these two sectors.

Here lies the linkage between cave $\delta^{18}O$ and the subtropical meridional insolation gradient. Orbitally induced changes of the (June, 30°N to 30°S) subtropical isolation gradient mimic the seasonal cycle of low-latitude insolation variations that drive monsoons. These orbitally induced changes can thus either enhance or diminish this seasonal cycle, depending on the phase. Modeling results (*14*, *32*, *33*) support this idea that orbital-induced changes of the low-latitude interhemispheric insolation gradient do strongly regulate the strength of the winter hemisphere Hadley cell, and that this may work via SST regulation of monsoon outflow to the subtropical highs (*33*) coupled with changes in the mean position of subtropical jet relative to the TP (*23*, *24*). In our mechanism, these changes produce an "intertropical seesaw" between Hadley and Walker circulation that regulates the ratio of EASM moisture derived from the ISM versus WNPSM, modulating Chinese cave $\delta^{18}O$ (Fig. 4). The effect on EASM intensity—as told by the ^{10}Be-proxy—has the same sign and phase but is weaker during cold interstadials because persistent summer snow cover on the TP inhibits onshore monsoon air mass transport and keeps the subtropical westerly jet positioned south of the TP, weakening the Tibetan high.

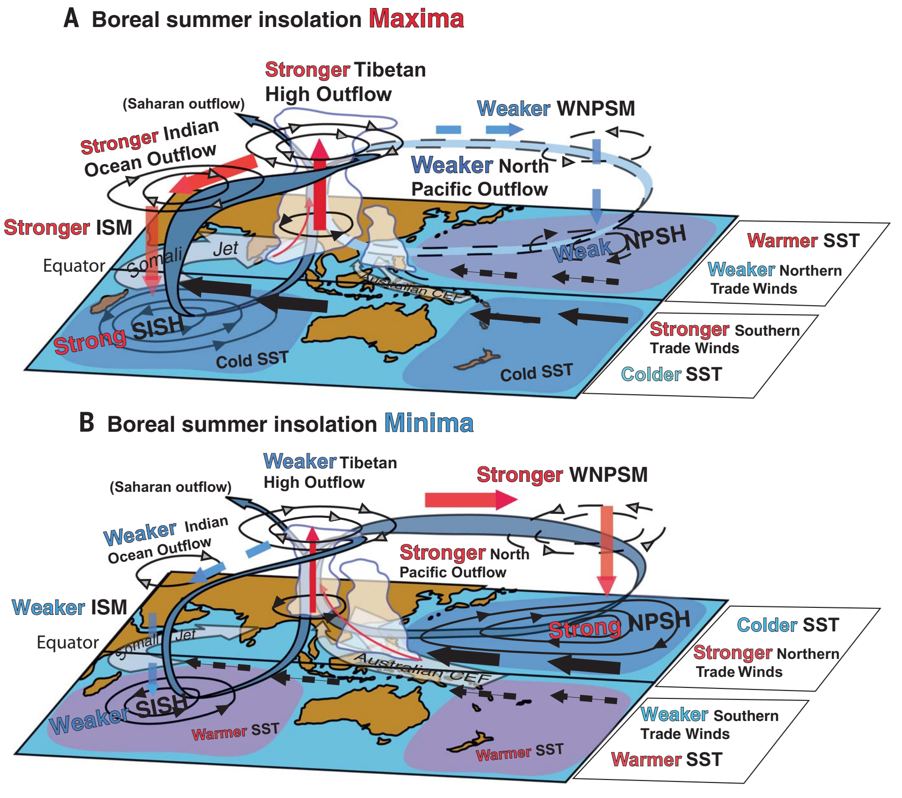


Fig. 4. Intertropical seesaw: Orbitally induced changes in solar insolation mimic the seasonal interhemispheric heating-cooling cycle in the tropics and subtropics. (**A**) During boreal summer insolation maxima, Tibetan high outflow to the SISH is increased in response to two factors: (i) colder SST in the South Indian ocean, which enhances the surface anticyclone there, and (ii) earlier seasonal transition of the upper-tropospheric subtropical zonal westerly jet (*24*, *25*) to the north side of the TP, strengthening the Tibetan high. Both factors enhance Hadley circulation and result in strengthening trade winds and Somali jet in the Indian ocean, strengthening the ISM. At the same time, outflow to the NPSH is reduced in consequence to warmer north Pacific SST, producing weaker northern trade winds and smaller onshore pressure gradient in the North Pacific, weakening the WNPSM. (**B**) During Boreal summer insolation minima, the opposite occurs.

High- and low-latitude coupling

The observation that variation in the Asian monsoon outflow to the SISH versus NPSH is coupled to trade wind and westerlies velocities in these two sectors reveals another linkage to high-northern-latitude ice volume via the Atlantic meridional overturning circulation (AMOC) (*34*–*38*). AMOC strongly affects the oceanic transport of heat to the high North Atlantic. AMOC in turn is strongly influenced by the Agulhas current leakage of high-salinity water masses from the southern Indian Ocean into the Atlantic. Modulating this source of high-density surface waters profoundly influences North Atlantic Deep Water production—a critical element of AMOC. But Agulhas current leakage is coupled to the strength and position of the SISH (the Mascarene High).

Modern observations between 1993 and 2009 (*34*) demonstrate that intensification of the SISH was indeed accompanied by both southern westerlies and trade wind intensification. The resulting increased wind stress coupling to the sea surface has led to enhanced South Equatorial Current and intensification of the Agulhas western boundary current system in the southwest Indian Ocean (fig. S5). These observations show that both mean kinetic energy and eddy current energy in the Agulhas system increased over this period. Of particular interest is the observation that both westward and southward propagation of mesoscale eddies in the southern Agulhas (32 to 36°S) also increased over this period. Westward propagation of these large eddies into the South Atlantic is the chief mechanism by which the Agulhas leakage occurs. Thus, modern evidence also shows that modulating southern trade wind velocity over the Indian ocean through changing SISH strength could strongly affect AMOC and the transport of heat to the high North Atlantic region.

Counteracting effects might also occur, however, because southern trade wind intensification can also lead to increased Agulhas mean velocity, resulting in Agulhas overshoot in the region south of South Africa. In this case, increased Agulhas momentum could carry the current past the zero-wind stress curl latitude into the region dominated by the westerlies, reducing Agulhas leakage and increasing retroflection back into the south Indian Ocean. Instabilities in either one of these two phenomena could produce Dansgaard-Oeschger cycles in the North Atlantic through modulation of AMOC, explaining the close correspondence between Greenland ice core and Chinese cave records on orbital or even suborbital time scales (*1*–*4*). There is evidence for strong increases in Agulhas leakage prefacing each of the past five glacial terminations (*38*), on the basis of the sudden increases of Indian Ocean phytoplankton species in South Atlantic cores that were either concurrent with or occurred just before the ice age terminations.

These observations illustrate how the cave isotope and ^{10}Be rainfall proxy records can be reconciled and show that there are mechanisms by which the Asian monsoon can directly influence high-northern-hemisphere climate, reopening the fundamental paleoclimate question of whether, or to what degree, glacial and interglacial climate is being forced from high versus low latitudes.

REFERENCES AND NOTES

1. H. Cheng *et al.*, *Nature* **534**, 640–646 (2016).
2. Y. J. Wang *et al.*, *Science* **294**, 2345–2348 (2001).
3. D. Yuan *et al.*, *Science* **304**, 575–578 (2004).
4. H. Cheng, A. Sinha, X. Wang, F. Cruz, R. Edwards, *Clim. Dyn.* **39**, 1045–1062 (2012).
5. S. Clemens, W. Prell, Y. Sun, *Paleoceanography* **25**, 1–19 (2010).
6. S. Caillet, P. Arpagaus, F. Monna, J. Dominik, *J. Environ. Radioact.* **53**, 241–256 (2001).
7. W. J. Zhou *et al.*, *Radiocarbon* **49**, 139–160 (2007).
8. K. M. Grant *et al.*, *Nat. Commun.* **5**, 5076 (2014).
9. L. E. Lisiecki, M. E. Raymo, *Paleoceanography* **20**, PA001071 (2005).
10. H. Cheng *et al.*, *Science* **326**, 248–252 (2009).
11. The phasing of cave isotopes in the Himalayas region of northeast India (which are in phase with Chinese lowland cave records, but outside the EASM district) may also be explained by this mechanism. In that region, the westward propagating

moisture vector carries relatively unfractionated moisture in part locally derived from the Bay of Bengal, although its transport is nonetheless still influenced by WNPSM dynamics (*12*).
12. S. Sengupta, A. Sarkar, *Earth Planet. Sci. Lett.* **250**, 511–521 (2006).
13. R. Shukla, B. Huang, *Clim. Dyn.* **46**, 1977–1990 (2016).
14. D. F. Mantsis *et al*., *J. Clim.* **27**, 5504–5516 (2014).
15. P. Webster, in *The Hadley Circulation: Present, Past and Future*, H. F. Diaz, R. S. Bradley, Eds. (Kluwer Academic Publishers, 2005), pp. 9–60.
16. A. E. Gill, *Quat. J. R. Met. Soc.* **106**, 447–462 (1980).
17. F. Liu, B. Wang, *Am. Met. Soc.* **26**, 1791–1806 (2013).
18. J. Schwendike *et al*., *J. Geophys. Res.* **119**, 1322–1339 (2014).
19. J. Walker, S. Bordoni, T. Schneider, *J. Clim.* **28**, 3731–3750 (2015).
20. Y. Huang, H. Wang, K. Fan, Y. Gao, *Clim. Dyn.* **44**, 2035–2047 (2015).
21. I. Dima, J. M. Wallace, I. Kraucunas, *Am. Met. Soc.* **62**, 2499–2513 (2005).
22. G. Basha *et al*., *Remote Sens. Environ.* **163**, 262–269 (2015).
23. J. Chiang *et al*., *Quat. Sci. Rev.* **108**, 111–129 (2014).
24. R. Schiemann, D. Lüthi, C. Schär, *J. Clim.* **22**, 2940–2957 (2009).
25. Y. Zhu, *Adv. Atmos. Sci.* **29**, 509–518 (2012).
26. S.-L. Lee, C. Mechoso, C. Wang, J. D. Neelin, *J. Clim.* **26**, 10193–10204 (2013).
27. T. Shaw, A. Voigt, *Nat. Geosci.* **8**, 560–566 (2015).
28. M. Hitchman, M. Rogal, *J. Geophys. Res.* **115**, D14118 (2010).
29. A. Broccoli, K. Dahl, R. Stouffer, *Geophys. Res. Lett.* **33**, L01702 (2006).
30. J. Cao, S. Gui, Q. Su, Y. Yang, *J. Clim.* **29**, 5027–5040 (2016).
31. C. Li, S. Li, *Am. Met. Soc.* **27**, 3966–3981 (2014).
32. M. Erb, A. Broccoli, A. Clement, *J. Clim.* **26**, 5897–5914 (2013).
33. M. Erb *et al*., *J. Clim.* **28**, 9258–9276 (2015).
34. B. Backeberg, P. Penven, M. Rouault, *Nat. Clim. Chang.* **2**, 608–612 (2012).
35. L. M. Beal, W. P. De Ruijter, A. Biastoch, R. Zahn, SCOR/WCRP/IAPSO Working Group 136, *Nature* **472**, 429–436 (2011).
36. L. Braby *et al*., *Geophys. Res. Lett.* **43**, 8143–8150 (2016).
37. W. Broeker, *Oceanography* **4**, 79–89 (1991).
38. F. J. Peeters *et al*., *Nature* **430**, 661–665 (2004).
39. A. Berger, M. F. Loutre, *Quat. Sci. Rev.* **10**, 297–317 (1991).

ACKNOWLEDGMENTS

This work was jointly supported by the U.S. National Science Foundation through EAR-0908709, EAR-0902633, and EAR-0929458, as well as the Ministry of Science and Technology of China, the National Science Foundations of China, and the Key Research Programs of Frontier Sciences, Chinese Academy of Sciences. We also gratefully acknowledge insightful editorial assistance from J. Russell and P. Goodman, University of Arizona Geosciences Department, and three anonymous reviewers. The data used in this work can be found in the associated supplementary materials.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/360/6391/877/suppl/DC1
Materials and Methods
Figs. S1 to S7
Tables S1 and S2
References (*40–54*)

15 December 2016; accepted 30 March 2018
10.1126/science.aam5825

NEURODEVELOPMENT

Evolution of pallium, hippocampus, and cortical cell types revealed by single-cell transcriptomics in reptiles

Maria Antonietta Tosches,*† Tracy M. Yamawaki,*‡ Robert K. Naumann,§ Ariel A. Jacobi,‖ Georgi Tushev, Gilles Laurent†

Computations in the mammalian cortex are carried out by glutamatergic and γ-aminobutyric acid–releasing (GABAergic) neurons forming specialized circuits and areas. Here we asked how these neurons and areas evolved in amniotes. We built a gene expression atlas of the pallium of two reptilian species using large-scale single-cell messenger RNA sequencing. The transcriptomic signature of glutamatergic neurons in reptilian cortex suggests that mammalian neocortical layers are made of new cell types generated by diversification of ancestral gene-regulatory programs. By contrast, the diversity of reptilian cortical GABAergic neurons indicates that the interneuron classes known in mammals already existed in the common ancestor of all amniotes.

Amniote vertebrates (mammals, reptiles, and birds) originated from a common ancestor about 320 million years ago. In all developing amniotes, the dorsal telencephalon, or pallium, is patterned by the same signaling molecules and subdivided in homologous embryonic regions (*1*). In adult amniote brains, however, the structures that arise from these homologous pallial regions have different morphologies and connectivity. For example, a six-layered neocortex exists only in mammals, and the dorsal ventricular ridge (DVR) is found only in birds and reptiles. The hippocampus may be the most conserved pallial region (*2*), but even there, it is uncertain whether all the subfields known in mammals (for example, the dentate gyrus) exist in nonmammals (*3*) (fig. S1).

Gene expression data can help reconstruct the evolution of brain regions and cell types (*4*, *5*). Here we used single-cell transcriptomics to study the evolution of neuronal diversity in the amniote telencephalon. Because cell types defined through transcriptomics match those defined by morphology, physiology, and connectivity (*6*–*8*), single-cell mRNA sequencing can be applied both for cell-type discovery and for cross-species comparisons (*9*). We focused on reptiles, because they, unlike birds, have a layered cortex, and asked the following: (i) Can we identify molecular similarities and homologies between reptilian and mammalian pallial regions? (ii) Are there reptilian cortical neurons homologous to mammalian hippocampal neurons? (iii) Can we link the reptilian three-layered cortex to the mammalian six-layered neocortex? (iv) Are mammalian γ-aminobutyric acid–releasing (GABAergic) interneuron types also found in the reptilian cortex?

Neuronal and glial cells in the reptilian pallium

Using Drop-seq (*8*), we obtained a comprehensive, unbiased classification of adult cell types sampled from the pallium of a turtle and a lizard species (dissections and sequencing statistics in figs. S1 and S2 and tables S1 and S2). After quality filtering (fig. S3), we used unsupervised graph-based clustering of the transcripts from 18,828 turtle and 4187 lizard cells (with a median of 2731.5 and 1918 transcripts per cell, respectively) and identified first-level clusters of neuronal and non-neuronal cells (Fig. 1, A to C). Among non-neuronal cells, we found clusters expressing prototypical markers of mammalian glial cells: mature oligodendrocytes and their precursors, ependymoglial cells, and microglia (Fig. 1, B to D). Ependymoglial cells coexpressed markers of mammalian astrocytes (*GFAP*), adult stem cells (*SOX9*), and ependymal cells (*FOXJ1*), suggesting a shared evolutionary origin of these cell types (fig. S4, A to C). Reptilian neuronal clusters included vesicular glutamate transporter type 1 ($VGLUT1^+$) glutamatergic excitatory neurons, glutamate decarboxylase 1 and 2 ($GAD1^+$ $GAD2^+$) GABAergic inhibitory interneurons, and neural progenitor cells (Fig. 1, A to D), consistent with the existence of adult neurogenesis in reptiles. To compare the transcriptomes of turtle, lizard, and mouse neuronal and non-neuronal cells (*6*), we selected one-to-one orthologs differentially expressed among these cell types, defined a specificity score for each gene in each cluster, and computed pairwise rank correlations (see methods). This analysis supports correspondence between major neuronal and non-neuronal cell classes (fig. S4, D to F) across turtle, lizard, and mouse.

Subclustering of the turtle and lizard neurons revealed finer distinctions between and within excitatory glutamatergic and inhibitory GABAergic neuron types (Fig. 1, E to G, and fig. S5). From these neuronal data sets, we selected differentially expressed genes as markers for histological validation and for unbiased comparisons with mammalian cell types. We focused on the turtle data and used the lizard data for corroboration.

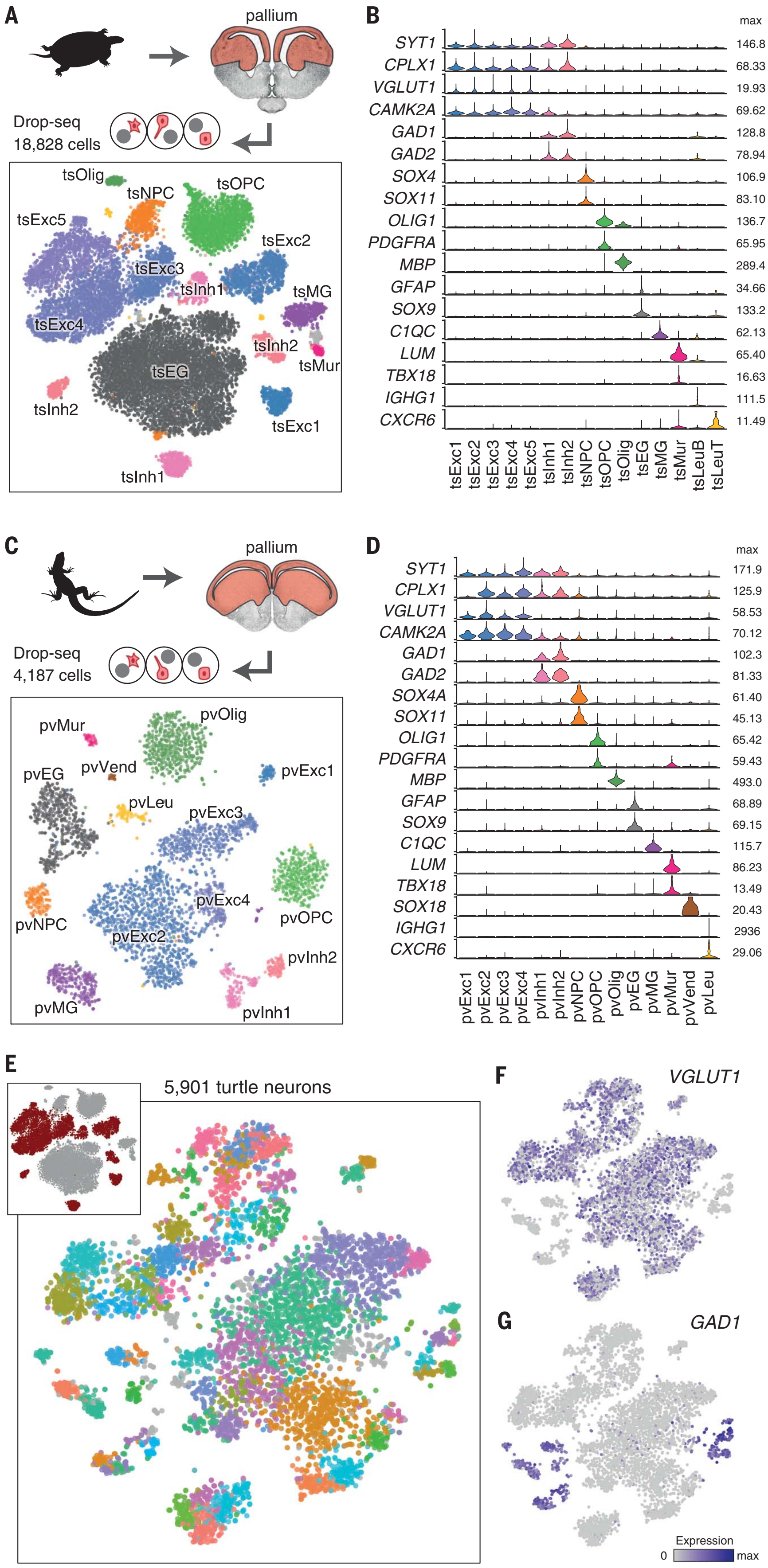
Spatial segregation of glutamatergic neurons: A molecular map of the reptilian pallium

Our dissections encompassed multiple regions of the reptilian pallium likely to contain heterogeneous populations of glutamatergic neurons. To establish the anatomical location of our turtle glutamatergic-cell clusters, we selected highly variable genes in the data set that were expressed in some clusters but not detected in others. These markers were expressed in distinct regions of the pallium (fig. S6). The combinatorial expression of markers defined "superclusters," seen also as groupings of adjacent clusters in t-distributed stochastic neighbor embedding (tSNE) plots (Fig. 2, A to B, and fig. S6A; compare Fig. 1E and Fig. 2B), in agreement with higher-level clustering analysis and with our tissue dissections (figs. S7 and S8D and tables S1 and S2). These superclusters correspond to defined anatomical regions: the medial and dorsomedial cortices, the anterior and posterior dorsal cortex, the pallial thickening, the anterior and posterior lateral cortex, and the anterior and posterior DVR (fig. S8, A to C). Cell-type similarity was generally higher within than between superclusters (fig. S8, F and G). Weighted gene correlation network analysis (WGCNA, see methods) indicates that the unique genetic signature of each region results from different combinations of gene modules (Fig. 2C). We also associated glutamatergic clusters to anatomically defined pallial regions in lizard (fig. S9). Corresponding regions in lizard and turtle have different relative sizes (for example, lizard anterior dorsal cortex is small) but are clearly delineated by the expression of the same developmental transcription factors such as *ZBTB20*, *SATB1*, *DACH2*, and *ETV1* (*ER81*) (Fig. 2D). These data define the molecular regionalization of the adult reptilian pallium.

Putative homologies between reptilian and mammalian pallial derivatives are disputed (*1*, *2*, *10*, *11*). Central to this debate is the anterior DVR, one of the derivatives of the ventral pallium, which is enlarged in birds and reptiles. This region has been proposed as a homolog of either ventral pallium derivatives [claustrum, endopiriform nucleus, and pallial amygdala (*1*, *2*, *12*)] or parts of the neocortex in mammals (*5*, *10*, *11*). The latter hypothesis suggests a dual evolutionary origin of the neocortex, either as separate regions—where medial and lateral neocortex are homologous to reptilian dorsal cortex and DVR, respectively (*11*)—or as intermixed cell types, where separate neocortical layers are homologous to reptilian dorsal cortex or DVR (*5*, *10*). We compared the turtle superclusters

---

Max Planck Institute for Brain Research, Max-von-Laue Strasse 4, 60438 Frankfurt am Main, Germany.
*These authors contributed equally to this work.
†Corresponding author. Email: gilles.laurent@brain.mpg.de (G.L.); maria.tosches@brain.mpg.de (M.A.T.) ‡Present address: Amgen, Inc., 1120 Veterans Blvd., South San Francisco, CA 94080, USA. §Present address: Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, 1068 Xueyuan Ave., Shenzhen University Town, Nanshan District, 518055 Shenzhen, China. ‖Present address: University of California, Davis, School of Medicine, 4610 X St., Sacramento, CA 95817, USA.



to mammalian pallial derivatives, using a human microarray data set as a reference [(*13*); see analysis in methods].

Our analysis using all differentially expressed genes (Fig. 2E) reveals similarities between turtle medial and dorsomedial cortices and human hippocampus, supporting previous hypotheses (*3*). The posterior dorsal cortex also showed positive correlations to human hippocampus and subiculum but low negative correlations with nonhippocampal cortices, consistent with the gene network analysis (Fig. 2C). Earlier studies recognized anatomical similarities between parts of the reptilian dorsal cortex and mammalian subiculum (*2*, *14*), where *ETV1* (*ER81*) is expressed (*15*). It may thus be that the posterior dorsal cortex relates to mammalian peri-hippocampal regions.

Turtle pallial thickening and mammalian claustrum shared expression of several claustrum-enriched markers (fig. S6B), consistent with anatomical and developmental data (*1*, *2*). We also found correlations between reptilian lateral cortex and mammalian piriform cortex and between posterior DVR and pallium-derived amygdalar nuclei, with the exception of the lateral amygdala. Individual posterior DVR clusters expressed markers of mammalian basolateral (*ETV1* and *FEZF2*) and cortical (*ZIC2* and *TBR1*) amygdala. These clusters mapped to distinct nuclei of the turtle posterior DVR, suggesting that these pallial amygdala subdivisions existed in the common ancestor of mammals and reptiles (fig. S10) (*16*).

Mammalian neocortex showed correlations with several turtle pallial regions (Fig. 2E). Reasoning that correlations based on all differentially expressed genes might not separate homology

Fig. 1. Reptilian single-cell data sets. (**A**) tSNE representation of 18,828 turtle (*Trachemys scripta elegans*, ts) single-cell transcriptomes, with cells color coded by cluster. (**B**) Violin plots showing expression of neuronal and non-neuronal markers in turtle clusters. (**C**) tSNE representation of 4187 lizard (*Pogona vitticeps*, pv) single-cell transcriptomes, color coded by cluster. (**D**) Violin plots showing expression of neuronal and non-neuronal markers in lizard clusters. In (B) and (D), for each gene, violin plots are scaled to the maximum number of transcripts (unique molecule identifiers) detected for that gene (max). Exc, glutamatergic excitatory neurons; Inh, GABAergic inhibitory interneurons; NPC, neural progenitor cells; Olig, mature oligodendrocytes; OPC, oligodendrocyte precursors; EG, ependymoglial cells; MG, microglia; Leu, leucocytes; Mur, mural cells; Vend, vascular endothelial cells. Subclusters of Exc and Inh cells in (A) and (C) are not matched by name or color. (**E**) tSNE plot and clusters of 5901 high-quality turtle neurons. Inset is the tSNE plot in (A), indicating neuronal and glial clusters in dark red and gray, respectively. (**F** and **G**) Expression of the glutamatergic marker *VGLUT1* (F) and GABAergic marker *GAD1* (G) in turtle neurons.

from convergent evolution (because functional convergence could recruit the same effector genes under different transcription factors), we next restricted our analysis to transcription factors. Under these conditions, only anterior dorsal cortex correlated with human neocortex (Fig. 2F). Anterior dorsal cortex and anterior DVR differed by the expression of those transcription factors that, in mammals, are also found in either the neocortex or pallial amygdala (for example, *NFIX*, *BCL11A*, and *SATB2* in turtle anterior dorsal cortex and mouse neocortex; *NR2F2* and *DACH2* in turtle anterior DVR and mouse amygdala) (fig. S8, H and I). Of the mammalian pallial amygdala subdivisions, only the lateral amygdala correlates with the anterior DVR. Our results suggest that (i) different combinations of transcription factors may regulate the expression of "neocortical" effector genes in anterior dorsal cortex and anterior DVR and (ii) transcription-factor expression reflects the developmental (and evolutionary) history of pallial neurons. We propose that reptilian anterior dorsal cortex and mammalian neocortex are homologous as dorsal pallium derivatives and that reptilian DVR and mammalian pallial amygdala are homologous as ventral pallium derivatives [see also (*1*)]. In reptiles, the expansion of the sensory-recipient anterior DVR led to the emergence of neuronal types specialized in processing different sensory modalities, recognizable as separate molecular, anatomical, and functional clusters (figs. S9 and S10) (*17*). According to this hypothesis, reptilian anterior DVR and mammalian neocortex acquired, by convergent evolution, the expression of similar sets of effector genes.

Conservation of hippocampal neuronal types and areal organization

Anatomical and developmental evidence suggest that the reptilian medial-most cortex is homologous to mammalian hippocampus (*3*, *18*). In line with this, turtle and lizard medial cortices express the mammalian pan-hippocampal transcription factor *ZBTB20* (Fig. 3A and fig. S11). Whether individual hippocampal subfields [dentate gyrus (DG), cornu ammonis (CA)3, CA2, and CA1] are present in reptiles and birds is less clear. Developmental evidence suggests so (*18*), but some describe mammalian DG, with its mossy cells and granule cells, as a mammalian novelty (*3*).

Reptilian *ZBTB20*-positive clusters could be further distinguished by the expression of mammalian DG or CA transcription factors: In turtle, *PROX1* and *MEF2C* (specifying mouse DG granule cells) labeled the medial cortex, and *ETV1*, *MEIS2*, and *LMO4* (CA) labeled the dorsomedial cortex (Fig. 3A). The expression of these genes in adjacent domains of turtle and lizard medial cortices (fig. S11) suggests the existence of DG- and CA-like neuronal types in reptiles. This was further supported by unbiased analyses of cell-type transcriptomes. Using WGCNA to identify and compare gene modules [mouse data from (*19*), see methods], we found statistically significant overlaps between mouse DG and turtle medial cortex modules and between mouse CA and turtle dorsomedial cortex modules (fig. S12). Mouse DG and turtle medial cortex modules shared genes coding for K^+-channel subunits or associated proteins (*KCNG1*, *KCNA1*, and *KCNIP4*), possibly accounting for electrophysiological similarities (*20*). Other shared genes included the cadherin *CDH8*, involved in the formation of DG-CA3 synapses, and the granule-cell specific regulators of synaptogenesis and AMPA receptor–mediated synaptic transmission *LRRTM4* and *CNIH3* (fig. S12). Hence, DG granule cells, including aspects of their membrane and synaptic physiology, are conserved across mammals and reptiles (*20*). [No mossy-cell marker (*21*) had cell-type specific expression in turtle medial cortex.]

Next, we computed pairwise cluster correlations using one-to-one orthologs differentially expressed among the *ZBTB20*$^+$ clusters. This revealed further heterogeneity among the *ZBTB20*$^+$ *ETV1*$^+$ cells, with turtle dorsomedial cortex clusters showing best mutual correlations to mouse CA3 or CA1 (Fig. 3B). Mammalian CA3 or CA1 markers (*19*, *21*) were differentially expressed in these turtle clusters (Fig. 3C). CA3- and CA1-like cells occupy adjacent positions [with CA3-like cells closest to medial cortex (DG)] and form morphologically distinct cell plates (Fig. 3A and fig. S11). We found no evidence for a reptilian CA2 cell type (Fig. 3B and fig. S11B).

In summary, reptilian hippocampus can be subdivided into adjacent territories similar to the mammalian DG, CA3, and CA1 fields. The developmental origin of these cells from the medial pallium (*18*) and their similar mediolateral distribution, connections (*22*), biophysical properties (*20*), and molecular identities (this study) support the hypothesis that mammalian-like hippocampal regions were already present as adjacent fields in the ancestor of all amniotes. The architecture of the mammalian hippocampus, with its signature infoldings, may thus result from the considerable enlargement of neocortex in the mammalian lineage.

Transcriptomic signatures of neocortical upper and deep layers in turtle dorsal cortex

Mammalian six-layered neocortex evolved from a presumed three-layered ancestral cortex in a stem amniote. What is the relationship between the neurons and layers in the mammalian neocortex and the reptilian cortex? Earlier studies suggested that L2/3 and L4 ("upper layer," UL) neurons are mammalian novelties; reptilian pyramidal neurons have thus been compared to the mammalian early born L5 and L6 ("deep layer," DL) neurons (*23*, *24*). By contrast, connectivity implies that reptilian cortex should harbor both L4 (that is, thalamorecipient) and L5 (cortico-fugal) types, and it has been suggested that these types, rather than occupying different layers, populate adjacent fields in turtle cortex: one in anterior dorsal cortex (*RORB*$^+$ L4-like neurons) and one in posterior dorsal cortex (*ETV1*$^+$ L5-like neurons) (*5*, *10*).

Our data indicate that anterior dorsal cortex is the only region comparable to mammalian neocortex (Fig. 2). We examined the expression of prototypical mammalian-layer markers (*7*, *25*, *26*) in the six turtle anterior dorsal cortex glutamatergic clusters (e07, e08, and e13 to e16). These cells expressed several mammalian UL and DL markers, but these genes were often coexpressed in the same clusters (Fig. 4A and fig. S13A). When we focused on the transcription factors that establish and maintain cell identity in the neocortex, we observed that, in the turtle, these genes were expressed in combinations that were never observed in differentiated mammalian cortical neurons. For example, all turtle anterior dorsal cortex cell types coexpress genes enriched in mammalian L2/3, L4, and L5a intratelencephalic neurons, including *SATB2*, *RORB*, and *RFX3*, as well as genes specifying L5b and L6 corticofugal projection neurons, such as *BCL11B* (*CTIP2*), *TBR1*, and *SOX5* (all clusters except e13) (*25*) (Fig. 4, A and B). In mouse neocortex, some of these genes are known to repress each other's expression in postmitotic cells (for example, *Satb2* and *Bcl11b*; *Tbr1* and *Bcl11b*) (*26*).

Extending the comparative analysis to all differentially expressed genes revealed that anterior dorsal cortex cell types correlated with either mammalian UL (e13 to e16) or DL (e07 and e08) neurons, independent of the neocortical area used for comparison (Fig. 4C and fig. S14, A to C). This trend was confirmed by gene network analysis (fig. S13, B and C). By contrast, anterior DVR clusters could not be grouped into UL- and DL-like classes (fig. S14, D to F).

As anticipated by the sequencing data, in situ hybridizations (ISHs) showed coexpression of mammalian UL and DL transcription factors throughout the turtle anterior dorsal cortex (Fig. 4B). Individual UL-like neuronal types (e13 to e16) were differentially distributed along the mediolateral and rostrocaudal axes of the anterior dorsal cortex, matching known anatomical subdivisions (fig. S15). By contrast, the DL-like cells e07 and e08, identified by parathyroid hormone–like hormone (*PTHLH*) expression, appeared interspersed throughout the anterior dorsal cortex. In the rostro-lateral dorsal cortex, DL-like cells were confined to the superficial part of L2 (Fig. 4D; additional markers in fig. S15, C to E). These markers thus identify two distinct sublayers in turtle L2: a superficial L2a with mostly DL-like neurons and a deeper L2b with mostly UL-like neurons. Retrograde tracing from the thalamus labels L2a cells (*27*, *28*), suggesting that these neurons, or a subset of them, are corticofugal and project to the thalamus.

In conclusion, our transcriptome-wide comparisons reveal the presence of global UL- and DL-like genetic signatures in distinct neuronal types of turtle anterior dorsal cortex but do not support, with the current resolution, homologies between turtle glutamatergic types and individual neocortical layers. In reptiles and mammals, UL and DL genetic signatures might have evolved independently from a stem amniote that lacked distinct UL- and DL-like neurons: Neurons with similar characteristics, such as input-output connectivity, would have acquired the expression of similar gene sets by convergent evolution. Alternatively, UL- and DL-like neurons may have existed already in the dorsal cortex of stem amniotes. If true, the emergence of the six layers that form mammalian neocortex would be a novelty [sensu (*29*)], with the evolution of new pyramidal cell types through extensive modifications of the genetic programs specifying ancestral UL- and DL-like types.

Conservation of GABAergic interneuron classes across amniotes

GABAergic interneurons in mammalian neocortex are diverse and participate in different circuit motifs and computations (*30*). Little is known about cortical interneurons outside of mammals. In several vertebrates, including reptiles, GABAergic interneurons are generated in conserved subpallial regions—the medial, caudal, and lateral ganglionic eminences (MGE, CGE, and LGE)—and migrate to the pallium (*31*, *32*). Reptiles, however, are thought to lack some mammalian GABAergic types, such as cortical vasoactive intestinal peptide (VIP) interneurons (*24*). We examined the turtle GABAergic clusters (i01 to i18) and the expression of transcription factors known to define mammalian GABAergic types.

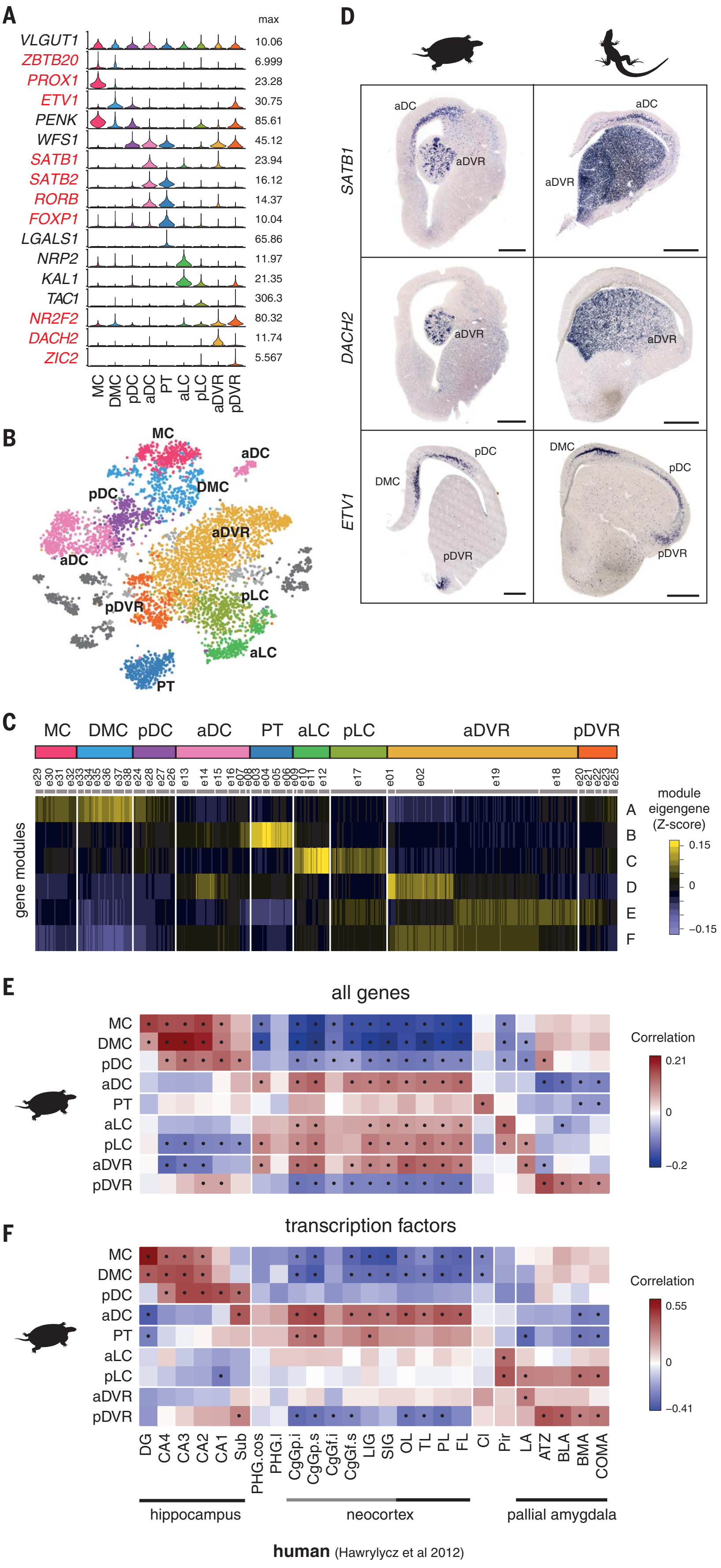


Fig. 2. Reptilian pallial regions. (**A**) Violin plots showing expression of genes discriminating between spatially segregated glutamatergic neurons in the turtle pallium. (**B**) tSNE showing turtle glutamatergic neurons colored by supercluster. Transcription factor names are in red. (**C**) Heatmap showing expression of module eigengenes calculated from turtle glutamatergic neurons. Pseudocells (averages of 4 to 5 cells used for WGCNA, see methods) are shown in columns, ordered by cluster and supercluster. (**D**) Expression of transcription factors that label corresponding pallial regions in turtle and lizard (ISH, frontal sections at different anterior-posterior levels; see also figs. S8 and S9). Scale bars, 1 mm. (**E** and **F**) Pairwise correlations of turtle glutamatergic superclusters and human pallial regions, calculated from all genes (E) or transcription factors (F) differentially expressed in turtle or human. In mammals, only some parts of the cingulate gyrus (CgG), long insular gyrus (LIG), and short insular gyrus (SIG) are classified as neocortex [see (*13*) for human data]. Dots indicate statistically significant correlations. MC, medial cortex; DMC, dorsomedial cortex; pDC and aDC, posterior and anterior dorsal cortex; PT, pallial thickening; aLC and pLC, anterior and posterior lateral cortex; aDVR and pDVR, anterior and posterior dorsal ventricular ridge; DG, dentate gyrus; CA, cornu ammonis; Sub, subiculum; PHG, parahippocampal gyrus; OL, occipital lobe; TL, temporal lobe; PL, parietal lobe; FL, frontal lobe; Cl, claustrum; Pir, piriform cortex; LA, lateral amygdala; ATZ, amygdalohippocampal transition zone; BLA, basolateral amygdala; BMA, basomedial amygdala; COMA, corticomedial amygdala.

From the combinatorial expression of these genes, we identified putative MGE-derived (i07 to i13: *LHX6*$^+$, *SOX6*$^+$, and *SATB1*$^+$), CGE-derived (i14 to i18: *NPAS1*$^+$, *SP8*$^+$, *NR2E1*$^+$, and *PROX1*$^+$), and LGE-derived (i01 and i04 to i06: *MEIS2*$^+$ *ZIC1*$^-$ and in subsets *TSHZ1*$^+$ *SIX3*$^+$ and/or *PBX3*$^+$) clusters (Fig. 5A). Corresponding MGE-, CGE-, and LGE-derived neurons could also be identified in lizard (fig. S16, A to C). The remaining GABAergic clusters correspond to cells dissected from the neighboring septum (i02 and i03) and striatum (i04). Stainings revealed the presence of LGE- and septum-derived interneurons in the amygdala (i04 and i05) and olfactory bulb (i01 and i02), indicating that LGE- and septum-derived GABAergic neurons have similar molecular identities and migratory trajectories in reptiles and mammals (Fig. 5A and fig. S16, D and E).

Turtle MGE- and CGE-derived interneurons could be further subdivided into somatostatin ("SST"), parvalbumin-like ("PV-like"), "HTR3A Reln," and "HTR3A VIP-like" classes, matching the classification of mammalian cortical GABAergic interneurons (*7*, *30*) (Fig. 5A). Pairwise cluster correlations (Fig. 5B) and gene network analysis (fig. S17) further supported the conclusion that PV-like, SST, and HTR3A interneuron classes are conserved in reptiles and mammals. [Note that transcriptomics identifies VIP- and PV-like interneurons even though *VIP* and *PVALB* are not reliably expressed in these cells (Fig. 5A).]

In turtle and lizard, markers of MGE- and CGE-derived interneurons were expressed in cells scattered throughout the pallium, including the dorsal cortex (Fig. 5C). As in mammals, where MGE- and CGE-derived interneurons have different densities across cortical layers, neurons expressing *ADARB2* (HTR3A) and *SST* (SST) were found preferentially in superficial and deep dorsal cortex, respectively; *NDNF* (HTR3A Reln) was expressed in rare dorsal cortex subpial cells, reminiscent of mammalian L1 neurogliaform cells (Fig. 5, D and E). Because the same classes of cortical GABAergic neurons exist in mammals and reptiles, they likely existed in their amniote ancestor. Our analysis did not detect similarities at a finer level: For example, we found no turtle

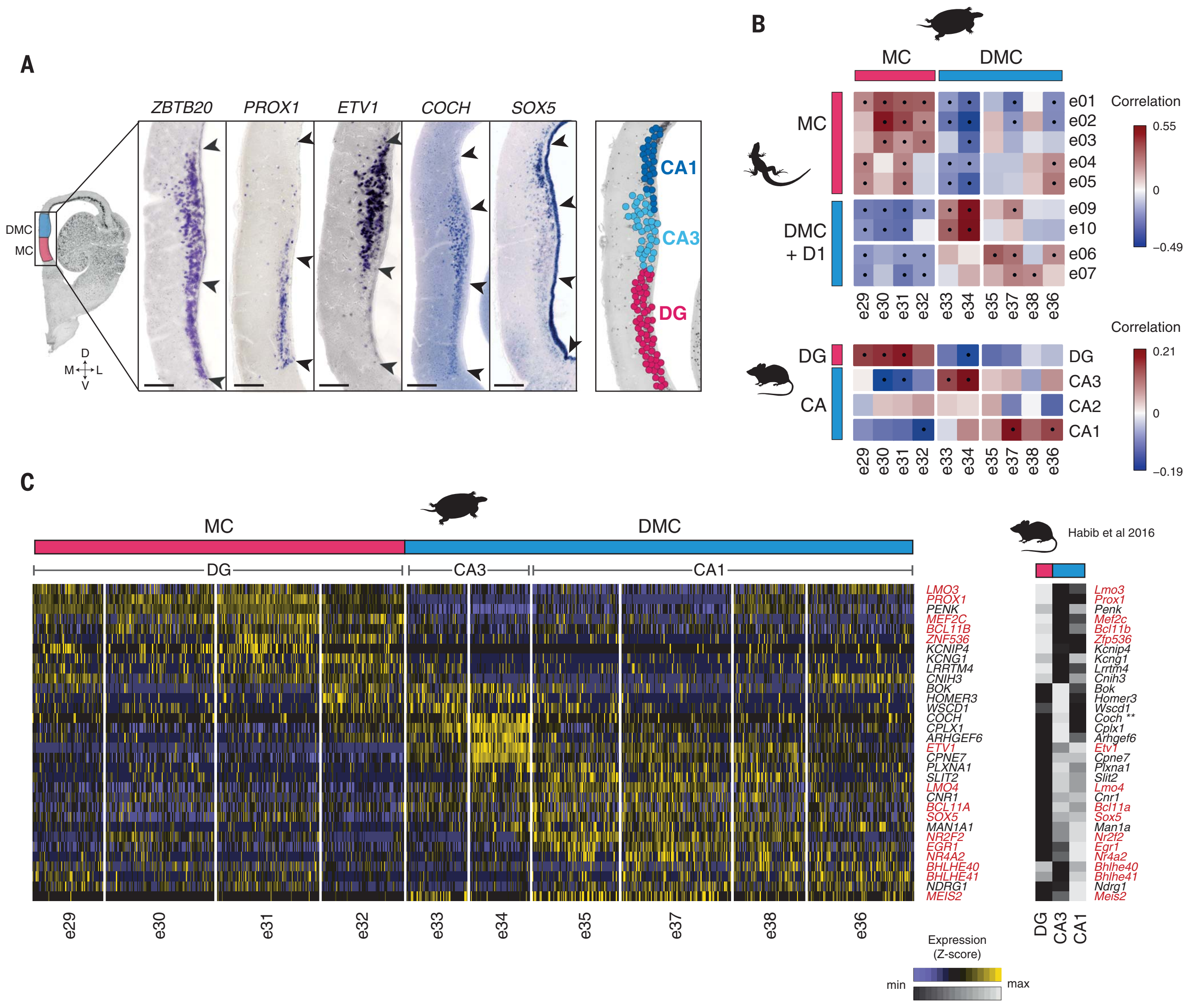


Fig. 3. Molecular architecture of the reptilian hippocampus. (**A**) Expression of hippocampal markers in turtle medial (MC) and dorsomedial (DMC) cortices. Arrowheads indicate boundaries between DG, CA3, and CA1 (compare to schematic at right). D, dorsal; V, ventral; M, medial; L, lateral. (**B**) Pairwise correlations of turtle and lizard (top) and turtle and mouse (bottom) hippocampal clusters. Dots indicate statistically significant correlations. In lizard, the *ZBTB20*$^+$ *ETV1*$^+$ medial cortices are termed DMC and D1 in the classical literature but correspond to turtle DMC. (**C**) On the left, expression of mammalian hippocampal markers in turtle single cells (columns, arranged by cluster) and, on the right, expression of the same markers in mouse [data from (*19*)]. Transcription factor names are in red. Scale bars, 200 μm.

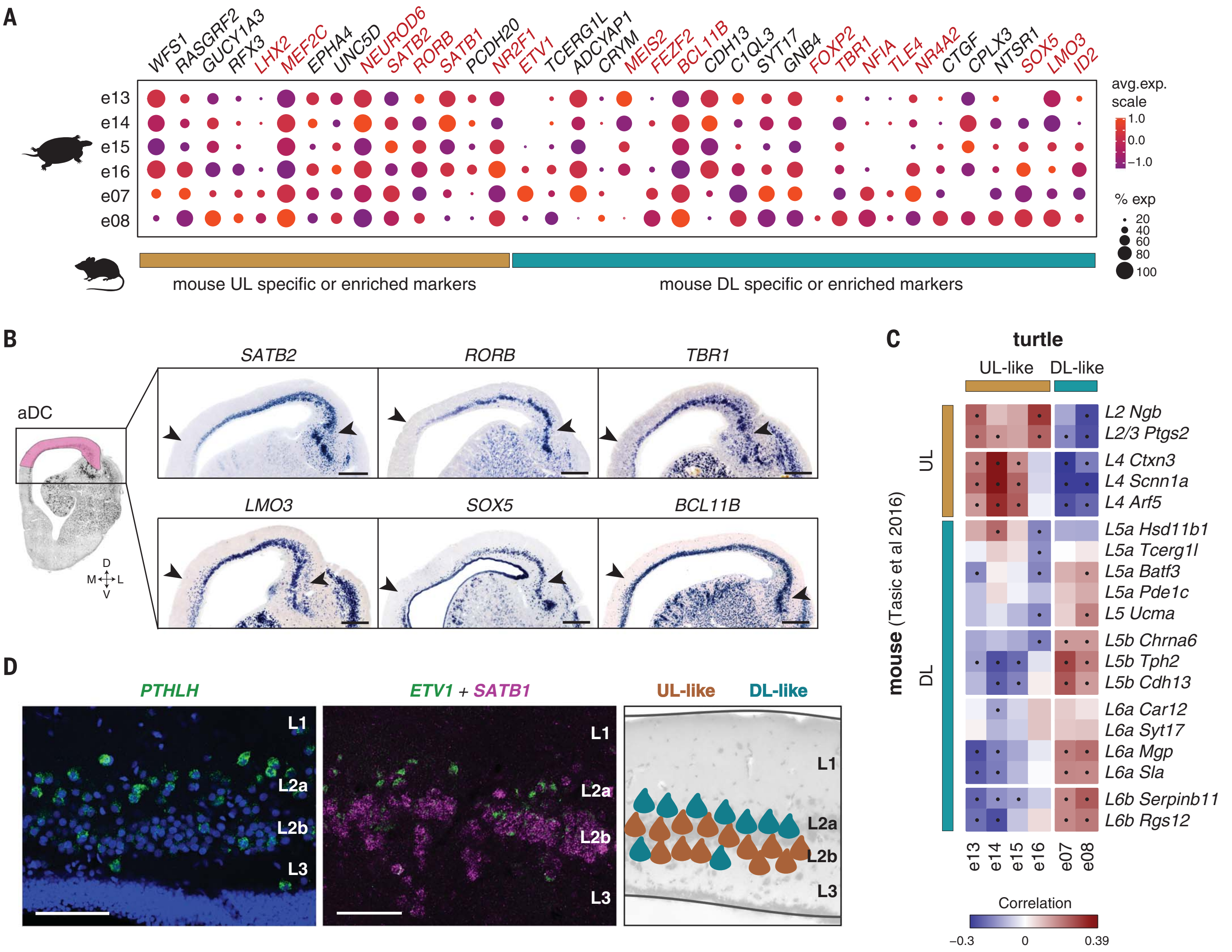


Fig. 4. UL- and DL-like neurons in turtle dorsal cortex. (**A**) Dot-plot showing expression of canonical mammalian layer markers (columns) in turtle anterior dorsal cortex (aDC) clusters (rows). Dot size indicates the percentage of cells in the cluster where the gene is detected; color indicates expression (exp) level. Names of transcription factors are in red. (**B**) ISHs showing expression of mammalian UL and DL transcription factors throughout the turtle aDC. Arrowheads indicate the medial and lateral boundaries of anterior dorsal cortex. Scale bar, 500 μm. (**C**) Pairwise correlations of turtle aDC (columns) and mouse neocortex (rows) glutamatergic types. Dots indicate statistically significant correlations. Mouse clusters are from (*7*). (**D**) Left panel, turtle DL-like cells, labeled by *PTHLH* (green), are preferentially found in superficial L2 [blue, 4′, 6-diamidino-2-phenylindole (DAPI)]. Middle panel, double ISH for *SATB1* (enriched in UL-like cells, magenta) and *ETV1* (e07 DL-like cells, green), showing the relative positions of these cells. Right panel, schematic summarizing the distribution of UL- and DL-like glutamatergic neurons (inverse arrangement to that observed in mammals). See also fig. S15. Scale bar, 100 μm.

clusters corresponding to mammalian Martinotti or basket cells. This suggests that species-specific subtypes of interneurons diversified independently in mammals and reptiles from ancestral PV-like, SST, and HTR3A (possibly VIP-like and Reln) cell classes.

Discussion

Our single-cell data provide molecular support to the existence of conserved regions and cell types in the amniote pallium. Homologs of the mammalian neocortex, "core" hippocampus (with dentate and CA fields), claustrum, and pallial amygdala probably existed already in stem amniotes. Glutamatergic neurons with an UL- or DL-like genetic signature exist in turtle anterior dorsal cortex, but one-to-one homologies of these cell types and individual mammalian layer types are not supported by our data. By contrast, cortical GABAergic interneuron classes (SST, PV-like, and HTR3A) are ancestral in amniotes.

Our analysis indicates that mammalian and reptilian brains diversified by expansion and independent evolution of different pallial territories and that these expansions coincided with the evolution of new neuronal types. The large reptilian anterior DVR (ventral pallium), for example, consists of spatially segregated neuronal types specialized in processing visual, auditory, or somatosensory stimuli (*17*). The same sensory pathways also reach the ventral pallium of mammals (for example, the lateral amygdala) but are not processed by dedicated neuronal populations there (*12*). This suggests that the elaboration of DVR neurons and circuits is specific to reptiles and birds.

The situation is reversed with the dorsal pallium. Mammalian neocortex emerged by expansion of a small dorsal pallial territory, homologous to the anterior dorsal cortex of reptiles and to the avian Wulst (*1*). This may have coincided with the generation of new glutamatergic cell types from the diversification of UL- and DL-like neurons. The sequential generation of neurons with distinct identities is a conserved feature of mammalian and reptilian (*33*) dorsal pallium progenitors, and corticothalamic neurons are early born neurons in mouse and turtle (*28*).

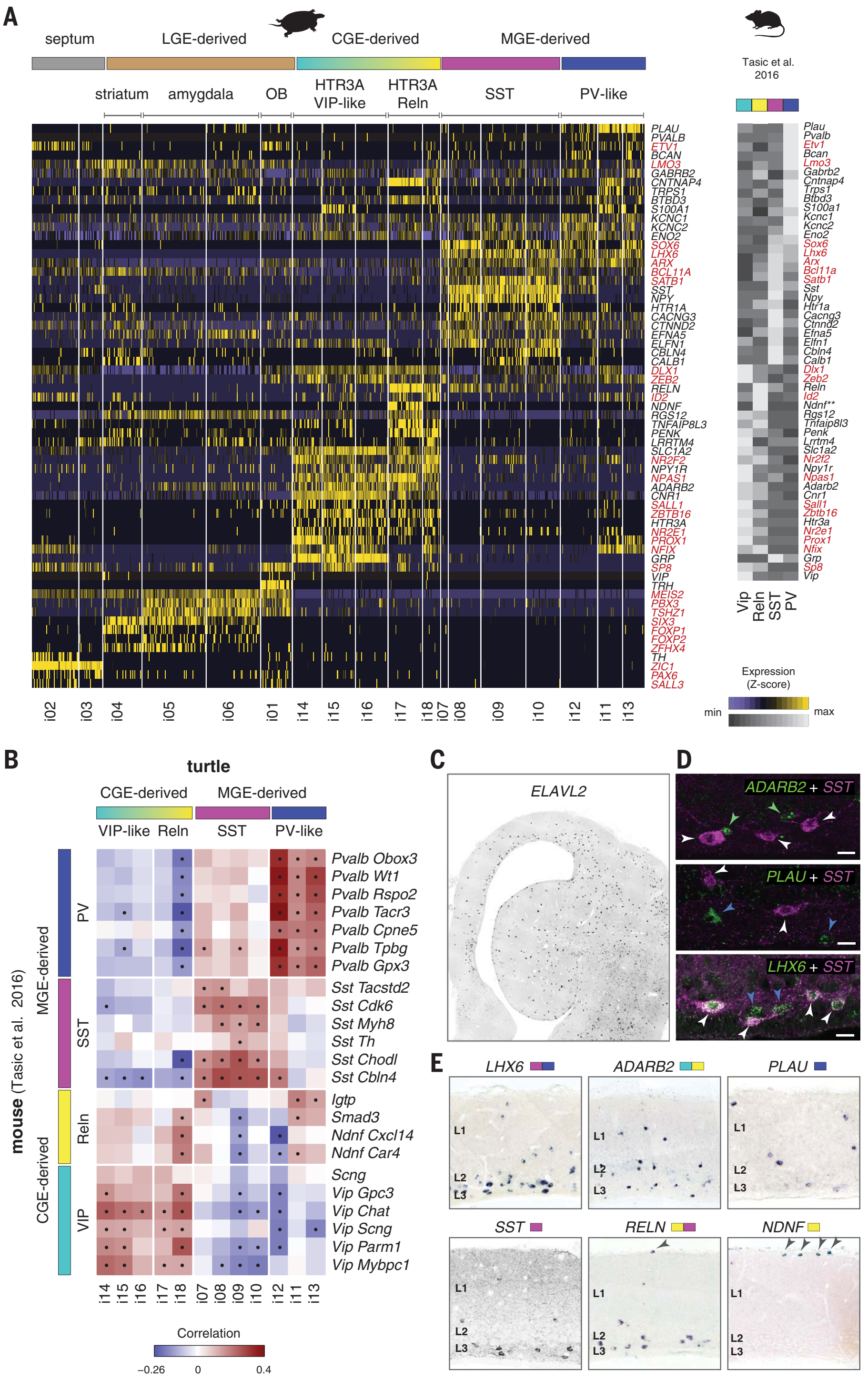


Fig. 5. Turtle GABAergic interneurons. (**A**) On left, heatmap showing expression of selected mammalian GABAergic markers (rows) in turtle single cells, grouped by cluster (column), and, on right, expression of the same markers in mouse cortical GABAergic cells [data from (*7*)]. Names of transcription factors are in red. OB, olfactory bulb. (**B**) Pairwise correlations of turtle (columns) and mouse (rows) MGE- and CGE-derived GABAergic clusters. Dots indicate statistically significant correlations. (**C**) GABAergic neurons, labeled by *ELAVL2*, in turtle. (**D**) Double labeling of *SST* (immunohistochemistry, magenta) and *ADARB2*, *PLAU*, or *LHX6* (ISH, green) in turtle dorsal cortex. *LHX6* labels both SST^+ (SST) and SST^- (PV-like) neurons. White arrowheads indicate SST neurons (SST^+ $LHX6^+$); blue arrowheads indicate PV-like neurons ($PLAU^+$ $LHX6^+$ SST^-); green arrowheads indicate HTR3A neurons ($ADARB2^+$). Scale bars, 20 μm. (**E**) ISHs of GABAergic markers in turtle dorsal cortex.

Neuronal birth order is the same in turtle and mammals (DL first, UL late) (*28*), and the superficial position of turtle DL-like neurons is consistent with the inversion of corticogenesis (outside-in in reptiles, inside-out in mammals) (*28*, *34*). Finally, in mammals, transcription factors specifying UL and DL fate are coexpressed in progenitors and acquire mutually exclusive expression only after cell-cycle exit (*35*). The coexpression of mammalian UL and DL fate specifiers in turtle neurons suggests that cortical layers may have evolved from the remodeling of regulatory interactions between these transcription factors, possibly through new repressive interactions. The temporal extension of neurogenesis (*33*) could have enabled the segregation of originally overlapping gene expression programs and thus the diversification of ancestral UL and DL types.

These molecular maps of turtle and lizard pallium provide a framework for future functional studies. How do similarities and differences in gene expression programs inform us about the evolution of brain function? Much of reptilian cortex, for example, is molecularly related to core mammalian hippocampus. Lesion experiments in reptiles suggest a role for medial cortex in navigation (*22*), but no chronic electrophysiological data have, as of yet, been obtained from these regions. Integrating our molecular maps with tracing, electron microscopy, and electrophysiological studies is needed to understand whether and how gene expression programs are linked to the evolution of connectivity and function.

Common principles of synaptic organization can be extrapolated from the comparison of three- and six-layered cortices (*36*). Are these similarities coincidental or do they reflect ancestral cortical circuits? In mammals, local circuit motifs involve GABAergic types with distinct molecular and functional features (*30*). Our study indicates that interneuron diversity is ancestral to both mammals and reptiles. The conservation of interneuron classes over at least 320 million years suggests not only that there is selective pressure for interneuron diversity but also that interneuron diversity itself may have evolved in stem vertebrates within primordial cortical circuits (*37*). Alternatively, the conservation of interneuron classes may result simply from developmental constraints on subpallium patterning. The combined molecular and functional investigation of circuit motifs in reptilian cortex may shed light on the ancestral design and computational logic of vertebrate cortices.

REFERENCES AND NOTES

1. L. Puelles *et al.*, in *Evolution of Nervous Systems*, J. H. Kaas, Ed. (Elsevier, ed. 2, 2017), vol. 1, pp. 519–555.
2. G. F. Striedter, *Brain Behav. Evol.* **49**, 179–194 (1997).
3. G. F. Striedter, *J. Comp. Neurol.* **524**, 496–517 (2016).
4. T. G. Belgard *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **110**, 13150–13155 (2013).
5. J. Dugas-Ford, J. J. Rowell, C. W. Ragsdale, *Proc. Natl. Acad. Sci. U.S.A.* **109**, 16974–16979 (2012).
6. A. Zeisel *et al.*, *Science* **347**, 1138–1142 (2015).
7. B. Tasic *et al.*, *Nat. Neurosci.* **19**, 335–346 (2016).

8. E. Z. Macosko *et al*., *Cell* **161**, 1202–1214 (2015).
9. G. La Manno *et al*., *Cell* **167**, 566–580.e19 (2016).
10. H. J. Karten, *Curr. Biol.* **23**, R12–R15 (2013).
11. A. B. Butler, Z. Molnár, *Brain Res. Bull.* **57**, 475–479 (2002).
12. L. L. Bruce, T. J. Neary, *Brain Behav. Evol.* **46**, 224–234 (1995).
13. M. Hawrylycz *et al*., *Nat. Neurosci.* **18**, 1832–1844 (2015).
14. P. V. Hoogland, E. Vermeulen-Vanderzee, *J. Comp. Neurol.* **285**, 289–303 (1989).
15. H. Yoneshima *et al*., *Neuroscience* **137**, 401–412 (2006).
16. N. Moreno, A. González, *J. Anat.* **211**, 151–163 (2007).
17. P. R. Manger, D. A. Slutsky, Z. Molnár, *J. Comp. Neurol.* **453**, 226–246 (2002).
18. L. Medina, A. Abellán, E. Desfilis, *Brain. Behav. Evol.* **90**, 25–40 (2017).
19. N. Habib *et al*., *Science* **353**, 925–928 (2016).
20. J. M. Shen, A. R. Kriegstein, *J. Neurophysiol.* **56**, 1626–1649 (1986).
21. M. S. Cembrowski, L. Wang, K. Sugino, B. C. Shields, N. Spruston, *eLife* **5**, e14997 (2016).
22. S. Reiter, H.-P. Liaw, T. M. Yamawaki, R. K. Naumann, G. Laurent, *Brain. Behav. Evol.* **90**, 41–52 (2017).
23. F. F. Ebner, in *Evolution of Brain and Behavior in Vertebrates*, R. B. Masterton, M. E. Bitterman, C. B. G. Campbell, N. Hotton, Eds. (Lawrence Erlbaum Associates, Hillsdale, NJ, 1976), pp. 115–167.
24. A. Reiner, *Comp. Biochem. Physiol. Comp. Physiol.* **104**, 735–748 (1993).
25. B. J. Molyneaux, P. Arlotta, J. R. L. Menezes, J. D. Macklis, *Nat. Rev. Neurosci.* **8**, 427–437 (2007).
26. L. T. Gray *et al*., *eLife* **6**, 1–30 (2017).
27. P. S. Ulinski, *J. Comp. Neurol.* **254**, 529–542 (1986).
28. M. G. Blanton, A. R. Kriegstein, *J. Comp. Neurol.* **310**, 558–570 (1991).
29. D. Arendt *et al*., *Nat. Rev. Genet.* **17**, 744–757 (2016).
30. R. Tremblay, S. Lee, B. Rudy, *Neuron* **91**, 260–292 (2016).
31. C. Métin *et al*., *Development* **134**, 2815–2827 (2007).
32. O. Marín, J. L. R. Rubenstein, *Nat. Rev. Neurosci.* **2**, 780–790 (2001).
33. T. Nomura, H. Gotoh, K. Ono, *Nat. Commun.* **4**, 2206 (2013).
34. A. M. Goffinet, C. Daumerie, B. Langerwerf, C. Pieau, *J. Comp. Neurol.* **243**, 106–116 (1986).
35. S. Lodato, P. Arlotta, *Annu. Rev. Cell Dev. Biol.* **31**, 699–720 (2015).
36. G. M. Shepherd, *Front. Neuroanat.* **5**, 30 (2011).
37. S. M. Suryanarayana, B. Robertson, P. Wallén, S. Grillner, *Curr. Biol.* **27**, 3264–3277.e5 (2017).

ACKNOWLEDGMENTS

The authors are grateful to V. Benes and P. Collier (EMBL GeneCore Facility, Heidelberg) for training and support on sequencing; A. Georges (University of Canberra) for prepublication access to the *Pogona* genome; C. Müller (MPI for Brain Research) for initial investigations on turtle GABAergic interneurons; M. Klinkmann, A. Arends, A. Macias Pardo, T. Manthey, C. Thum, and J. Knop for technical assistance; S. Masneri for help with scientific computing; and E. Schuman, C. Müller, S. Reiter, and the Laurent laboratory for feedback during the course of this work and on the manuscript. **Funding:** This research was funded by the Max Planck Society and the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013)/ERC grant agreement n° 322705 (G.L.). **Author contributions:** Single-cell experiments: M.A.T. and T.M.Y.; bioinformatics: M.A.T., T.M.Y., and G.T.; anatomy and histology: M.A.T., R.K.N., and A.A.J.; data analysis: M.A.T., T.M.Y., R.K.N., A.A.J., G.T., and G.L.; project management and supervision: G.L.; and manuscript writing: M.A.T. and G.L., with input from T.M.Y. and R.K.N. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** Sequencing data have been deposited in the NCBI Sequence Read Archive (BioProject PRJNA408230), and processed gene expression data can be explored at https://brain.mpg.de/research/laurent-department/software-techniques.html. The code used for analysis is available at https://github.molgen.mpg.de/MPIBR/ReptilePallium/.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/360/6391/881/suppl/DC1
Materials and Methods
Figs. S1 to S17
Tables S1 and S2
References (*38*–*55*)

6 November 2017; accepted 12 March 2018
Published online 3 May 2018
10.1126/science.aar4237

REPORT

CATALYSIS

Cobalt-catalyzed asymmetric hydrogenation of enamides enabled by single-electron reduction

Max R. Friedfeld,[1] Hongyu Zhong,[1] Rebecca T. Ruck,[2] Michael Shevlin,[2]* Paul J. Chirik[1]*

Identifying catalyst activation modes that exploit one-electron chemistry and overcome associated deactivation pathways will be transformative for developing first-row transition metal catalysts with performance equal or, ideally, superior to precious metals. Here we describe a zinc-activation method compatible with high-throughput reaction discovery that identified scores of cobalt-phosphine combinations for the asymmetric hydrogenation of functionalized alkenes. An optimized catalyst prepared from (*R*,*R*)-Ph-BPE {Ph-BPE, 1,2-bis[(2*R*,5*R*)-2,5-diphenylphospholano]ethane} and cobalt chloride [$CoCl_2 \cdot 6H_2O$] exhibited high activity and enantioselectivity in protic media and enabled the asymmetric synthesis of the epilepsy medication levetiracetam at 200-gram scale with 0.08 mole % catalyst loading. Stoichiometric studies established that the cobalt (II) catalyst precursor (*R*,*R*)-Ph-BPECoCl$_2$ underwent ligand displacement by methanol, and zinc promoted facile one-electron reduction to cobalt (I), which more stably bound the phosphine.

Asymmetric catalysis with soluble metal complexes has transformed the preparation of single enantiomers in the pharmaceutical, fragrance, and fine-chemical industries (*1*, *2*). Because different antipodes of chiral molecules often exhibit distinct biological properties, the U.S. Food and Drug Administration has strict requirements for single-enantiomer drugs, and the importance of asymmetric transformations in the pharmaceutical industry will continue to grow. Beginning with Knowles's synthesis of the Parkinson's medication L-dopa by rhodium-catalyzed asymmetric alkene hydrogenation (*3*), catalysis by homogeneous catalysts containing precious metals with tunable ligands has revolutionized the approach to single-enantiomer active pharmaceutical ingredients (APIs).

The widespread application of asymmetric hydrogenation, particularly in the pharmaceutical industry, has motivated efforts to identify catalysts based on earth-abundant first-row transition metals rather than traditionally used precious metals (*4*). In alkene hydrogenation catalysis, rhodium and iridium catalysts operate by predictable, two-electron cycles involving oxidative addition and reductive elimination [for example, M(I)-M(III)]. However, compared to their heavier congeners, first-row transition metals have kinetically and thermodynamically accessible oxidation states separated by one electron, often to the detriment of catalytic chemistry (*5*). Although considerable progress has been made (*6*, *7*), state-of-the-art catalysts with iron, cobalt, and nickel lack many of the favorable properties associated with precious metal catalysts that facilitate scale up. Alkene hydrogenation catalysts with earth-abundant metals are typically air- and moisture-sensitive, requiring rigorously dried solvents; are intolerant of many polar functional groups found in APIs; and have insufficient activity to be applied industrially.

Asymmetric hydrogenation and other enantioselective metal-catalyzed reactions often rely on the successful relay of stereochemical information from a chiral ligand to the substrate (*8*). Therefore, understanding and controlling ligand coordination and dissociation equilibria are key to enabling catalyst stability and communicating stereochemical information. Unlike other tactics for improving catalyst performance, rational control of catalyst activation and deactivation are challenging and often circumvented by increased catalyst loadings (*9*–*11*). With classic transition metal catalysts such as Wilkinson's $(Ph_3P)_3RhCl$ (*12*) and $(Ph_3P)_3RuCl_2$ (*13*) (Ph, phenyl), catalyst activation and, ultimately, performance is limited by phosphine dissociation equilibria and halide coordination (Fig. 1, A and C). These limitations are overcome with weakly coordinating anions and hydrogenation of ancillary diene or triene ligands to create open coordination sites (Fig. 1, B and D), as exemplified by the Schrock-Osborn–type catalysts [(P,P)Rh(diene)][X] (*14*), where X is any noncoordinating anion, and the cationic ruthenium catalyst [(P-P)Ru(H)(triene)][BF_4] (*15*). These design principles have proven useful for fast, reliable catalyst activation and are widely used for new catalyst discovery.

Our laboratory has reported that two-carbon-bridged, C_2-symmetric (bis)phosphines support highly active and enantioselective cobalt catalysts for the asymmetric hydrogenation of simple dehydro-α–amino acid derivatives (*6*). Although these catalysts are state-of-the-art among first-row metals and provide an important demonstration of the promise of earth-abundant metals in asymmetric alkene hydrogenation, major limitations include the use of pyrophoric activators such as $LiCH_2SiMe_3$, extreme air sensitivity of the catalyst, and lack of reactivity among many classes of phosphines, likely owing to catalyst deactivation by irreversible loss of ligand (Fig. 1E). Isolated organometallic compounds such as (*R*,*R*)-QuinoxP*Co(CH_2SiMe_3)$_2$ and (*R*,*R*)-^{i}Pr-DuPhosCo (CH_2SiMe_3)$_2$ (^{i}Pr, isopropyl), albeit more active, require multistep organometallic syntheses and special handling that are likely impractical for industrial application. Here we describe advances in cobalt-catalyzed asymmetric alkene hydrogenation, where mechanistic insights into ligand dissociation equilibria and the unique properties of the first-row transition metals are leveraged to address fundamental limitations of existing catalysts. Two sequential single-electron reductions of substitutionally labile Co(II) complexes result in formation of more robust catalysts in situ. This advance in catalyst activation enabled the discovery of scores of effective metal-ligand combinations for asymmetric hydrogenation, culminating in the use of low catalyst loadings for the practical, pilot-scale synthesis of an API.

The hydrogenation of dehydro-levetiracetam (**1**) was selected for initial catalyst development studies to highlight the challenges associated with API synthesis. The corresponding chiral product, levetiracetam (Keppra), is a widely used medication for epilepsy (*16*). In one patented route (*17*), levetiracetam was prepared by asymmetric hydrogenation using an optimized condition of 0.5 mole % (mol %) of in situ–generated [(*S*,*S*)-Et-DuPhosRh(COD)][OTf] in dichloromethane (Et, ethyl; COD, 1,5-cyclooctadiene; OTf, triflate). The relatively high catalyst loading and use of a noncoordinating, chlorinated solvent reflects the challenges associated with hydrogenation of **1** as a poorly coordinating substrate (*17*) with limited conformational flexibility for achieving two-point binding (*18*).

Initial studies on hydrogenation of **1** relied on high-throughput experimentation to evaluate reaction variables, including solvents, cobalt sources, activators, temperature, and catalyst loadings (tables S1 to S6). A remarkable solvent dependence was identified: Protic solvents such as methanol (MeOH), ethanol, and trifluoroethanol provided the highest yields and enantiomeric excesses (tables S1 to S4). These unexpected results indicated that cobalt hydrides competent for enantioselective alkene hydrogenation could, despite their anticipated hydricity, be formed in protic solvents. The use of alcohol solvents has been prevalent since the discovery of enantioselective

---

[1] Department of Chemistry, Princeton University, Princeton, NJ 08544, USA. [2] Department of Process Research and Development, Merck Research Laboratories, Rahway, NJ 07065, USA.
*Corresponding author. Email: michael_shevlin@merck.com (M.S.); pchirik@princeton.edu (P.J.C.)

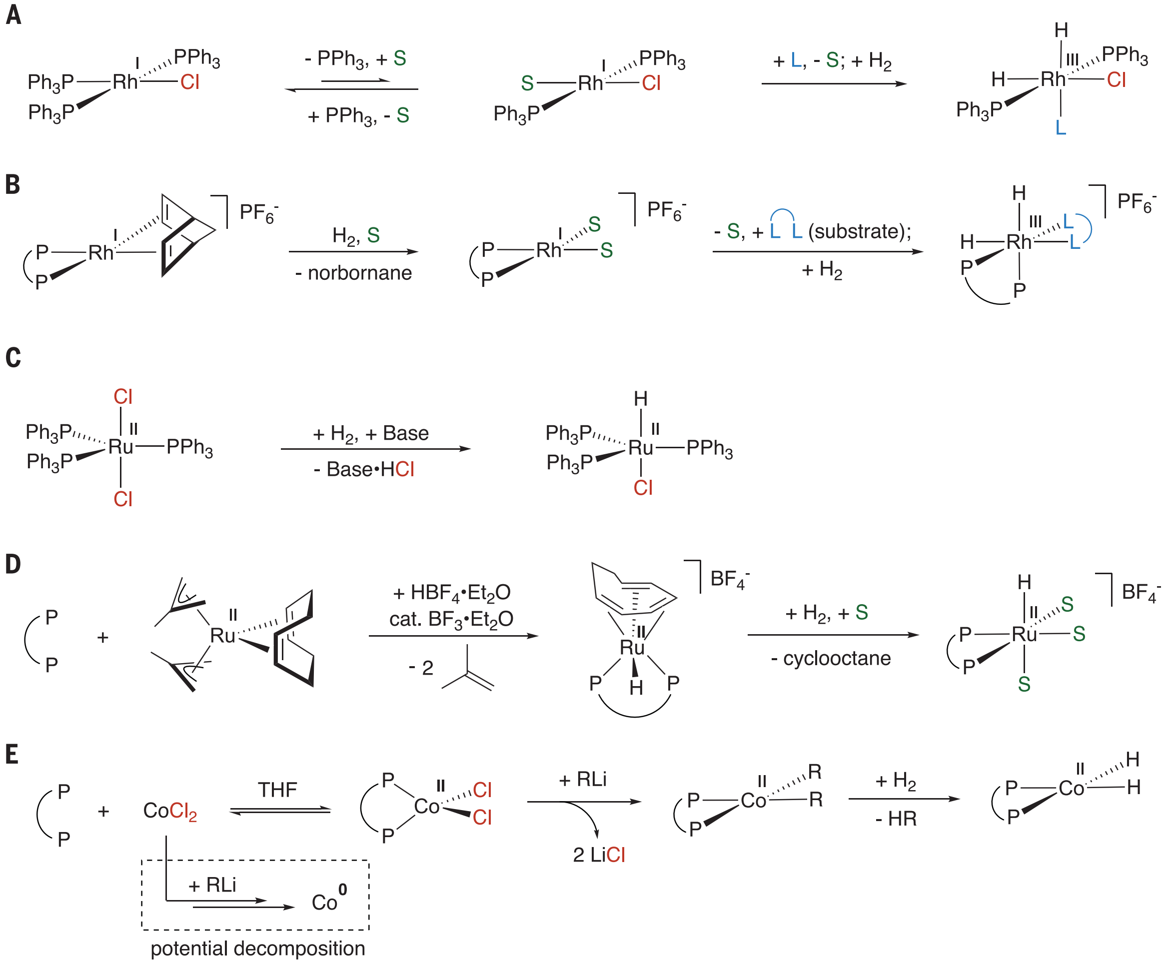


Fig. 1. Catalyst-activation strategies for alkene hydrogenation catalysts. (**A**) Activation of Wilkinson's catalyst $(Ph_3P)_3RhCl$ is limited by unfavorable PPh_3 dissociation equilibrium and strongly coordinating Cl^-. (**B**) Schrock-Osborn–type catalysts are paired with a weakly coordinating anion such as PF_6^- and rely on irreversible hydrogenation of diene ligands to open coordination sites. (**C**) Wilkinson's ruthenium catalyst is activated by base in H_2 to form ruthenium monohydride but suffers from limited coordination sites. (**D**) Cationic ruthenium catalysts are designed to open coordination sites. (**E**) Previous work on cobalt catalysts relied on activation by alkyl lithium reagents and posited formation of cobalt dihydride as an active catalyst. S, solvent molecule; L, neutral ligand; cat., catalyst.

homogeneous hydrogenation with transition metal catalysts (*19*). Catalysts containing Rh(I), Ir(I), and Ru(II) are usually highly active in MeOH, in part, because the solvent can act as a stabilizing ligand occupying open coordination sites on the metal but be readily displaceable by alkene substrates during catalysis (Fig. 1, A, B, and D). Also highly preferred in process scale reactions, MeOH is a green solvent (*20*) with excellent solubility for both H_2 and drug intermediates with polar functional groups.

The proposed mechanism of cobalt-catalyzed hydrogenation of dehydro-α–amino acid derivatives involves formation of a Co(II) dihydride as the active species that promotes alkene insertion (Fig. 1E) (*21*). We wondered whether this potential intermediate could be accessed instead from H_2 oxidative addition to Co(0), a pathway reminiscent of Rh(I) and Ir(I) catalysis. This strategy, in turn, would require a convenient means to reduce cobalt precursors from higher oxidation states.

Mild reducing agents compatible with MeOH, such as readily available and easily handled Zn, Mn, Mg, and Fe powders, were examined for in situ activation of Co(II) phosphine complexes for hydrogenation of **1**. Both high yields and enantiomeric excesses (ee's) of levetiracetam were obtained after optimization with most in situ metallic reductants (table S7), suggesting the generation of reduced Co species in MeOH, which activated H_2 and were effective for the enantioselective reduction of **1**. Filtration of the catalyst suspension to remove excess reductant and the accompanying metal salt had minimal effect on catalyst performance, which simplified the reaction setup. Among the metal powders examined, Zn dust provided the best results, with respect to both product yields and enantioselectivity.

Given the sensitivity of ligand-substrate interactions to subtle structural manipulations, the broad application of earth-abundant transition metal catalysts for asymmetric alkene hydrogenation will depend on the realization of a practical, rapid, and reproducible method for evaluating various metal-ligand combinations for the ever-changing structures of alkenes found in the drug pipeline. Accordingly, the optimized reaction conditions, $CoCl_2$ with Zn dust as an activator in MeOH, were applied to a high-throughput evaluation of an expanded library of 216 chiral bidentate ligands. Zinc activation enabled high activity and enantioselectivity across a broad range of ligands, including bis(phosphine), secondary phosphine oxide–phosphine, phosphine-oxazoline, and phosphine-thioether with varied chelate sizes (Fig. 2C and fig. S13). The number of ligands effective under Zn activation was remarkably higher than the corresponding compatibilities with the previously reported (*6*) organolithium-activation method (Fig. 2A; for Zn activation, only the 192-ligand library results are plotted). The high efficacy of the Zn-activation method was not limited to the asymmetric hydrogenation of **1**. Scores of cobalt-phosphine combinations were identified for the successful asymmetric

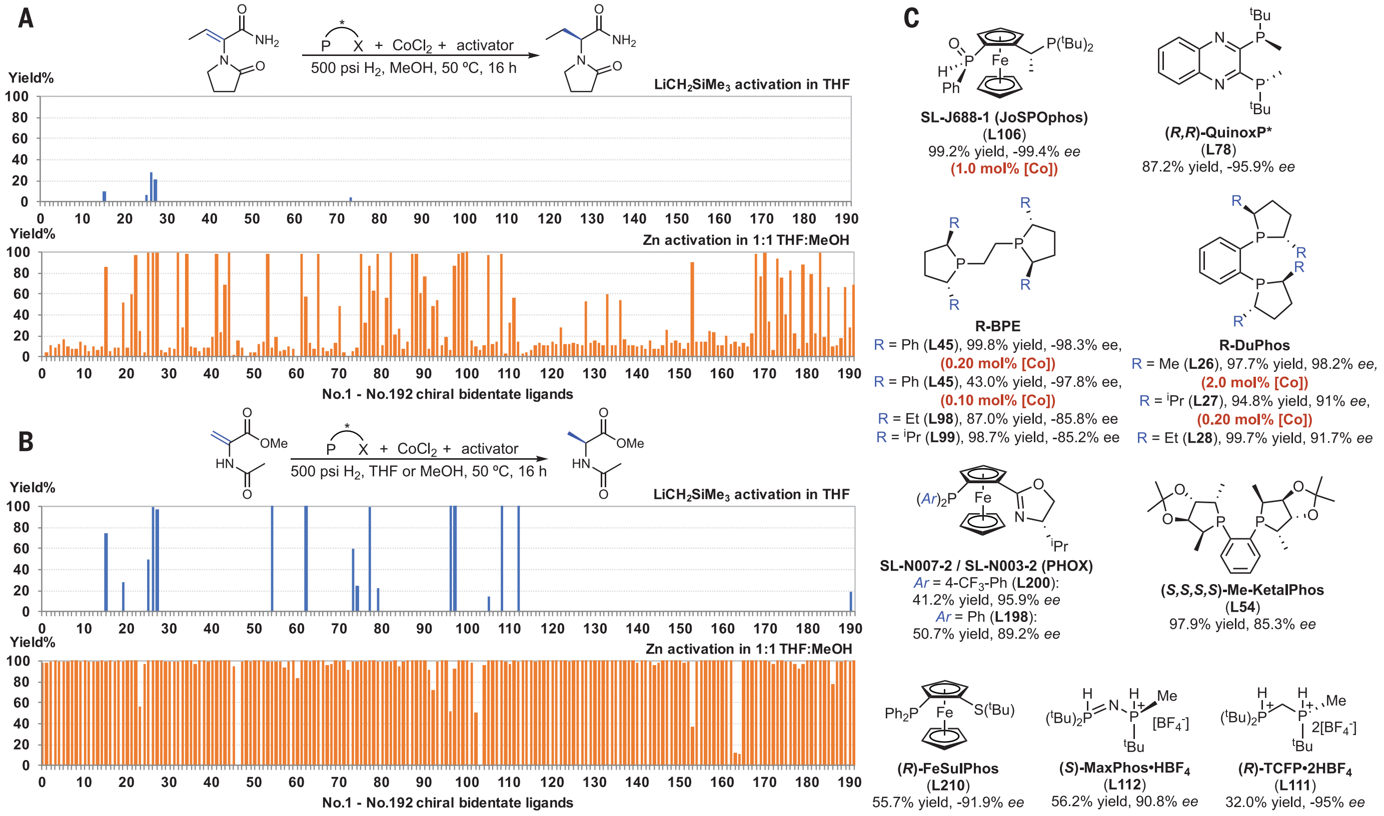


Fig. 2. High-throughput evaluation of chiral phosphine ligands. (**A**) Comparing results of 192 chiral ligands for the asymmetric hydrogenation of dehydro-levetiracetam by using $LiCH_2SiMe_3$ (top, table S9) and Zn (bottom, table S8). (**B**) Ligand compatibility with $LiCH_2SiMe_3$ (top; see also table S19) and Zn (bottom; see also table S20) for hydrogenation of a representative alkene methyl-2-acetamidoacrylate. (**C**) Ligands in the expanded 216-ligand library that give highest enantioselectivity for dehydro-levetiracetam hydrogenation. Catalyst loadings were 10 mol % unless otherwise noted. Positive and negative ee values correspond to (*S*) and (*R*) enantiomers, respectively. Me, methyl; ^{t}Bu, *tert*-buty.

hydrogenation of *N*-(1-phenylvinyl)acetamide, methyl 2-acetamidoacrylate, and methyl (*Z*)-2-acetamido-3-phenylacrylate (figs. S16 to S18 and tables S21 to S24), highlighting the improved versatility over the previously reported $LiCH_2SiMe_3$-activation method (*6*). The hydrogenation of methyl 2-acetamidoacrylate was particularly notable (Fig. 2B), as essentially all phosphines in the library produced quantitative conversion to product. The highly successful Zn-MeOH reduction is likely attributable to enhanced activation and suppressed deactivation of catalyst: Low oxidation-state Co species with open coordination sites are generated owing to the high Cl^- binding affinity of Zn^{2+}, and formation of metallic Co(0) is likely suppressed owing to favorable ligand coordination to reduced Co species. The robustness of the Zn method was also demonstrated by the successful hydrogenation of **1** in the presence of 5 volume % of water in MeOH. Among the many successful ligands identified, (*R,R*)-Ph-BPE {Ph-BPE, 1,2-bis[(2*R*,5*R*)-2,5-diphenylphospholano]ethane} and (*R,R*)-^{i}Pr-DuPhos were selected for subsequent studies because of their outstanding performance at low (0.20 mol %) catalyst loadings (Fig. 2C and table S17). Although Zn activation of bis(phosphine)-cobalt catalysts has been reported in C–C bond–forming reactions (*22*, *23*), the protocol was not applied to hydrogenation in protic solvents.

Stoichiometric experiments were conducted to determine which oxidation states of Co were formed under catalytic conditions and to understand the origin of the improved catalyst fidelity. Compound (*R,R*)-Ph-BPECoCl$_2$ (**2a**) was isolated after addition of the phosphine to a tetrahydrofuran (THF) solution of $CoCl_2 \cdot 6H_2O$ in 96% yield and crystallographically characterized (fig. S50). A high spin, S = 3/2 ground state for **2a** was determined by electron paramagnetic resonance (EPR) spectroscopy (fig. S9) and solution magnetometry. Surprisingly, dissolution of the dark purple **2a** in MeOH resulted in complete dissociation of the (bis)phosphine from the substitutionally labile Co(II) center to form a proposed solvento complex $[Co(MeOH)_6][Cl]_2$ along with precipitation of free ligand, as judged by ultraviolet-visible spectroscopy and ^{31}P nuclear magnetic resonance spectroscopy (figs. S40 to S44). Phosphine dissociation proved reversible, as removing the MeOH solvent and dissolving the residue in THF reconstituted **2a**. Cobalt dichloride, (*R,R*)-^{i}Pr-DuPhosCoCl$_2$ (**2b**), underwent partial ligand dissociation in MeOH (fig. S45). These findings suggest that unfavorable ligand dissociation from Co(II) by MeOH displacement could be a catalyst-deactivation pathway.

Further studies revealed how the Zn-MeOH method overcomes the unfavorable ligand loss for productive catalysis. Under conditions resembling those of the catalytic reaction, an equimolar mixture of (*R,R*)-Ph-BPE and $CoCl_2 \cdot 6H_2O$ was stirred with excess Zn in 1:1 THF:MeOH. A chloride-bridged Co(I) dimer [(*R,R*)-Ph-BPECo(μ-Cl)]$_2$ (**3a**) was isolated and crystallographically characterized (fig. S51) as the Zn-reduction product. The analogous ^{i}Pr-DuPhos complex, **3b**, was prepared and characterized (fig. S53) by using the same method. Both Co(I) complexes **3a** and **3b** exhibit S = 1 ground states, as judged by their paramagnetic, EPR-silent nature and solution magnetic moments. Phosphine displacement by MeOH was not observed with **3a** or **3b**, indicating that reduction of Co(II) to Co(I) reduces phosphine lability and enhances inertness to substitution by MeOH. Consistent with these observations, Zn reduction of $[Co(MeOH)_6][Cl]_2$ also yielded **3a** after addition of an equimolar amount of (*R,R*)-Ph-BPE. MeOH also facilitated reduction of **2a** by Zn, as the reaction was complete in 15 min at 23°C; an identical procedure in THF required 12 hours. The Zn and MeOH work synergistically to impart catalyst

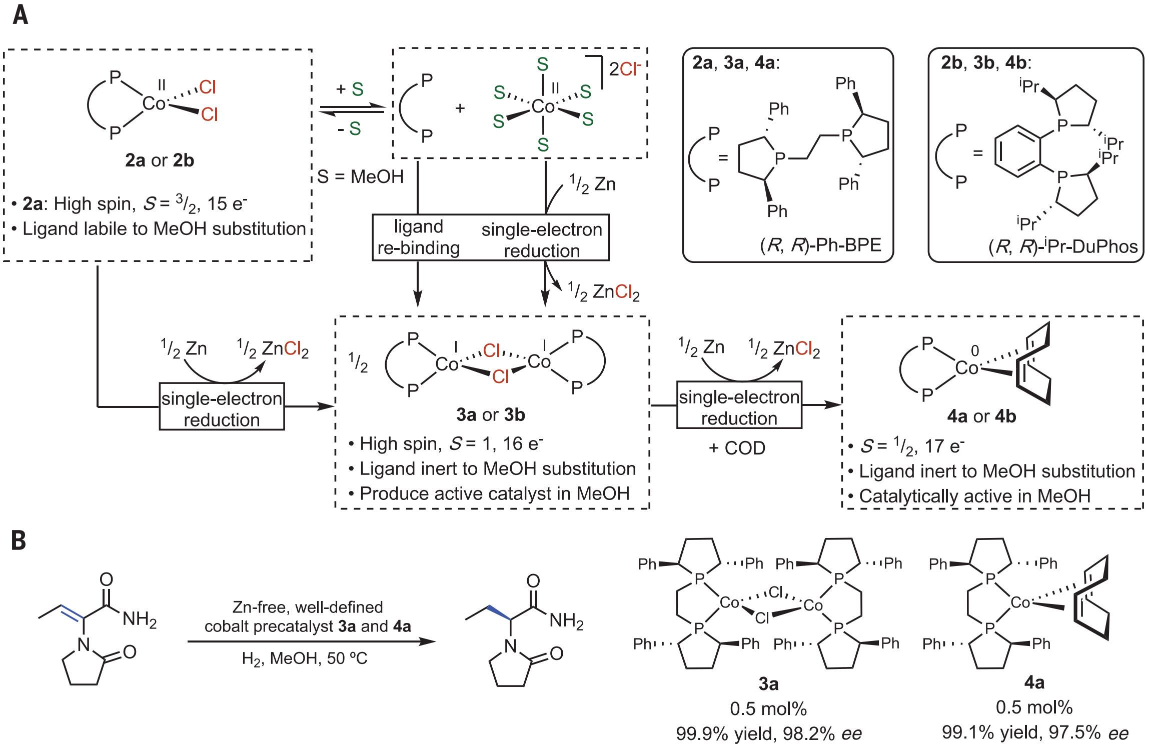


Fig. 3. Cobalt-catalyst activation enabled by single-electron reduction. (**A**) Properties and reactivity of Co(II), Co(I), and Co(0) species; (P,P)$CoCl_2$; $[(P,P)CoCl]_2$; and (P,P)CoCOD under catalytic conditions (Zn-MeOH). (**B**) Single-component, Zn-free Co(I) and Co(0) precatalysts are competent for hydrogenation.

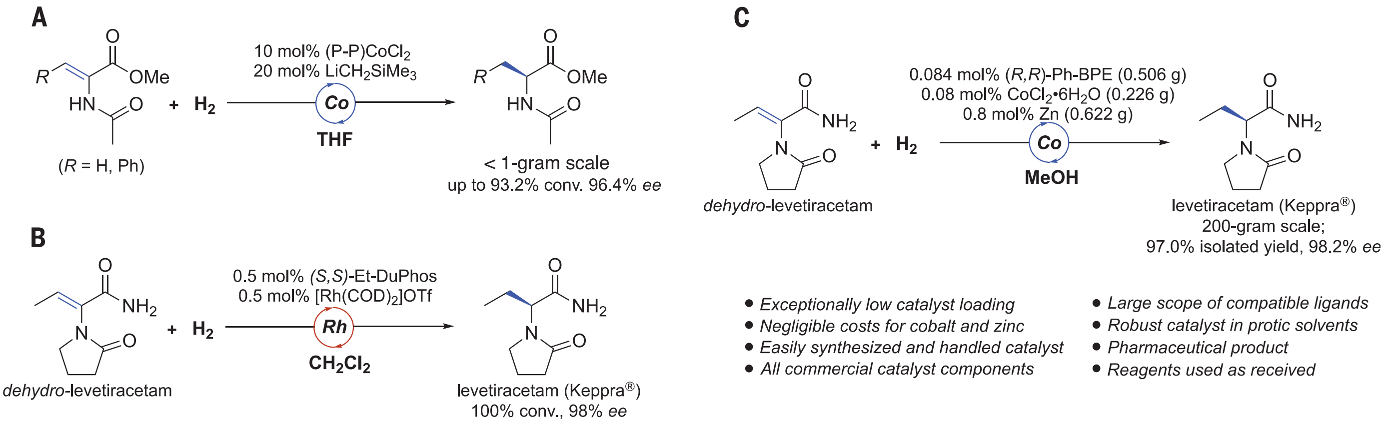


Fig. 4. Applications at large scale. (**A**) Previous work on cobalt-catalyzed asymmetric hydrogenation of dehydro-α–amino acid derivatives. (**B**) Patented route for asymmetric hydrogenation for levetiracetam synthesis by a rhodium catalyst in dichloromethane solvent. (**C**) Industrially relevant cobalt-catalyzed asymmetric hydrogenation for levetiracetam synthesis in MeOH solvent.

fidelity. Phosphine dissociation, a deleterious catalyst-deactivation pathway, is corrected by the fast Zn reduction in MeOH to regenerate and preserve the chiral environment on the metal. Abstraction of chloride by Zn also opens coordination sites on the Co for H_2 activation and alkene coordination.

Cobalt (0) complexes were also prepared by two sequential one-electron reductions upon prolonged treatment with Zn-MeOH in the presence of ancillary ligands (fig. S1). The diene and arene Co(0) derivatives, (*R*,*R*)-Ph-BPECo(COD) (**4a**) and (*R*,*R*)-Ph-BPECo(η^6-C_6H_6) (**5**) were synthesized from Zn reduction of either Co(II) complex **2a** or from Co(I) complex **3a**. Complexes **4a** and **5** were characterized by EPR (figs. S10 and S11), establishing an S = 1/2 ground state, and **5** was also characterized by x-ray diffraction (fig. S52). Compound (*R*,*R*)-^{i}Pr-DuPhosCo(COD) (**4b**) was also prepared from Zn reduction of either **2b** or **3b** and characterized by EPR and x-ray diffraction (figs. S12 and S54). These observations suggest that both Co(I) and Co(0) are accessible via one-electron reductions using the Zn-MeOH protocol.

Alternative routes other than Zn reduction to complexes **3a**, **3b**, **4a**, **4b**, and **5** were also developed (fig. S1). The Co(II) dialkyl complex (*R*,*R*)-Ph-BPECo(CH_2SiMe_3)$_2$ (**6**) was synthesized by reported procedures (*6*), and hydrogenation of an equimolar mixture of **2a** and **6** yielded the monochloride dimer **3a** along with $SiMe_4$. Hydrogenation of **6** in the presence of diene or arene produced the Co(0) complexes **4a** and **5**. The corresponding Co(I) and Co(0) complexes with (*R*,*R*)-^{i}Pr-DuPhos were also prepared from the cobalt dialkyl.

To demonstrate the catalytic performance of the isolated Co(I) and Co(0) sources, reactions with preformed **3a** or **4a** were conducted. Hydrogenation of **1** in the presence of 0.5 mol % of Zn-free **3a** in MeOH with 500 psi (34 atm) of H_2 at 50°C produced levetiracetam in 99.9% yield and 98.2% ee (Fig. 3B), demonstrating that single-component, Zn-free Co(I) complexes are also effective for synthesis of the API. EPR experiments (figs. S47 to S49) support a pathway whereby the substrate **1** induces disproportionation of the Co(I) monochloride **3a** into Co(II) dichloride **2a** and a Co(0) alkene complex, with the latter undergoing reaction with H_2 to enter the catalytic cycle. A notable pressure dependence was observed with **3a**, as hydrogenation of **1** at 55 psi (3.7 atm) of H_2 resulted in only 38.5% yield after 72 hours, highlighting that H_2 is key to favor equilibria to generate and trap a Co(II) dihydride from **3a** (*21*, *24*). Consistent with this hypothesis, the cobalt diene precursor **4a** was effective at the lower H_2 pressure of 55 psi (3.7 atm) and produced levetiracetam in 99.1% yield and 97.5% ee (Fig. 3B). Hydrogenation of the diene ligand liberated cyclooctane and likely generated the Co(II) dihydride, which, in the absence of chloride, is free from inhibitive binding equilibria. Similar results were obtained with **5**. Deuterium labeling experiments were conducted with the Zn-activated catalyst **3a** and D_2 gas in natural abundance MeOH and furnished 1,2-d_2-levetiracetam (d, deuterium), supporting homolytic H_2 cleavage and formation of cobalt dihydride, distinct from previous systems where heterolytic H_2 cleavage by metal carboxylates and protonation by solvent are the proposed mechanism. (*7*, *25*).

To examine practicality in an industrial setting, the Co catalysts were applied to larger scale hydrogenations. Zinc activation proved optimal and most convenient and was therefore selected for these experiments. Quantities ranging from 1.3 to 20 g of **1** were successfully hydrogenated with 0.06 mol % each of (*R*,*R*)-Ph-BPE and $CoCl_2{\cdot}6H_2O$ in MeOH and 0.6 mol % Zn, producing levetiracetam in up to 100% yield and 97.5% ee in solution. Monitoring H_2 uptake established a highly reactive catalyst in the first few minutes after H_2 was introduced, and the reaction was completed in 4 hours (fig. S15); special care must be taken with higher loadings to avoid heat generation from the exothermic reaction. Immediate introduction of H_2 after catalyst injection results in the best catalyst performance. By using these optimized conditions, a 200-g-scale hydrogenation was carried out with a slightly higher loading of 0.08 mol %, and levetiracetam was obtained in 97% isolated yield and 98.2% ee (Fig. 4C), demonstrating that earth-abundant metals are competent, if not superior to precious metals, for the synthesis of high-value single-enantiomer APIs at an industrial scale in environmentally benign solvent. These findings highlight the benefits of first-row transition metals in catalysis where oxidation states separated by one electron offer distinct strategies for improving catalyst performance and lifetime.

REFERENCES AND NOTES

1. C. S. Shultz, S. W. Krska, *Acc. Chem. Res.* **40**, 1320–1326 (2007).
2. L. A. Saudan, *Acc. Chem. Res.* **40**, 1309–1319 (2007).
3. W. S. Knowles, *Angew. Chem. Int. Ed.* **41**, 1998–2007 (2002).
4. P. J. Chirik, *Acc. Chem. Res.* **48**, 1687–1695 (2015).
5. S. Ge, R. A. Green, J. F. Hartwig, *J. Am. Chem. Soc.* **136**, 1617–1627 (2014).
6. M. R. Friedfeld *et al.*, *Science* **342**, 1076–1080 (2013).
7. M. Shevlin *et al.*, *J. Am. Chem. Soc.* **138**, 3562–3569 (2016).
8. T. P. Yoon, E. N. Jacobsen, *Science* **299**, 1691–1693 (2003).
9. R. H. Crabtree, *Chem. Rev.* **115**, 127–150 (2015).
10. D. Heller, A. H. M. de Vries, in *The Handbook of Homogeneous Hydrogenation*, J. G. de Vries, C. J. Elsevier, Eds. (Wiley-VCH, 2007), pp. 1483–1516.
11. J. M. Hoyt *et al.*, *Organometallics* **33**, 5781–5790 (2014).
12. J. A. Osborn, F. H. Jardine, J. F. Young, G. Wilkinson, *J. Chem. Soc. A* **0**, 1711–1732 (1966).
13. P. S. Hallman, D. Evans, J. A. Osborn, G. Wilkinson, *Chem. Commun. (Camb.)* **0**, 305–306 (1967).
14. R. R. Schrock, J. A. Osborn, *J. Am. Chem. Soc.* **98**, 2134–2143 (1976).
15. D. A. Dobbs *et al.*, *Angew. Chem. Int. Ed.* **39**, 1992–1995 (2000).
16. B. Abou-Khalil, *Neuropsychiatr. Dis. Treat.* **4**, 507–523 (2008).
17. J. Surtees, V. Marmon, E. Differding, V. Zimmerman, PCT international application patent WO 2001064637 A1 (2001).
18. C. R. Landis, J. Halpern, *J. Am. Chem. Soc.* **109**, 1746–1754 (1987).
19. B. D. Vineyard, W. S. Knowles, M. J. Sabacky, G. L. Bachman, D. J. Weinkauff, *J. Am. Chem. Soc.* **99**, 5946–5952 (1977).
20. D. Prat *et al.*, *Org. Process Res. Dev.* **17**, 1517–1525 (2013).
21. X. Ma, M. Lei, *J. Org. Chem.* **82**, 2703–2712 (2017).
22. G. Hilt, *Eur. J. Org. Chem.* **2012**, 4441–4451 (2012).
23. D. K. Kim, J. Riedel, R. S. Kim, V. M. Dong, *J. Am. Chem. Soc.* **139**, 10208–10211 (2017).
24. M. R. Friedfeld, G. W. Margulieux, B. A. Schaefer, P. J. Chirik, *J. Am. Chem. Soc.* **136**, 13178–13181 (2014).
25. T. J. Korstanje, J. I. van der Vlugt, C. J. Elsevier, B. de Bruin, *Science* **350**, 298–302 (2015).

ACKNOWLEDGMENTS

We acknowledge J. Hill and M. Maust for assistance with high-pressure reactions conducted on large scale, A. Chin for determining solubility data for product isolation, and J. Gouker for assistance with chiral method development. **Funding:** Financial support was from a U.S. National Science Foundation (NSF) Grant Opportunities for Academic Liaison with Industry (GOALI) grant (CHE-1265988) between Princeton and Merck. M.R.F. acknowledges the NSF for a Graduate Research Fellowship (DGE-1148900). **Author contributions:** M.R.F., H.Z., and M.S. conducted all of the experimental work, interpreted results, and prepared the manuscript. R.T.R. and P.J.C. supervised the research and assisted with data analysis and manuscript preparation. **Competing interests:** None declared. **Data and materials availability:** Crystallographic data for compounds **2a**, **3a**, **3b**, **4b**, and **5** are available free of charge from the Cambridge Crystallographic Data Centre under reference numbers CCDC 1529292, 1529294, 1586348, 1586349, and 1529293, respectively. All other characterization data are reported in the supplementary materials.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/360/6391/888/suppl/DC1
Materials and Methods
Figs. S1 to S54
Tables S1 to S30
References (*26*–*29*)

28 November 2017; resubmitted 13 February 2018
Accepted 3 April 2018
10.1126/science.aar6117

VALLEYTRONICS

Imaging of pure spin-valley diffusion current in WS_2-WSe_2 heterostructures

Chenhao Jin,[1]* Jonghwan Kim,[1,2]* M. Iqbal Bakti Utama,[1,3] Emma C. Regan,[1,4] Hans Kleemann,[1] Hui Cai,[5] Yuxia Shen,[5] Matthew James Shinner,[1] Arjun Sengupta,[1] Kenji Watanabe,[6] Takashi Taniguchi,[6] Sefaattin Tongay,[5] Alex Zettl,[1,7,8] Feng Wang[1,7,8]†

Transition metal dichalcogenide (TMDC) materials are promising for spintronic and valleytronic applications because valley-polarized excitations can be generated and manipulated with circularly polarized photons and the valley and spin degrees of freedom are locked by strong spin-orbital interactions. In this study we demonstrate efficient generation of a pure and locked spin-valley diffusion current in tungsten disulfide (WS_2)–tungsten diselenide (WSe_2) heterostructures without any driving electric field. We imaged the propagation of valley current in real time and space by pump-probe spectroscopy. The valley current in the heterostructures can live for more than 20 microseconds and propagate over 20 micrometers; both the lifetime and the diffusion length can be controlled through electrostatic gating. The high-efficiency and electric-field–free generation of a locked spin-valley current in TMDC heterostructures holds promise for applications in spin and valley devices.

Transition metal dichalcogenides (TMDCs) offer a promising platform for applications in spintronics and valleytronics because of their distinctive electronic structure and strong spin-orbital interactions (*1–13*). At the K and K′ points of the Brillouin zone in TMDCs, two degenerate but inequivalent valleys are present, and these valleys can be used to encode binary information. In addition, spin-orbital coupling leads to locked spin and valley degrees of freedom in TMDC monolayers. The locked spin-valley polarization of charge carriers can exhibit a very long lifetime because intervalley scattering requires a large momentum change and a flip of the spin simultaneously and is therefore a rare event (*14–18*).

An outstanding challenge in spintronics and valleytronics is to efficiently generate, transport, and detect pure spin-valley current, which will be crucial not only for understanding novel spin-valley physics but also for potential applications in charge-current–free devices with low power consumption (*19*, *20*). In traditional spintronic systems, a transverse spin current is generated through the spin Hall effect in materials with a large spin Hall angle, such as tantalum and tungsten (*21–23*). Similarly, a transverse valley current has also been realized through the valley Hall effect in TMDC materials, but with low efficiency caused by a small valley Hall angle (*13*). In both cases a strong driving electric field is necessary, resulting in a dominant and unavoidable longitudinal electric current.

In this study we exploited TMDC heterostructures for efficient optical generation of a pure valley diffusion current without an external electric field, which is accompanied by a pure spin diffusion current because of the spin-valley locking in TMDCs.

Figure 1, A and B, shows an optical microscopy image and a side-view illustration of a representative heterostructure device, respectively. A WSe_2-WS_2 heterostructure (black dashed box in Fig. 1A) is encapsulated in two hexagonal boron nitride (hBN) flakes, with the ~40-nm-thick bottom hBN also serving as the gate dielectric. A few-layer graphene (FLG) back gate is used to tune the carrier concentration in the WSe_2-WS_2 heterostructure, and two FLG source and drain contacts (yellow dashed boxes in Fig. 1A) are symmetrically placed on two sides of the heterostructure. All of the two-dimensional materials were first mechanically exfoliated from bulk crystals and then stacked together by a dry transfer method with a polyethylene terephthalate stamp (*24*). The whole stack was then transferred onto a 90-nm SiO_2-Si substrate.

We first characterized the gate-dependent optical transitions in the heterostructure by reflection spectroscopy. The results are summarized in Fig. 1C for carrier concentrations ranging from $-5 \times 10^{12}/cm^2$ (hole doping) to $5 \times 10^{12}/cm^2$ (electron doping). Two prominent resonance features appear around 1.72 and 2.05 eV, corresponding to the A exciton energies in the WSe_2 and WS_2 monolayers, respectively. However, the WSe_2 and WS_2 resonances exhibit distinctively different gate dependencies. Resonances from WSe_2 vary substantially on the hole-doping side, showing both a decrease in the exciton response and an emergence of the trion response, consistent with the behavior of hole-doped monolayer WSe_2 (*25*, *26*). Meanwhile, the WS_2 resonance shows only a slight red shift. The behavior on the electron-doping side is the opposite: only a slight red shift occurs in the WSe_2 resonance, but the WS_2 exciton and trion transitions vary markedly. This peculiar gate dependence can be understood from the type II band alignment (*27–29*) between WSe_2 and WS_2 (Fig. 1D): The conduction and valence bands in WS_2 are lower than the corresponding bands in WSe_2. Consequently, electrostatically induced electrons and holes will stay only in the WS_2 and WSe_2 layers, respectively.

We then generated a spin-valley current in the hole-doped heterostructures by local circularly polarized laser excitation (Fig. 2, A to D). This mechanism exploits the selective coupling of valley excitons to photon helicity, the ultrafast interlayer charge transfer process, and the ultralong valley hole lifetime in WS_2-WSe_2 heterostructures. First, left-handed circularly polarized photons at 1.80 eV selectively excite K valley excitons in WSe_2. The excited electrons transfer to the conduction band of WS_2 within the first ~100 fs (*28*, *29*), turning off the exciton valley depolarization channel (Maialle-Silva-Sham mechanism) (*30*) and creating excess valley-polarized holes in the WSe_2 K valley (Fig. 2A) (*18*). Afterward, electrons in WS_2 recombine with holes in WSe_2 within ~100 ns. As discussed below, the recombinations with K valley and K′ valley holes in WSe_2 have almost equal probabilities (Fig. 2B), which leads to an excess of K valley holes and a deficiency of K′ valley holes in WSe_2 with an ultralong lifetime of many microseconds (Fig. 2C). The local imbalance of valley-polarized holes excited by a focused laser light can drive a pure spin-valley diffusion current (with zero net charge current) through diffusion in the heterostructure (Fig. 2D).

We performed spatial-temporal imaging of the spin-valley current in the WS_2-WSe_2 heterostructure by tracking the valley-polarized holes in the device with space- and time-resolved circular dichroism spectroscopy. An elliptically shaped pump beam at 1.80 eV efficiently generates spin- and valley-polarized holes in the heterostructure, and a second elliptically shaped probe beam at 1.70 eV probes their evolution in space and time by varying both the spatial separation and the temporal delay between the pump and probe pulses (*24*). Figure 2E displays the profile of the elliptically shaped pump and probe beams at the sample, each with a half width of ~1.5 μm. The temporal delay Δt between the pump and probe pulses is electronically generated by a data acquisition card with 12.5-ns resolution, whereas the spatial separation Δx between the pump and probe pulses

---

[1]Department of Physics, University of California at Berkeley, Berkeley, CA 94720, USA. [2]Department of Materials Science and Engineering, Pohang University of Science and Technology, Pohang 790-784, Korea. [3]Department of Materials Science and Engineering, University of California at Berkeley, Berkeley, CA 94720, USA. [4]Graduate Group in Applied Science and Technology, University of California at Berkeley, Berkeley, CA 94720, USA. [5]School for Engineering of Matter, Transport and Energy, Arizona State University, Tempe, AZ 85287, USA. [6]National Institute for Materials Science, 1-1 Namiki, Tsukuba, 305-0044, Japan. [7]Division of Material Science, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA. [8]Kavli Energy NanoSciences Institute at the University of California at Berkeley and Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA.
*These authors contributed equally to this work.
†Corresponding author. Email: fengwang76@berkeley.edu

along the width direction is controlled with ~0.2-μm resolution by tuning the angle of a mirror (24). The circular dichroic reflection contrast of the probe pulse directly measures the valley-polarized hole density in the heterostructure at a specific probe position and time (18). All experiments were carried out at 10 K.

Figure 2F shows the measured evolution of the valley-polarized hole density in real space and real time for the hole-doped WS_2-WSe_2 heterostructure at an initial electrostatic hole doping $p_0 = 1 \times 10^{12}/cm^2$. The horizontal and vertical axes represent the temporal and spatial separation between the pump and probe pulses, respectively; the colors represent the amplitude of the circular

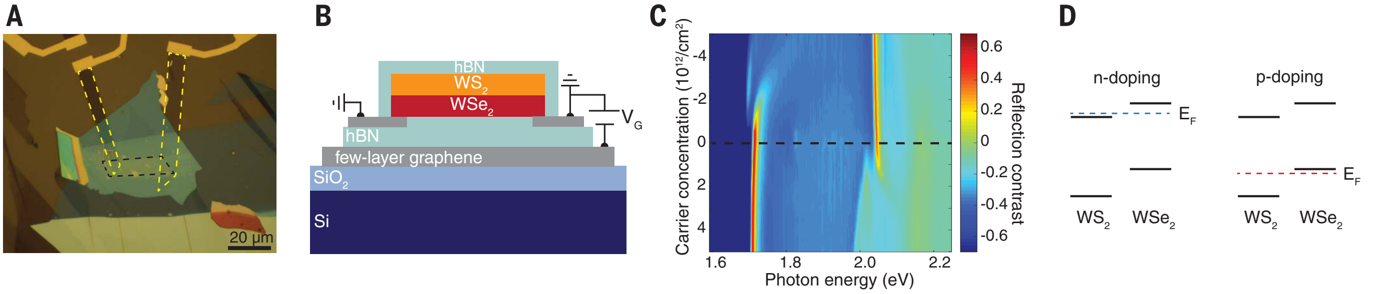


Fig. 1. Gate-dependent optical transitions in the WSe_2-WS_2 heterostructure. (**A** and **B**) Optical microscope image (A) and side-view illustration (B) of the heterostructure device. Black and yellow dashed boxes in (A) mark the heterostructure region and FLG source and drain electrodes, respectively. V_G, gate voltage. (**C**) Doping-dependent reflection contrast of the WSe_2-WS_2 heterostructure. The dashed line indicates charge neutrality. The two prominent resonances around 1.72 and 2.05 eV correspond to the A exciton energies in the WSe_2 and WS_2 monolayers, respectively, which exhibit distinctively different doping dependencies because of the type II band alignment of the heterostructure. (**D**) Type II band alignment of the WSe_2 and WS_2 interface. The conduction band minimum and the valence band maximum reside in WS_2 and WSe_2, respectively. Electrostatically doped electrons stay only in WS_2 for the n-doping case (left panel), whereas holes stay only in WSe_2 for the p-doping case (right panel). E_F, Fermi energy.

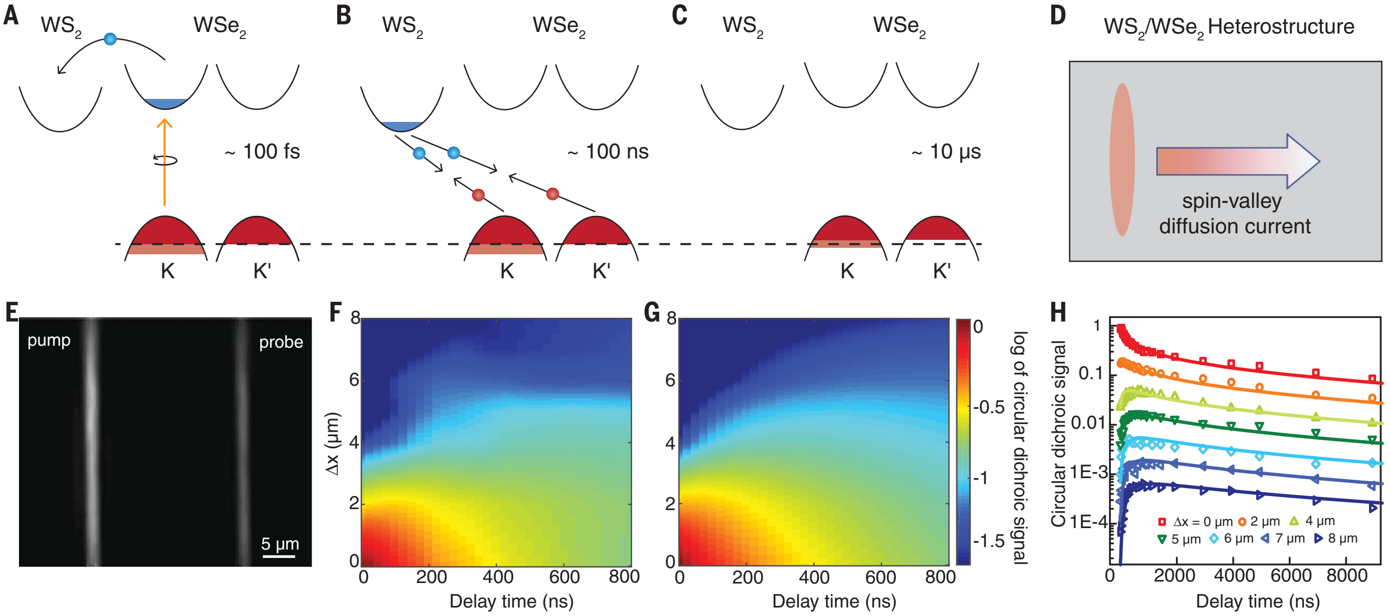


Fig. 2. Pure spin-valley diffusion current in hole-doped WSe_2-WS_2 heterostructures. (**A** to **D**) Experimental scheme for valley current generation in hole-doped WSe_2-WS_2 heterostructures. The dashed line represents the Fermi level before optical excitation. (A) Excitons in the K valley of WSe_2 are selectively excited by circularly polarized pump light and efficiently converted into excess holes within ~100 fs through the ultrafast interlayer charge transfer process. (B) Electrons in WS_2 recombine with K valley and K′ valley holes in WSe_2 with almost equal probabilities, resulting in an excess of K valley holes and a deficiency of K′ valley holes in WSe_2. (C) This valley-polarized hole population persists for tens of microseconds. (D) Local excitation of such valley-polarized holes can drive a pure valley diffusion current (with zero net charge current) through diffusion in the heterostructure. This pure valley diffusion current will be unidirectional if holes are excited at one edge of the device. (**E** to **H**) Spatial-temporal imaging of the pure spin-valley current. (E) Elliptically shaped pump and probe beam profiles, each with a ~1.5-μm half width. The long axis of the elliptical light (>50 μm) is much longer than the sample width (~10 μm). Therefore, the illumination is nearly homogeneous along the long axis direction. (F) The spatial-temporal evolution of spin- and valley-polarized holes in the heterostructure with an initial hole doping $p_0 = 1 \times 10^{12}/cm^2$. The horizontal and vertical axes represent the temporal and spatial separation between the pump and probe pulses, respectively. The colors represent the amplitude of the circular dichroism signal on a logarithmic scale, which is proportional to the valley-polarized hole density. The fast decrease in signal at small Δx values and the increase in signal at large Δx values signify the propagation of a valley diffusion current, which can be well captured by simulation from a diffusion-decay model (G). (H) Horizontal cuts from (F) (symbols) and fits from the diffusion-decay model (curves) over a longer time scale. The whole region within Δx = 8 μm becomes homogeneous after 2 μs because of the diffusive valley current and shows an overall slow decay afterward.

dichroism signal on a logarithmic scale. The signal at the origin (corresponding to spatially and temporally overlapped pump and probe pulses) is normalized to 1. At zero time delay, the valley-polarized hole population is localized at the origin with a distribution matching the pump profile (half width of ~1.5 μm), and the circular dichroism signal is negligible at pump-probe separations larger than 3 μm. After a finite delay time, valley-polarized holes start to diffuse out of the excitation region, generating a pure spin-valley diffusion current. This leads to a strong decrease in signal close to the origin but an increase in signal in the region with large pump-probe spatial separation. In particular, a finite signal emerges at $\Delta x > 3$ μm, where no valley polarization is present from the direct pump excitation. As illustrated in Fig. 2F, the valley current can readily propagate to a distance of more than 8 μm within a time period of 800 ns.

We can compare the measured spatial-temporal image of the valley-polarized hole density to diffusion theory and extract important dynamic transport parameters. The spatial-temporal evolution of the one-dimensional valley-polarized hole density $\Delta p_v(x, t)$ is described by a simple diffusion-decay model (*24*)

$$\Delta p_v(x,t) = \frac{\Delta p_0}{\sqrt{\pi(\sigma_0^2 + 4Dt)}} e^{-\frac{x^2}{\sigma_0^2+4Dt}} e^{-\frac{t}{\tau}} \qquad (1)$$

where Δp_0 is the total number of pump-induced spin- and valley-polarized holes, σ_0 is the half width of the pump beam, D is the hole diffusion constant, and τ is the lifetime of valley-polarized holes (limited by a finite intervalley scattering time). The pump-probe signal $S(\Delta x, \Delta t)$ can be calculated by convolving $\Delta p_v(x, \Delta t)$ with the probe beam intensity $T(x)$ at a given pump-probe separation Δx (*24*). Figure 2G shows the simulated pump-probe signal with a hole diffusion constant of $D = 0.2$ cm^2/s and a valley lifetime of $\tau = 20$ μs. The excellent agreement between the experimental and theoretical results confirms the diffusion-dominated dynamics of valley-polarized holes. Figure 2H shows several horizontal line cuts of Fig. 2F over a longer time scale (different curves are shifted vertically and successively by 40% of the major tick separation for visual clarity; fig. S2 illustrates vertical line cuts). We observed that the whole region within $\Delta x = 8$ μm became homogeneous after 2 μs because of the diffusive valley current and exhibited an overall slow decay afterward. This dynamic is well captured by the diffusion-decay model in Eq. 1 (curves in Fig. 2H). The valley lifetime of $\tau = 20$ μs extracted here is more than one order of magnitude longer than the earlier reported values (*14*–*18*). From the valley hole diffusion constant and valley polarization lifetime, we obtain a valley diffusion length of $l = \sqrt{D\tau} = 20$ μm. In addition, the optical generation can produce a very high valley current density, reaching 2×10^7 A/m^2 in the ~1-nm-thick heterostructure at a valley-polarized hole density of 10^{12}/cm^2 (*24*).

The photo-generated valley diffusion current can be modulated substantially through electrostatic gating. Figure 3, A and C, shows the spatial-temporal mapping of the valley-polarized hole density with the initial hole doping at ~0 and 2.8×10^{12}/cm^2, respectively. We found that the valley-polarized hole density decays over time but maintains the same spatial profile for near-zero initial doping (Fig. 3A). This indicates that at the low-doping limit valley-polarized holes are largely localized and there is no valley current. In contrast, the valley diffusion current flow becomes quite large for the highly hole-doped heterostructure at 2.8×10^{12}/cm^2 (Fig. 3C), leading to a peak spin-valley diffusion current density of 1.2×10^8 A/m^2 at a valley-polarized hole density of 10^{12}/cm^2 (*24*).

Figure 3, B and D, shows theoretical modeling of the valley hole diffusion at $p_0 = 0$ and 2.8×10^{12}/cm^2, respectively. The extracted diffusion constants are summarized in Fig. 3E. From the diffusion constant, we can directly obtain the density-dependent hole mobility using the Einstein relation for a Fermi liquid at 10 K (*24*). The holes are almost completely localized at a doping level close to the charge neutrality point. Therefore, both the diffusion constant and the mobility are negligibly small. At larger carrier concentrations, the hole diffusion constant and the mobility both increase substantially, reaching $D = 1.2$ cm^2/s and $\mu = 160$ cm$^2/(\mathrm{V} \cdot \mathrm{s})$ at $p_0 = 2.8 \times 10^{12}$/cm^2. The localization of holes at low carrier density signifies a mobility edge in TMDC materials, which may be related to Anderson localization from potential fluctuations (*31*). Traditional low-temperature transport measurements at such low carrier concentrations are quite challenging or almost impossible because of an exponentially large contact resistance. In this study the spatially and temporally resolved optical imaging provided a contact-free approach to measure the intrinsic charge transport in TMDCs, and it can be a powerful and general tool for characterizing the transport behavior of two-dimensional systems at low carrier concentrations.

Like the hole diffusion constant, the valley-polarized hole lifetime also depends strongly on the electrostatic gating. Notably, the valley-polarized hole population $(\Delta p_v \equiv \Delta p_{\mathrm{K}} - \Delta p_{\mathrm{K'}})$, where Δp_{K} and $\Delta p_{\mathrm{K'}}$ are photo-generated excess hole densities in the K and K′ valleys of WSe$_2$, respectively, can exhibit very different behavior from the total excess hole population $(\Delta p_{\mathrm{tot}} \equiv \Delta p_{\mathrm{K}} + \Delta p_{\mathrm{K'}})$. The different dynamic behaviors of Δp_{tot} and Δp_v can be probed from the sum and difference responses of left and right circularly polarized probe light, respectively (*18*). Figure 4A shows the population decay of the total excess holes in the heterostructure at $\Delta x = 0$ for different initial carrier concentrations, with the decay constants summarized in Fig. 4B (triangles). The decay lifetime of Δp_{tot} ranges from 20 to 500 ns and becomes shorter at both greater electron doping and greater hole doping of the heterostructures. In contrast, the valley-polarized hole lifetime (circles in Fig. 4B) is the same as the total hole population lifetime for charge-neutral and electron-doped heterostructures but decouples and becomes much longer for hole-doped cases.

The unusual dynamics of doping-dependent valley lifetime arises from the distinctive interlayer electron-hole recombination process in the heterostructure, as illustrated in Fig. 4, C and D (*24*). For electron-doped or charge-neutral heterostructures (Fig. 4C), all of the holes in WSe$_2$ are pump-generated "excess holes." Therefore, when Δp_{tot} decays to zero because of interlayer electron-hole recombination, no holes—and certainly no valley-polarized holes—remain in the WSe$_2$. In these cases, the valley lifetime is limited by the total excess hole lifetime. The scenario is completely different for hole-doped heterostructures (Fig. 4D). If the original hole density is much larger than the photo-generated one,

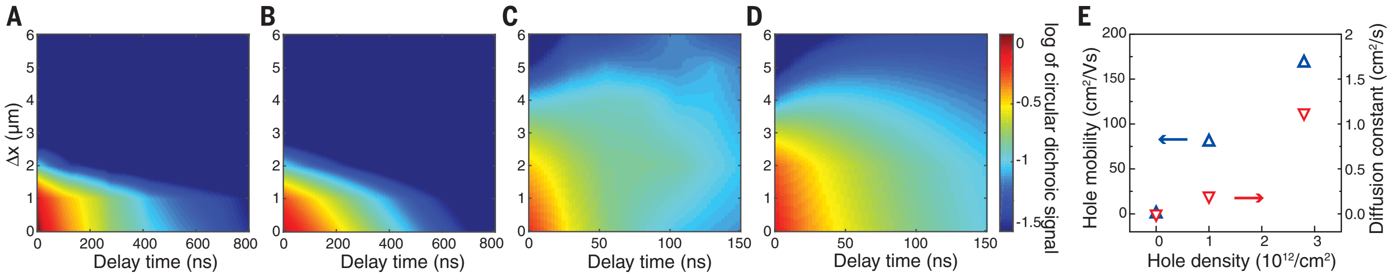


Fig. 3. Strong modulation of spin-valley current with electrostatic gating. (**A** and **C**) Spatial-temporal imaging of spin-valley current at initial hole concentrations of near zero (A) and 2.8×10^{12}/cm^2 (C). (**B** and **D**) Corresponding simulation results from the diffusion-decay model. The valley-polarized holes have negligible diffusion near the charge-neutral point (A and B). In contrast, holes diffuse efficiently at large hole doping (C and D), creating a large spin-valley current density up to 1.2×10^8 A/m^2 at an optical excitation density of 10^{12}/cm^2. (**E**) Extracted hole diffusion constant (red) and mobility (blue) show a strong dependence on the initial hole concentration.

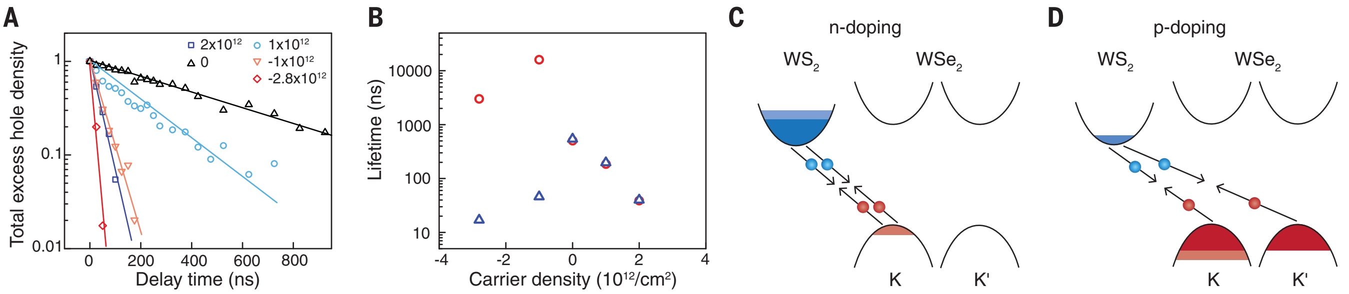


Fig. 4. Doping-dependent lifetime of total excess hole and valley-polarized hole populations. (**A**) Decay dynamics of the total excess hole density at different initial carrier concentrations. The decay lifetime is longest near charge neutrality and decreases with both electron doping (positive carrier concentration) and hole doping (negative carrier concentration). (**B**) Comparison between valley-polarized hole lifetime (circles) and total excess hole lifetime (triangles). The valley lifetime is limited by the total population lifetime in charge-neutral and electron-doped heterostructures; in hole-doped heterostructures the valley lifetime can be orders of magnitude longer than the total excess hole lifetime. (**C**) For electron-doped or charge-neutral heterostructures, interlayer electron-hole recombination directly reduces the valley-polarized hole density. Dark and light blue colors represent the electrostatically injected electrons and photo-generated electrons, respectively. (**D**) If the initial hole doping density (dark red) is much larger than the photo-generated one (light red), electrons in WS_2 (blue) will recombine with holes at K and K′ valleys of WSe_2 with almost equal probabilities, reducing only the total excess hole population but not the valley-polarized hole density.

excess electrons in WS_2 will recombine with holes from both valleys of WSe_2 with almost equal probabilities (*24*). Consequently, the valley-polarized hole density Δp_v decouples from the net excess hole population Δp_{tot} and persists long after the latter decays to zero, enabling the generation of a pure spin-valley diffusion current.

The generation efficiency of a pure spin-valley diffusion current in our heterostructure-based device can be very high because of the near-perfect conversion from optical excitation to the pure valley diffusion current (*24*). Because of the lack of any associated charge current in our device, the pure spin-valley diffusion current density scales linearly with the optical excitation power, as opposed to the square-root scaling with power for electrically generated current. Therefore, the power efficiency of heterostructure devices can be favorable for generating large spin and valley current. Along with the ultralong spin-valley lifetimes and diffusion lengths, TMDC heterostructures provide exciting opportunities to realize future spintronic and valleytronic devices.

REFERENCES AND NOTES

1. D. Xiao, G. B. Liu, W. Feng, X. Xu, W. Yao, *Phys. Rev. Lett.* **108**, 196802 (2012).
2. X. Xu, W. Yao, D. Xiao, T. F. Heinz, *Nat. Phys.* **10**, 343–350 (2014).
3. T. Cao *et al.*, *Nat. Commun.* **3**, 887 (2012).
4. H. Zeng, J. Dai, W. Yao, D. Xiao, X. Cui, *Nat. Nanotechnol.* **7**, 490–493 (2012).
5. K. F. Mak, K. He, J. Shan, T. F. Heinz, *Nat. Nanotechnol.* **7**, 494–498 (2012).
6. Y. Li *et al.*, *Phys. Rev. Lett.* **113**, 266804 (2014).
7. E. J. Sie *et al.*, *Nat. Mater.* **14**, 290–294 (2015).
8. D. MacNeill *et al.*, *Phys. Rev. Lett.* **114**, 037401 (2015).
9. J. Kim *et al.*, *Science* **346**, 1205–1208 (2014).
10. A. Srivastava *et al.*, *Nat. Phys.* **11**, 141–147 (2015).
11. Z. L. Ye, D. Z. Sun, T. F. Heinz, *Nat. Phys.* **13**, 26–29 (2017).
12. G. Aivazian *et al.*, *Nat. Phys.* **11**, 148–152 (2015).
13. K. F. Mak, K. L. McGill, J. Park, P. L. McEuen, *Science* **344**, 1489–1492 (2014).
14. L. Yang *et al.*, *Nat. Phys.* **11**, 830–834 (2015).
15. W.-T. Hsu *et al.*, *Nat. Commun.* **6**, 8963 (2015).
16. X. Song, S. Xie, K. Kang, J. Park, V. Sih, *Nano Lett.* **16**, 5010–5014 (2016).
17. P. Rivera *et al.*, *Science* **351**, 688–691 (2016).
18. J. Kim *et al.*, *Sci. Adv.* **3**, e1700518 (2017).
19. S. Murakami, N. Nagaosa, S. C. Zhang, *Science* **301**, 1348–1351 (2003).
20. I. Žutić, J. Fabian, S. Das Sarma, *Rev. Mod. Phys.* **76**, 323–410 (2004).
21. J. Sinova, S. O. Valenzuela, J. Wunderlich, C. H. Back, T. Jungwirth, *Rev. Mod. Phys.* **87**, 1213–1260 (2015).
22. L. Liu *et al.*, *Science* **336**, 555–558 (2012).
23. C. F. Pai *et al.*, *Appl. Phys. Lett.* **101**, 122404 (2012).
24. See supplementary materials.
25. K. F. Mak *et al.*, *Nat. Mater.* **12**, 207–211 (2013).
26. Z. Wang, J. Shan, K. F. Mak, *Nat. Nanotechnol.* **12**, 144–149 (2017).
27. J. Kang, S. Tongay, J. Zhou, J. B. Li, J. Q. Wu, *Appl. Phys. Lett.* **102**, 012111 (2013).
28. X. Hong *et al.*, *Nat. Nanotechnol.* **9**, 682–686 (2014).
29. F. Ceballos, M. Z. Bellus, H.-Y. Chiu, H. Zhao, *ACS Nano* **8**, 12717–12724 (2014).
30. M. Z. Maialle, E. A. de Andrada e Silva, L. J. Sham, *Phys. Rev. B* **47**, 15776–15788 (1993).
31. P. W. Anderson, *Phys. Rev.* **109**, 1492–1505 (1958).

ACKNOWLEDGMENTS

Funding: This work was supported primarily by the director, Office of Science, Office of Basic Energy Sciences, Materials Sciences and Engineering Division of the U.S. Department of Energy, under contract no. DE-AC02-05-CH11231 (van der Waals heterostructures program, KCWF16). The device fabrication was supported by the NSF EFRI program (EFMA-1542741). The growth of hBN crystals was supported by the Elemental Strategy Initiative conducted by the MEXT, Japan, and JSPS KAKENHI grant JP15K21722. S.T. acknowledges support from NSF DMR CAREER award 1552220 for the growth of MoS_2 and WSe_2 crystals. H.K. acknowledges fellowship support from the Deutsche Forschungsgemeinschaft (KL 2961/1-1). **Author contributions:** F.W. and C.J. conceived of the research. C.J., J.K., and E.C.R. carried out optical measurements. C.J., F.W., and E.C.R. performed theoretical analysis. J.K., C.J., M.I.B.U., H.K., M.J.S., A.S., and A.Z. fabricated van der Waals heterostructures. H.C., Y.S., and S.T. grew WSe_2 and WS_2 crystals. K.W. and T.T. grew hBN crystals. All authors discussed the results and wrote the manuscript. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** Experimental data files are available from the Open Science Framework at https://osf.io/8jpqg/.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/360/6391/893/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S4
References (*32*, *33*)

10 July 2017; accepted 11 April 2018
10.1126/science.aao3503

ORGANIC ELECTRONICS

Long-range exciton transport in conjugated polymer nanofibers prepared by seeded growth

Xu-Hui Jin,[1]* Michael B. Price,[2]* John R. Finnegan,[1] Charlotte E. Boott,[1] Johannes M. Richter,[2] Akshay Rao,[2] S. Matthew Menke,[2] Richard H. Friend,[2]† George R. Whittell,[1]† Ian Manners[1]†

Easily processed materials with the ability to transport excitons over length scales of more than 100 nanometers are highly desirable for a range of light-harvesting and optoelectronic devices. We describe the preparation of organic semiconducting nanofibers comprising a crystalline poly(di-*n*-hexylfluorene) core and a solvated, segmented corona consisting of polyethylene glycol in the center and polythiophene at the ends. These nanofibers exhibit exciton transfer from the core to the lower-energy polythiophene coronas in the end blocks, which occurs in the direction of the interchain π-π stacking with very long diffusion lengths (>200 nanometers) and a large diffusion coefficient (0.5 square centimeters per second). This is made possible by the uniform exciton energetic landscape created by the well-ordered, crystalline nanofiber core.

The ability to transport excitation energy over length scales comparable to the optical absorption depth (100 nm and beyond) is central to the function of a range of devices, including solar cells. Most thin-film organic semiconductor structures, such as those formed from conjugated polymers, show short exciton diffusion lengths (L_D) of around 10 nm (*1*) that are primarily constrained by energetic disorder. As a result, these are fabricated as blends of electron donor and acceptor materials with length scales for the partly demixed materials targeted around 10 nm, so that all excitons can reach the charge-generating hetero-interface (*1*, *2*). In contrast, diffusion ranges for singlet excitons in purified single crystals are known to be considerably larger (up to 220 nm) (*1*, *3*–*6*). However, in these cases the materials are generally polydisperse in dimensions and are problematic to incorporate into useful devices. Device development is thus dependent on the ability to develop uniform nanostructures that are amenable to processing and able to support long-range exciton diffusion.

Self-assembly of molecular and polymeric amphiphiles in solution has recently emerged as a promising route to core-corona nanoparticles (micelles) suitable for device applications. For instance, long-range exciton diffusion has been observed in dye molecule H and J fiber-like aggregates (where H and J characterize a shift to shorter or longer wavelength upon aggregation, respectively) by single-molecule spectroscopy (*7*, *8*) and in more detailed investigations performed on single conjugated polymers isolated in matrices or dilute solution (*9*, *10*). These latter studies, however, necessarily exclude the role of interchain transport in exciton diffusion. Such transport is possible in solution-processed bulk conjugated polymers, which have been reported to show a singlet exciton diffusion length of 70 nm (*11*), but this is still short of the lengths exhibited in molecular crystals. Here, we report exceptional exciton diffusion ranges in uniform nanofibers formed from the seeded-solution self-assembly of block copolymers with a crystallizable π-conjugated poly(di-*n*-hexylfluorene) (PDHF) block.

To facilitate the study of exciton diffusion within π-conjugated crystalline PDHF, we prepared segmented nanofibers comprising a continuous PDHF core with a discrete region of energy-accepting quaternized polythiophene (QPT) covalently attached as a corona at each end (Fig. 1, A and B). The central segment contained an electronically insulating poly(ethylene glycol) (PEG) corona, and the length of this region was varied to afford a size series of nanofibers. These segmented nanostructures were prepared using living crystallization–driven self-assembly, a recently developed seeded-growth method for producing one-dimensional (1D) and 2D objects of controlled dimensions (*12*, *13*). The resulting structures are generally typified by highly crystalline cores, which can comprise a number of building blocks (*14*–*17*). We, Faul, and co-workers have previously demonstrated the use of this method with polythiophene-containing block copolymers to yield fibers with a π-conjugated core, which form ensembles with promising transport properties but weak fluorescence due to aggregation-induced quenching (*18*).

PDHF exhibits bright blue fluorescence in the solid state (*19*), in electrospun microfibers (*20*), and in polydisperse nanofibers self-assembled from triblock copolymers in solution (*21*). The uniform examples used in the present study, however, were prepared in a multistep approach, so as to achieve the desired segmented coronal structure and varied length of the PEG component. Initially, nanofibers ~5 to 10 μm in length, derived from the block copolymer PDFH$_{14}$-*b*-PEG$_{227}$ (where the numbers refer to the number-average degree of polymerization of each block; see figs. S1, S3 to S5, and S10), were prepared by homogeneous nucleation in a 10:8 mixture of tetrahydrofuran and methanol (THF:MeOH). Analysis by transmission electron microscopy (TEM), atomic force microscopy (AFM), and wide-angle x-ray scattering (WAXS) revealed the presence of a crystalline PDHF core with a rectangular cross section (number-average width W_n = 12.9 nm, height = 4.5 nm) surrounded by a PEG corona (figs. S10 to S12). A solution of micelle seeds (number-average length L_n = 30 nm; polydispersity L_w/L_n = 1.03, where L_w is the weight-average length; height = 4.5 nm) was then prepared by sonication of the multimicrometer-long fibers (figs. S13, A and B, and S14A). Subsequent addition of different volumes of a solution of unimeric (molecularly dissolved) PDHF$_{14}$-*b*-PEG$_{227}$ copolymer in THF led to the formation of uniform nanofibers of controlled length (Fig. 1A and figs. S14 and S15).

To create nanofibers with a segmented corona, we prepared an all-π-conjugated donor-acceptor diblock copolymer, PDHF$_{14}$-*b*-QPT$_{22}$, comprising the same PDHF core-forming block and a QPT corona–forming block (figs. S2 and S6 to S9). This material was added in a molecularly dissolved unimeric state (in THF:MeOH 3:1) to the PDHF$_{14}$-*b*-PEG$_{227}$ nanofibers in THF:MeOH 1:1, leading to growth from the two PDHF core termini. The resulting uniform nanofibers had a B-A-B structure of controlled overall and segment length with a crystalline PDHF core present over the entire length, but with a π-conjugated corona-forming block located only on the terminal (B) segments (Fig. 1A). This was used to produce a size series of near-uniform nanofibers with L_n = 180 ± 40 nm, 300 ± 70 nm, 505 ± 100 nm, 945 ± 240 nm, and 1840 ± 540 nm, with each QPT segment comprising 35 to 120 nm of this length, as measured by TEM and supported by AFM (figs. S13, C and D, S16, and S17 and table S1). The dimensional control is illustrated further in Fig. 2A (and fig. S18), which shows a laser scanning confocal microscopy (LSCM) image of model nanofibers with central A-segments of 1.6 μm and QPT-corona–containing segments of 1.8 μm in length.

Figure 1C shows the absorption and photoluminescence (PL) spectra of unsegmented PDHF$_{14}$-*b*-PEG$_{227}$ nanofibers, the absorption spectrum of the QPT homopolymer (see fig. S19 for emission spectrum), and the energy levels of the two conjugated species. The overlap of absorption and emission in the PDHF indicates a small Stokes shift, whereas the PL illustrates a high degree of vibronic structure. The good overlap of QPT absorption and PDHF emission gives a PDHF-to-QPT Förster transfer radius of 4 nm.

---

[1] School of Chemistry, University of Bristol, Bristol BS8 1TS, UK. [2] Cavendish Laboratory, Department of Physics, University of Cambridge, Cambridge CB3 0HE, UK.
*These authors contributed equally to this work.
†**Corresponding author. Email: rhf10@cam.ac.uk (R.H.F.); g.whittell@bristol.ac.uk (G.R.W.); ian.manners@bristol.ac.uk (I.M.)**

We performed further steady-state optical measurements on the PDHF$_{14}$-*b*-PEG$_{227}$ as both nanofiber (of average length 435 nm) and unimer to probe the nature of exciton transport in the PDHF core. The nanofibers show more vibrational structure (Fig. 1C) than the unimer (fig. S20), and there is no pronounced red or blue shift upon aggregation. This is consistent with a π-stacked polymer possessing a transition dipole along the polymer backbone (*22*), which would be expected to exhibit both H- and J-like aggregate characteristics (*23*, *24*). The ratio of the I_{0-0} to I_{0-1} PL bands is larger in the nanofiber than in the unimer (fig. S20), indicating a decreased Huang-Rhys parameter and hence a smaller configurational relaxation in the excited state (*19*).

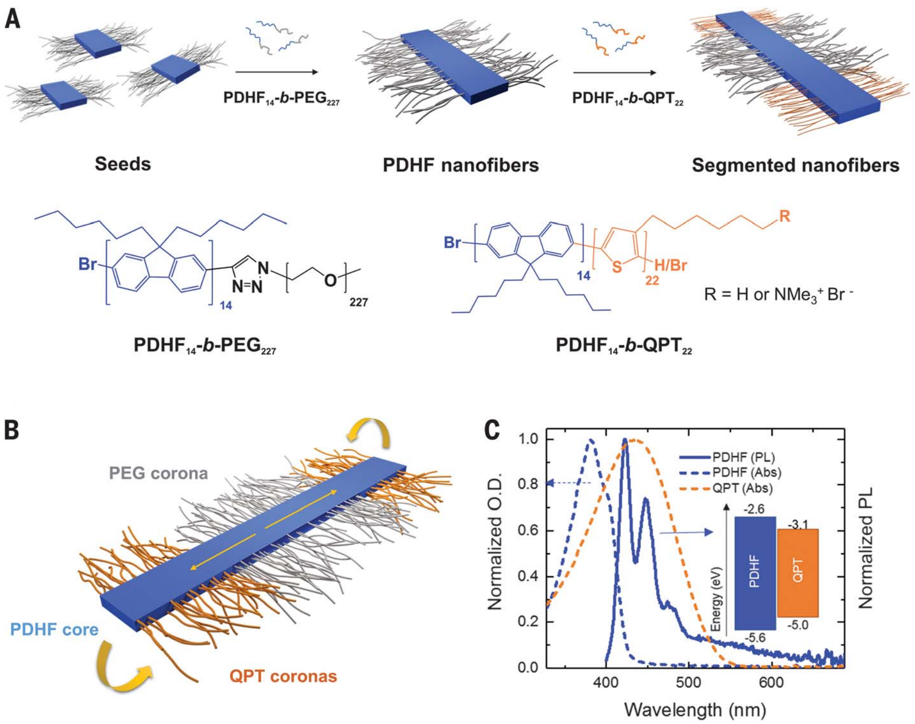


Fig. 1. Formation of segmented PDHF nanofibers by multistep self-assembly. (**A**) Schematic diagram illustrating the seeded growth process and the structures of PDHF$_{14}$-*b*-PEG$_{227}$ and PDHF$_{14}$-*b*-QPT$_{22}$. (**B**) Illustration of the segmented B-A-B nanofiber structure with separate donor and acceptor domains. (**C**) Normalized absorption of QPT homopolymer in THF:MeOH (1:1) (orange dashes), and unsegmented PDHF nanofibers (L_n = 1605 nm) (blue dashes), and photoluminescence (PL) emission of unsegmented PDHF nanofibers (L_n = 1605 nm) in the same solution (solid blue line). The I_{0-0} and I_{0-1} peaks in the PL are at 425 nm and 455 nm, respectively. The inset shows the energy levels of the PDHF and QPT.

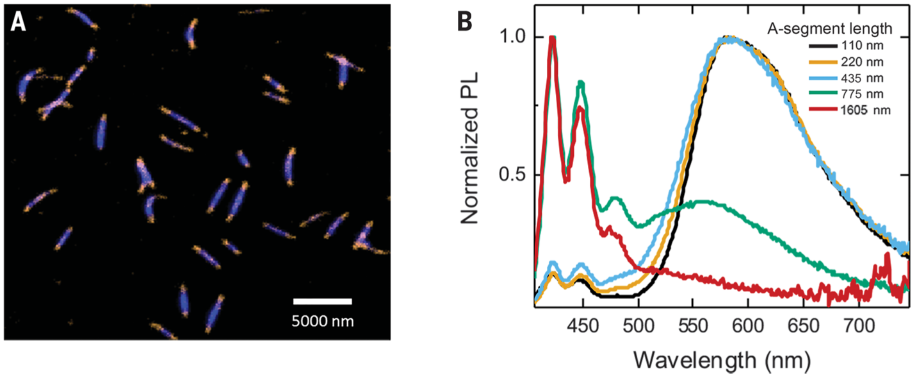


Fig. 2. Photoluminescence of segmented PDHF B-A-B nanofibers in solution. (**A**) LSCM image of the uniform segmented PDHF nanofibers with a crystalline PDHF core (blue emission) and two terminal segments with QPT coronas (orange emission in THF:MeOH 1:1). Values of L_n for the central and terminal segments were 1.6 μm and 0.9 μm, respectively. (**B**) PL spectra of segmented PDHF nanofibers with different A-segment lengths, normalized to peak maxima. In each case, the solutions of segmented nanofibers (~0.5 mg/ml) in THF:MeOH 1:1 were excited at the PDHF absorption peak of 380 nm. Emission arising from direct excitation of the QPT in the 1605-nm sample was unresolved.

We studied the energy transfer from PDHF to QPT for the aforementioned size series of segmented nanofibers; Fig. 2B shows the normalized PL when the PDHF is selectively excited (at 380 nm). We observed quenching of the PDHF peak relative to the QPT peak for average A-segment lengths below 775 ± 150 nm. This indicates energy transfer from the A-segment PDHF core to the B-segment QPT corona, although reabsorption of the PDHF PL is expected at this concentration. Further evidence for energy transfer comes from the excitation PL scan (fig. S21), which maps the absorption (fig. S22), and from the change in PL quantum efficiency with A-segment length (fig. S23). The latter matches that of pure QPT (13 ± 5%) below the critical length (775 nm) and approaches that for the unsegmented structure (73 ± 10%, a very high value) beyond this length.

Time-resolved PL measurements enable better quantification of energy transfer, because donor quenching kinetics are not affected by reabsorption effects. We used time-correlated single photon counting (TCSPC; instrument response time ~300 ps) to access low-excitation density regimes, and transient grating PL spectroscopy to probe the 1- to 100-ps time scale (see supplementary materials).

From TCSPC, the predominant natural lifetime of the unsegmented PDHF$_{14}$-*b*-PEG$_{227}$ nanofibers (with average length 435 nm) in solution is 430 ps (fig. S24), shorter than for the corresponding unimer (700 ps), as expected for a J-aggregate. Figure 3A shows the transient grating PL time slices of the segmented nanofiber with an A-segment core length of 775 nm. We see a reduction of the PDHF emission (~480 nm) and a concurrent rise in the broader QPT emission at longer wavelengths (530 to 630 nm). Figure 3B shows the kinetics of this transfer. We note that the QPT emission must be integrated over a broad wavelength range to account for the slow energy transfer that excitations undergo within the QPT itself (fig. S25).

We modeled the kinetics of the PDHF and QPT PL with a 1D diffusion model (*25*) that includes a contribution from Förster resonance energy transfer at the ends of the PDHF nanofiber into the QPT (see supplementary materials). For the TCSPC data (fig. S26) on the 775-nm A-segment length nanofiber solution, the best fit gives a diffusion length of L_{D} = 210 ± 100 nm and a diffusion constant of $D = L_{\mathrm{D}}^2/2\tau$ = 0.5 ± 0.2 cm^2 s^{-1}, with errors estimated from the polydispersity of the nanofiber solution and robustness of the fit. Such an exciton diffusion constant is higher than any currently reported for organic semiconductors (*1*).

Figure 4, A and B, shows the transient grating PL kinetics for the PDHF and QPT for a range of shorter segmented nanofibers as a function of A-segment length. We see faster quenching of PDHF emission and faster rises in the QPT emission as the A-segment length is decreased. To fit the transient grating data, and hence a larger sample of quenching lengths than are available to TCSPC measurements, we add a time-dependent exciton-exciton annihilation term to the diffusion model (see supplementary materials and figs. S27 and S28) to account for second-order decay at the higher excitation densities used in this measurement (between ~5 × 10^{17} and ~10 × 10^{17} cm^{-3}). Using a global fit over multiple PDHF_{14}-*b*-PEG_{227} segment lengths and fluences (fig. S29), we obtain a best fit to our data for an exciton diffusion length of 380 nm. A residual analysis (fig. S30) confirms the robustness of these values for L_D greater than ~150 nm and less than ~600 nm. This agrees within error with the TCSPC result, adding support to our initial observation of large diffusion coefficients. We consider that the diffusion length most likely falls toward the smaller end of these values, because of the possibility of a small amount of exciton-charge annihilation in the higher fluence measurements, as well as the potential for a small amount of intermolecular energy transfer (fig. S31). Nonetheless, these results taken together indicate an exceptional diffusion constant for a conjugated polymer structure.

In conjugated polymers, efficient exciton diffusion is correlated with an increased degree of structural order (*4*). Our results are consistent with this, as corroborated by the WAXS data (fig. S12), which show appreciable structural order in the unsegmented PDHF_{14}-*b*-PEG_{227} nanofibers. This is supported by the pronounced vibronic structure and narrow spectral linewidths in the PL. Ultrafast transient grating PL measurements show evidence of both excitonic movement (during the instrument response time, ~200 fs) and some small degree of localization [in the first 700 fs after photoexcitation (Fig. 4C and fig. S32)]. We believe that this arises from rapid migration of excitons from the disordered to the ordered regions (evident from the WAXS data) and that exciton diffusion then occurs within these ordered regions in the nanofiber core.

Further understanding can be gained from the self-Förster radius, which is a useful tool for quantifying exciton diffusion lengths (*26*, *27*), although the description has limitations (*28*, *29*). We calculate a self-Förster radius in our PDHF fibers of 2.5 ± 0.2 nm (see supplementary materials). This is large in part because the parallel alignment of polymer chains gives ideal dipolar orientations for energy transfer (we have set κ^2, our dipole orientation factor, to 1). From this radius we can calculate (*27*) a diffusion length, $L_D = R_0^3/d^2 = 75 \pm 15$ nm, where d is the nearest-neighbor distance of 0.46 nm (fig. S12). Applying a Förster theory beyond the point-dipole approximation (*29*) would likely reduce this number slightly; however, this gives an estimate of expected exciton diffusion length for an incoherent, nearest-neighbor hopping regime. The fact that we measure diffusion lengths beyond this value is clear evidence that excitons are not hindered by the presence of localization sites, nor by local energy minima (*30*) in their transport. This provides evidence of a remarkably uniform energy landscape in these materials, enabled by their structural order. This uniformity would also benefit transport properties through a narrowing of the excitonic density of states (*31*).

Our measurement of a diffusion length greater than our estimate for nearest-neighbor Förster transport implies that we must look beyond this picture to explain our data. An example of a model that includes some degree of coherence but reduces to diffusive Förster transport as a limiting case is that proposed by Barford and Duffy (*31*). Their model uses a modified Redfield equation to show that a degree of interchain coherence in a polyfluorene film leads to an increase in the mean exciton hopping range, and hence gives a larger diffusion length. This requires interchain overlap, and the factor of 2.5 increase needed to account for our measured diffusion range gives a physically reasonable energy transfer integral of around 20 meV. Therefore, our measured exciton diffusion rate is high because of a combination of a physical packing structure optimizing the self-Förster radius, a lack of energetic trap sites, and a coherent component to exciton motion.

We have synthesized segmented nanofibers of controlled length with a PDHF core and containing a QPT corona at each terminus. Spectroscopic measurements in solution show that these nanostructures exhibit long-range exciton transport on the critical length scale comparable to

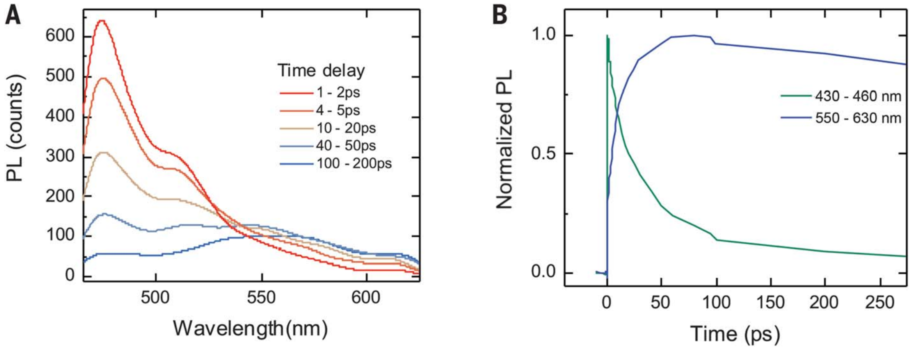


Fig. 3. Transient grating PL spectra and kinetics of segmented PDHF B-A-B nanofibers. (**A**) Transient grating PL time slices of a segmented nanofiber solution (0.5 mg/ml) with an average A-segment length of 775 nm, showing energy transfer from the PDHF to the QPT acceptor corona. There is a decay of the core PDHF I_{0-1} peak (which appears at 480 nm because of a filter cutting off the blue edge of the spectrum) due to exciton annihilation and quenching to the acceptor. This is accompanied by a concurrent rise in the broad QPT PL peak from 550 to 630 nm in the first tens of picoseconds (materials are excited with a 200-fs laser pulse at 400 nm with an equivalent excitation density of ~5 × 10^{17} cm^{-3}). (**B**) Normalized PL kinetics of PDHF decay and rise of QPT signal for the spectra shown in (A). The green line shows the PDHF signal (integrated from 430 to 460 nm); the blue line shows the QPT PL (integrated from 530 nm to 630 nm). The solvent used was THF:MeOH 1:1.

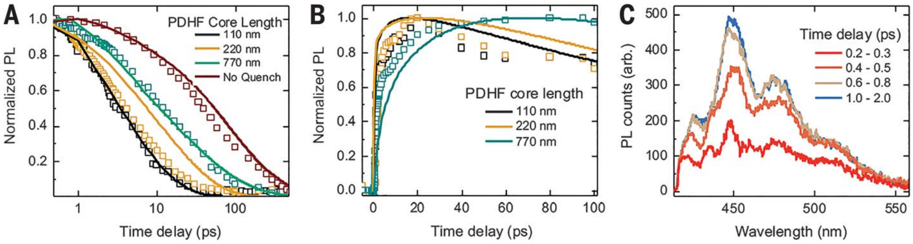


Fig. 4. Size-dependent transient PL kinetics and corresponding diffusion length model fits. (**A**) Transient grating PL kinetics (squares) of PDHF PL signal (integrated from 430 to 460 nm) in segmented PDHF B-A-B nanofibers of different A-segment lengths. The PL decay time decreases with decreasing segment length, showing efficient transfer. Solid lines are example fits of a 1D diffusion model with a diffusion length of L_D = 340 nm. The system was excited with a 400-nm, 200-fs laser pulse, at ~10^{18} cm^{-3} equivalent excitation densities for the samples in solution. (**B**) The corresponding PL kinetics (squares) of the rise in the QPT signal in the segmented PDHF B-A-B nanofibers, fitted with the same 1D exciton diffusion model (solid lines) and diffusion length in the PDHF of L_D = 340 nm. (**C**) Transient grating spectra of PDHF-*b*-PEG nanofibers in solution. Time slices show spectra at early times after excitation, with excitation density of ~5 × 10^{17} cm^{-3}. Spectral red-shifting is present until ~200 fs, and the ratio of the first and second vibronic peaks continues to reduce until ~700 fs (fig. S32). The solvent used was THF:MeOH 1:1.

the optical absorption length in conjugated polymers, and that this is enabled by the high degree of structural order in the PDHF core. In context, a 200-nm-thick film (the depth of our diffusion length) of conjugated polymer of average absorption coefficient (*32*) would absorb 98% of incoming photons. Such diffusion lengths could enable light-harvesting devices that use these polymer structures as antennae coupled to photodetector materials of limited absorption [such as monolayer transition metal dichalcogenides (*33*)] and would also enable much simpler bilayer design of organic photovoltaics relative to those based on the bulk heterojunction.

REFERENCES AND NOTES

1. O. V. Mikhnenko, P. W. M. Blom, T.-Q. Nguyen, *Energy Environ. Sci.* **8**, 1867–1888 (2015).
2. J. J. M. Halls *et al.*, *Nature* **376**, 498–500 (1995).
3. M. Pope, C. E. Swenberg, *Electronic Processes in Organic Crystals and Polymers* (Oxford Univ. Press, ed. 2, 1999).
4. R. R. Lunt, J. B. Benziger, S. R. Forrest, *Adv. Mater.* **22**, 1233–1236 (2010).
5. V. Bulović, S. R. Forrest, *Chem. Phys. Lett.* **238**, 88–92 (1995).
6. P. Peumans, A. Yakimov, S. R. Forrest, *J. Appl. Phys.* **93**, 3693–3723 (2003).
7. A. T. Haedler *et al.*, *Nature* **523**, 196–199 (2015).
8. H. Lin *et al.*, *Nano Lett.* **10**, 620–626 (2010).
9. F. Dubin *et al.*, *Nat. Phys.* **2**, 32–35 (2006).
10. E. Collini, G. D. Scholes, *Science* **323**, 369–373 (2009).
11. J. Vogelsang, T. Adachi, J. Brazard, D. A. Vanden Bout, P. F. Barbara, *Nat. Mater.* **10**, 942–946 (2011).
12. X. Wang *et al.*, *Science* **317**, 644–647 (2007).
13. H. Qiu *et al.*, *Science* **352**, 697–701 (2016).
14. D. Tao *et al.*, *J. Am. Chem. Soc.* **139**, 7136–7139 (2017).
15. M. C. Arno *et al.*, *J. Am. Chem. Soc.* **139**, 16980–16985 (2017).
16. W. Zhang *et al.*, *Science* **334**, 340–343 (2011).
17. S. Ogi, K. Sugiyasu, S. Manna, S. Samitsu, M. Takeuchi, *Nat. Chem.* **6**, 188–195 (2014).
18. X. Li *et al.*, *Nat. Commun.* **8**, 15909 (2017).
19. A. Monkman, C. Rothe, S. King, F. Dias, *Adv. Polym. Sci.* **212**, 187–225 (2008).
20. C. C. Kuo, C. H. Lin, W. C. Chen, *Macromolecules* **40**, 6959–6966 (2007).
21. C. H. Lin, Y. C. Tung, J. Ruokolainen, R. Mezzenga, W. C. Chen, *Macromolecules* **41**, 8759–8769 (2008).
22. S. M. King, H. L. Vaughan, A. P. Monkman, *Chem. Phys. Lett.* **440**, 268–272 (2007).
23. F. C. Spano, C. Silva, *Annu. Rev. Phys. Chem.* **65**, 477–500 (2014).
24. N. J. Hestand, F. C. Spano, *Acc. Chem. Res.* **50**, 341–350 (2017).
25. S. R. Scully, M. D. McGehee, *J. Appl. Phys.* **100**, 034907 (2006).
26. S. M. Menke, R. J. Holmes, *ACS Appl. Mater. Interfaces* **7**, 2912–2918 (2015).
27. J. D. A. Lin *et al.*, *Mater. Horiz.* **1**, 280–285 (2014).
28. H. Wiesenhofer *et al.*, *Adv. Funct. Mater.* **15**, 155–160 (2005).
29. G. D. Scholes, *Annu. Rev. Phys. Chem.* **54**, 57–87 (2003).
30. S. Kilina *et al.*, *ACS Nano* **2**, 1381–1388 (2008).
31. W. Barford, C. D. P. Duffy, *Phys. Rev. B* **74**, 075207 (2006).
32. M. S. Vezie *et al.*, *Nat. Mater.* **15**, 746–753 (2016).
33. M. Bernardi, M. Palummo, J. C. Grossman, *Nano Lett.* **13**, 3664–3670 (2013).

ACKNOWLEDGMENTS

We thank A. Chin for helpful discussions, R. Harniman for AFM measurements, and D. Hayward and O. Gould for WAXS experiments. **Funding:** Supported by EPSRC grants EP/K017799/1 (I.M.), EP/K016520/1 (R.H.F.), and EP/M005143/1 (R.H.F. and M.B.P.) and by a KAUST Competitive Research Grant (S.M.M. and R.H.F.). **Author contributions:** The project was devised by X.-H.J., M.B.P., R.H.F., G.R.W., and I.M., and these authors contributed to discussion of the data and writing of the manuscript; S.M.M. and A.R. contributed to the manuscript and project conception; X.-H.J., M.B.P., J.M.R., J.R.F., and C.E.B. performed the experimental work. **Competing interests:** The authors have no competing interests. **Data and materials availability:** All data are available in the manuscript or the supplementary materials.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/360/6391/897/suppl/DC1
Materials and Methods
Figs. S1 to S32
Table S1
References (*34–53*)

20 December 2017; accepted 13 April 2018
10.1126/science.aar8104

COLD MOLECULE PHYSICS

Building one molecule from a reservoir of two atoms

L. R. Liu,[1,2,3] J. D. Hood,[1,3] Y. Yu,[1,2,3] J. T. Zhang,[1,2,3] N. R. Hutzler,[1,2,3]*
T. Rosenband,[2] K.-K. Ni[1,2,3]†

Chemical reactions typically proceed via stochastic encounters between reactants. Going beyond this paradigm, we combined exactly two atoms in a single, controlled reaction. The experimental apparatus traps two individual laser-cooled atoms [one sodium (Na) and one cesium (Cs)] in separate optical tweezers and then merges them into one optical dipole trap. Subsequently, photoassociation forms an excited-state NaCs molecule. The discovery of previously unseen resonances near the molecular dissociation threshold and measurement of collision rates are enabled by the tightly trapped ultracold sample of atoms. As laser-cooling and trapping capabilities are extended to more elements, the technique will enable the study of more diverse, and eventually more complex, molecules in an isolated environment, as well as synthesis of designer molecules for qubits.

Chemical reactions proceed through individual collisions between atoms or molecules. However, when performed in stochastic ensembles, the individual reaction probabilities are observed as averages. Crossed molecular beams reduce the thermal velocity dispersion to probe elementary reaction processes based on single collision events, illuminating many aspects of reaction dynamics (*1*–*4*). In quantum degenerate gases, cooled to temperatures <1 μK, the quantum motional degrees of freedom play a critical role in the reaction (*5*–*7*). Comparisons of such experimental reaction rates with theoretical models currently underpin our understanding of reactions at the most elementary level (*8*–*10*).

To further improve the specificity and precision of reaction steps (*11*–*13*), individual particle control is needed, similar to pioneering atom-positioning experiments with scanning tunneling microscopes (*14*), but untethered from surfaces. By controlling individual particles via laser cooling and optical trapping, molecules may be constructed atom by atom while maintaining specific internal and external quantum states.

We realized chemistry in the minimum number regime, where precisely two atoms are brought together to form one molecule with the aid of a photon. We achieved this using movable optical tweezers, in which individual atoms of different elements [in this work, sodium (Na) and cesium (Cs)] are isolated, cooled, manipulated, and eventually combined into a single optical tweezer. With exactly two atoms in an optical tweezer, we can observe their collisions. We can also perform single-molecule spectroscopy in the gas phase by optically exciting the atom pair on a molecular transition, realizing the chemical reaction Na + Cs → NaCs*. Subsequent imaging of Na and Cs fluorescence distinguishes between four possible experimental outcomes: Both, only one, or no atoms are detected in the tweezer, the latter indicating that a reaction has occurred. We chose NaCs for the demonstration because it possesses a large molecular fixed-frame dipole moment of 4.6 D (*15*), making it a strong candidate for a molecular qubit in a future quantum computing architecture.

We began by preparing laser-cooled Na and Cs atoms at a few hundred microkelvin in overlapped magneto-optical traps (MOTs) in a vacuum chamber (10^{-8} Pa). The MOTs serve as cold atom reservoirs for loading single atoms into tightly focused optical tweezer traps (*16*). After loading, the MOTs are extinguished. A schematic of the apparatus is shown in Fig. 1A. The numerical aperture = 0.55 microscope objective focuses two different wavelengths of light, 700 and 976 nm, to waists of 0.7 and 0.8 μm radius. Because of the difference in Na and Cs polarizabilities, the 700-nm-wavelength light attracts Na and repels Cs, whereas 976-nm-wavelength light attracts Cs five times more strongly than does Na (*17*), enabling us to manipulate Na and Cs independently (Fig. 2A). A typical trap depth of 1 mK is achieved for 5 mW of tweezer power.

When tightly confined identical atoms are illuminated with near-atomic-resonant light, light-assisted pairwise collisions result in either zero or one final atom in the trap (*16*, *18*). Single-atom loading succeeds approximately half of the time (*19*). However, the large light shifts for Na in a 700-nm-wavelength tweezer would normally prevent atom cooling and, consequently, efficient atom loading. We eliminated this light shift for Na by alternating the tweezer and cooling beams at a rate of 3 MHz (*20*). Subsequently, Na followed by Cs were imaged, and the polarization gradient was cooled to 70 and 10 μK, respectively. To determine whether an atom is in the optical tweezer, the fluorescence photoelectron counts from each atom in a region of interest (Fig. 1B) are compared with a threshold (Fig. 1C). The fluorescence histograms indicate that the cases of zero or one atom can be distinguished with a fidelity better than 99.97%. We found that in 33% of cases, we loaded a single Na and a single Cs atom side by side. In 18% of cases, no atoms were loaded, and the rest of the time, either a single Na or Cs atom was loaded (Fig. 1B). The experiment, which repeats at 3 Hz, records initial and final fluorescence images in order to determine survival probabilities for different stages of the molecule-formation process.

Once single atoms have been loaded in separate traps, they need to be transported to the same location for molecule formation. Optical tweezers have been used to move single atoms while maintaining atomic internal state coherence (*21*) and to merge two indistinguishable atoms through coherent tunneling into one tweezer (*22*). We adiabatically transported and merged two different atoms, Na and Cs, into the same tweezer by using optical tweezers at two different wavelengths (Fig. 2A). The trap depths were adjusted by changing the beam intensities, and the positions were steered by applying different radio frequencies to the respective acousto-optic deflectors (AODs) (Fig. 1A).

For the merge sequence, the 700-nm tweezer containing Na was kept stationary, while the 976 nm tweezer containing Cs was moved to overlap the atoms (Fig. 2A, I to III). After the merge, the 700-nm tweezer was extinguished adiabatically so as to leave both atoms in the 976-nm tweezer (Fig. 2A, IV). We designed this merge trajectory so that (i) Cs is deeply confined at all times and (ii) the double-well potential imposed on Na is sufficiently asymmetric to avoid a near-degenerate ground state. This process is time-reversible, which enables us to image the atoms separately and determine survival probability.

Because the 700-nm tweezer is extinguished for 1 ms after the merge, while the 976-nm tweezer is always active, the Na atom escapes, unless the two tweezers are overlapped at the end of the merge sequence, whereas the Cs atom is always trapped. The result obtained when scanning the endpoint of the 976-nm tweezer trajectory is shown in Fig. 2B. The height of the Na survival peak at 0 μm of 94(1)% is near the reimaging survival probability of 96% (numbers in parentheses indicate the Wilson score interval).

Having demonstrated adiabatic transport and merging of two species into a tight tweezer, we turned to their collisions. Isolated collisions between two atoms do not usually result in molecule formation because of the need to simultaneously conserve momentum and energy. However, the atoms can change their hyperfine states after colliding, and the exothermic hyperfine-spin-changing collisions impart enough kinetic energy (≈100 mK) to the atoms to eject them from the tweezer (≈1 mK depth) (*23*).

---

[1] Department of Chemistry and Chemical Biology, Harvard University, Cambridge, MA 02138, USA. [2] Department of Physics, Harvard University, Cambridge, MA 02138, USA. [3] Harvard-MIT Center for Ultracold Atoms, Cambridge, MA 02138, USA.
*Present address: Division of Physics, Mathematics, and Astronomy, California Institute of Technology, Pasadena, CA 91125, USA
†Corresponding author. Email: ni@chemistry.harvard.edu

Generally, a given initial trap occupancy can evolve into four possible outcomes after an experiment: (i) both atoms, (ii) no atoms, (iii) only Cs, and (iv) only Na remain in the trap. Single-atom images from each repetition allow us to post-select on any of these cases and separate one- and two-body processes, giving both lifetimes from a single data set (Fig. 3). For example, when Na and Cs are both present (effective pair density of $n_2 = 2 \times 10^{12}$ cm^{-3}) (*24*), and prepared in a mixture of hyperfine spin states, they are both rapidly lost: $\tau_{\text{loss}} = 8(1)$ ms, where τ_{loss} is the 1/*e* time of exponential decay. This yields a loss rate constant $\beta = 5 \times 10^{-11}$ cm^3/s. By contrast, if the atoms are both optically pumped into the lowest-energy hyperfine levels, conservation of energy prevents the change of hyperfine states, and the atom lifetime increases to 0.63(1) s, which is similar to the rate of hyperfine-state relaxation for Cs owing to off-resonant scattering of the tweezer light (*25*). When only one atom is present, one-body loss due to collisions with background gas limits the lifetime to 5 s.

Because of the rapid two-body loss for mixed hyperfine states, we optically pumped each atom into its lowest-energy hyperfine state in order to maintain a long-lived sample of cotrapped Na and Cs atoms. We then performed photoassociation (PA) of the atoms in order to form an excited state molecule, realizing a single instance of the chemical reaction Na + Cs → NaCs*. When illuminating the atoms with resonant PA light, an electronically excited state molecule may form (Fig. 4A) and then rapidly decay to the ground state. The molecule does not scatter imaging light, causing molecule formation to manifest as simultaneous loss of both Na and Cs atoms. These loss resonances are shown in Fig. 4B, bottom, as the frequency of the PA light is scanned below the dissociation threshold.

Our optical tweezer architecture offers a number of advantages for PA measurements over previous methods with bulk samples (*26*). The ability to precisely define the initial reagents eliminates contributions from other reaction processes such as Cs_2 formation or three-body loss. The combination of the high effective pair density $n_2 = 3 \times 10^{12}$ cm^{-3} (*24*), afforded by the tweezer confinement, and high-PA light intensity of 3 kW/cm^2 yields fast PA rates. The high-contrast measurements of single-atom loss result in near-unity molecule detection efficiency and avoid the need for ionization detection (*26*).

We scanned the 200-MHz frequency–broadened PA light from 30 to 250 GHz below the Cs atomic D2 line ($6S_{1/2} - 6P_{3/2}$). We took steps of 200 MHz with 100-ms pulse duration and took an average of approximately 100 repetitions at each data point. An absolute accuracy of 1 GHz was set by the wavemeter. During PA, the Cs atom could be promoted into the upper hyperfine level because of off-resonant scattering of the PA beam, which would lead to spin-changing collisional loss. We counteracted this effect by simultaneously optically pumping Cs into the lower hyperfine level with a separate beam.

The ability to detect molecule formation via atom loss with high efficiency allowed us to probe NaCs* vibrational levels near the dissociation threshold, including resonances that have not been previously observed (Fig. 4). According to ab initio calculations of NaCs* with spin-orbit coupling (*27*), five molecular potentials converge to the Cs ($6P_{3/2}$) + Na ($3S_{1/2}$) asymptote (Fig. 4A): $B^1\Pi_1$, $c^3\Sigma^+_{\Omega=0,1}$, and $b^3\Pi_{\Omega=0,2}$. Of these, only the $c^3\Sigma^+_1$ levels have previously been observed in the near-threshold regime (*28*), and our measurement agrees to within 1 GHz. To identify the vibrational progressions, we fit the LeRoy-Bernstein (LB) dispersion model (*29*) to our observed resonances. Near threshold, the vibrational quantum number v' ($v' = -1$ is the highest bound state) is related to the binding energy by

$$E_{v'} = -\frac{1}{C_6^{1/2}}\left[2\hbar\left(\frac{2\pi}{\mu}\right)^{1/2}\frac{\Gamma(7/6)}{\Gamma(2/3)}(v' - v'_0)\right]^3 \tag{1}$$

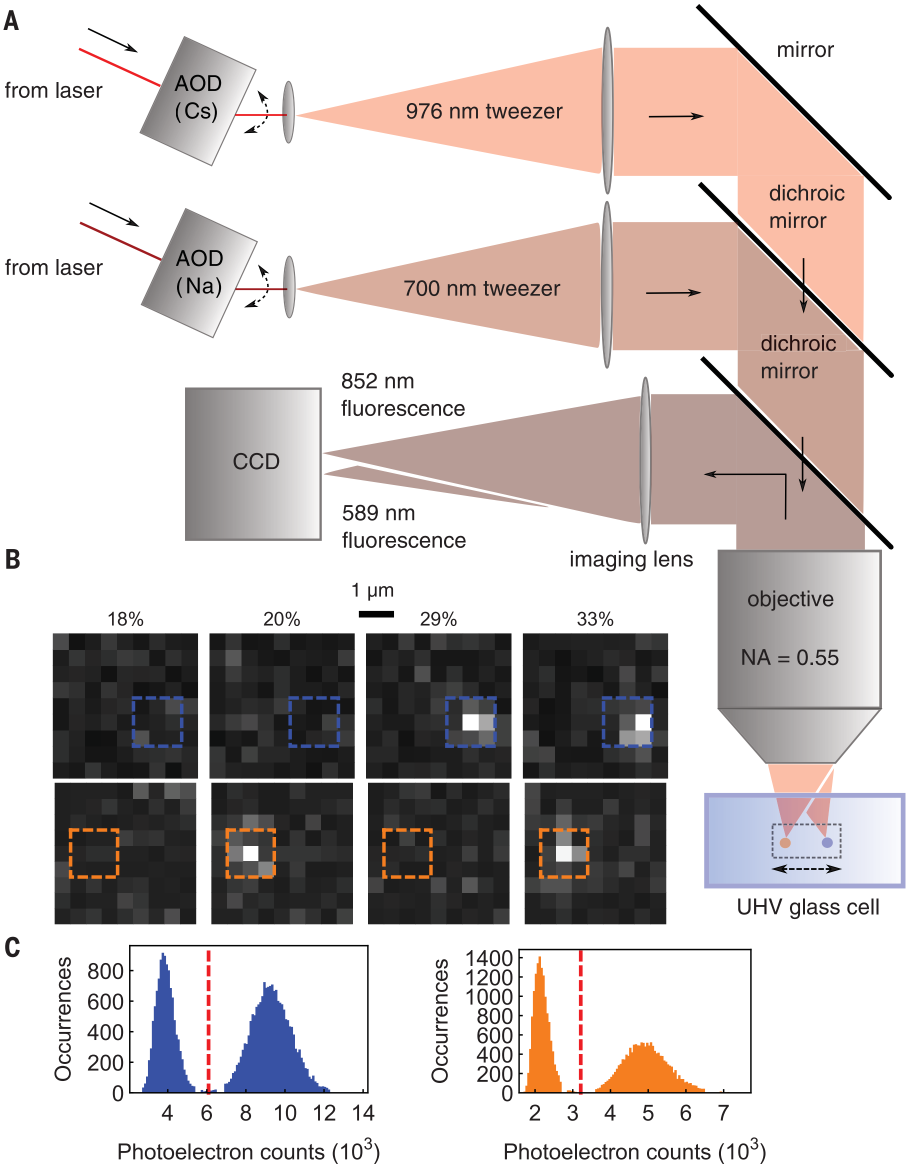


Fig. 1. Dual-species single-atom trapping and imaging. (**A**) Schematic of the setup. Optical tweezer atom-trapping beams (700 and 976 nm wavelengths) are independently steered by AODs, expanded by telescopes, and then combined on a dichroic mirror before being focused by the objective into a glass cell. Fluorescence from trapped Na and Cs atoms is collected through the objective onto the charge-coupled device camera. (**B**) Fluorescence images of single Na and Cs atoms. Length scale of 1 μm is indicated. Cs (top) and Na (bottom) are imaged sequentially in the same field of view. The four possible cases are shown with their initial loading probabilities: no atoms, a single Na atom, a single Cs atom, or both Na and Cs atoms. Dashed blue (Cs) and orange (Na) boxes indicate the regions of interest for determining the presence of atoms. (**C**) Histogram of Cs (blue) and Na (orange) fluorescence. The bimodal distribution shows clear separation between zero- and one-atom peaks. Red dashed lines indicate the threshold that is used to determine the presence of an atom.

Fig. 2. Merging single Na and Cs atoms, which are initially separated by 3 μm, into one tweezer. (**A**) One-dimensional cuts of the combined, time-varying 700- and 976-nm tweezer potentials for both atoms during the merge sequence. Na and Cs are represented by dots that track the minima of their potentials (orange, Na; blue, Cs). Overlaid are graphics of the optical tweezers. Radial trap frequencies are labeled in the first and last panels (axial trap frequencies are roughly 6 times smaller). I to III depict the merging process. In IV, the 700-nm tweezer has been extinguished, and only the 976-nm tweezer remains. (**B**) Measured survival probability of Na and Cs after the sequence depicted in (A), followed by separating the tweezers through a reverse sequence to image the atoms. The two atoms are merged into the same tweezer at the survival maximum for Na. Error bars denote the Wilson score interval. The dashed lines represent the survival rates due to imperfect reimaging.

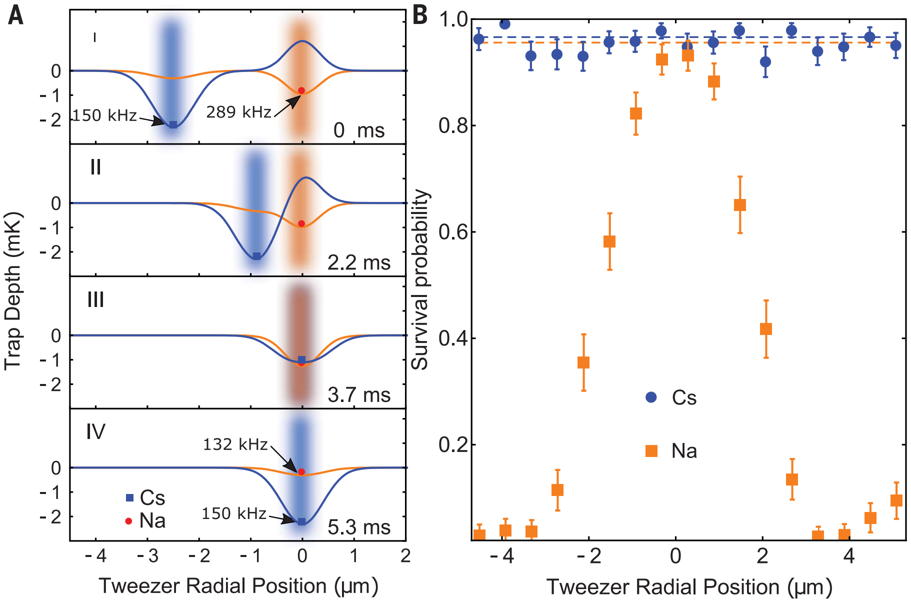


Fig. 3. Collisions of Na and Cs. The hold time in the merged trap is varied so as to measure the evolution of trap occupancy owing to various collision mechanisms. Post-selection on initial and final trap occupancies allows us to distinguish one- and two-body processes. The fastest time scales are indicated next to the thick fitted curves. The fits are explained in the supplementary materials. (**Left**) For both atoms in a mixture of hyperfine states, the loss is dominated by rapid two-body hyperfine-state-changing collision-induced loss. (**Middle**) For both atoms in their lowest hyperfine states, the loss is explained by two-body hyperfine-state-changing collisions that follow off-resonant scattering of trap light. In both the left and middle, different markers denote the final trap occupancy. (**Right**) One-body loss gives background gas limited lifetime of ~5 s for both atoms. We post-select on empty final tweezers, and markers denote initial trap occupancy.

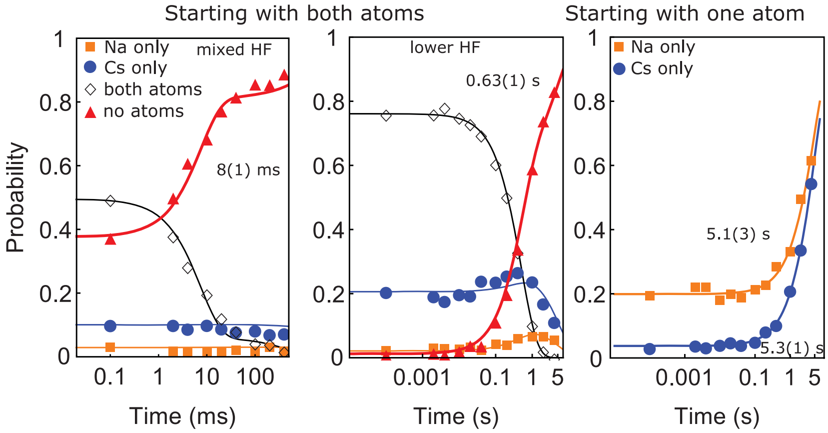
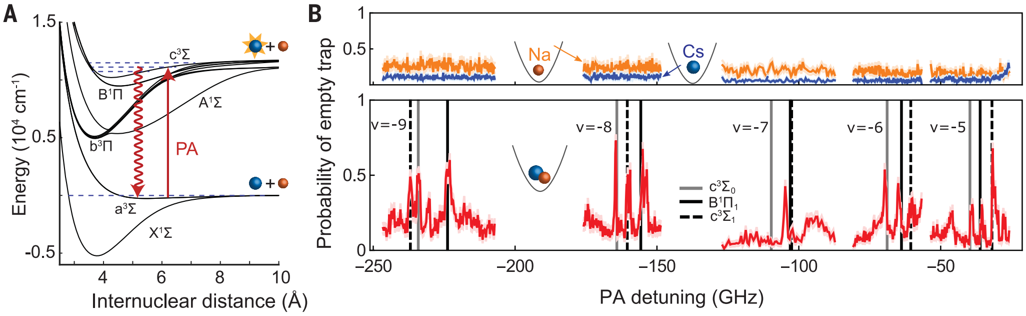


Fig. 4. PA spectroscopy of NaCs*. (**A**) NaCs molecular potentials as a function of internuclear distance (27). PA light excites the ground-state atoms to vibrational levels of the NaCs* excited molecular potentials, from which they mostly decay to vibrationally excited electronic ground-state molecules (red wavy line). The long-range asymptotes of the excited state potentials (dominated by van der Waals interactions in the heteronuclear molecules) correspond to one of two cases: ground-state Na colliding with excited Cs in either the lower-energy $6P_{1/2}$ (D1 line) or higher-energy $6P_{3/2}$ state (D2 line). (**B**) The probability of single Na (orange), single Cs (blue), and joint Na$^+$Cs (red) atoms evolving to the "no atoms" detection channel as the PA light is detuned from the Cs D2 line dissociation threshold at 351,730 GHz. When both atoms are initially loaded into the tweezer (bottom), two-body loss resonances appear because of molecule formation. As a validation of our method, we checked that no loss resonances were observed when only one atom was present (top). The positions of the loss resonances are fitted with the LB dispersion model in Eq. 1 in order to identify three different potentials and fit the respective C_6 dispersion coefficients. The expected resonance positions based on these fits are marked by vertical lines, as indicated in the legend. Except for $v = -7$, the root mean square deviation of the fitted dispersion curve from the measured frequencies are 0.3, 0.6, and 0.8 GHz for the $c^3\Sigma_1$, $c^3\Sigma_0$, and $B^1\Pi_1$ states, respectively. At $v = -7$, a crossing of molecular energy levels causes the measured spectrum to deviate from the prediction according to Eq. 1. Unassigned lines in the spectrum are likely due to rotational and hyperfine structure and predissociating potentials.



where μ is the reduced mass, and $\hbar$ is the reduced Planck's constant. We extracted the C_6 dispersion coefficients that characterize the $1/r^6$ component of the potentials and v'_0, which is an offset between –1 and 0.

Fitting to the positions of our observed $c^3\Sigma_1^+$ resonances gives $v_0 = -0.79$ and $C_6 = 8.5(6) \times 10^3$ au (atomic units), which is in agreement with the theoretical value $C_6 = 7.96 \times 10^3$ au (*30*). From the remaining loss resonances, we identified two additional progressions ($B^1\Pi_1$ and $c^3\Sigma_0^+$) with $C_6 = 1.42(33) \times 10^4$ au and $C_6 = 1.47(26) \times 10^4$ au (Fig. 4B). Both values are near the theoretical value of $C_6 = 1.83 \times 10^4$ au (*30*). Our state labels correspond to the molecular wave functions in the near-threshold regime and differ from the labels in (*30*) because of an avoided crossing, as noted in (*31*). Here, the assignment of the $c^3\Sigma_1^+$ progression is based on previous observation of the same resonances (*28*), whereas $B^1\Pi_1$ continues a previously observed sequence (*31*). The remaining progression corresponds to $c^3\Sigma_0^+$ because this is the only other compatible state. We interpret the PA spectrum as clear evidence for molecule formation because the resonance peaks appear exclusively as simultaneous loss of Na and Cs, and the resonance frequencies agree with independent measurements.

Our technique can in principle be extended beyond the simple bialkalis demonstrated here and to produce deeply bound molecules. Molecules in a single-quantum state could be created through coherent transfer (*32*, *33*) of atoms prepared in the motional ground state (*34*–*37*). Dipolar molecules trapped in a configurable array of optical tweezers (*38*, *39*) would constitute a new type of qubit for quantum information processing (*40*) and an important resource with which to explore quantum phases (*41*, *42*).

REFERENCES AND NOTES

1. D. R. Herschbach, *Angew. Chem. Int. Ed. Engl.* **26**, 1221–1243 (1987).
2. Y. T. Lee, *Angew. Chem. Int. Ed. Engl.* **26**, 939–951 (1987).
3. A. B. Henson, S. Gersten, Y. Shagam, J. Narevicius, E. Narevicius, *Science* **338**, 234–238 (2012).
4. W. E. Perreault, N. Mukherjee, R. N. Zare, *Science* **358**, 356–359 (2017).
5. S. Ospelkaus *et al.*, *Science* **327**, 853–857 (2010).
6. K.-K. Ni *et al.*, *Nature* **464**, 1324–1328 (2010).
7. M. H. G. de Miranda *et al.*, *Nat. Phys.* **7**, 502–507 (2011).
8. K. Liu, *Annu. Rev. Phys. Chem.* **52**, 139–164 (2001).
9. X. Yang, *Annu. Rev. Phys. Chem.* **58**, 433–459 (2007).
10. A. Klein *et al.*, *Nat. Phys.* **13**, 35–38 (2016).
11. L. Ratschbacher, C. Zipkes, C. Sias, M. Köhl, *Nat. Phys.* **8**, 649–652 (2012).
12. S. A. Moses *et al.*, *Science* **350**, 659–662 (2015).
13. P. Puri *et al.*, *Science* **357**, 1370–1375 (2017).
14. D. M. Eigler, E. K. Schweizer, *Nature* **344**, 524–526 (1990).
15. P. J. Dagdigian, L. Wharton, *J. Chem. Phys.* **57**, 1487–1496 (1972).
16. N. Schlosser, G. Reymond, I. Protsenko, P. Grangier, *Nature* **411**, 1024–1027 (2001).
17. M. S. Safronova, B. Arora, C. W. Clark, *Phys. Rev. A* **73**, 022505 (2006).
18. A. Fuhrmanek, R. Bourgain, Y. R. P. Sortais, A. Browaeys, *Phys. Rev. A* **85**, 062708 (2012).
19. P. Sompet, A. V. Carpentier, Y. H. Fung, M. McGovern, M. F. Andersen, *Phys. Rev. A* **88**, 051401 (2013).
20. N. R. Hutzler, L. R. Liu, Y. Yu, K.-K. Ni, *New J. Phys.* **19**, 023007 (2017).
21. J. Beugnon *et al.*, *Nat. Phys.* **3**, 696–699 (2007).
22. A. M. Kaufman *et al.*, *Science* **345**, 306–309 (2014).
23. B. Ueberholz, S. Kuhr, D. Frese, D. Meschede, V. Gomer, *J. Phys. At. Mol. Opt. Phys.* **33**, L135–L142 (2000).
24. Materials and methods are available as supplementary materials.
25. R. A. Cline, J. D. Miller, M. R. Matthews, D. J. Heinzen, *Opt. Lett.* **19**, 207 (1994).
26. K. M. Jones, E. Tiesinga, P. D. Lett, P. S. Julienne, *Rev. Mod. Phys.* **78**, 483–535 (2006).
27. M. Korek, S. Bleik, A. R. Allouche, *J. Chem. Phys.* **126**, 124313 (2007).
28. A. Grochola *et al.*, *Phys. Rev. A* **84**, 012507 (2011).
29. R. J. LeRoy, R. B. Bernstein, *J. Chem. Phys.* **52**, 3869–3879 (1970).
30. M. Marinescu, H. Sadeghpour, *Phys. Rev. A* **59**, 390–404 (1999).
31. A. Grochola, P. Kowalczyk, W. Jastrzebski, *Chem. Phys. Lett.* **497**, 22–25 (2010).
32. K. Bergmann, H. Theuer, B. W. Shore, *Rev. Mod. Phys.* **70**, 1003–1025 (1998).
33. L. R. Liu *et al.*, arXiv:1701.03121 [physics.atom-ph] (2017).
34. C. Monroe *et al.*, *Phys. Rev. Lett.* **75**, 4011–4014 (1995).
35. X. Li, T. A. Corcovilos, Y. Wang, D. S. Weiss, *Phys. Rev. Lett.* **108**, 103001 (2012).
36. A. M. Kaufman, B. J. Lester, C. A. Regal, *Phys. Rev. X* **2**, 041014 (2012).
37. Y. Yu *et al.*, arXiv:1708.03296 [physics.atom-ph] (2017).
38. D. Barredo, S. de Léséleuc, V. Lienhard, T. Lahaye, A. Browaeys, *Science* **354**, 1021–1023 (2016).
39. M. Endres *et al.*, *Science* **354**, 1024–1027 (2016).
40. D. DeMille, *Phys. Rev. Lett.* **88**, 067901 (2002).
41. N. Y. Yao, M. P. Zaletel, D. M. Stamper-Kurn, A. Vishwanath, *Nat. Phys.* **14**, 405–410 (2018).
42. B. Sundar, B. Gadway, K. R. Hazzard, *Sci. Rep.* **8**, 3422 (2018).

ACKNOWLEDGMENTS

We thank R. González-Férez, P. Julienne, D. DeMille, and C. Regal for discussions. K.-K.N. thanks D. S. Jin for encouragement to pursue the research presented here. **Funding:** This work is supported by the Arnold and Mabel Beckman Foundation as well as the AFOSR Young Investigator Program, the NSF through the Harvard-MIT CUA, and the Alfred P. Sloan Foundation. **Author contributions:** L.R.L., J.D.H., Y.Y., J.T.Z., N.R.H., T.R., and K.-K.N. performed the experiment. L.R.L., J.D.H., T.R., and K.-K.N. analyzed the data and wrote the manuscript. **Competing interests:** None declared. **Data and materials availability:** All data are supplied in the main paper and the supplementary materials.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/360/6391/900/suppl/DC1
Materials and Methods
Supplementary Text
Fig. S1
Table S1
Reference (*43*)

15 January 2018; accepted 3 April 2018
Published online 12 April 2018
10.1126/science.aar7797

PHOTOVOLTAICS

Flexo-photovoltaic effect

Ming-Min Yang,* Dong Jik Kim,* Marin Alexe†

It is highly desirable to discover photovoltaic mechanisms that enable enhanced efficiency of solar cells. Here we report that the bulk photovoltaic effect, which is free from the thermodynamic Shockley-Queisser limit but usually manifested only in noncentrosymmetric (piezoelectric or ferroelectric) materials, can be realized in any semiconductor, including silicon, by mediation of flexoelectric effect. We used either an atomic force microscope or a micrometer-scale indentation system to introduce strain gradients, thus creating very large photovoltaic currents from centrosymmetric single crystals of strontium titanate, titanium dioxide, and silicon. This strain gradient–induced bulk photovoltaic effect, which we call the flexo-photovoltaic effect, functions in the absence of a p-n junction. This finding may extend present solar cell technologies by boosting the solar energy conversion efficiency from a wide pool of established semiconductors.

Since its first observation in the 19th century, the photovoltaic (PV) effect has been studied intensively for scientific interest and as a sustainable energy source to replace fossil fuels and reduce carbon emissions (*1*–*3*). In 1954, the first high-power modern silicon solar cells—in which the photoexcited carriers were separated by a built-in electric field developed at a p-n junction—were invented (*4*). By virtue of the appropriate bandgap energies of the semiconductors in a p-n junction solar cell, sunlight is efficiently absorbed, resulting in considerable power conversion efficiency. At present, silicon PV cells are the mainstay of the modern solar industry, contributing more than 1% of the global electricity supply. Nevertheless, PV cells based on p-n junctions have a photovoltage limited by the bandgap energy of the constituent semiconductors and an ultimate efficiency constrained by the Shockley-Queisser (S-Q) limit (*5*).

Ferroelectric materials exhibit a PV effect, called the bulk photovoltaic (BPV) effect (*6*, *7*), that is distinct from that of p-n junctions. Under uniform illumination, a homogeneous ferroelectric material gives rise to a current under zero bias [short-circuit current (I_{SC})] that depends on the polarization state of the incident light and produces an anomalously large photovoltage well exceeding its bandgap energy. This peculiar PV effect originates from the asymmetric distribution of photoexcited nonequilibrium carriers in momentum space, caused by the absence of centrosymmetry in the material (*7*). Owing to its distinctive charge separation mechanism, solar cells based on the BPV effect can, in principle, exceed the S-Q limit (*8*). However, a substantial BPV effect is generally found in wide-bandgap noncentrosymmetric materials such as ferroelectric $BaTiO_3$ (*8*) and $BiFeO_3$ (*9*), leading to an overall extremely low device efficiency under solar illumination. Hence, one way to enhance the efficiency may be to realize the BPV effect in semiconductors with more favorable bandgaps, regardless of their crystalline symmetry.

Flexoelectricity is an electromechanical property that reflects a coupling between an electric polarization and a strain gradient (*10*, *11*). In centrosymmetric materials, a strain gradient breaks its inversion symmetry, resulting in a polarization with a preferred direction and enabling a piezoelectric composite containing no piezoelectric elements (*11*, *12*). Likewise, one can hypothesize that a strain gradient allows the manifestation of the BPV effect in materials that were originally centrosymmetric. In this study, we propose and demonstrate that the BPV effect can be induced in any semiconductor by mediation of the flexoelectric effect. Given that flexoelectricity is a universal property of all materials, ranging from biomaterials (*13*) to semiconductors and dielectrics (*14*) to two-dimensional materials such as graphene (*15*), this strain gradient–induced BPV effect, termed here the flexo-photovoltaic (FPV) effect, is possible for all symmetry classes. Thus, devices designed on the basis of the FPV effect can be fabricated with silicon or any other semiconductors.

To demonstrate our idea, we explored the PV effect induced by a point force exerted onto the surface of centrosymmetric materials, including a $SrTiO_3$ single crystal and a rutile TiO_2 single crystal [section S1 of (*16*)]. The point force was exerted by the tip of an atomic force microscope (AFM), inducing local inhomogeneous strain at the tip-surface contact area and, therefore, a local breaking of centrosymmetry (*17*, *18*). In our experiments, we used a custom-made photoelectric atomic force microscope (Ph-AFM) consisting of an AFM-based system equipped with a customized current amplifier-filter system and an optical system (*19*). The optical system allowed us to illuminate a sample surface with a 405-nm laser properly polarized by a half-wavelength plate. A conductive AFM tip applied a local force on a sample surface and simultaneously collected the resultant PV current. A brief schematic of the Ph-AFM setup is shown in Fig. 1, A and B.

$SrTiO_3$ single crystals are ideal for studying the flexoelectric effect, owing to their simple cubic centrosymmetric lattices and large dielectric permittivities (*14*, *20*). Unlike its sister material $BaTiO_3$, ordinary crystals of $SrTiO_3$ do not exhibit the BPV effect because they have a center of inversion symmetry. However, with illumination around the contact area on a (001) face of a $SrTiO_3$ crystal, we observed that I_{SC} exhibits a large transient peak as the loading force is increased from 1 to 18 μN (Fig. 1C). This peak is reproducible, proven by repeated exertion and withdrawal of the loading force.

To show that the pronounced enhancement of I_{SC} by the point force is not confined to $SrTiO_3$ crystals or a cubic structure, we investigated the force-induced PV current in a single crystal of rutile TiO_2, which is well known for its photoelectric applications such as dye-sensitized PV cells and photocatalysis. As in the case of $SrTiO_3$, large negative I_{SC} appears once a large force is exerted on a TiO_2 (100) face. Figure 1D shows a stable current under the 15-μN force, confirming that the force-induced PV effect is not a transient effect in this material. Although it depends on contact conditions and locations, the point force exerted by the AFM tip gives rise to a substantial current density (up to ~1 A/cm^2) at the nanoscale contact area. This current density is more than three orders of magnitude higher than the I_{SC} value (0.3 mA/cm^2) obtained from a Schottky junction between TiO_2 and Pt under the same illumination condition [section S6 of (*16*)].

The current density increases by more than a factor of 100 when the loading force is increased from 1 to 15 μN (Fig. 1E). This phenomenon cannot be explained by expansion of the contact area with the loading force, because a contact area increases by, at most, a factor of ~6 with a 15-fold increase in loading force in a simple elastic sphere contact model [section S2 of (*16*)]. Moreover, a 100-fold increase in contact area as a result of force applied by an AFM tip is not realistic. Notably, I_{SC} has a negative value on a (100) face of the TiO_2 crystal but becomes positive when the conductive AFM tip is loaded on a (001) face (Fig. 1F). The fact that the direction of the PV current depends on the crystallographic orientation of the TiO_2 crystal indicates that the observed PV effect cannot merely be attributed to a probable Schottky contact formed by the TiO_2 surface and the Pt coating of the AFM tip.

The BPV effect is a potential origin of this PV current enhancement. As hypothesized, it is likely that a point force exerted on a crystal surface generates a local strain gradient, resulting in local centrosymmetry breaking and, thus, a local BPV current under illumination—i.e., the FPV effect—in the absence of a p-n junction and an appropriate band alignment. Given that both the flexoelectric response and the BPV effect depend on the crystallographic orientation (*19*–*21*), our results suggest that the strain gradient and the resultant FPV effect have a predominant role in the enhanced local PV current.

The strain gradient induced by a sphere contact has a complex spatial distribution with very

---

Department of Physics, University of Warwick, Coventry CV4 7AL, UK.
*These authors contributed equally to this work.
†Corresponding author. Email: m.alexe@warwick.ac.uk

large strain gradient values in an elastic material, as described in section S2 of (*16*). An AFM tip apex can be approximated as a hemisphere, and the distribution of strain gradient induced by the AFM tip can be calculated analytically with the Hertzian model and the Boussinesq's calculation (*22*). Figure 2 shows the spatial distributions of the *z* component of the calculated strain and its derivative with respect to *z* under ~15.7 μN of force with a 10-nm–radius contact area at the origin. The strain gradient is as large as 10^7 m^{-1}. We found that the I_{SC} density and the volume beneath the AFM tip, which is subject to a strain gradient larger than 1×10^6 m^{-1}, show a similar dependence on the exerted force [(Fig. 1E); the detailed process to obtain the relative volume can be found in fig. S9]. It is expected that the very large strain gradients induced by the AFM tip will break local symmetry, thus leading to the manifestation of the BPV effect locally under illumination. However, deep theoretical considerations are required to understand the intricate relationship between the complex distribution of the strain gradient and the BPV properties.

The main characteristic of the BPV effect is a periodic dependence on the angle between the PV current and the light polarization, stemming from its tensorial nature (*21*). For the present case, this dependence is predicted to be

$$I_{FPV} = \frac{\pi}{2} I_0 (A_z + B_z \cos 2\alpha) \qquad (1)$$

Where I_{FPV} is the PV current, I_0 is a light intensity, A_z and B_z are effective BPV coefficients of the locally deformed crystal, and α is the polarization angle of the incident light with respect to the top surface edge as described in section S3 of (*16*). The FPV effect should inherit the distinctive feature of the BPV effect and exhibit a sinusoidal dependence on the incident light polarization angle with a period of 180°. Indeed, I_{SC} measured by a conductive AFM tip in the configuration illustrated in Fig. 1B on the $SrTiO_3$ (010) and TiO_2 (001) surfaces exhibits a light polarization dependence in accord with Eq. 1, as seen in Fig. 3. The sinusoidal behavior upon rotating the light polarization provides strong evidence that the underlying mechanism of the force-induced PV effect is the BPV effect generated by local symmetry breaking due to inhomogeneous strain—namely, the FPV effect.

The FPV effect should be neither confined to ionic crystals nor restricted to nanoscale geometries. First, we performed the same experiment (as depicted in Fig. 1) on an HF-passivated surface of a commercial p-type Si (001) crystal. As in our initial experiment, I_{SC} increased by two orders of magnitude from ~5 pA with the 1-μN loading force to ~0.5 nA with the 15-μN loading force [Fig. 4A and section S7 of (*16*)]. Second, we demonstrated the FPV effect by using a homebuilt indentation system that deforms a semiconductor by means of a conductive tungsten probe needle with a radius of ~10 μm (fig. S1). Figure 4B shows crystallographic orientation–

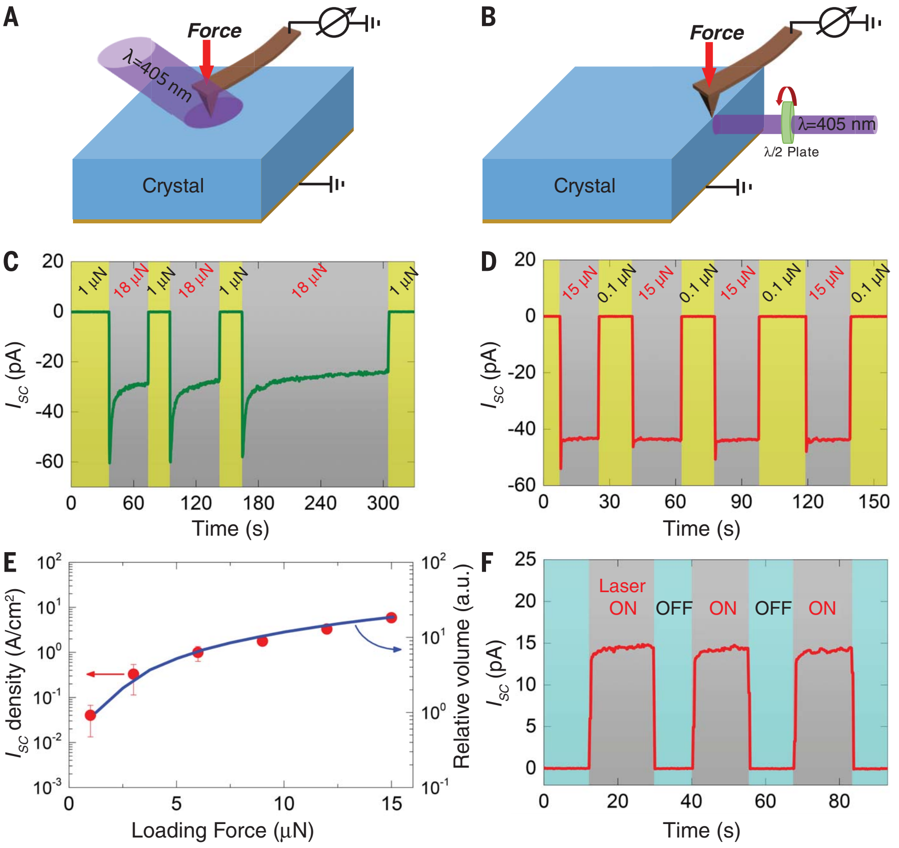


Fig. 1. Force-induced PV effect in centrosymmetric $SrTiO_3$ and TiO_2 single crystals. (**A**) Setup for illumination around the contact area. The tip loading force is controlled by the feedback loop of an AFM. (**B**) Setup for illumination on the side surface. This illumination geometry was chosen to avoid the effect of Fresnel reflection and to ensure that light absorption would be independent of the light polarization. (**C** and **D**) Evolution of the photocurrent induced and collected by a conductive AFM tip with a high loading force on (C) a $SrTiO_3$ (001) face and (D) a TiO_2 (100) face. (**E**) Loading force dependence of the induced photocurrent density and the relative volume subject to a strain gradient higher than 1×10^6 m^{-1} (fig. S9). a.u., arbitrary units. (**F**) Positive photocurrent measured on a TiO_2 (001) face with a 15-μN force applied by the AFM tip.

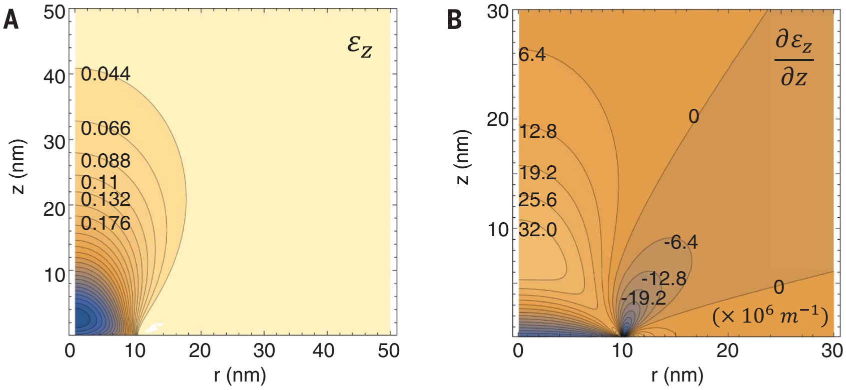


Fig. 2. Spatial distributions of strain and strain gradient induced by an ideal spherical indenter. (**A**) The *z* component of strain (ε) and (**B**) its partial derivative with respect to *z* are shown. The contact area with a radius (*r*) of 10 nm is centered at the origin, and the force is pointing upward. The positive (or negative) value of the strain means a tensile (or compressive) strain along the *z* axis.

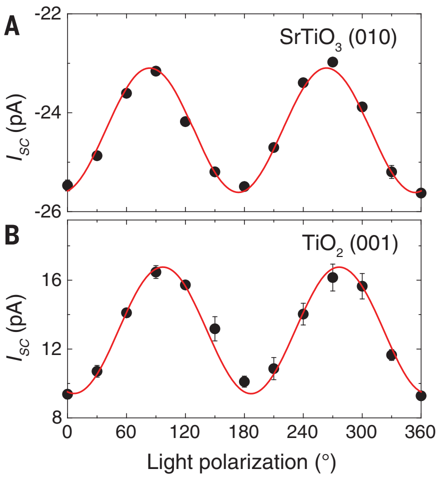


Fig. 3. Light polarization dependence of the force-induced photocurrent. (**A** and **B**) Sinusoidal dependence of the photocurrent measured on (A) a $SrTiO_3$ (010) face and (B) a TiO_2 (001) face under 405-nm laser illumination. The red lines are fits to Eq. 1. Error bars indicate the SD of the PV current measured 25 times.

dependent photocurrent-voltage (*I-V*) characteristics of the $SrTiO_3$ crystal acquired under 4 N of mechanical force exerted by the probe needle and 405-nm laser illumination directly to the probe contact area, as in Fig. 1A. These linear *I-V* characteristics are similar to those of ferroelectrics under illumination (*23*). The oscillating PV current with the rotating light polarization, which is well fitted to Eq. 1 (Fig. 4C), demonstrates the FPV effect under the indentation force on the micrometer scale. The crystallographic orientation–dependent PV current is also observed when the TiO_2 crystal is deformed by the indentation system (fig. S2). The persistence of the FPV effect from the nanoscale (AFM) to the micrometer level is promising for future device design and potential applications. The FPV effect is not related to a plasmonic effect found in the tip-enhanced Raman scattering: Only an atomically sharp tip coated with Ag or Au shows the plasmonic enhancement in the visible range (*24*), whereas we used platinum or tungsten probes ranging from the nanoscale to the micrometer scale. Likewise, a potential cubic-to-tetragonal phase transition induced in a $SrTiO_3$ single crystal under a large hydrostatic pressure (>6 GPa) should not play a large role because of the centrosymmetric nature of the induced tetragonal phase (*25*).

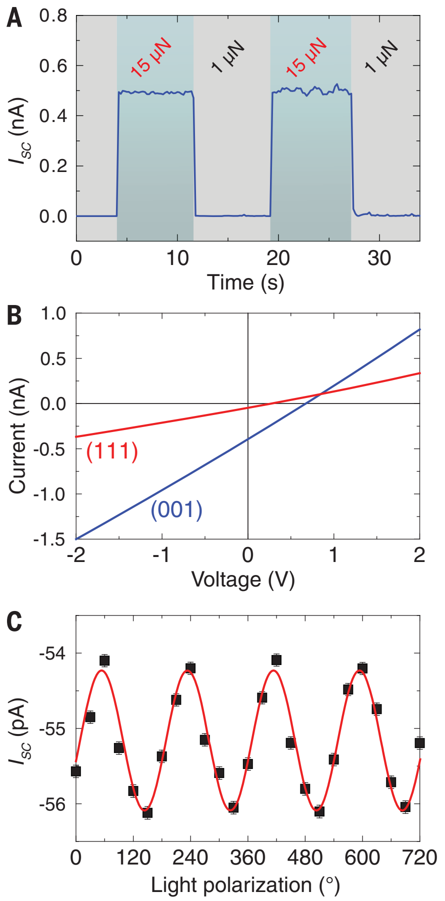


Fig. 4. FPV effect extended to covalent crystals and to the micrometer scale. (**A**) Force-induced photocurrent on a (001) Si crystal, as measured by Ph-AFM under illumination of the top indented surface. (**B**) Current-voltage characteristics measured on (001) and (111) faces of a $SrTiO_3$ crystal by the microindenter applying a 4-N force under illumination of the top indented surface. (**C**) Light polarization dependence of the photocurrent on a $SrTiO_3$ (010) face measured by the microindenter under illumination on the side surface. The red line is the fit of experimental data to Eq. 1. Because the polarization angle origin of the microindenter is not coincident with the Ph-AFM setup, the oscillating I_{SC} has a phase shift compared with that depicted in Fig. 3B. Error bars indicate the SD of the PV current measured 50 times.

We emphasize here four main features of the FPV effect. First, the separation of the photoexcited carriers in the FPV effect is controlled by the local symmetry and the resultant local BPV effect, in which the power conversion efficiency can, in principle, exceed the S-Q limit (*8*). Second, to obtain a high photocurrent from any semiconductor, only a strain gradient generator, such as a sharp probe with a sufficient loading force is necessary. This should be distinguished from the previous reports that a strain gradient modifies a bandgap, but the charge separation in the photoelectric process still requires a proper band alignment (*26*) or a nanostructure (*27*, *28*). Third, whereas the BPV effect is possible only in noncentrosymmetric materials, the FPV effect is universal. It is allowed by symmetry in all materials, owing to the universal nature of the strain gradient–induced centrosymmetry breaking. The FPV effect can be realized in ionic crystals ($SrTiO_3$ and TiO_2) and covalent crystals (Si) but is also relevant to any semiconductors, ranging from organic-inorganic hybrid perovskites to semiconducting polymers and even topological insulators. For instance, the topological insulator Bi_2Te_3 with a centrosymmetric structure exhibits the BPV effect by means of the flexoelectric effect (*29*). Finally, given that the BPV effect consists of asymmetric quantum mechanical processes such as photoexcitation, relaxation, recombination, and scattering, we demonstrate that one can readily control the quantum mechanical processes by macroscopic tools such as an AFM tip and a probe needle.

The configuration of our PV indentation system is very simple, and the FPV effect can be increasingly substantial with material dimensions decreasing into the nanoscale where flexoelectricity is more important (*11*). Thus, an valuable strain engineering route for improving the performance of solar cells and optoelectronic devices is now open. For example, a tandem solar cell can be easily fabricated by combining an array of indenters and a conventional p-n junction, enabling a higher efficiency because the FPV effect can be designed to boost the existing PV current generated by the buried p-n junction. Given that the lattice mismatch at the interfaces and crystallographic disorders in epitaxial and polycrystalline thin film solar cells produce substantial strain gradients (*30*–*32*), the associated FPV effect would considerably enhance the performance of these solar cells; however, this topic remains largely unexplored. In addition to inorganic solar cells, the FPV effect is also likely to play an important role in flexible and stretchable electronics based on organic and polymeric semiconductors. Both the bending of flexible organic devices at the macroscopic level (*33*) and the folding and entanglement of the polymeric chains at the nanoscale would generate sizable strain gradients (*34*), which redistribute the electron cloud of π molecular orbits, modifying electronic transport and inducing the FPV effect under illumination (*15*, *35*). But the details of how strain gradient and FPV effects influence nanoscale electronic properties remain a matter of debate.

REFERENCES AND NOTES

1. W. Siemens, *Lond. Edinb. Dublin Philos. Mag. J. Sci.* **19**, 315–316 (1885).
2. M. A. Green, S. P. Bremner, *Nat. Mater.* **16**, 23–34 (2017).
3. A. Polman, M. Knight, E. C. Garnett, B. Ehrler, W. C. Sinke, *Science* **352**, aad4424 (2016).
4. A. Fahrenbruch, R. Bube, *Fundamentals of Solar Cells: Photovoltaic Solar Energy Conversion* (Elsevier, 2012).
5. W. Shockley, H. J. Queisser, *J. Appl. Phys.* **32**, 510–519 (1961).
6. A. Chynoweth, *Phys. Rev.* **102**, 705–714 (1956).
7. B. I. Sturman, V. M. Fridkin, *The Photovoltaic and Photorefractive Effects in Noncentrosymmetric Materials* (Gordon and Breach Science Publishers, 1992).
8. J. E. Spanier *et al.*, *Nat. Photonics* **10**, 611–616 (2016).
9. A. Bhatnagar, A. R. Chaudhuri, Y. H. Kim, D. Hesse, M. Alexe, *Nat. Commun.* **4**, 2835 (2013).
10. A. K. Tagantsev, *Phys. Rev. B* **34**, 5883–5889 (1986).
11. P. Zubko, G. Catalan, A. K. Tagantsev, *Annu. Rev. Mater. Res.* **43**, 387–421 (2013).
12. J. Fousek, L. Cross, D. Litvin, *Mater. Lett.* **39**, 287–291 (1999).

13. K. D. Breneman, W. E. Brownell, R. D. Rabbitt, *PLOS ONE* **4**, e5201 (2009).
14. U. K. Bhaskar *et al.*, *Nat. Nanotechnol.* **11**, 263–266 (2016).
15. S. V. Kalinin, V. Meunier, *Phys. Rev. B* **77**, 033403 (2008).
16. See supplementary materials.
17. J. Očenášek *et al.*, *Phys. Rev. B* **92**, 035417 (2015).
18. H. Lu *et al.*, *Science* **336**, 59–61 (2012).
19. M. M. Yang, A. Bhatnagar, Z. D. Luo, M. Alexe, *Sci. Rep.* **7**, 43070 (2017).
20. P. Zubko, G. Catalan, A. Buckley, P. R. Welche, J. F. Scott, *Phys. Rev. Lett.* **99**, 167601 (2007).
21. V. Fridkin, *Crystallogr. Rep.* **46**, 654–658 (2001).
22. A. C. Fischer-Cripps, *Introduction to Contact Mechanics* (Springer, 2000).
23. V. M. Fridkin, *Photoferroelectrics*, vol. 9 (Springer, 2012).
24. M. D. Sonntag, E. A. Pozzi, N. Jiang, M. C. Hersam, R. P. Van Duyne, *J. Phys. Chem. Lett.* **5**, 3125–3130 (2014).
25. M. Guennou, P. Bouvier, J. Kreisel, D. Machon, *Phys. Rev. B* **81**, 054115 (2010).
26. J. Feng, X. Qian, C.-W. Huang, J. Li, *Nat. Photonics* **6**, 866–872 (2012).
27. X. Fu *et al.*, *ACS Nano* **8**, 3412–3420 (2014).
28. J. Greil, S. Birner, E. Bertagnolli, A. Lugstein, *Appl. Phys. Lett.* **104**, 163901 (2014).
29. Y. Liu *et al.*, *APL Mater.* **4**, 126104 (2016).
30. M. S. Bennett, J. J. Kramer, *J. Appl. Phys.* **54**, 7159–7165 (1983).
31. D. Lee *et al.*, *Phys. Rev. Lett.* **107**, 057602 (2011).
32. M.-W. Chu *et al.*, *Nat. Mater.* **3**, 87–90 (2004).
33. M. A. Reyes-Martinez, A. J. Crosby, A. L. Briseno, *Nat. Commun.* **6**, 6948 (2015).
34. Q. Tang *et al.*, *Small* **7**, 189–193 (2011).
35. T. Dumitrică, C. M. Landis, B. I. Yakobson, *Chem. Phys. Lett.* **360**, 182–188 (2002).

ACKNOWLEDGMENTS

We thank J. Lloyd-Hughes for grammatical revision of the manuscript and B. Tao for assistance with schematic drawing. **Funding:** M.-M.Y. acknowledges the University of Warwick for a Chancellor's International Scholarship. M.A. acknowledges the Wolfson Research Merit and Theo Murphy Blue Skies awards from The Royal Society. This work was partly supported by the EPSRC (UK) through grants EP/M022706/1, EP/P031544/1, and EP/P025803/1. **Author contributions:** D.J.K. initiated the project. M.-M.Y., D.J.K., and M.A. conceived of and developed the ideas. M.-M.Y. and D.J.K. designed and conducted the experiments, analyzed the data, and wrote the manuscript. All authors contributed to the discussion of the results and the manuscript revision. M.A. supervised the project. **Competing interests:** M.-M.Y., D.J.K., and M.A. are authors of British Patent Application GB1702466.2 filed in the UK by the University of Warwick on 15 February 2017 that covers PV devices based on strain gradient inducers. **Data and materials availability:** All scientific data related to this paper are available at the University of Warwick open access research repository (http://wrap.warwick.ac.uk/100429).

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/360/6391/904/suppl/DC1
Supplementary Text
Figs. S1 to S15
References (*36*–*43*)

28 March 2017; accepted 6 April 2018
Published online 19 April 2018
10.1126/science.aan3256

EVOLUTION

High parasite diversity accelerates host adaptation and diversification

A. Betts,[1]* C. Gray,[2] M. Zelek,[1] R. C. MacLean,[1]† K. C. King[1]*†

Host-parasite species pairs are known to coevolve, but how multiple parasites coevolve with their host is unclear. By using experimental coevolution of a host bacterium and its viral parasites, we revealed that diverse parasite communities accelerated host evolution and altered coevolutionary dynamics to enhance host resistance and decrease parasite infectivity. Increases in parasite diversity drove shifts in the mode of selection from fluctuating (Red Queen) dynamics to predominately directional (arms race) dynamics. Arms race dynamics were characterized by selective sweeps of generalist resistance mutations in the genes for the host bacterium's cell surface lipopolysaccharide (a bacteriophage receptor), which caused faster molecular evolution within host populations and greater genetic divergence among populations. These results indicate that exposure to multiple parasites influences the rate and type of host-parasite coevolution.

Coevolution between hosts and parasites is hypothesized to be central to numerous biological phenomena, from rapid evolutionary change (*1*) to the maintenance of diversity (*2*) and sexual reproduction (*3*, *4*). Host-parasite coevolution is traditionally investigated in a pairwise framework where a single parasite species infects one host species (*5*). However, in nature, these pairwise relationships rarely exist in isolation (*6*, *7*); hosts are often under attack by multiple parasite species (*8*), and parasites must compete for hosts (*9*–*12*). Host-parasite coevolution thus operates within a network of species interactions (*13*). It nevertheless remains unknown whether and how diverse parasite communities can shape coevolutionary patterns and processes.

Coevolution with highly diverse parasite communities should impose stronger selection for host resistance if this diversity yields more infections and higher mortality than single parasites. This outcome could ultimately drive faster host evolution and increased divergence among host populations. To test these predictions, we applied an experimental coevolution approach to an in vitro bacterium-phage system. We exposed the host bacterium, *Pseudomonas aeruginosa*, to communities of one to five lytic viral parasites (bacteriophages PEV2, LUZ19, LUZ7, 14-1, and LMA2) that are obligate lethal parasites. These parasites each infect their hosts by attaching to one of three specific cell surface receptors [lipopolysaccharides (LPSs), Ton-B–dependent receptors, or type IV pili], and hosts typically evolve resistance by modifying or deleting these attachment sites (*14*). Parasites, in turn, can evolve reciprocal adaptations to circumvent host resistance (*15*). We experimentally coevolved host populations independently in cell culture tubes with replicates of a five-parasite community (high diversity, n = 30 cultures), with every possible pairwise parasite combination (medium diversity, n = 30), and in a single-parasite treatment with replicates of each parasite in monoculture (low diversity, n = 30), alongside a parasite-free control (n = 30). We tested for effects on the tempo and mode of host-parasite coevolution, conducting phenotypic assays of host resistance and parasite infectivity, as well as using deep sequencing to measure changes in the genomic compositions of host and parasite populations through time. Because we were able to experimentally manipulate parasite diversity, our work differs from the approaches used to investigate coevolution in natural communities, where diversity is not controlled exogenously (*16*, *17*).

As with predators and prey, parasite densities are expected to track host densities, producing time-lagged cycles in population sizes over time (*18*). In our experiment, given that the host was initially susceptible, we might have expected large amplitude oscillations in population size, unless host resistance rapidly evolved (*19*). By contrast, we observed that host densities increased throughout the experiment while parasite densities remained stable (Fig. 1A and supplementary text), indicating that underlying coevolutionary dynamics interfered with the expected ecological dynamics (*20*).

We determined whether coevolution took place in our experiment by performing 13,500 time-shift assays. This approach assesses the infectivity of parasites toward their past, present, and future hosts and, reciprocally, the resistance of hosts to their past, present, and future parasites (*21*). Resistance was measured by inhibition assays in which individual host genotypes were challenged with parasite communities and resultant growth or inhibition was analyzed to compare the mean levels of resistance among treatments. Infectivity and resistance varied widely across individual host-parasite communities over time (fig. S1). Despite this complexity, two clear and significant patterns ($F_{2,13044}$ = 1766.21, $P < 0.0001$) emerged. First, parasites from the future were more infective than parasites from the present or past (Fig. 1B). Likewise, hosts evolved increased resistance over time (Fig. 1C). This pattern is typical of arms race dynamics, commonly seen with in vitro bacteria-phage coevolution (*5*, *22*). Second, bacterial resistance increased with parasite diversity and parasite infectivity decreased ($F_{2,85}$ = 9.7, $P < 0.001$), suggesting that parasite communities impose stronger selection on their host than single parasites alone.

To test the hypothesis that parasite diversity accelerated coevolution at a molecular level, we sequenced entire host and parasite populations longitudinally at multiple time points (ancestor, midpoint, and end point) for the high- (n = 12 cultures), medium- (n = 10), and low-diversity (n = 15) and control (n = 10) treatments. We could thus detect any mutations that arose during coevolution and estimate their frequency changes within populations. In the hosts, we found 474 nonsynonymous and 75 synonymous polymorphisms across 173 genes and 133 intergenic mutations. Parallel evolution with mutations in the same genes or genes in the same pathways was common across all treatments: >50% of mutations occurred in <5% of these genes. Notably, we observed parallel evolution in genes that are known phage receptors (table S1), including genes for LPS (n = 190 mutations), type IV pili (n = 69 mutations), and the Ton-B–dependent receptor (n = 55 mutations). To measure the rate of bacterial molecular evolution, we calculated the Euclidean genetic distance of each population from the ancestor (Fig. 2A). We found that higher levels of parasite diversity drove greater host divergence from the ancestor and consequently faster host evolution (fig. S2) [analysis of variance (ANOVA), $F_{2,34}$ =10.5, $P < 0.001$; Tukey post hoc test, $P < 0.05$].

Two mechanisms could explain why diverse parasite communities accelerate the rate of host evolution. On one hand, diverse communities are more likely, by chance alone, to contain parasites that drive rapid evolution. On the other hand, it is possible that parasite diversity per se accelerates host evolution by increasing selection for resistance. To discriminate between these mechanisms, we calculated the contribution of parasite diversity to the speed of host evolution by adopting an approach developed to determine the contribution of biodiversity to ecosystem function, represented in this study by the host evolutionary rate (*23*). Using this approach, we found that the combination of all five parasites contributed to faster host evolution overall than phage monocultures (test for transgressive overyielding $D_{\max} > 0$; one-sample t tests with Bonferroni adjustment, $P < 0.05$). Thus, the diversity of interactions among parasites within communities, and not merely the presence of a particular parasite, imposed stronger selection.

---

[1]Department of Zoology, University of Oxford, South Parks Road, Oxford OX1 3PS, UK. [2]Department of Life Sciences, Imperial College London, Silwood Park Campus, Ascot SL5 7PY, UK.
***Corresponding author. Email: alex.betts@evobio.eu (A.B.); kayla.king@zoo.ox.ac.uk (K.C.K.)**
†These authors contributed equally to this work.

Whole-genome sequencing also revealed profound changes in parasite communities. At an ecological level, the composition of communities changed markedly over time (fig. S3). Parasite diversity decreased in both the medium- and high-diversity treatments, but diversity remained higher in the high-diversity treatment than in the medium-diversity treatment (fig. S4). At a genetic level, we found evidence for rapid molecular evolution of phage, which supports the results of our phenotypic assays. In total, we detected 533 single-nucleotide polymorphisms (SNPs) across all five parasite types that reached a frequency of >10% in the parasite populations (table S2). The proportions of nonsynonymous SNPs were very high, exceeding 0.7 for each parasite (table S3), and parallel evolution was common, implying that these SNPs were predominantly beneficial mutations. To investigate the role of parasite diversity in phage evolution, we focused our analysis on communities that included the parasite PEV2 because this phage was well-represented at all levels of parasite diversity and because, compared with the genomes of the other phages used in this study, the PEV2 genome is well annotated. For example, all but two of the genes in the parasite 14-1 genome code for hypothetical proteins. We detected 225 mutations in the PEV2 genome, and the rate of PEV2 evolution was independent of parasite diversity (fig. S5). Thirty-nine percent of the PEV2 mutations (n = 88) occurred in a single tail fiber gene, *gp52*. The tail fibers are the means by which the parasite recognizes and binds to its receptor and are known to undergo coevolution (*24*). We detected mutations in *gp52* and at least one host LPS mutation (usually in *wzz* or *migA*) in every community containing PEV2 across all levels of parasite diversity (figs. S6 and S7). Collectively, these results provide genetic confirmation for reciprocal coevolution between host and parasite across all levels of parasite diversity, and they provide further evidence that parasite attachment to LPS is a key mechanism underpinning coevolution between PEV2 and *P. aeruginosa*.

Coevolutionary dynamics are usually described as being one of two types. Arms races occur when selection results in directional increases in host resistance and parasite infectivity (*25*, *26*), and with Red Queen dynamics, negative frequency-dependent selection causes fluctuations in allele frequencies (*27*, *28*). It has been previously shown that each parasite follows a different coevolutionary trajectory when interacting with *P. aeruginosa* in a pairwise fashion (fig. S8) (*29*). Using a population genomic approach, we therefore sought to examine whether parasite diversity altered the mode of coevolution.

We examined the type of coevolution occurring in each community by comparing the changes in frequency of each host allele between the mid- and end points of the experiment (*29*, *30*). These time frames were previously shown to be suitable for observing coevolutionary dynamics in *P. aeruginosa*–phage interactions (*29*). We focused our analysis on the host because changes in the frequency of phage alleles in

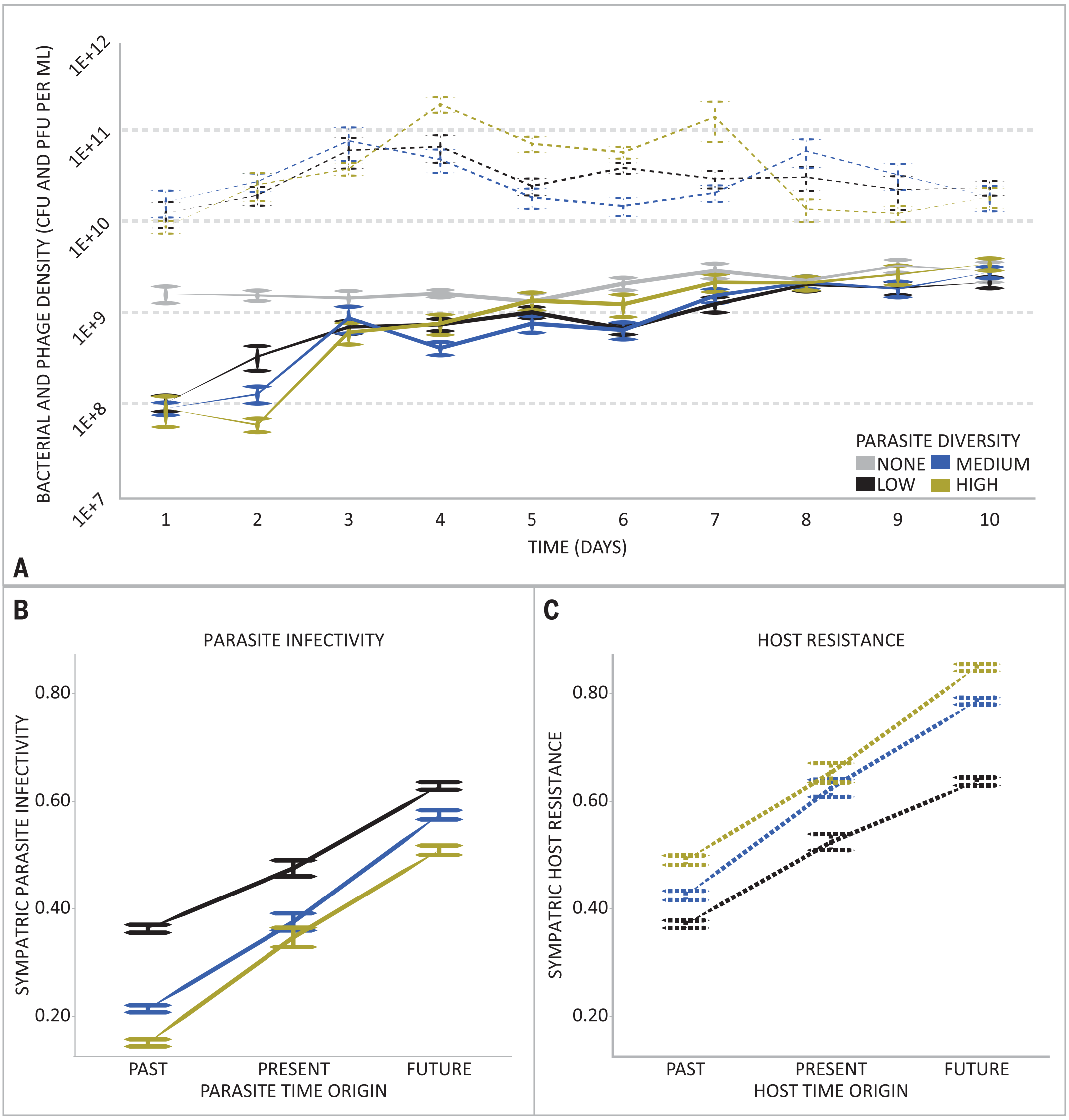


Fig. 1. Ecological and coevolutionary dynamics of host-parasite interactions. (**A**) We tracked the population dynamics of hosts (solid lines) and parasites (dashed lines) for high-, medium-, and low-diversity populations alongside the parasite-free control. All plotted points show the mean population density ± SE. Bacterial populations challenged with phage rapidly recovered their density. We did not observe cyclical oscillations in host and parasite densities. CFU, colony-forming units; PFU, plaque-forming units. (**B** and **C**) To directly test for coevolution, we used 13,500 time-shift assays to measure changes in phage infectivity and bacterial resistance to phage. Lower levels of parasite infectivity (B) and higher levels of host resistance (C) evolved with increasing parasite diversity ($F_{2,85}$ = 9.7, P < 0.001). Furthermore, hosts were more susceptible to parasites from the future than to contemporary parasites, and they were most resistant to parasites from the past. Likewise, future parasites were more infective than past or contemporary parasites ($F_{2,13044}$ = 1766.21, P < 0.0001). Error bars show 1 SE.

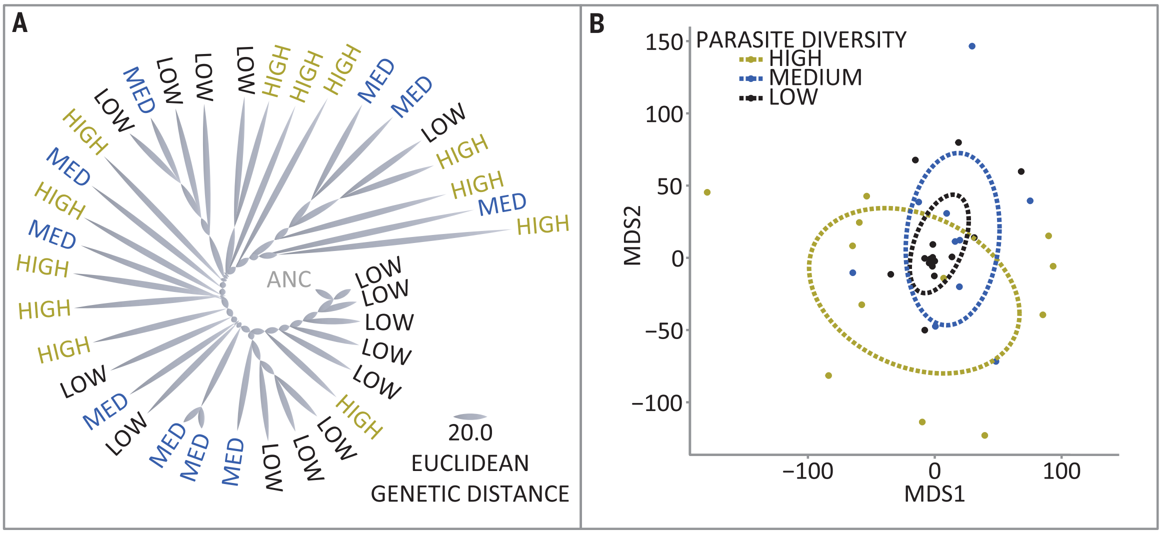


Fig. 2. Parasite diversity accelerates host evolution. (**A**) Host allele frequencies after 10 days of coevolution with the high-, medium-, and low-parasite-diversity treatments were used to calculate pairwise Euclidean distances between the ancestral sequence (ANC) and each coevolved population. Increasing parasite diversity accelerated the rate of host evolution. (**B**) The genetic distances between coevolved populations were ordinated by nonmetric multidimensional scaling (MDS1 and MDS2). The ellipses represent a 95% confidence bubble around the means for the different treatments. We found evidence of divergence between populations within treatments (ANOSIM, $R = 0.11$, $P < 0.01$), and the greatest within-treatment diversification was observed in the high-parasite-diversity treatment.

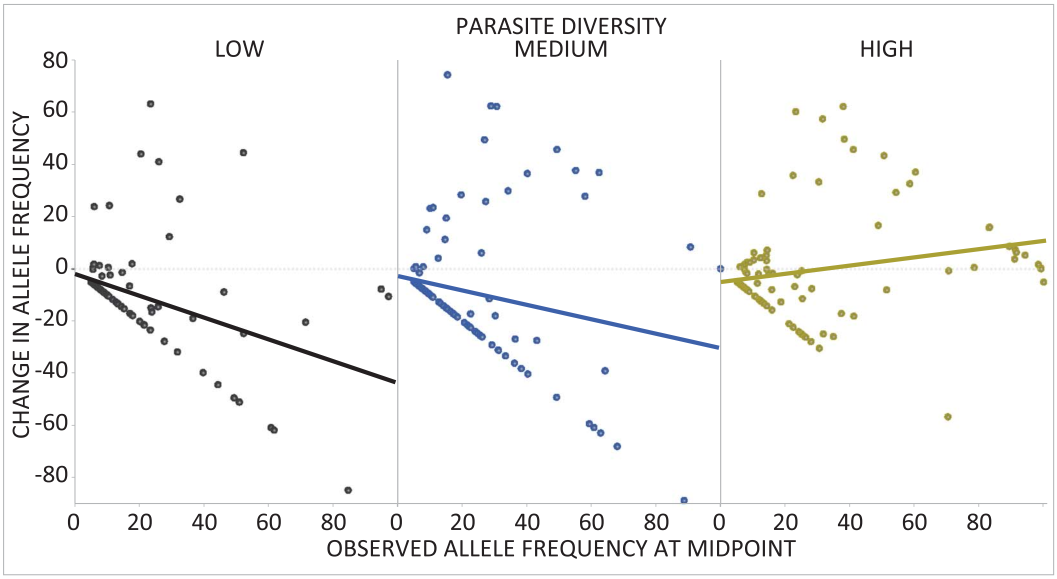


Fig. 3. Red Queen coevolution is more common in pairwise host-parasite interactions. We tested for Red Queen dynamics by regressing the change in host allele frequency (%) from day 5 to day 10 (*y* axis) with the observed frequency (%) on day 5 (*x* axis). We found evidence for negative frequency-dependent selection on host alleles under low and medium parasite diversity, but not at high parasite diversity. The string of points forming a straight downward slope from zero in all three panels represents alleles observed on day 5 that had subsequently decreased in frequency to below our ability to detect them at day 10.

diverse communities might additionally be driven by interspecific competition among parasites (*14*). On average, host alleles with higher frequencies tended to decline in frequency over time in the low- and medium-parasite-diversity treatments, consistent with negative frequency-dependent selection on host resistance genes (Fig. 3). This mode of selection became less predominant at higher levels of parasite diversity [analysis of covariance (ANCOVA), $F_{2,335} = 12.64$, $P < 0.0001$], suggesting that the occurrence of Red Queen dynamics was reduced in these communities.

In contrast to Red Queen dynamics, arms races should result in recurrent selective sweeps of resistance mutations (*5*). The effects of directional selection in host-parasite interactions have been studied in a variety of systems and can be important for coevolution of human pathogens with human populations (*31*–*34*). To test for selective sweeps in host populations, we calculated F_{ST}, which measures the genetic divergence per mutation between each coevolved host population and a sequenced ancestor (*35*), and we set a conservative threshold F_{ST} above which any mutation was deemed to be under strong directional selection (Fig. 4A). Strong directional selection was more common as parasite diversity increased ($\chi^2 = 20$, df = 3, $P < 0.001$), such that the number of fixed alleles in the high-diversity treatment was more than two times that observed under medium diversity (Fig. 4B). Of the 29 alleles approaching fixation, 23 were in genes relating to LPS biosynthesis. LPS is very abundant on the bacterial cell surface, and phage may initially bind to LPS to anchor themselves to the cell before forming an irreversible attachment to a final receptor (*36*). Mutations that alter LPS biosynthesis may provide a general phage resistance mechanism. We have previously shown that mutations in the LPS biosynthesis gene *wzy* provide resistance to four of the phages used in this study (*14*). Red Queen and arms race dynamics fall at two ends of a continuum (*21*): Tight interaction specificity is required for the former, and for the latter, general host resistance mechanisms should evolve to allow future hosts to resist all previous parasite genotypes. Our data point to the availability of generalist LPS biosynthesis

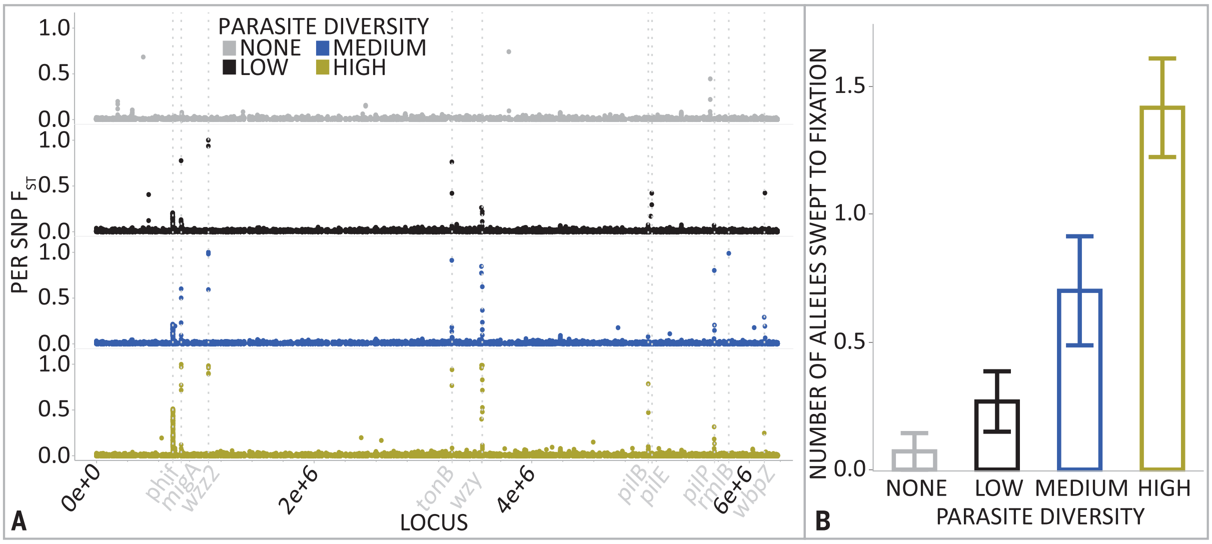


Fig. 4. Parasite diversity leads to more arms race coevolutionary dynamics. (**A**) To test for arms race dynamics, we calculated F_{ST} for all host SNPs across the *P. aeruginosa* genome for the high-, medium-, and low-parasite-diversity treatments and the parasite-free control. Each plotted point in the panel represents a single SNP, and host genes that are known parasite targets, including genes involved in LPS biosynthesis, type IV pilus biosynthesis, and the Ton-B–dependent receptor, are shown in italics. (**B**) We used a conservative F_{ST} cutoff to identify selective sweeps of SNPs per host population (± SE), and we found that increasing parasite diversity increased the rate of fixation of SNPs (χ^2 = 20, df = 3, $P < 0.001$).

resistance mutations as being key to the transition from Red Queen to arms race dynamics as parasite diversity increases. That both dynamics can arise from host-parasite community coevolutionary interactions is consistent with studies of bacteria-phage systems conducted in natural communities (*16*, *17*).

Parasites have been shown to play a key role in evolutionary diversification within (*37*–*39*) and among (*40*) host populations. Specifically, parasite-mediated selection against common host genotypes is predicted to stably maintain genetic diversity (*31*, *41*). Host population genetic diversity should be greatest during pairwise coevolution, where we observed that Red Queen dynamics are relatively stronger. Surprisingly, we found that levels of genetic diversity within host populations were not affected by parasite diversity (fig. S9) despite the different modes of selection acting across treatments. Parasite diversity, conversely, had a profound effect on allopatric diversification among host populations [analysis of similarity (ANOSIM), R = 0.11, $P < 0.01$]. The greatest divergence was found in the high-parasite-diversity treatment (Fig. 2B) (ANOVA, $F_{2,213}$ = 122.15, $P < 0.001$; Tukey post hoc test, $P < 0.05$). The simplest explanation for this result is that rapid evolution within host populations leads to accelerated divergence among populations.

Our study reveals that parasite communities can form hotbeds of rapid antagonistic coevolution. Diverse parasite communities imposed stronger selection on host populations than single parasites, causing faster selective sweeps of generalist resistance mutations and higher levels of host resistance. The multispecies coevolutionary process ultimately accelerated the rate of molecular evolution within host populations and increased the genomic divergence among populations. Host resistance to parasite infection can be specialized (*26*). However, broad-spectrum defense strategies that confer protection against different enemies are widespread in animals (*42*–*44*) and plants (*45*, *46*), as well as bacteria (*47*–*49*). We should thus consider looking beyond the pairwise framework toward parasite communities as drivers of coevolution and rapid host evolution in nature. This focus could be particularly informative in biodiversity hotspots with elevated evolutionary and speciation rates (*27*), where a multitude of parasites might be the fuel.

REFERENCES AND NOTES

1. S. Paterson *et al.*, *Nature* **464**, 275–278 (2010).
2. J. B. S. Haldane, *Ric. Sci.* **19**, 68–76 (1949).
3. G. Bell, *The Masterpiece of Nature: The Evolution and Genetics of Sexuality* (Croom Helm, 1982).
4. W. D. Hamilton, *Oikos* **35**, 282 (1980).
5. M. A. Brockhurst *et al.*, *Proc. R. Soc. London Ser. B* **281**, 20141382 (2014).
6. D. J. Marcogliese, *Int. J. Parasitol.* **35**, 705–716 (2005).
7. P. J. Hudson, A. P. Dobson, K. D. Lafferty, *Trends Ecol. Evol.* **21**, 381–385 (2006).
8. S. Telfer *et al.*, *Science* **330**, 243–246 (2010).
9. E. C. Griffiths, A. B. Pedersen, A. Fenton, O. L. Petchey, *J. Infect.* **63**, 200–206 (2011).
10. J. Lello, *Sci. Transl. Med.* **5**, 191fs24 (2013).
11. S. Alizon, J. C. de Roode, Y. Michalakis, *Ecol. Lett.* **16**, 556–567 (2013).
12. F. Bashey, H. Hawlena, C. M. Lively, *Evolution* **67**, 900–907 (2013).
13. A. Betts, C. Rafaluk, K. C. King, *Trends Parasitol.* **32**, 863–873 (2016).
14. A. Betts, D. R. Gifford, R. C. MacLean, K. C. King, *Evolution* **70**, 969–978 (2016).
15. J. E. Samson, A. H. Magadán, M. Sabri, S. Moineau, *Nat. Rev. Microbiol.* **11**, 675–687 (2013).
16. B. Koskella, N. Parr, *Philos. Trans. R. Soc. London Ser. B* **370**, 20140297 (2015).
17. P. Gómez, A. Buckling, *Science* **332**, 106–109 (2011).
18. P. J. Hudson, A. P. Dobson, D. Newborn, *Science* **282**, 2256–2258 (1998).
19. A. Mougi, Y. Iwasa, *Proc. R. Soc. London Ser. B* **277**, 3163–3171 (2010).
20. M. H. Cortez, J. S. Weitz, *Proc. Natl. Acad. Sci. U.S.A.* **111**, 7486–7491 (2014).
21. S. Gandon, A. Buckling, E. Decaestecker, T. Day, *J. Evol. Biol.* **21**, 1861–1866 (2008).
22. A. Buckling, Y. Wei, R. C. Massey, M. A. Brockhurst, M. E. Hochberg, *Proc. R. Soc. London Ser. B* **273**, 45–49 (2006).
23. M. Loreau *et al.*, *Science* **294**, 804–808 (2001).
24. J. S. Weitz, H. Hartman, S. A. Levin, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 9535–9540 (2005).
25. A. Buckling, P. B. Rainey, *Proc. R. Soc. London Ser. B* **269**, 931–936 (2002).
26. A. Agrawal, C. Lively, *Evol. Ecol. Res.* **4**, 79–90 (2002).
27. M. D. Pirie *et al.*, *BMC Evol. Biol.* **16**, 190 (2016).
28. E. Decaestecker *et al.*, *Nature* **450**, 870–873 (2007).
29. A. Betts, O. Kaltz, M. E. E. Hochberg, *Proc. Natl. Acad. Sci. U.S.A.* **111**, 11109–11114 (2014).
30. B. Koskella, C. M. Lively, *Evolution* **63**, 2213–2221 (2009).
31. R. D. Schulte, C. Makus, B. Hasert, N. K. Michiels, H. Schulenburg, *Proc. Natl. Acad. Sci. U.S.A.* **107**, 7359–7364 (2010).
32. L. Wilfert, F. M. Jiggins, *Biol. Lett.* **6**, 666–668 (2010).
33. B. Roche *et al.*, *Trends Parasitol.* **33**, 21–29 (2017).
34. C. Eizaguirre, T. L. Lenz, M. Kalbe, M. Milinski, *Ecol. Lett.* **15**, 723–731 (2012).
35. S. Wright, *Evolution and the Genetics of Populations*, vol. 2 (Univ. of Chicago Press, 1969).
36. S. T. Abedon, *The Bacteriophages* (Oxford, ed. 2, 2006).
37. K. C. King, L. F. Delph, J. Jokela, C. M. Lively, *Curr. Biol.* **19**, 1438–1441 (2009).
38. M. F. Marston *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **109**, 4544–4549 (2012).
39. L. T. Morran, O. G. Schmidt, I. A. Gelarden, R. C. Parrish II, C. M. Lively, *Science* **333**, 216–218 (2011).
40. J. Thompson, *Geographic Mosaic of Coevolution* (Univ. of Chicago Press, 2005).
41. V. M. D'Costa *et al.*, *Nature* **477**, 457–461 (2011).
42. J. Klein, C. O'Huigin, J. Deutsch, *Philos. Trans. R. Soc. London Ser. B* **346**, 351–358 (1994).
43. A. B. Duncan, S. Fellous, O. Kaltz, *Evolution* **65**, 3462–3474 (2011).
44. G. Tetreau, R. Stalinski, J. P. David, L. Després, *Arch. Insect Biochem. Physiol.* **82**, 71–83 (2013).
45. V. Nicaise, *Front. Plant Sci.* **5**, 660 (2014).
46. L. G. Barrett, F. Encinas-Viso, J. J. Burdon, P. H. Thrall, *Front. Plant Sci.* **6**, 761 (2015).
47. F. Rodriguez-Valera *et al.*, *Nat. Rev. Microbiol.* **7**, 828–836 (2009).

48. S. Avrani, O. Wurtzel, I. Sharon, R. Sorek, D. Lindell, *Nature* **474**, 604–608 (2011).
49. S. Doron *et al*., *Science* **359**, eaar4120 (2018).

ACKNOWLEDGMENTS

We thank M. Brockhurst, D. Dahan, K. Foster, O. Lewis, and S. Paterson for their helpful comments. We also thank the High-Throughput Genomics Group at the Wellcome Trust Centre for Human Genetics, funded by Wellcome Trust grant reference 090532/Z/09/Z and Medical Research Council Hub grant G0900747 91070, for generation of the high-throughput sequencing data. **Funding:** A.B. was supported by D.Phil. funding from the Biotechnology and Biological Sciences Research Council (BBSRC) (grant number BB/J014427/1). R.C.M. was funded by the Royal Society, European Research Council grant 281591, and Wellcome Trust grant 106918/Z/15/Z. K.C.K. was supported by Leverhulme research project grant RPG-2015-165. **Author contributions:** A.B., R.C.M., and K.C.K. conceived and designed the study. A.B. and M.Z. collected the data. A.B. and C.G. conducted the data analysis. A.B., R.C.M., and K.C.K. drafted the article, with critical revisions provided by all authors. **Competing interests:** None declared. **Data and materials availability:** The host and parasite sequences reported in this paper have been deposited in the National Center for Biotechnology (NCBI) database (www.ncbi.nlm.nih.gov; accession numbers PRJNA448325 and PRJNA448370, respectively). All other data are available in the supplementary materials.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/360/6391/907/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S9
Tables S1 to S4
References (*50–56*)
Data S1

18 October 2017; resubmitted 18 February 2018
Accepted 16 April 2018
10.1126/science.aam9974

MIGRATORY BEHAVIOR

From local collective behavior to global migratory patterns in white storks

Andrea Flack,[1,2]*† Máté Nagy,[2,3,4]*† Wolfgang Fiedler,[1,2]
Iain D. Couzin,[2,3] Martin Wikelski[1,2]

Soaring migrant birds exploit columns of rising air (thermals) to cover large distances with minimal energy. Using social information while locating thermals may benefit such birds, but examining collective movements in wild migrants has been a major challenge for researchers. We investigated the group movements of a flock of 27 naturally migrating juvenile white storks by using high-resolution GPS and accelerometers. Analyzing individual and group movements on multiple scales revealed that a small number of leaders navigated to and explored thermals, whereas followers benefited from their movements. Despite this benefit, followers often left thermals earlier and at lower height, and consequently they had to flap considerably more. Followers also migrated less far annually than did leaders. We provide insights into the interactions between freely flying social migrants and the costs and benefits of collective movement in natural populations.

In many animal taxa, migrations are performed by large social groups (*1*), providing various benefits to individual group members (*2*–*6*). In heterogeneous groups, coordinated movements frequently generate leader-follower patterns (*7*–*10*), and individuals may adopt these behavioral strategies, presumably without knowing their own or their group members' roles (*11*). Given interindividual variation within migratory groups, it is necessary to quantify the relationships between each individual to fully understand group dynamics, social influences, and the resulting overall migration patterns. Despite existing knowledge from theoretical work (*12*–*14*), studying collective movement in wild migratory species has been a major challenge (*15*) because it is nearly impossible to record the simultaneous movements of freely flying animals in large, natural groups with appropriate spatiotemporal resolution.

We approached this question by tagging large numbers of juvenile white storks (*Ciconia ciconia*) with high-resolution tracking devices (Fig. 1, A to E). We recorded the trajectories of 27 GPS-tagged naturally migrating juvenile storks flying in a flock (together with untagged birds; table S1) over ~1000 km during the first 5 days of their migratory journey. Owing to some birds eventually leaving the flock, the number of tagged birds in the flock was 27, 22, 21, 20, and 17 for the first 5 days, respectively. Using solar GSM (Global System for Mobile Communications)–GPS–accelerometer loggers, we recorded triaxial acceleration (at 10.54 Hz for 3.8 s every 10 min) and high-frequency GPS locations (at 1 Hz for 2 or 5 min every 15 min, synchronized in time between individuals; henceforth, GPS bursts) of each individual during the group flights (Fig. 1, F to I). After these 5 days, we continued to monitor each bird's movements throughout their entire lifetime, using GPS and accelerometer recordings at lower resolution (fig. S1).

Similar to other large-bodied soaring migrants (*16*–*19*), white storks try to reduce the amount of energetically costly flapping flight by exploiting their atmospheric surroundings (*20*). When comparing movement activity among our tagged juvenile birds of the same flock, we found large differences in the amount of costly flapping. For each bird, we calculated a quantitative measure of animal activity from triaxial acceleration data (henceforth, flapping activity; see the methods) (Fig. 1F) (*21*, *22*). Although storks flew in close proximity (figs. S2 and S3), flapping activity ranged from 0.8 to 1.8. Thus, to cover the same distance during the same time, some individuals performed considerably more flapping flight than did others. Flapping activity was not influenced by individual features (e.g., body measures or sex) or conditions before fledging (general linear model, $F_{17,8} = 0.798$, $P = 0.671$; table S2). Within-individual differences in flapping activity were stable across the different migration days (table S3).

First, we examined how these differences in flapping activity relate to birds' positions within the group. Exploring group structure in detail is challenging because of the different flight modes of soaring migrants (*23*, *24*). To examine flock organization during all flight modes, we developed a metric that quantifies time advances or delays (Δt) between each pair of birds, allowing us to measure the time that separates two individuals—i.e., how much time a bird needs to reach the current location of the other bird (figs. S4 and S5). Storks with low flapping activity flew ahead of other flock members on average, whereas storks with high flapping activity flew behind (Pearson's $r = -0.778$, $n = 27$, $P = 1.7 \times 10^{-6}$; Fig. 2A). Next, we found that an individual's position within the flock (Δt) correlated with leadership (*25*) during the gliding segments (Pearson's $r = 0.846$, $n = 27$, $P = 2.7 \times 10^{-8}$; Fig. 2B). Further, following birds tended to have higher flapping activity than did leaders (Pearson's $r = -0.770$, $n = 27$, $P = 2.6 \times 10^{-6}$; figs. S6 and S7). Because leader and follower roles are respectively reflected in the front and back positions in the flock, we refer to birds that are ahead of the flock on average as leaders and those behind as followers (supplementary text and figs. S8 and S9).

Followers not only spent considerably more time flapping their wings, but also spent less time thermalling than did leaders (Pearson's $r = -0.688$, $n = 27$, $P = 7.2 \times 10^{-5}$; Fig. 2C). Followers finished thermalling earlier, at a lower altitude, likely to avoid being isolated from others—thus seemingly failing to exploit the full potential of thermals (fig. S10). In addition, followers flew farther behind, and at lower altitudes, than leaders during glides (Fig. 2D and figs. S11 to S13). Given that the tagged juveniles migrated together with untagged storks, it is likely that the motion of the observed leaders was in fact affected by other, possibly more experienced, adult birds. Juveniles have higher flight costs than adults, but their ability to use thermals effectively improves throughout their journey (*26*). Collective movements may also partly arise from identical reactions to the same environmental features, but in this study we cannot distinguish between responses to environmental and social cues (*27*).

Leaders and followers differed in their path "tortuosity" while flying within the thermals. Leading birds showed irregular circling while thermalling (calculated as the absolute value of the time derivative of the horizontal curvature, $|d\kappa/dt|$), demonstrating that they make considerable adjustments to their flight paths, consistent with a need to locate the center of the complex thermal structures. In contrast, followers circled more regularly, indicating that, as theory has suggested (*24*), followers can benefit from social information to reach the center of thermals; Δt correlates highly with $|d\kappa/dt|$ for individual averages (Pearson's $r = 0.570$, $n = 27$, $P = 0.002$; Fig. 3A). Because every bird spent some time in the front and back half of the flock, we determined each individual's thermalling performance when ahead and behind the center of the flock (Fig. 3B). Almost all exhibited more regular circling and faster

---

[1] Department of Migration and Immuno-Ecology, Max Planck Institute for Ornithology, Radolfzell, Germany. [2] Department of Biology, University of Konstanz, Konstanz, Germany. [3] Department of Collective Behaviour, Max Planck Institute for Ornithology, Konstanz, Germany. [4] MTA-ELTE Statistical and Biological Physics Research Group, Hungarian Academy of Sciences, Budapest, Hungary.
***Corresponding author. Email: aflack@orn.mpg.de (A.F.); mnagy@orn.mpg.de (M.N.)** †These authors contributed equally to this work.

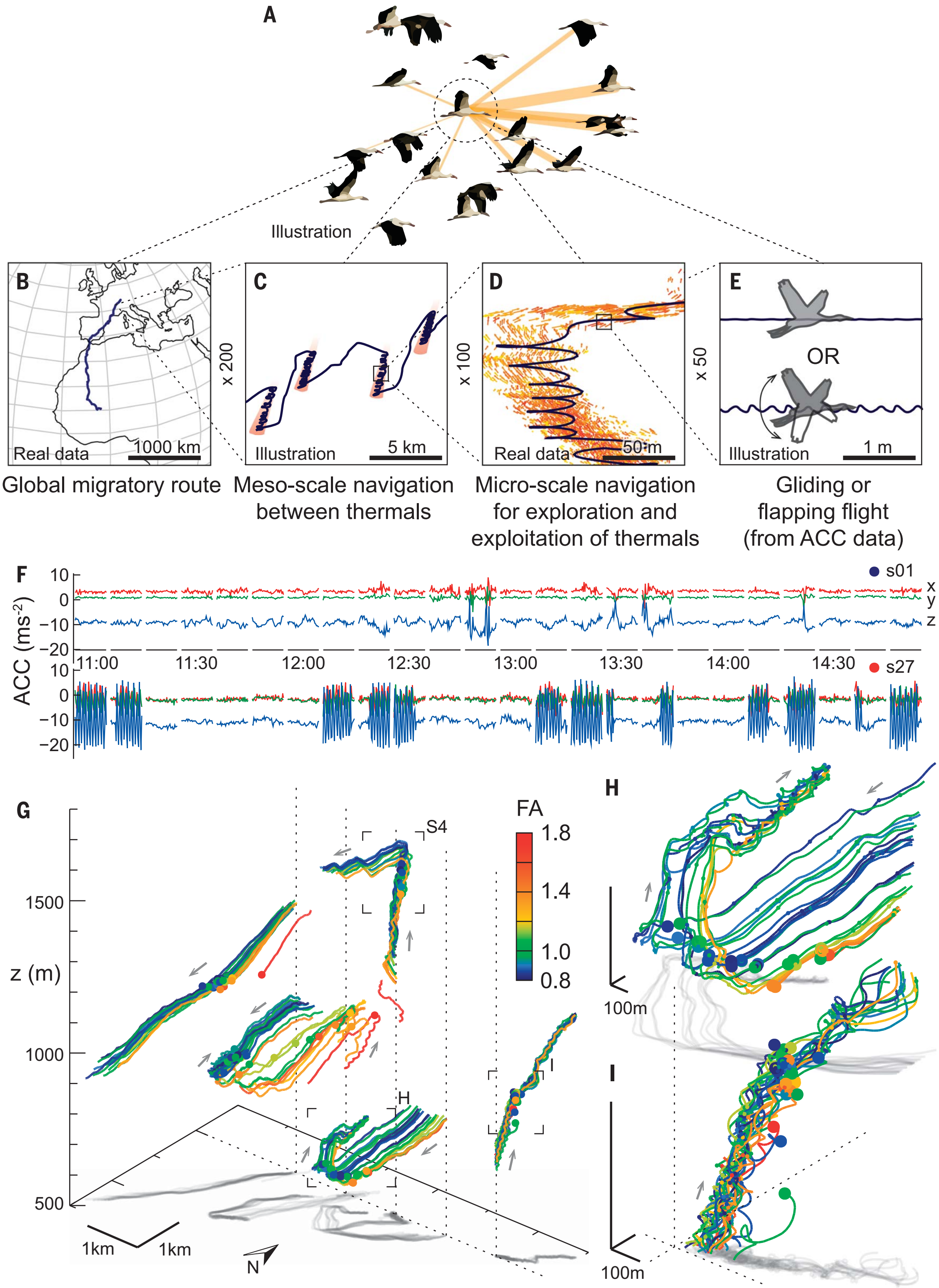


Fig. 1. Collective migration on different spatial scales, recorded using accelerometers and high-resolution GPS. (**A** to **E**) Social interactions during migration (A) shape the global migratory route (B) by influencing small-scale navigational decisions [flight behavior between (C) and within (D) thermals] and individual flight performance (E). Arrow color and size in (D) represent coarse-grained local air velocities estimated from the birds' tracks (*29*). ACC, accelerometer. (**F**) Sample of triaxial accelerometer data used to calculate flapping activity, defined as the standardized mean of daily overall dynamic body acceleration. Plots show data for the birds with the lowest (s01; top) and highest (s27; bottom) flapping activity. (**G**) Five flock trajectories (1-Hz GPS bursts) of migrating storks during thermalling and gliding flight. Bursts are shifted by 1 km for visualization. Gray arrows indicate flight direction. Filled circles show the positions of all individuals at 2 min. Track color corresponds to flapping activity (FA). (**H** and **I**) Enlarged view of the tracks marked in (G). The third area marked in (G) is shown in fig. S4.

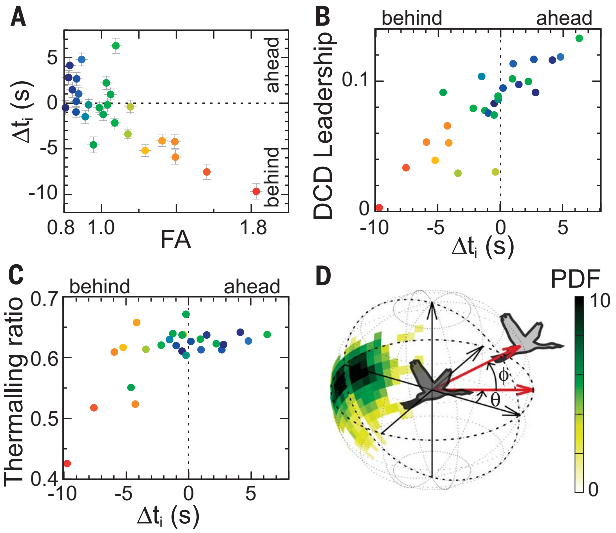


Fig. 2. Relationship of time advance or time delay (Δ*t*) to leadership and flapping activity and relative positions of leaders and followers. (**A**) Individual Δ*t* (Δt_i) averaged over all pairs and bursts, plotted against flapping activity. Bars represent the standard error of the mean (SEM). In (A) to (C), color-coding indicates flapping activity. (**B**) Directional correlation delay (DCD) leadership during gliding, plotted against Δ*t*. (**C**) Average ratio of thermalling time to overall flight time in the first 2 min of the bursts, plotted against Δ*t*. (**D**) Relative position of followers in a spherical coordinate system during gliding. We placed the focal bird (dark gray; one of the five birds with the lowest flapping activity) at the center of the coordinate system and measured the relative position of another bird (light gray; one of the five birds with the highest flapping activity) using the polar (θ) and elevation angle (ϕ) measured from the focal bird's horizontal flight direction. The probability density function (PDF) shows the relative locations of these birds when both were gliding and Δ*t* ∈ (2.5 s, 7.5 s).

climb rates in thermals when following others than when flying ahead (paired *t* test, $n = 22$, $P = 0.030$ and 0.018, respectively; Fig. 3C and figs. S13 and S14).

Examining the complete migratory paths of the 27 birds (at lower temporal resolution) revealed considerable differences in migratory distance, with some birds remaining within Europe and others traveling several thousand kilometers to Africa (Fig. 4). Migratory distance was strongly correlated with the birds' migratory flight behavior; birds that exhibited a high proportion of (costly) flapping activity migrated less far than birds that occupied frontal positions and exhibited low flapping activity when within the flock (Pearson's $r = -0.66$, $n = 20$, $P = 0.001$; Fig. 4, inset). These differences in long-term migration behaviors can be predicted using only a few minutes of movement data from the flock's first migration day (supplementary text and fig. S15). Furthermore, flight time before migration (i.e., total number of GPS bursts in which each bird was found to be flying, before migrating) was also highly correlated with flapping activity (Pearson's $r = -0.648$, $n = 27$, $P = 2.6 \times 10^{-4}$; fig. S16) and migratory distance (Pearson's $r = 0.619$, $n = 20$, $P = 0.004$). The differences in flight performance between leaders and followers suggest that juvenile storks may differ in their aerodynamic features and/or their behavioral strategies, which may affect their migration and group behavior over multiple scales. Nevertheless, birds can compensate for

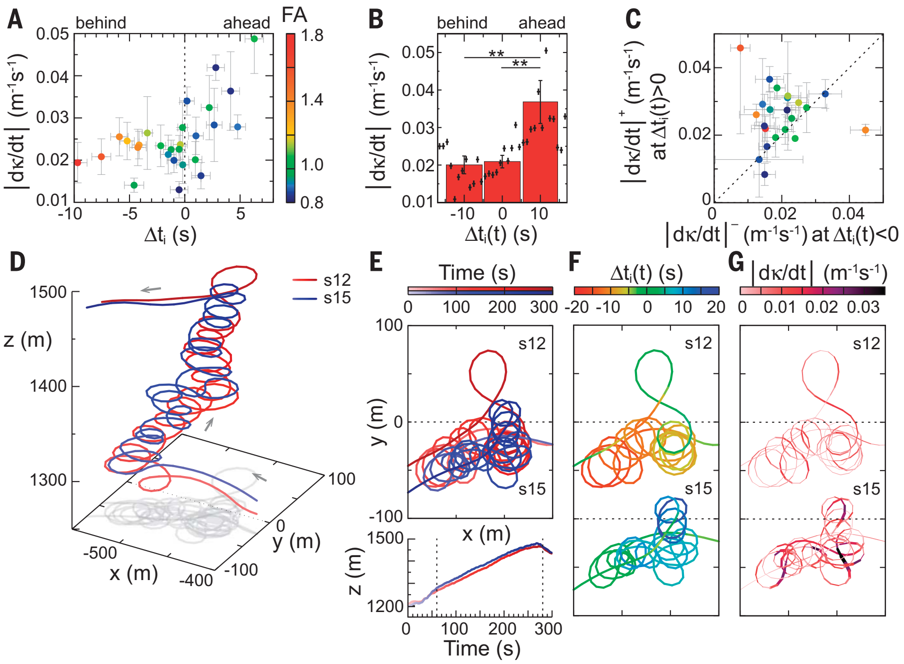


Fig. 3. Derivative of curvature during thermalling flight for leaders and followers. (**A**) Relationship between the absolute value of the time derivative of horizontal path curvature (|d*κ*/d*t*|) and Δ*t*, averaged for each individual across the first migration day. Error bars, SEM. (**B**) |d*κ*/d*t*|, calculated using Δ*t* as 1-s (crosses) and 10-s (bars) bins. Error bars represent the standard deviation of the mean (two-tailed *t* test, $^{**}P < 0.01$, $n = 27$). (**C**) Relationship between |d*κ*/d*t*| while flying ahead of (Δ*t* > 0) and behind (Δ*t* < 0) the average of the flock. Line, $y = x$. Error bars, SEM. (**D**) Example trajectories for illustrating the derivative of curvature. Of all bursts that contained thermalling and had a small wind drift, we chose one random example and depicted the most leading (blue) and following (red) individuals, identified on the basis of their Δ*t* (highest and lowest value of that burst, respectively). (**E**) Horizontal (top) and vertical (bottom) movement components of these tracks. (**F** and **G**) Tracks color-coded to show Δ*t* (F) and |d*κ*/d*t*| (G) (s15 is shifted down for better visualization). |d*κ*/d*t*| is also indicated by line width.

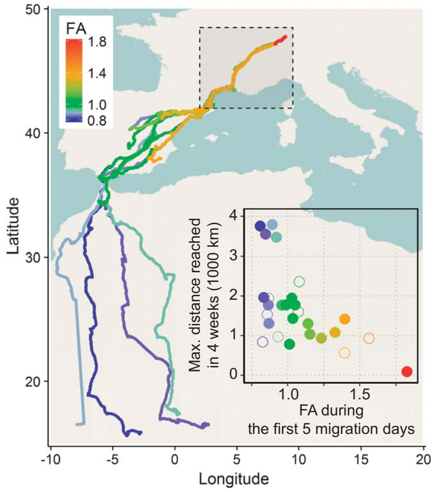


Fig. 4. Relationship between migratory distance and flapping activity. Migration routes of storks during the first 4 weeks of migration. Tracks are color-coded on the basis of the overall flapping activity measured during the high-resolution data acquisition period (black dashed rectangle). The inset shows the relationship between flapping activity and maximum distance reached within 4 weeks. Color corresponds to flapping activity; open circles show birds that died within the first 4 weeks.

their inferior flight skills [e.g., lower glide ratio (ratio of forward speed to sink speed) and more flapping flight] by following others, which enables them to rise faster within thermals (figs. S13 and S14).

Unlike storks, which form large groups with spatiotemporally dynamic structures, other species have been suggested to improve social information usage by flying in V-formation (*28*). Although the number of studies that use advanced tracking technologies to examine collective migration is increasing (*3*, *4*, *29*), the consequences of social behavior and social organization are still largely unknown, especially in wild, freely moving animals. We identified two different behavioral strategies in a flock of migrating white storks, a finding that agrees with theoretical predictions (*2*). We unraveled mechanisms of collective migration in a natural environment by showing how local-scale leader-follower strategies emerge through a differential exploitation of the atmosphere. We suggest that integrating intraspecific interactions into the study of animal movements will enable a better, more mechanistic understanding of broad-scale ecological processes.

REFERENCES AND NOTES

1. E. J. Milner-Gulland, J. M. Fryxell, A. R. E. Sinclair, *Animal Migration: A Synthesis* (Oxford Univ. Press, 2011).
2. V. Guttal, I. D. Couzin, *Proc. Natl. Acad. Sci. U.S.A.* **107**, 16172–16177 (2010).
3. S. J. Portugal *et al*., *Nature* **505**, 399–402 (2014).
4. T. Mueller, R. B. O'Hara, S. J. Converse, R. P. Urbanek, W. F. Fagan, *Science* **341**, 999–1002 (2013).
5. N. Chernetsov, P. Berthold, U. Querner, *J. Exp. Biol.* **207**, 937–943 (2004).
6. A. Berdahl, C. J. Torney, C. C. Ioannou, J. J. Faria, I. D. Couzin, *Science* **339**, 574–576 (2013).
7. J. W. Jolles, N. J. Boogert, V. H. Sridhar, I. D. Couzin, A. Manica, *Curr. Biol.* **27**, 2862–2868.e7 (2017).
8. I. D. Couzin, J. Krause, N. R. Franks, S. A. Levin, *Nature* **433**, 513–516 (2005).
9. A. Strandburg-Peshkin, D. R. Farine, I. D. Couzin, M. C. Crofoot, *Science* **348**, 1358–1361 (2015).
10. A. Flack, B. Pettit, R. Freeman, T. Guilford, D. Biro, *Anim. Behav.* **83**, 703–709 (2012).
11. A. J. King, D. D. P. Johnson, M. Van Vugt, *Curr. Biol.* **19**, R911–R916 (2009).
12. A. M. Simons, *Trends Ecol. Evol.* **19**, 453–455 (2004).
13. E. A. Codling, J. W. Pitchford, S. D. Simpson, *Ecology* **88**, 1864–1870 (2007).
14. C. Torney, Z. Neufeld, I. D. Couzin, *Proc. Natl. Acad. Sci. U.S.A.* **106**, 22055–22060 (2009).
15. R. Kays, M. C. Crofoot, W. Jetz, M. Wikelski, *Science* **348**, aaa2478 (2015).
16. H. Weimerskirch, C. Bishop, T. Jeanniard-du-Dot, A. Prudor, G. Sachs, *Science* **353**, 74–78 (2016).
17. S. Sherub, G. Bohrer, M. Wikelski, R. Weinzierl, *Biol. Lett.* **12**, 20160432 (2016).
18. G. Bohrer *et al*., *Ecol. Lett.* **15**, 96–103 (2012).
19. C. M. Bishop *et al*., *Science* **347**, 250–254 (2015).
20. A. Hedenstrom, *Philos. Trans. R. Soc. London B Biol. Sci.* **342**, 353–361 (1993).
21. A. C. Gleiss, R. P. Wilson, E. L. C. Shepard, *Methods Ecol. Evol.* **2**, 23–33 (2011).
22. A. Flack *et al*., *Sci. Adv.* **2**, e1500931 (2016).
23. C. J. Pennycuick, *Ibis* **111**, 525–556 (1969).
24. Z. Ákos, M. Nagy, T. Vicsek, *Proc. Natl. Acad. Sci. U.S.A.* **105**, 4139–4143 (2008).
25. M. Nagy, Z. Ákos, D. Biro, T. Vicsek, *Nature* **464**, 890–893 (2010).
26. S. Rotics *et al*., *J. Anim. Ecol.* **85**, 938–947 (2016).
27. N. W. Bode *et al*., *Am. Nat.* **179**, 621–632 (2012).
28. B. Voelkl, J. Fritz, *Philos. Trans. R. Soc. London B Biol. Sci.* **372**, 20160235 (2017).
29. M. Nagy, I. D. Couzin, W. Fiedler, M. Wikelski, A. Flack, *Philos. Trans. R. Soc. London B Biol. Sci.* **373**, 20170011 (2018).

ACKNOWLEDGMENTS

We thank all the people who helped during fieldwork, especially W. Schäfle, R. van Noordwijk, B. Kranstauber, D. Piechowski, B. Eid, and Y. Flack. We thank W. Heidrich and F. Kümmeth (e-obs, Munich, Germany) for their suggestions on logger programming and S. Davidson for setting up the Movebank Data Repository. **Funding:** We acknowledge funding from the Max Planck Institute for Ornithology. A.F. was supported by the German Aerospace Center (DLR) and the Christiane Nüsslein-Volhard Stiftung. M.N. received funding from the Royal Society Newton Alumni scheme. I.D.C. acknowledges support from the NSF (IOS-1355061), the ONR (N00014-09-1-1074 and N00014-14-1-0635), the ARO (W911NG-11-1-0385 and W911NF14-1-0431), the Struktur- und Innovationsfunds für die Forschung of the State of Baden-Württemberg, and the Max Planck Society. **Author contributions:** A.F. and M.W. conceived the idea and designed the project; A.F., W.F., and M.W. conducted fieldwork and collected the data; M.N. designed the collective trajectory analyses and visualizations; A.F. and M.N. designed the data analyses and analyzed the data with contributions from I.D.C.; and A.F. and M.N. wrote the text with contributions from I.D.C. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** The data that were used for this study are deposited at the Movebank Data Repository (https://www.movebank.org/node/15294; DOI, 10.5441/001/1.bj96m274).

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/360/6391/911/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S16
Tables S1 to S3
Movies S1 to S3

25 August 2017; accepted 17 April 2018
10.1126/science.aap7781

SYNTHETIC BIOLOGY

An ingestible bacterial-electronic system to monitor gastrointestinal health

Mark Mimee,[1,2]* Phillip Nadeau,[3]*† Alison Hayward,[4,5] Sean Carim,[2] Sarah Flanagan,[3] Logan Jerger,[2,6,7] Joy Collins,[5] Shane McDonnell,[5] Richard Swartwout,[3] Robert J. Citorik,[1,2] Vladimir Bulović,[3] Robert Langer,[5,8] Giovanni Traverso,[5,8,9] Anantha P. Chandrakasan,[3]‡ Timothy K. Lu[2,3,10]‡

Biomolecular monitoring in the gastrointestinal tract could offer rapid, precise disease detection and management but is impeded by access to the remote and complex environment. Here, we present an ingestible micro-bio-electronic device (IMBED) for in situ biomolecular detection based on environmentally resilient biosensor bacteria and miniaturized luminescence readout electronics that wirelessly communicate with an external device. As a proof of concept, we engineer heme-sensitive probiotic biosensors and demonstrate accurate diagnosis of gastrointestinal bleeding in swine. Additionally, we integrate alternative biosensors to demonstrate modularity and extensibility of the detection platform. IMBEDs enable new opportunities for gastrointestinal biomarker discovery and could transform the management and diagnosis of gastrointestinal disease.

Microorganisms living on and in the human body constantly interrogate their biochemical surroundings and alter gene expression to adapt to changing environments. Synthetic biology enables the robust engineering of living cells with increasingly complex genetic circuits to sense biological inputs and control gene expression (*1*). Whole-cell biosensors harness this sensing ability to detect analytes associated with human health (*2*) or environmental contamination (*3*). Owing to their innate robust functionality in complex physiological environments, biosensors have been developed to sense clinically relevant biomarkers in serum or urine ex vivo (*4*), as well as gut biomolecules supplemented in diet (*5*–*7*) or generated during disease (*8*–*10*). However, despite their promise as noninvasive diagnostics, biosensors have yet to be employed for clinically compatible testing in an unobtrusive, real-time, and user-friendly way. Current research applications of ingestible biosensors in animal models rely on cumbersome analysis of bacterial gene expression or DNA in stool samples (*5*–*10*), rather than real-time reporting from within the body.

By contrast, the impressive scaling of semiconductor microelectronics over the past few decades has delivered sophisticated, highly miniaturized platforms for ultra–low-power sensing, computation, and wireless communication (*11*–*13*). In turn, these have enabled the recording of patient compliance and the evaluation of the gastrointestinal tract by using optical images, gases, temperature, and pH (*14*–*17*). However, the ability of electronics to directly and selectively sense biomolecules in vivo is limited by the availability of labile biochemical transducers and the size of the power-demanding circuits required to sense them.

Here, we describe an ingestible micro-bio-electronic device (IMBED) that combines engineered probiotic sensor bacteria together with ultra–low-power microelectronics to enable in situ detection of gastrointestinal biomolecules associated with health or disease (fig. S1). By partitioning sensing to biological systems and computation and communication to electrical devices, IMBEDs leverage the natural advantages of each approach to enable ingestible gastrointestinal diagnostics. In an IMBED, biosensor probiotics lie adjacent to readout electronics in individual wells separated from the outside environment by a semipermeable membrane that confines cells in the device and allows for diffusion of small molecules. Sensing of target biomarkers by the bacteria generates light, which is detected by photodetectors embedded in the electronics. These electrical signals are processed by an integrated bioluminescence detection circuit (*18*) and are transmitted wirelessly from the device to an external radio or cellular phone for convenient readout.

As a proof-of-concept IMBED for a clinically relevant biomarker, we developed a biosensor for gastrointestinal bleeding events via heme liberated from lysed red blood cells. Although cost-effective fecal occult-blood testing is available (*19*), diagnosis of acute bleeding in the upper gastrointestinal tract often requires endoscopic observation (*20*), and IMBEDs could offer a rapid, minimally invasive, and cost-effective means of detection. The heme biosensor was based on a synthetic promoter ($P_{L(HrtO)}$) (fig. S2A), regulated by the *Lactococcus lactis* heme-responsive transcriptional repressor, HrtR (*21*), and ChuA, an outer-membrane transporter from *Escherichia coli* O157:H7 that allows for the transit of extracellular heme through the cell envelope (Fig. 1A) (*22*). *Photorhabdus luminescens luxCDABE* was used as the output of the genetic circuit as it functions at body temperature and encodes all components necessary for intracellular substrate production (*23*). The resultant prototype biosensor in laboratory *E. coli* (MG1655 V1) responded to increasing heme input with luminescence output with a signal-to-noise ratio (SNR) of 5.9 and a Hill function threshold (K_D) of 1 μM heme (fig. S2B). Luminescence production was also induced by whole horse blood (fig. S2C), and lysis of red blood cells in simulated gastric fluid greatly improved sensitivity by liberating heme [K_D = 115 parts per million (ppm) blood] (Fig. 1B and fig. S2D). Next, the prototype genetic circuit was iteratively optimized with the goal of improving SNR without compromising maximum luminescence output (MG1655 V2; Fig. 1B and fig. S3). The final gene circuit was transferred to the probiotic *E. coli* Nissle 1917 (Nissle V2), and the resultant strain responded rapidly to lysed horse blood (SNR = 310; K_D = 95 ppm; $t_{1/2}$ = 45 min) (Fig. 1B and fig. S4) as well as human blood (fig. S5).

To examine functionality of the bacterial blood sensor in vivo, we evaluated whether bacterial blood sensors passing through the gut could detect upper gastrointestinal bleeding elicited by oral indomethacin administration (Fig. 1C) (*24*). At baseline, administration of the biosensors did not lead to detectable luminescence activity in stool, indicating that basal heme concentrations in the murine gut are insufficient to activate the gene circuit (fig. S6). After oral administration of indomethacin, blood sensor bacteria demonstrated 18-fold higher luminescence values in fecal pellets as compared to controls (Fig. 1C). Our designed biosensor can thus effectively detect the presence of gastrointestinal bleeding in vivo.

We next sought to integrate our bacterial biosensor with an electronic sensor and wireless transmission platform. Prior demonstrations of sensitive bioluminescence detection electronics have required external wiring and have been limited to bench-top assays (*18*, *25*, *26*), whereas a low-power wireless solution could enable convenient in vivo detection. We developed a miniaturized, fully integrated, wireless readout

---

[1]Microbiology Program, Massachusetts Institute of Technology (MIT), Cambridge, MA 02139, USA. [2]Synthetic Biology Center, MIT, Cambridge, MA 02139, USA. [3]Department of Electrical Engineering and Computer Science, MIT, Cambridge, MA 02139, USA. [4]Division of Comparative Medicine, MIT, Cambridge, MA 02139, USA. [5]Koch Institute for Integrative Cancer Research, MIT, Cambridge, MA 02139, USA. [6]Division of Pediatric Gastroentrology, Hepatology, and Nutrition, Department of Pediatrics, MassGeneral Hospital for Children, Boston, MA 02114, USA. [7]Harvard Medical School, Boston, MA 02115, USA. [8]Department of Chemical Engineering, MIT, Cambridge, MA 02139, USA. [9]Division of Gastroenterology, Brigham and Women's Hospital, Boston, MA 02115, USA. [10]Department of Biological Engineering, MIT, Cambridge, MA 02139, USA.
*These authors contributed equally to this work. †Present address: Analog Devices, Boston, MA 02110, USA.
‡Corresponding author. Email: timlu@mit.edu (T.K.L.); anantha@mtl.mit.edu (A.P.C.)

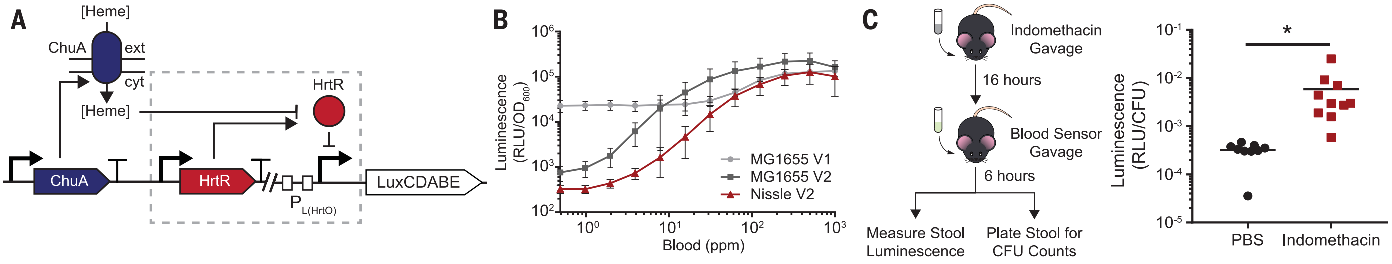


Fig. 1. Probiotic *E. coli* can be engineered to sense blood in vitro and in vivo. (**A**) Schematic of the blood sensor gene circuit. Extracellular heme is internalized through the outer-membrane transporter ChuA and interacts with the transcriptional repressor HrtR to allow for expression of the bacterial luciferase operon *luxCDABE*. (**B**) Dose-response curves of prototype (V1) and optimized (V2) heme-sensing genetic circuits in laboratory (MG1655) and probiotic (Nissle) strains of *E. coli*. Error bars represent SEM of three independent biological replicates. (**C**) In vivo blood sensor performance. C57BL/6J mice were administered indomethacin (10 mg/kg) to induce gastrointestinal bleeding or vehicle (PBS, phosphate-buffered saline) and inoculated with blood sensor *E. coli* Nissle cells the following day. Normalized luminescence values of fecal pellets were significantly higher in mice administered indomethacin compared to control animals (*P = 0.04; Student's t test; N = 10 mice). CFU, colony-forming units; RLU, relative luminescence units.

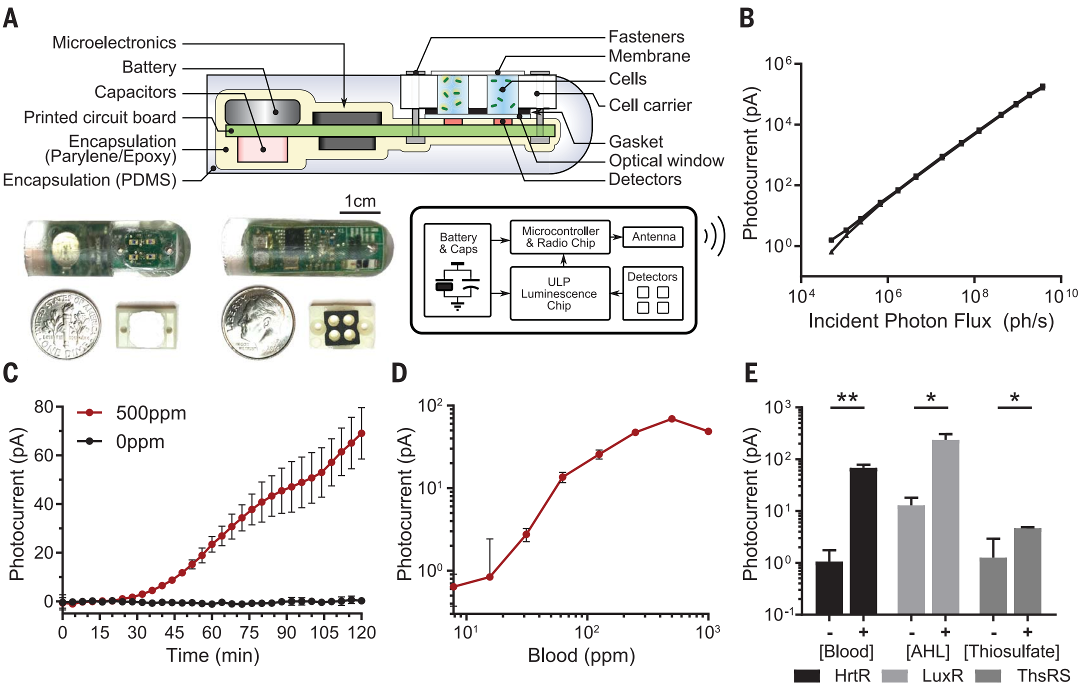


Fig. 2. Design and in vitro evaluation of IMBED for miniaturized wireless sensing with cellular biosensors. (**A**) Cross section, electrical system diagram, and front- and back-side photos of the device. PDMS, polydimethylsiloxane. (**B**) System photocurrent response measured without cells. The incident photon flux was supplied by green light-emitting diode (wavelength λ = 525 nm) and calibrated with an optical power meter (individual traces shown for N = 3 devices). (**C**) Kinetic response of blood sensor IMBED in bacterial growth media supplemented with 0 and 500 ppm blood. (**D**) Dose-response of blood sensor IMBEDs in bacterial growth media containing different blood concentrations 2 hours after exposure. The leftmost data point represents the background response in the absence of blood. (**E**) Detection of multiple gut-relevant small molecules with IMBEDs. HrtR-, LuxR- and ThsRS-containing *E. coli* Nissle strains in IMBEDs were exposed to 500 ppm blood, 100 nM acyl-homoserine lactone (AHL), or 10 mM thiosulfate for 2 hours. In (C) to (E), error bars denote the SEM for three independent biological replicates conducted with different IMBEDs. *P < 0.05, **P < 0.01, Student's t test.

capsule for targeted sensing of small molecules in the gastrointestinal tract (Fig. 2A). The device combined our prior nanowatt-level time-based luminometer chip (*18*) with a microprocessor, wireless transmitter, and a set of phototransistors inside a molded capsule. Bioluminescence from activated cells was detected by phototransistors located below each cavity. The detected luminescence was converted to a digital code by the low-power luminometer chip and transmitted wirelessly outside the body for calibration, display, and recording. The small button-cell battery inside the capsule (5 mAh) powered the device, and the measured power consumption (table S1) suggests a nominal device shelf-life of >9 months and active operation time of 1.5 months on a full charge. The electronic system was highly sensitive and captured photon flux as low as 5 × 10^4 photons/s incident on the 0.29-mm^2 area of the detectors [white-noise–limited coefficient of variation (CV) 13%, Fig. 2B and fig. S8A]. The mean channel mismatch was less than 6% (CV) (fig. S7A), and mean temperature-induced drift across 5°C variation was less than 2 pA (fig. S7B). Additionally, the electronic system was stable in simulated gastric fluid for up to 36 hours (fig. S7C), providing sufficient time to perform an ingestible measurement during gastrointestinal transit.

To demonstrate integration of the ingestible luminometer capsule and engineered biosensors, we tested the probiotic blood sensor strains in an IMBED in vitro. Upon exposure to 500 ppm blood, induced bioluminescence could be observed in as soon as 30 min (Fig. 2C). The dose-response curve of blood sensor IMBEDs was similar to that of plate-reader measurements (SNR = 76; K_D = 135 ppm; compare Fig. 1B and 2D), with saturation achieved at 250 ppm and significant detection as low as 32.5 ppm blood (fig. S8) (Student's t test; P = 0.03). By combining cellular sensors with ultra–low-power electronic readout, IMBEDs serve as a flexible platform for sensitive detection of bleeding in fluidic environments.

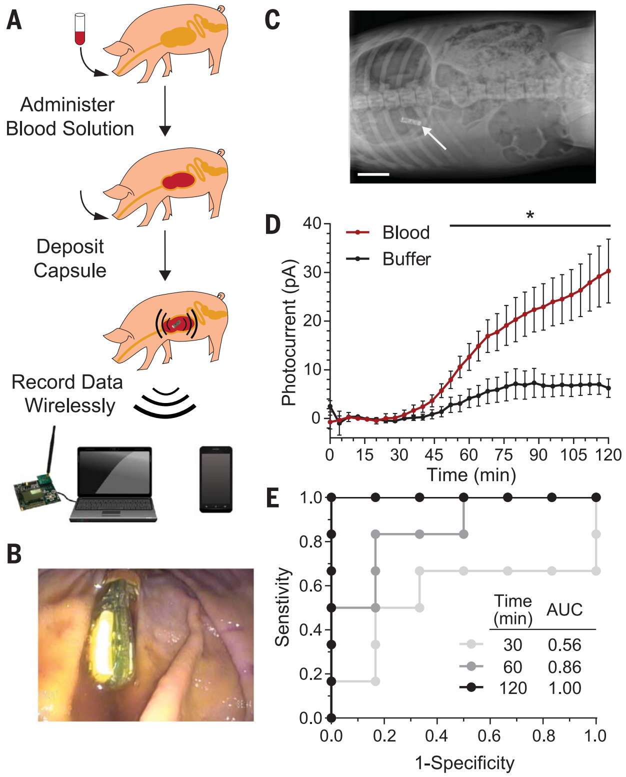


Fig. 3. IMBEDs can rapidly detect porcine gastric bleeding. (**A**) Schematic of the experimental flow, which consisted of blood administration in neutralization solution, capsule deposition, and wireless transmission to a commercial receiver connected to a laptop or a cellular phone. Representative endoscopic (**B**) and x-ray (**C**) images illustrate the location of the device in the stomach at the conclusion of our 2-hour experiments, just before device removal [scale bar (C), 5 cm]. (**D**) Kinetic response of blood sensor IMBED in a porcine model of gastric bleeding. IMBEDs deposited in gastric cavity can rapidly discriminate between pigs administered blood versus buffer control. Error bars denote SEM for six IMBED experiments (three animals on different days, two capsules per animal). $^*P < 0.05$, Student's *t* test. (**E**) Receiver operating characteristic (ROC) of IMBED sensing over time. Perfect detection is achieved at $t = 120$ min. AUC, area under the curve.

In addition to blood sensing, we adapted IMBEDs to sense alternative biomarkers, thiosulfate and acyl-homoserine lactone (AHL). Thiosulfate could serve as a biomarker of gut inflammation as it is elevated in murine models of colitis (*8*). AHLs are molecular signatures of particular bacteria, and their detection could indicate the presence of commensal or infectious agents in the gut microbiota (*27*, *28*). Thiosulfate- and AHL-inducible genetic circuits were introduced into *E. coli* Nissle, and exposure to inducer elicited a dose-dependent bioluminescence response (fig. S9). These alternative biosensors were integrated with IMBEDs, and different analytes were readily detectable in a fluidic environment (Fig. 2E). As additional biosensors of clinically relevant gut biomarkers continue to be developed, we anticipate that the breadth of potential analytes of the IMBED platform will continue to expand.

To examine wireless in situ detection of biomolecules with biosensors, we deployed blood sensor IMBEDs in a porcine model of gastrointestinal bleeding. Prior to device deposition, pigs were administered a bicarbonate-glucose neutralization solution with or without 0.25 ml of blood (Fig. 3A). The neutralization solution helped buffer the low pH of the porcine gastric fluid as acidic environments degrade the functionality of the biosensor (fig. S10). The blood sensor IMBED was subsequently deposited into the stomach via orogastric tube and remained resident and stable in the gastric cavity for the entire duration of the experiment (Fig. 3, B and C). Photocurrent data were wirelessly transmitted from the stomach over the course of 2 hours and logged by both a laptop computer and an Android phone equipped with a custom application for real-time data processing and visualization (figs. S11 and S12). The presence of blood in the porcine gastric environment could be observed as early as 52 min (Student's *t* test; $P < 0.05$) and led to a fivefold increase in photocurrent after 120 min as compared to animals given buffer alone (Fig. 3D and fig. S13). Luminescence production was not detected in biosensors lacking the ChuA heme transporter or the luciferase operon, indicating that observed light production was dependent on a functional genetic circuit activated in the presence of heme (fig. S14). The receiver operating characteristic of the blood-sensing IMBED improved over time, with a sensitivity and specificity of 83.3% at 60 min and 100% at 120 min (Fig. 3E). IMBEDs can thus detect small amounts of analyte in the harsh gastric environmental with high specificity and sensitivity.

By combining the environmental resilience and natural sensing properties of bacterial cells with the complex data processing and wireless transmission afforded by ultra–low-power microelectronics, we developed a device capable of in vivo biosensing in harsh, difficult-to-access environments. Using gastrointestinal bleeding as a proof-of-concept model system, we demonstrate strategies for genetic circuit design and optimization, fabrication of an ingestible low-power, wireless luminometer, and validation of integrated system functionality both in vitro and in a large animal model. As the field of whole-cell biosensors matures, newly developed sensors of clinically relevant biomarkers could be rapidly integrated into an IMBED to perform minimally invasive detection in the gastrointestinal tract. With a test panel of candidate biomolecules, IMBEDs could enable studies of the biochemistry of anatomical regions that are traditionally difficult to access and could lead to the discovery of new clinical biomarkers associated with health or disease. The in situ detection afforded by IMBEDs could also allow sensing of labile gut or microbiota-derived biomolecules that would otherwise be degraded before excretion in stool. Further integration of electronic modules, such as photodetectors, microprocessor, and transmitter, in a single integrated circuit could allow for further miniaturization of IMBEDs as well as lower power consumption. Additional measurement channels would also enable more precise biochemical readings, as the response of replicate biosensors within the same device could be averaged to mitigate the inherent variance of biological sensors as well as the heterogeneity of the complex gastrointestinal environment. Improved preparation of bacterial cultures for long-term storage, such as lyophilization, could be implemented to extend the shelf-life of fully assembled IMBEDs. Furthermore, the devices could be equipped with new orally delivered encapsulation technologies to enable long-term residency, monitoring, and anatomic localization in the gastrointestinal tract (*29*, *30*). This integration of biological engineering and semiconductor electronics offers opportunities to transform diagnosis, management, and monitoring of health and disease.

REFERENCES AND NOTES

1. J. A. N. Brophy, C. A. Voigt, *Nat. Methods* **11**, 508–520 (2014).
2. S. Slomovic, K. Pardee, J. J. Collins, *Proc. Natl. Acad. Sci. U.S.A.* **112**, 14429–14435 (2015).
3. C. Roggo, J. R. van der Meer, *Curr. Opin. Biotechnol.* **45**, 24–33 (2017).
4. A. Courbet, D. Endy, E. Renard, F. Molina, J. Bonnet, *Sci. Transl. Med.* **7**, 289ra83 (2015).
5. J. W. Kotula *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **111**, 4838–4843 (2014).
6. M. Mimee, A. C. Tucker, C. A. Voigt, T. K. Lu, *Cell Syst.* **1**, 62–71 (2015).
7. B. Lim, M. Zimmermann, N. A. Barry, A. L. Goodman, *Cell* **169**, 547–558.e15 (2017).
8. K. N.-M. Daeffler *et al.*, *Mol. Syst. Biol.* **13**, 923 (2017).
9. D. T. Riglar *et al.*, *Nat. Biotechnol.* **35**, 653–658 (2017).
10. J. M. Pickard *et al.*, *Nature* **514**, 638–641 (2014).
11. B. Otis, B. Parviz, Introducing our smart contact lens project. *Google Off. Blog* (2014); available at https://googleblog.blogspot.com/2014/01/introducing-our-smart-contact-lens.html.
12. H. Wang, *IEEE Microw. Mag.* **14**, 110–130 (2013).
13. H. Norian, R. M. Field, I. Kymissis, K. L. Shepard, *Lab Chip* **14**, 4076–4084 (2014).
14. G. Iddan, G. Meron, A. Glukhovsky, P. Swain, *Nature* **405**, 417–417 (2000).
15. K. Kalantar-Zadeh *et al.*, *Nat. Electron. 2017 11.* **1**, 79 (2018).
16. P. J. van der Schaar *et al.*, *Gastrointest. Endosc.* **78**, 520–528 (2013).
17. H. Hafezi *et al.*, *IEEE Trans. Biomed. Eng.* **62**, 99–109 (2015).

18. P. Nadeau, M. Mimee, S. Carim, T. K. Lu, A. P. Chandrakasan, Nanowatt circuit interface to whole-cell bacterial sensors. *2017 IEEE International Solid-State Circuits Conference (ISSCC)*, San Francisco, CA, 2017, pp. 352–353.
19. D. C. Rockey, J. Koch, J. P. Cello, L. L. Sanders, K. McQuaid, *N. Engl. J. Med.* **339**, 153–159 (1998).
20. A. Barkun, M. Bardou, J. K. Marshall, *Ann. Intern. Med.* **139**, 843–857 (2003).
21. D. Lechardeur *et al.*, *J. Biol. Chem.* **287**, 4752–4758 (2012).
22. C. L. Nobles, J. R. Clark, S. I. Green, A. W. Maresso, *J. Microbiol. Methods* **118**, 7–17 (2015).
23. D. Close *et al.*, *Sensors (Basel)* **12**, 732–752 (2012).
24. A. Lanas, F. K. L. Chan, *Lancet* **390**, 613–624 (2017).
25. H. Eltoukhy, K. Salama, A. El Gamal, *IEEE J. Solid-State Circuits* **41**, 651–662 (2006).
26. R. R. Singh, L. Leng, A. Guenther, R. Genov, *IEEE J. Solid-State Circuits* **47**, 2822–2833 (2012).
27. I. Y. Hwang *et al.*, *Nat. Commun.* **8**, 15028 (2017).
28. M. Schuster, D. J. Sexton, S. P. Diggle, E. P. Greenberg, *Annu. Rev. Microbiol.* **67**, 43–63 (2013).
29. A. M. Bellinger *et al.*, *Sci. Transl. Med.* **8**, 365ra157 (2016).
30. A. R. Kirtane *et al.*, *Nat. Commun.* **9**, 2 (2018).

ACKNOWLEDGMENTS

The thiosulfate sensor plasmids were a gift from J. Tabor (Rice University). The *luxCDABE* genes were obtained from pAKlux2, a gift from A. Karsi (Addgene plasmid no. 14080). We thank D. Glettig for help with pilot experiments in rodents and M. Goulamaly for preparing initial parylene samples. **Funding:** Supported by Texas Instruments (A.P.C.); the Hong Kong Innovation and Technology Fund (ITS/195/14FP) (A.P.C. and T.K.L.); Office of Naval Research (N00014-13-1-0424) (T.K.L.); NSF Biological Computing (1522074) (T.K.L.); the Center for Microbiome Informatics and Therapeutics (15127713) (T.K.L.); Division of Gastroenterology, Brigham and Women's Hospital (G.T.); NIH (EB-000244) (R.L.); the Qualcomm Innovation Fellowship (M.M. and P.N.); the HHMI International Student Fellowship (M.M.); and the Natural Sciences and Engineering Council of Canada Fellowship (P.N.). We also thank the TSMC University Shuttle Program for chip fabrication. **Author contributions:** M.M., P.N., A.H., G.T., R.L., A.P.C., T.K.L. conceived and designed the research; M.M. and S.C. designed and performed in vitro biological experiments; P.N. designed the electronic capsule and packaging strategy; M.M. and P.N. performed in vitro capsule experiments; S.F. wrote the mobile phone application; R.S. performed Parylene C deposition; M.M., L.J., and R.J.C. performed in vivo mouse experiments; M.M., P.N., L.J., A.H., J.C., and S.M. conducted in vivo pig experiments; M.M., P.N., A.H., L.J., R.S., V.B., R.L., G.T., A.P.C., and T.K.L. analyzed the data and wrote the manuscript. **Competing interests:** M.M. and T.K.L. have filed a patent application based on the microbial heme sensor (US20170058282A1) with the U.S. Patent and Trademark Office. M.M., P.N., A.P.C., and T.K.L have filed a patent application based on the integrated device (PCT/US2018/027904) with the U.S. Patent and Trademark Office. **Data and materials availability:** Genetic constructs are available in Addgene.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/360/6391/915/suppl/DC1
Materials and Methods
Tables S1 to S3
Figs S1 to S14
References (*31–35*)

7 January 2018; accepted 20 April 2018
10.1126/science.aas9315

MOLECULAR BIOLOGY

RNA buffers the phase separation behavior of prion-like RNA binding proteins

Shovamayee Maharana,[1] Jie Wang,[1]* Dimitrios K. Papadopoulos,[1,2]* Doris Richter,[1] Andrey Pozniakovsky,[1] Ina Poser,[1] Marc Bickle,[1] Sandra Rizk,[1,3] Jordina Guillén-Boixet,[1] Titus M. Franzmann,[1] Marcus Jahnel,[1,4] Lara Marrone,[5] Young-Tae Chang,[6,7] Jared Sterneckert,[5] Pavel Tomancak,[1] Anthony A. Hyman,[1]† Simon Alberti[1]†

Prion-like RNA binding proteins (RBPs) such as TDP43 and FUS are largely soluble in the nucleus but form solid pathological aggregates when mislocalized to the cytoplasm. What keeps these proteins soluble in the nucleus and promotes aggregation in the cytoplasm is still unknown. We report here that RNA critically regulates the phase behavior of prion-like RBPs. Low RNA/protein ratios promote phase separation into liquid droplets, whereas high ratios prevent droplet formation in vitro. Reduction of nuclear RNA levels or genetic ablation of RNA binding causes excessive phase separation and the formation of cytotoxic solid-like assemblies in cells. We propose that the nucleus is a buffered system in which high RNA concentrations keep RBPs soluble. Changes in RNA levels or RNA binding abilities of RBPs cause aberrant phase transitions.

The intracellular environment is organized into membraneless compartments that have been termed biomolecular condensates because they form by liquid-liquid phase separation (*1*, *2*). These condensates often contain RNA binding proteins (RBPs) with distinctive domains, so-called prion-like domains, which are structurally disordered and contain polar amino acids (*3*) (Fig. 1A). Interactions between prion-like domains and additional interactions between RNAs and RNA binding domains drive the assembly of prion-like RBPs by phase separation (*4*, *5*). However, several prion-like RBPs, such as FUS, TDP43, and hnRNPA1, can also undergo an aberrant transition from a liquid-like state into solid aggregates that has been linked to neurodegenerative diseases such as amyotrophic lateral sclerosis (ALS) (*4*–*6*). One important aspect of these diseases is that aggregate formation is strongly associated with the subcellular location of the proteins. Aggregates in patient neurons are usually found in the cytoplasm, whereas the nucleus is usually devoid of the aggregating proteins (*7*–*10*), although there are some noteworthy exceptions (*11*). Disease-causing mutations frequently affect the nuclear partitioning of prion-like RBPs (*12*, *13*), highlighting the importance of cytoplasmic mislocalization in disease. Protein mislocalization to the cytoplasm causes loss of function for the nucleus and gain of function in the cytoplasm, phenotypes that are thought to underlie disease (*14*–*17*). Importantly, genetic relocalization of FUS to the nucleus in yeast strongly decreases FUS toxicity (*18*). This suggests that the localization of FUS to the nuclear environment suppresses its pathological behavior, which raises two important questions: What prevents prion-like RBPs from forming solid-like aggregates in the nucleus? And why do these RBPs form aggregates in the cytoplasm?

To answer these questions, we investigated the phase behavior of several prion-like RBPs (Fig. 1A). First, we determined the nuclear concentrations of these proteins. The values ranged from 0.2 μM for TAF15 to 42.3 μM for hnRNPA1 (Fig. 1, B to D, and supplementary methods). Next, we purified these proteins as green fluorescent protein (GFP) fusions and added them to a physiological buffer. At a concentration similar to the nuclear concentration (7.6 μM), FUS phase-separated into droplets (Fig. 1, E and F). This behavior contrasted with that in living cells, where only 1% of the nuclear FUS protein was contained in condensates (Fig. 1F), which are paraspeckles (*19*). The remaining 99% of nuclear FUS protein was diffusely localized. Similar observations were made for TDP43, EWSR1, TAF15, and hnRNPA1 (Fig. 1G, bottom panels). These results suggest that although the protein concentration is high enough for phase separation in the nucleus, an additional nuclear factor prevents phase separation.

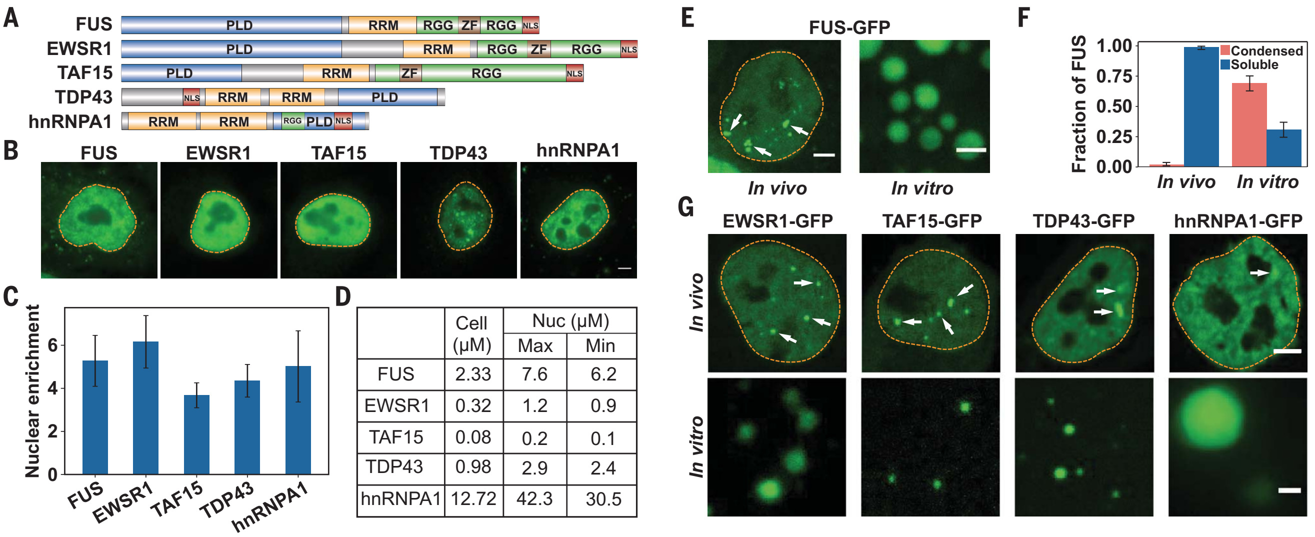
We hypothesized that nuclear RNA could regulate the phase behavior of prion-like RBPs. To test this idea, we performed an in vitro phase separation assay with FUS in the presence of total RNA (Fig. 2A). In agreement with previous work (*20*–*22*), we found that small amounts of RNA promoted liquid droplet formation (Fig. 2B and fig. S1, A to D). RNA-containing droplets contained a higher FUS concentration than RNA-free droplets, and they appeared slightly more viscous (fig. S2, A to C). However, upon further increase in the RNA/protein concentration ratio, the droplets became smaller and finally dissolved (Fig. 2, A and B, and fig. S3). The addition of RNase A resulted in droplet reappearance (Fig. 2D and figs. S4A, panels on the right, and S5), indicating that droplet solubilization depends on intact RNA. Similar results were obtained for EWSR1, TAF15, hnRNPA1, and TDP43 (Fig. 2C). Thus, we conclude that high RNA/protein ratios prevent phase separation and that low ratios promote phase separation.

We next tested whether different types of RNAs differ with respect to their abilities to dissolve FUS droplets. Individually, ribosomal RNA, tRNA, and a noncoding RNA that is known to bind to FUS (Neat1) were all able to solubilize FUS droplets, suggesting a general effect, but smaller RNAs were more potent than larger ones (fig. S4, A to D). Secondary structure was important for enriching FUS in droplets, consistent with results in previous work (*20*), but secondary structure (fig. S4, A to E) and binding affinity (fig. S6) affected droplet solubilization only slightly. We next asked whether the cellular RNA concentration is high enough to suppress phase separation of FUS. We estimated that the nuclear RNA concentration is ~10.6 times as high as that required for droplet dissolution in vitro (fig. S7 and supplementary methods). However, ~1% of nuclear FUS formed condensates (paraspeckles) (Fig. 1E) by binding to the noncoding RNA Neat1 (*19*). To test whether Neat1 could nucleate FUS droplets in the presence of a high background concentration of RNA, we added Neat1 RNA to a FUS sample that had been solubilized with tRNA. This led to a reappearance of FUS droplets (Fig. 2E and fig. S4F). We attribute this result to the ability of Neat1 to form large RNA assemblies (fig. S4C), which subsequently recruit FUS. This observation suggests that highly structured RNAs such as Neat1 act as scaffolds that promote the nucleation of condensates in the high–RNA concentration environment of the nucleus. A similar scenario may apply for stress granules in the cytoplasm, which contain large amounts of structured polyadenylated mRNA (fig. S8).

To test experimentally whether the high nuclear RNA concentration keeps FUS soluble, we microinjected ribonuclease A (RNase A) into the nuclei of HeLa cells. Immediately after RNase A injection, FUS-GFP condensed into many liquid-like droplets (Fig. 3A, fig. S9, and movie S1), and this effect was not due to a general loss of nuclear integrity (figs. S10 and S11). As an alternative approach to decrease the RNA/protein ratio, we injected purified FUS-GFP into the nucleus, which led to an immediate increase in the number and size of nuclear FUS assemblies (fig. S12). RNase A microinjection

---

[1]Max Planck Institute of Molecular Cell Biology and Genetics, Pfotenhauerstraße 108, 01307 Dresden, Germany. [2]MRC Human Genetics Unit, Institute of Genetics and Molecular Medicine, University of Edinburgh, Crewe Road, Edinburgh EH4 2XU, UK. [3]B Cube–Center for Molecular Bioengineering, Technische Universität Dresden, Arnoldstraße 18, 01307 Dresden, Germany. [4]Biotechnology Center, Technische Universität Dresden, Tatzberg 47/49, 01307 Dresden, Germany. [5]Technische Universität Dresden–Center for Molecular and Cellular Bioengineering (CMCB), DFG–Center for Regenerative Therapies Dresden, 01307 Dresden, Germany. [6]Center for Self-Assembly and Complexity, Institute for Basic Science (IBS), Pohang 37673, Republic of Korea. [7]Department of Chemistry, Pohang University of Science and Technology (POSTECH), Pohang 37673, Republic of Korea.
*These authors contributed equally to this work.
†Corresponding author. Email: alberti@mpi-cbg.de (S.A.); hyman@mpi-cbg.de (A.A.H.)



	Cell (μM)	Nuc (μM)	
		Max	Min
FUS	2.33	7.6	6.2
EWSR1	0.32	1.2	0.9
TAF15	0.08	0.2	0.1
TDP43	0.98	2.9	2.4
hnRNPA1	12.72	42.3	30.5

Fig. 1. Prion-like RBPs phase-separate at their physiological concentrations. (**A**) Domain structure. PLD, prion-like domain; RRM, RNA recognition motif; RGG, arginine- and glycine-rich region; ZF, zinc finger; NLS, nuclear localization sequence. (**B**) Representative images of immunostained HeLa cells. Dashed lines indicate the nuclear boundary. Scale bar, 5 μm. (**C**) Quantification of the nuclear enrichment of RBPs. Error bars represent SD. (**D**) Calculated cellular and nuclear (Nuc) concentrations of RBPs in HeLa cells. (**E**) Left, live HeLa cell nucleus expressing GFP-tagged FUS from a bacterial artificial chromosome (BAC). Arrows point to paraspeckles. Right, FUS-GFP phase-separated in vitro at 7.5 μM. Scale bars, 2 μm. (**F**) Quantification of the fractions of FUS present in condensed and soluble states in vivo and in vitro. Error bars represent SD. (**G**) Top, HeLa cell nuclei expressing GFP-tagged RBPs from BACs. White arrows indicate condensates. Bottom, purified RBPs phase-separate at their respective nuclear concentrations. Scale bars, 2 μm.

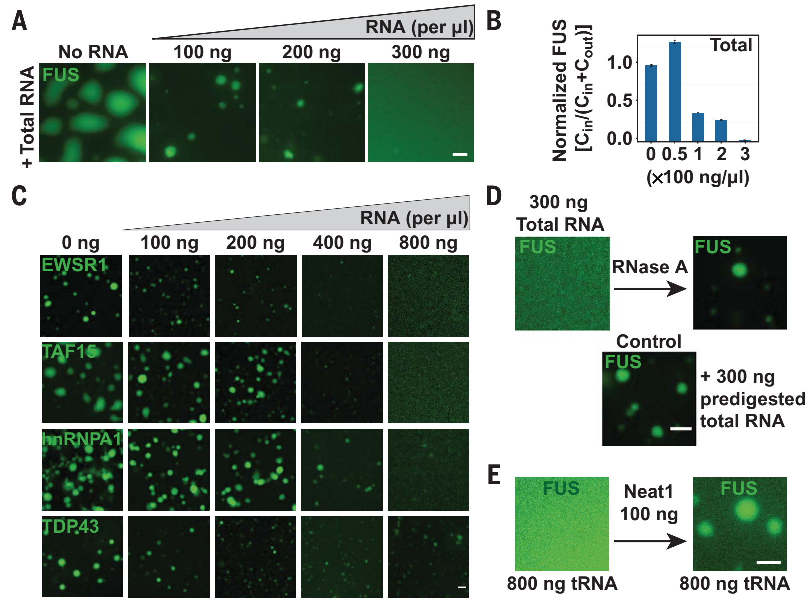


Fig. 2. RNA regulates the phase behavior of prion-like RBPs. (**A**) Representative images of purified FUS-GFP (5 μM) in vitro in the presence of total RNA. (**B**) Quantification of the fraction of condensed FUS-GFP. C_{in}, fraction of total protein in the droplets; C_{out}, fraction of total protein in the soluble phase outside the droplets. The value of FUS enrichment in the droplet phase in the absence of RNA was normalized to 1. Error bars represent SD. (**C**) In vitro phase separation assay with EWSR1, TAF15, hnRNPA1, or TDP43 in the presence of total RNA. (**D**) Addition of RNase A to a sample of FUS-GFP (5 μM) solubilized with 300 ng/μl of total RNA. (**E**) Left, FUS-GFP (5 μM) solubilized with 800 ng/μl tRNA in vitro. Right, FUS phase separation triggered by the addition of 100 ng/μl Neat1 RNA in the presence of tRNA (800 ng/μl). Scale bars in (A), (C), (D), and (E), 2 μm.

into the nucleus also triggered rapid phase separation of hnRNPA1, EWSR1, TDP43, and TAF15 (figs. S13 and S14). To investigate whether FUS forms complexes with RNA in living cells, we used fluorescence correlation spectroscopy (FCS). We identified two populations of FUS, one slow moving and one fast moving (details are in supplementary methods). We estimate that the amount of slow FUS in the nucleus is 10 times as high as that in the cytoplasm (Fig. 3, B to D; fig. S15, A to E; and supplementary methods). The fraction of slow FUS in the nucleus was decreased by the mutation of RNA binding domains in FUS (generating variants FUS-mutRRM/ZnF and FUS-mutRGG) and was further decreased by the removal of all RNA binding domains (generating variant FUS-PLD) (Fig. 3, E and F; and figs. S15, F to I, S16, and S17). These results indicate that a large fraction of nuclear FUS is complexed with RNA. To further investigate the solubilizing role of RNA, we performed genetic experiments with transfected FUS-GFP–encoding plasmids. We observed that the number of nuclear FUS assemblies was directly proportional to the nuclear FUS concentration (Fig. 3G). We further found that FUS variants with a weaker capacity to bind RNA generally formed a higher number of assemblies (Fig. 3, H to J, and figs. S16 and S17). Thus, reduced RNA binding directly affects the solubility and decreases the saturation concentration at which FUS phase-separates.

We showed previously that FUS in vitro initially forms liquid-like assemblies, but these mature into more solid-like gels and fibrils over time (*4*). These solid-like states are reminiscent of pathological aggregates in ALS (*8*, *9*). Thus, we next tested whether the addition of RNA prevents the formation of fibrils in vitro. The addition of RNA kept the droplets in a soluble state, and fibers were not seen (Fig. 4A). We next investigated whether RNA also changes the material properties of FUS assemblies in vivo. We set up an in vivo aging assay in which we microinjected RNase A into HeLa cells and then monitored the dynamics of the liquid-like drops. After about 30 min, the FUS drops no longer fused (Fig. 4, B and C, and movie S2) but stuck together in large clusters, similar to phenotypes seen previously

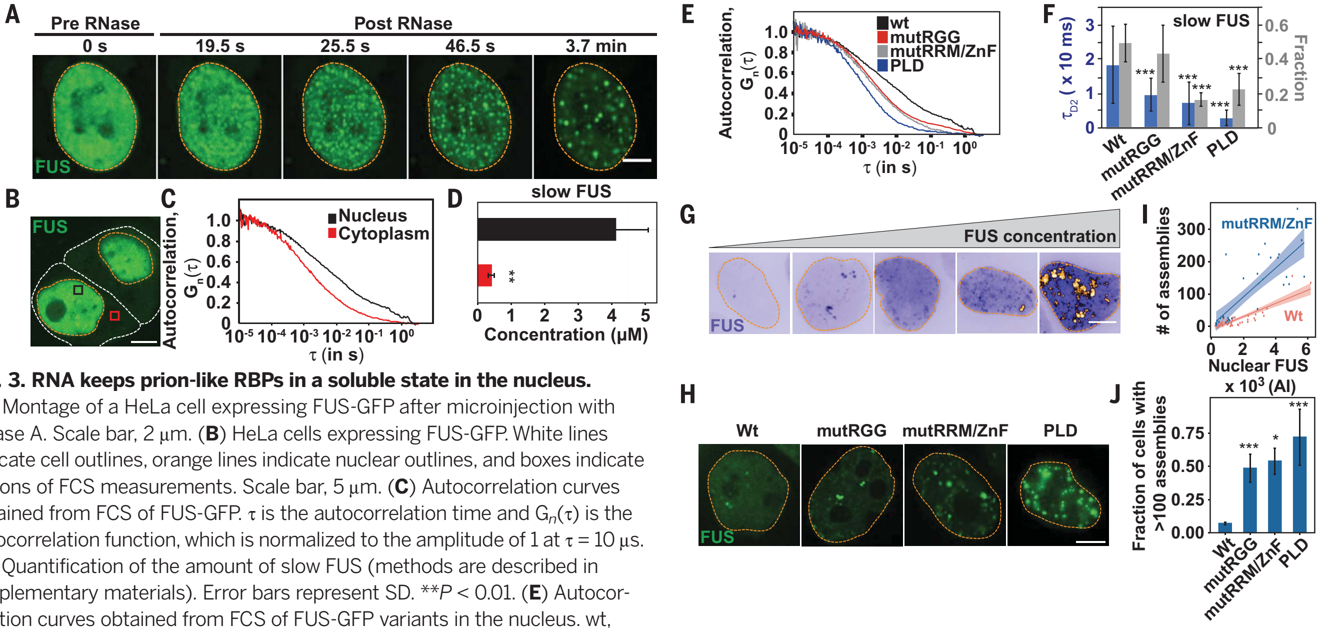


Fig. 3. RNA keeps prion-like RBPs in a soluble state in the nucleus. (**A**) Montage of a HeLa cell expressing FUS-GFP after microinjection with RNase A. Scale bar, 2 μm. (**B**) HeLa cells expressing FUS-GFP. White lines indicate cell outlines, orange lines indicate nuclear outlines, and boxes indicate regions of FCS measurements. Scale bar, 5 μm. (**C**) Autocorrelation curves obtained from FCS of FUS-GFP. τ is the autocorrelation time and $G_n(\tau)$ is the autocorrelation function, which is normalized to the amplitude of 1 at $\tau = 10$ μs. (**D**) Quantification of the amount of slow FUS (methods are described in supplementary materials). Error bars represent SD. $^{**}P < 0.01$. (**E**) Autocorrelation curves obtained from FCS of FUS-GFP variants in the nucleus. wt, wild type; mutRGG, mutations in the first RGG; mutRRM/ZnF, mutations in the RRM and zinc finger (ZnF); PLD, lacks all the RNA binding domains. (**F**) Quantification of slow FUS in the nucleus, obtained from two-component fits of the curves in (E) (methods are described in supplementary materials). Error bars represent SD. τ_{D2}, decay time for slow FUS. (**G**) HeLa cells showing variable FUS-GFP expression. Scale bar, 5 μm. (**H**) HeLa cells expressing different FUS-GFP variants with mutations in RNA binding domains (fig. S16). Scale bar, 5 μm. (**I**) Number of nuclear FUS-GFP assemblies per cell ($n > 30$) as a function of mean protein intensity (AI). Shading represents the confidence interval of the fitted linear regression model, which is plotted as a solid line. (**J**) Number of cells with more than 100 nuclear assemblies. $n > 100$ cells. Error bars represent SD. In (F) and (J), $^{*}P < 0.05$ and $^{***}P < 0.01$ in comparison with the wild type.

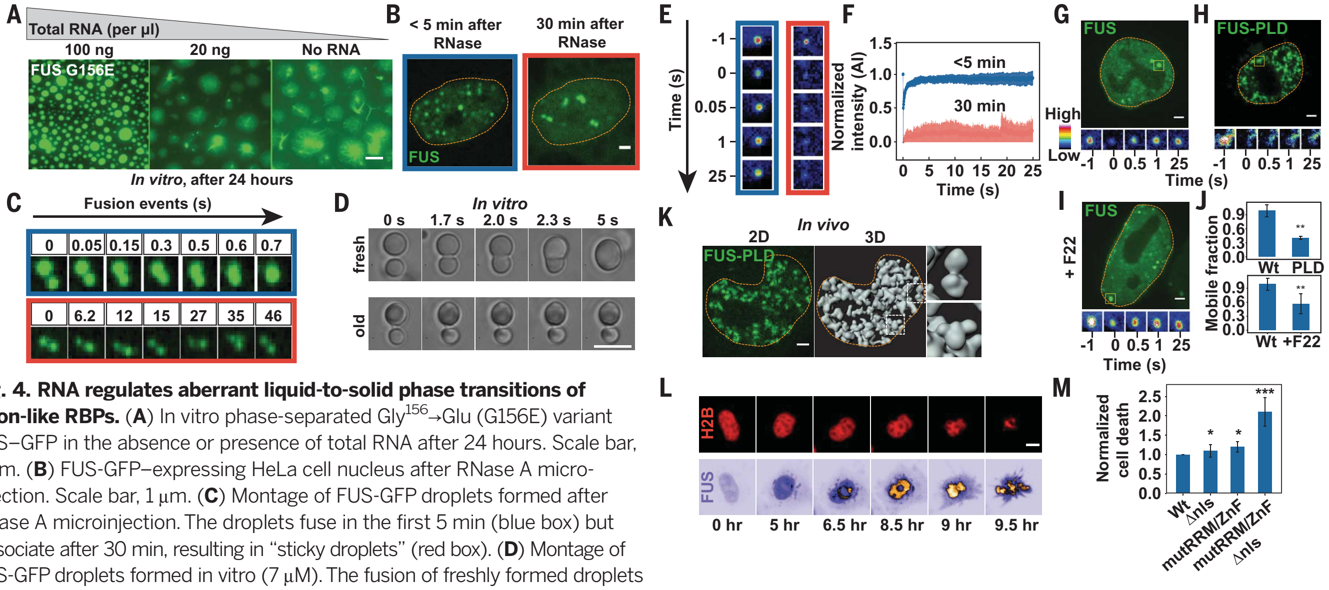


Fig. 4. RNA regulates aberrant liquid-to-solid phase transitions of prion-like RBPs. (**A**) In vitro phase-separated $\text{Gly}^{156}\rightarrow\text{Glu}$ (G156E) variant FUS–GFP in the absence or presence of total RNA after 24 hours. Scale bar, 2 μm. (**B**) FUS-GFP–expressing HeLa cell nucleus after RNase A microinjection. Scale bar, 1 μm. (**C**) Montage of FUS-GFP droplets formed after RNase A microinjection. The droplets fuse in the first 5 min (blue box) but dissociate after 30 min, resulting in "sticky droplets" (red box). (**D**) Montage of FUS-GFP droplets formed in vitro (7 μM). The fusion of freshly formed droplets is compared with 3-hour-old droplets. Scale bar, 5 μm. (**E**) Fluorescence recovery after photobleaching (FRAP) of nuclear FUS-GFP assemblies less than 5 min (blue box) or more than 30 min (red box) after RNase A microinjection. (**F**) FRAP of nuclear FUS-GFP assemblies in HeLa cells after RNA degradation as shown in (B) and (E) ($n > 10$ cells). (**G** to **I**) FRAP of nuclear assemblies in HeLa cells expressing full-length FUS [(G) and (I)] or FUS-PLD (H). The cell in (I) was also treated with F22. Scale bars, 1 μm. (**J**) Mobile fraction of photobleached assemblies in (G) to (I) ($n > 15$ cells). Error bars represent SD. (**K**) Three-dimensional (3D) rendering of FUS-PLD nuclear assemblies. The insets show aberrant "sticky droplets." Scale bar, 1 μm. (**L**) Time series to track the lifetime of FUS-GFP HeLa cells. H2B-mCherry was used to detect cell death. Scale bar, 5 μm. (**M**) Quantification of the fraction of cells undergoing cell death. Error bars represent SD. $^{*}P < 0.05$, $^{**}P < 0.01$, and $^{***}P < 0.001$ in comparison with the wild type.

in vitro (Fig. 4D and movie S3). A change in the material properties was also evident from photobleaching experiments (Fig. 4, E and F, and fig. S18, A and B). Similar results were obtained for TDP43, but the transition was much faster (fig. S19 and movie S4). We next used a genetic approach to test how RNA binding affects the dynamics of FUS in vivo. Complete abrogation of RNA binding resulted in a marked decrease of mobile FUS (Fig. 4, G, H, and J, and fig. S18C) and the formation of sticky droplet clusters (Fig. 4K). Lastly, we used a chemical approach with the dye F22 to reduce RNA binding (*23*). In F22-treated cells, the fraction of RNA-bound FUS was strongly diminished (fig. S20 and supplementary methods), and this caused a strong reduction in the mobile fraction of FUS (Fig. 4, I and J, and fig. S18, E and F). Together, these findings show that RNA keeps condensates formed by prion-like RBPs in a dynamic state and prevents the formation of solid assemblies that can cause disease.

To investigate how reduced RNA binding affects cell viability, we transiently transfected HeLa cells with wild-type and mutant FUS and monitored cell survival. Expressing a nuclear variant with reduced RNA binding (FUS-mutRRM/ZnF) affected the rate of cell death only slightly (Fig. 4, L and M, and fig. S16), presumably because the high nuclear RNA concentration compensated for the genetic defect. However, targeting the very same variant to the cytoplasm by removing the nuclear localization sequence (NLS; generating FUS-mutRRM/ZnFΔNLS) led to a strong increase in cell death, which was likely caused by the high propensity of this variant to form solid aggregates (figs. S21 to S23). Importantly, this increase was not observed for a cytosolic variant of FUS with normal RNA binding (FUSΔNLS). Thus, we conclude that excessive phase separation in the cytoplasm owing to low RNA levels induces a pathological state that leads to cell death.

One of the key questions in protein misfolding diseases caused by prion-like RBPs is why these proteins aggregate in the cytoplasm rather than the nucleus. In this study, we have shown that this pattern is due in part to different RNA concentrations in the cytoplasm and the nucleus. More specifically, the higher RNA concentration in the nucleus suppresses phase separation of prion-like RBPs, and the lower concentration in the cytoplasm stimulates phase separation. Therefore, by keeping the proteins in the nucleus, the cell ensures that they are in a soluble and nontoxic state, shuttling them out of the nucleus only upon stress. After the removal of stress, the proteins shuttle back into the nucleus, where they are again kept in a soluble and well-mixed state. The consequence is that any insult that prolongs the stress will tend to increase the propensity for aggregation because it prolongs the time that these proteins spend in the cytoplasm (fig. S24).

Our data also have important implications for the control of phase separation in cells. We find that paraspeckles are likely induced by locally concentrating Neat1 RNA, which has a strong affinity for FUS. Similar phenomena have been seen for nucleoli, which depend on local production of ribosomal RNA (*24*). Therefore, at least in the nucleus, local production of RNAs with high affinity for specific RBPs may provide the specificity to induce phase separation in a system buffered by nonspecifically interacting RNA. Thus, the phase behavior of FUS in the nucleus is likely controlled by many different types of specific and nonspecific RNAs. This situation does not apply to the cytoplasm. There, the RNA concentration is only slightly higher than the concentration required to suppress phase separation in vitro and there is no buffering of phase separation by RNA. This environment results in a much higher propensity of FUS to phase-separate. However, it also increases the tendency of FUS to form cytotoxic solid-like aggregates. Large amounts of RNA have been shown to suppress the toxicity of prion-like RBPs (*25*–*28*). Moreover, there are many cases of familial ALS in which mutated prion-like RBPs mislocalize to the cytoplasm and form cytotoxic aggregates. For example, mutations in FUS have been shown to increase its cytoplasmic concentration, thus causing the formation of aberrant solid-like aggregates (*8*, *9*, *29*–*31*). We predict that local changes in RNA levels or RNA binding abilities of proteins are frequent causes of age-related protein misfolding diseases.

REFERENCES AND NOTES

1. S. F. Banani, H. O. Lee, A. A. Hyman, M. K. Rosen, *Nat. Rev. Mol. Cell Biol.* **18**, 285–298 (2017).
2. Y. Shin, C. P. Brangwynne, *Science* **357**, eaaf4382 (2017).
3. Z. M. March, O. D. King, J. Shorter, *Brain Res.* **1647**, 9–18 (2016).
4. A. Patel *et al*., *Cell* **162**, 1066–1077 (2015).
5. A. Molliex *et al*., *Cell* **163**, 123–133 (2015).
6. D. Mateju *et al*., *EMBO J.* **36**, 1669–1687 (2017).
7. J. P. Taylor, R. H. Brown Jr., D. W. Cleveland, *Nature* **539**, 197–206 (2016).
8. C. Vance *et al*., *Science* **323**, 1208–1211 (2009).
9. T. J. Kwiatkowski Jr. *et al*., *Science* **323**, 1205–1208 (2009).
10. M. Neumann *et al*., *Science* **314**, 130–133 (2006).
11. M. Neumann *et al*., *Acta Neuropathol.* **117**, 137–149 (2009).
12. H. Deng, K. Gao, J. Jankovic, *Nat. Rev. Neurol.* **10**, 337–348 (2014).
13. D. Dormann *et al*., *EMBO J.* **29**, 2841–2857 (2010).
14. S. J. Barmada *et al*., *J. Neurosci.* **30**, 639–649 (2010).
15. D. C. Diaper *et al*., *Hum. Mol. Genet.* **22**, 1539–1557 (2013).
16. J. Scekic-Zahirovic *et al*., *EMBO J.* **35**, 1077–1097 (2016).
17. A. Sharma *et al*., *Nat. Commun.* **7**, 10465 (2016).
18. Z. Sun *et al*., *PLOS Biol.* **9**, e1000614 (2011).
19. A. H. Fox, S. Nakagawa, T. Hirose, C. S. Bond, *Trends Biochem. Sci.* **43**, 124–135 (2017).
20. S. Saha *et al*., *Cell* **166**, 1572–1584.e16 (2016).
21. S. Elbaum-Garfinkle *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **112**, 7189–7194 (2015).
22. H. Zhang *et al*., *Mol. Cell* **60**, 220–230 (2015).
23. Q. Li *et al*., *Chem. Biol.* **13**, 615–623 (2006).
24. J. Berry, S. C. Weber, N. Vaidya, M. Haataja, C. P. Brangwynne, *Proc. Natl. Acad. Sci. U.S.A.* **112**, E5237–E5245 (2015).
25. A. Kitamura *et al*., *Sci. Rep.* **6**, 19230 (2016).
26. T. Ishiguro *et al*., *Neuron* **94**, 108–124.e7 (2017).
27. Y.-C. Huang *et al*., *PLOS ONE* **8**, e64002 (2013).
28. Y. M. Ayala *et al*., *J. Cell Sci.* **121**, 3778–3785 (2008).
29. M. Sabatelli *et al*., *Hum. Mol. Genet.* **22**, 4748–4755 (2013).
30. S. Dini Modigliani, M. Morlando, L. Errichelli, M. Sabatelli, I. Bozzoni, *Nat. Commun.* **5**, 4335 (2014).
31. Y. Shang, E. J. Huang, *Brain Res.* **1647**, 65–78 (2016).

ACKNOWLEDGMENTS

We thank the members of MPI-CBG for discussions; B. Borgonovo (chromatography facility); R. Wegner (protein expression and purification facility); A. Bogdanova for providing vectors; M. Leuschner and A. Ssykor for preparing BAC lines; J. Peychl, B. Nitzsche, and B. Schroth-Diez (light microscopy facility); C. Andree and C. Möbius (Technology Development Studio); J. Jarrells (DNA microarray facility); the FACS facility; and D. Dormann and E. Bogaert for providing reagents. **Funding:** We acknowledge funding from the Max Planck Society, the ERC (nos. 725836 and 643417), the BMBF (01ED1601A and 031A359A), and the JPND (CureALS). S.M. was supported by a fellowship of the Humboldt Foundation (3.5-INI/1155756 STP), L.M. by the Hans und Ilse Breuer Stiftung, and J.G.-B. by an EMBO fellowship (ALTF 406-2017). **Author contributions**: S.M. and S.A. designed and coordinated the project. J.W. performed in vitro experiments with RBPs, D.K.P. performed FCS experiments, J.G.-B. performed the RNA immunoprecipitation, and T.M.F. performed the RNA binding assay. D.R., A.P., and S.R. generated plasmids. I.P. generated HeLa lines. L.M. and J.S. provided induced pluripotent stem cell–derived cell lines. M.B. analyzed the cell viability assay. M.J. performed optical tweezer experiments. Y.-T.C. provided the dye F22. S.M., A.A.H., and S.A. drafted the manuscript with input from P.T. All authors contributed to data analysis and interpretation. **Competing interests:** The dye F22 was covered by U.S. patent US 7790896 B2 awarded to Y.-T.C. The other authors declare no competing interests. **Data and materials availability:** All data are available in the main text or the supplementary materials.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/360/6391/918/suppl/DC1
Materials and Methods
Figs. S1 to S24
References (*32*–*52*)
Movies S1 to S4

12 December 2017; accepted 4 April 2018
Published online 12 April 2018
10.1126/science.aar7366

MOLECULAR BIOLOGY

mRNA structure determines specificity of a polyQ-driven phase separation

Erin M. Langdon,[1] Yupeng Qiu,[2] Amirhossein Ghanbari Niaki,[2] Grace A. McLaughlin,[1] Chase A. Weidmann,[3] Therese M. Gerbich,[1] Jean A. Smith,[1] John M. Crutchley,[1] Christina M. Termini,[4] Kevin M. Weeks,[3] Sua Myong,[2] Amy S. Gladfelter[1,5]*

RNA promotes liquid-liquid phase separation (LLPS) to build membraneless compartments in cells. How distinct molecular compositions are established and maintained in these liquid compartments is unknown. Here, we report that secondary structure allows messenger RNAs (mRNAs) to self-associate and determines whether an mRNA is recruited to or excluded from liquid compartments. The polyQ-protein Whi3 induces conformational changes in RNA structure and generates distinct molecular fluctuations depending on the RNA sequence. These data support a model in which structure-based, RNA-RNA interactions promote assembly of distinct droplets and protein-driven, conformational dynamics of the RNA maintain this identity. Thus, the shape of RNA can promote the formation and coexistence of the diverse array of RNA-rich liquid compartments found in a single cell.

Formation of non–membrane-bound organelles through the condensation of macromolecules is a recently appreciated mechanism of intracellular organization. These liquid-like condensates form through liquid-liquid phase separation (LLPS) and are found in the cytoplasm and nucleus (*1*, *2*). A fundamental unsolved problem is how liquid droplets recruit distinct constituents and retain independent identities. RNA can drive LLPS and modulates the material properties of droplets (*3*–*6*), but it is unknown whether RNA controls the identity and maintenance of coexisting liquid compartments. Here, we show that mRNA secondary structure is required for droplet identity through directing interactions between mRNAs and RNA binding proteins.

Whi3, a polyQ-containing, RNA binding protein first identified in *Saccharomyces cerevisiae* (*7*), functions in morphogenesis, memory of mating, and stress responses, where it forms aggregates and associates with RNA-processing bodies (*8*–*11*). The homolog in the filamentous fungus *Ashbya gossypii* has an RNA recognition motif (RRM) and an expanded polyQ tract (fig. S1A), and both regions promote self-assembly. In vitro, Whi3 polyQ-dependent LLPS is driven by specific RNAs encoding regulators of either the cell cycle (e.g., the cyclin *CLN3*) or actin (e.g., the formin *BNI1* and *SPA2*) (*3*). Distinct types of Whi3 droplets form in *Ashbya* cells: perinuclear *CLN3* droplets and *BNI1* droplets at sites of polarized growth at cell tips (*12*, *13*) (Fig. 1A and movie S1). These two types of droplets have different Whi3 levels and Whi3 incorporation rates (Fig. 1, B and C), suggesting that they are structurally distinct.

The distinct droplet properties may depend on extrinsic features of the local cytosolic microenvironment or arise due to different droplet constituents. *CLN3* and *BNI1* mRNAs minimally colocalized in the cytoplasm by single-molecule (sm) RNA fluorescence in situ hybridization (smFISH), although they were occasionally coexpressed by the same nucleus (Fig. 1, D and F). The lack of colocalization suggests there are intrinsic, compositional differences between droplets. In contrast, mRNA of the polarity regulator *SPA2*, significantly colocalized with *BNI1* mRNAs, especially at growth sites (Fig. 1, E and F). Thus, mRNAs encoding functionally related proteins colocalize, whereas functionally unrelated mRNAs do not. How can distinct Whi3 binding mRNAs segregate to different droplets in a common cytoplasm?

To address this question, we employed a reconstitution system to investigate whether mRNA sequence was sufficient to generate droplet individuality (Fig. 2A). In vitro, as in cells, droplets composed of *BNI1* mRNA displayed higher Whi3 to RNA molar ratios than droplets made with *CLN3* mRNA (fig. S1B). Notably, when *CLN3* mRNA was added to Whi3 droplets made with *BNI1* mRNA, *CLN3* preferentially assembled into new droplets, rather than incorporating into *BNI1* droplets (Fig. 2, B and C, and fig. S1C). In contrast, *BNI1* mRNA readily incorporated into preformed droplets (Fig. 2, B and C). Notably, *SPA2* mRNA incorporated into *BNI1* droplets (Fig. 2, B and C), and *CLN3* did not incorporate into *SPA2* droplets (fig. S1D). Thus, as in cells, cyclin and polarity mRNAs assemble into distinct and immiscible droplets in vitro, indicating that droplet identity is encoded by the mRNA.

mRNA sequences could influence droplet identity by favoring homotypic or specific heterotypic interactions between RNA molecules. To test for specific RNA-RNA interactions, we used a protein-free system to induce electrostatic-mediated phase transitions of the mRNA (*14*), where all mRNAs were capable of homotypic assembly into liquid or gel-like droplets (Fig. 2D). Strikingly, *CLN3* mRNAs had minimal colocalization with *BNI1* or *SPA2* mRNAs, whereas *BNI1* and *SPA2* were significantly more colocalized (Fig. 2, E to G). Thus, sequence-encoded features of the mRNA can underpin the assembly of distinct, immiscible structures.

We next investigated which features of the mRNA sequence generate specificity. An mRNA with scrambled *CLN3* coding sequence (*cln3 scr*) with intact Whi3 binding sites formed Whi3 droplets (fig. S1E) but no longer showed specificity (Fig. 3, A and C). Because the length, nucleotide composition, and Whi3 binding sites were identical, we hypothesized that the secondary structure could promote specificity. *CLN3* mRNA heated to 95°C to disrupt secondary structure also readily incorporated into Whi3-*BNI1* droplets (Fig. 3, A and C). Melted *CLN3* mRNA that was slowly refolded (*CLN3* refold) showed significantly less recruitment than melted but more than native *CLN3* (Fig. 3, A and C). Mixing between melted *CLN3* and melted *BNI1* occurred in the presence of Whi3 and in RNA-only reactions, suggesting that mixing is initiated by RNA-RNA interactions (fig. S2). Thus, specificity information in *CLN3* mRNA can be eliminated by disrupting secondary structure.

To identify what features of *CLN3* mRNA secondary structure promote specificity, we performed selective 2′-hydroxyl acylation analyzed by primer extension and mutational profiling (SHAPE-MaP), which identifies highly flexible regions in RNA (*15*), to determine secondary structure changes on native, refolded, and scrambled *CLN3* mRNA (Fig. 3D and fig. S3, A and B). The first 400 nucleotides in the *CLN3* sequence exhibited especially low SHAPE reactivity (fig. S3C, purple shaded regions), suggesting many paired nucleotides and a highly folded structure. Refolded *CLN3* had a significant increase in SHAPE reactivity compared with native *CLN3* (fig. S3A) ($P < 0.001$, Wilcoxon rank sum test), indicating a transition to a more unstructured state (Fig. 3, D and E). Melting and refolding thus allows the RNA to sample different conformations from those formed during transcription. As expected, *cln3 scr* showed a different SHAPE profile with dramatically altered secondary structure (Fig. 3, D and E, and fig. S3B).

We hypothesized that secondary structure influences mRNA sorting, because stem loops may selectively display or mask sequences capable of hybridizing with other RNAs. *CLN3* contains five complementary regions to *BNI1* (fig. S4A),

---

[1]Department of Biology, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA. [2]Department of Biophysics, Johns Hopkins University, Baltimore, MD 21218, USA. [3]Department of Chemistry, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA. [4]Division of Hematology/Oncology, Department of Medicine, University of California, Los Angeles, Los Angeles, CA 90095, USA. [5]Marine Biological Laboratory, Woods Hole, MA 02543, USA.
***Corresponding author. Email: amyglad@unc.edu**

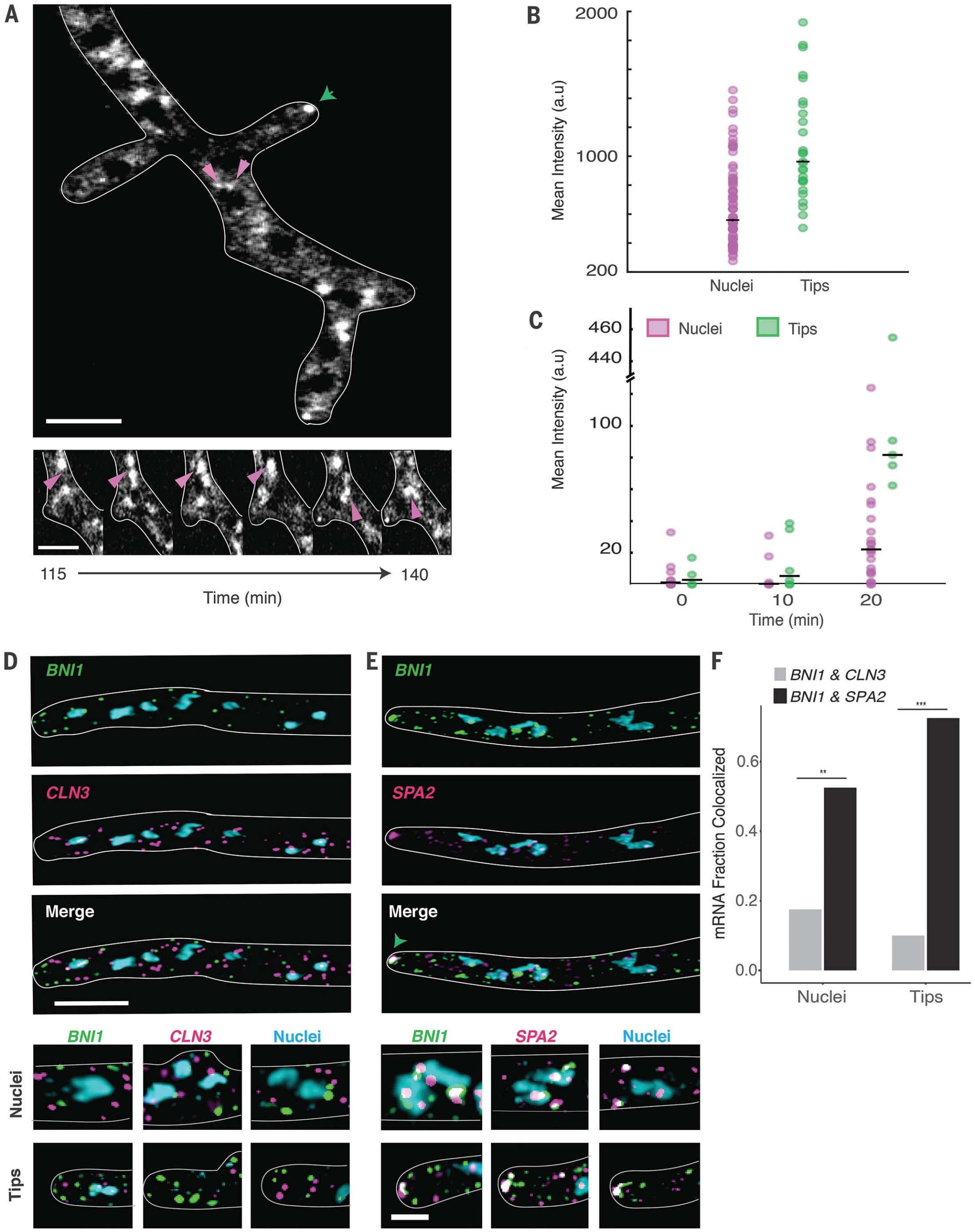


Fig. 1. Cyclin and polarity complexes are spatially and physically distinct within the cell. (**A**) (Top) Whi3 forms liquid droplets in *Ashbya gossypii*. (Bottom) Whi3 droplets accumulate and fuse around nuclei. Green arrow denotes polarity droplets. Pink arrows denote perinuclear droplets. Scale bars, 5 μm. (**B**) Mean intensity of Whi3-tomato is higher in tip droplets (green) than perinuclear droplets (pink). (**C**) Rate of Whi3 incorporation is higher in tip compared to perinuclear droplets. (**D**) smFISH images show that *BNI1* (green) and *CLN3* (pink) mRNAs are spatially distinct. Nuclei are in blue. Scale bar, 5 μm. (**E**) smFISH images show that *BNI1* (green) mRNAs colocalize with polarity mRNA *SPA2* (pink). Nuclei are in blue. The green arrow marks where the RNAs overlap at the tip. Inset scale bar, 2 μm. (**F**) *BNI1* and *SPA2* are significantly more colocalized than *BNI1* and *CLN3*. $^{***}P < 0.001$ for tips and $^{**}P < 0.01$ for nuclei (Fisher's exact test). $n = 40$ nuclei and tips for ≥30 cells.

most of which had low SHAPE reactivity and therefore were more structured (fig. S4B), suggesting that these regions are inaccessible for hybridizing with *BNI1*. We hypothesize that these regions became available to pair with *BNI1* when *CLN3* is melted, causing the structure-dependent loss of droplet specificity. To test this hypothesis, oligonucleotides (i.e., oligos) complementary to these regions were added to melted *CLN3* and significantly decreased the coassembly with *BNI1*, restoring the formation of distinct *CLN3* droplets (Fig. 3, B and C). Additionally, *cln3sm*, a mutant perturbing structure and exposing complementarity, colocalized with *BNI1* transcripts in vitro and at polarity sites in cells (Fig. 3F and fig. S5) (>60% tips colocalized). Thus, secondary structure can regulate RNA sorting into distinct droplets through altering the capacity to form intermolecular interactions.

We next asked whether exposed complementarity explains coassembly of *BNI1* and *SPA2* into the same droplets. Indeed, SHAPE-MaP analysis of *BNI1* and *SPA2* showed complementary regions between these colocalizing mRNAs having significantly higher SHAPE reactivity and less structure compared with the *CLN3*/*BNI1* regions (figs. S4 and S6) ($P < 0.002$, t test). Addition of complementary oligos to these regions disrupted colocalization in the presence of Whi3 and in RNA-only reactions (fig. S7, A and B). We predicted that *CLN3* may self-assemble and indeed *cln3 codon*, a *CLN3* mutant whose codons have been randomized but Whi3 binding sites remain intact, does not colocalize with endogenous *CLN3* mRNA in cells, further supporting RNA-RNA interactions in coassembly of related RNAs (fig. S7C). These data suggest that RNA-RNA interactions based on intermolecular hybridization direct RNAs into the same or different droplets.

Does Whi3 protein influence the identity of droplets? The majority of Whi3 binding sites are exposed on stem loops in *CLN3*, *BNI1*, and *SPA2* (Fig. 3E and figs. S8 and S9). Notably, refolding or scrambling the *CLN3* sequence rearranges the presentation of Whi3 binding sites (Fig. 3E). Therefore, RNA secondary structure may influence Whi3 binding and contribute to droplet composition and immiscibility in addition to RNA complexing. SHAPE-MaP analysis of *CLN3* mRNA in the presence of Whi3 supports that Whi3 binding sites are occupied by protein (Fig. 4A

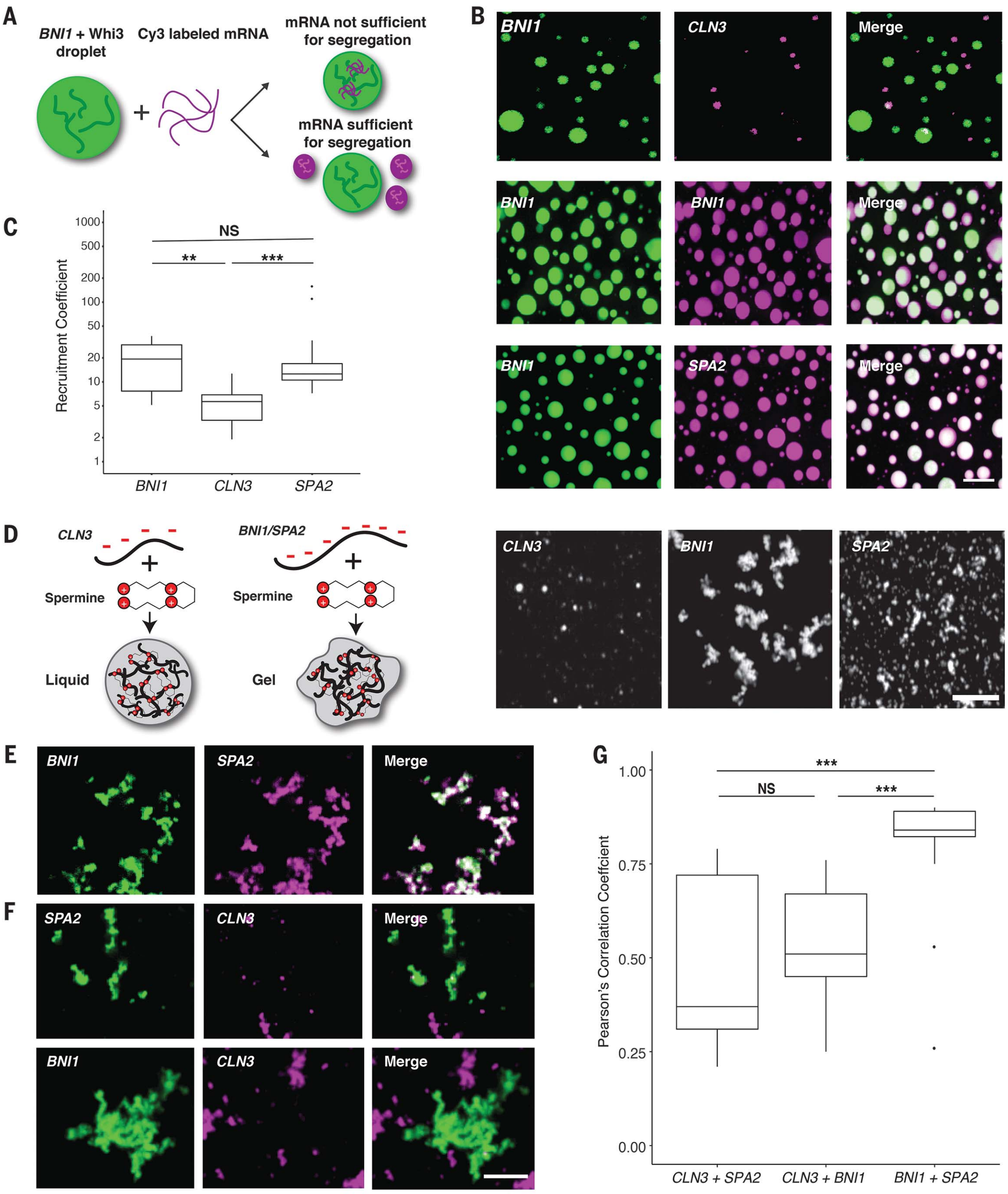


Fig. 2. Polarity and cyclin complexes segregate in vitro. (**A**) Experimental schematic of in vitro droplet recruitment assay. (**B**) *CLN3* mRNA (pink) is not efficiently recruited, but *BNI1* or *SPA2* mRNA (pink) are recruited into preformed Whi3-*BNI1* droplets (green) based on fluorescence microscopy. Scale bar, 10 μm. (**C**) Recruitment coefficients of mRNA from (B). Boxes indicate interquartile range, line is median, whiskers contain points within three times the interquartile range, and outliers are indicated with dots. NS, not significant, $P > 0.05$; $**P < 0.01$; $***P < 0.001$ (*t* test). $n \geq 500$ droplets for $N \geq 3$ biological replicates. (**D**) Cartoon schematic and representative images showing in vitro RNA-only droplet assay where *CLN3*, *BNI1*, and *SPA2* mRNAs assemble into liquid or gel-like droplets. Scale bar, 5 μm. (**E**) Fluorescence microscopy images showing that *BNI1* RNA (green) colocalizes with SPA2 RNA (pink) in droplets. (**F**) Fluorescence microscopy images showing that *CLN3* RNA (pink) does not colocalize with *SPA2* (green) and *BNI1* (green) droplets. Scale bar, 5 μm. (**G**) Quantification of colocalization between *BNI1* and *SPA2*, *SPA2* and *CLN3*, or *BNI1* and *CLN3* RNAs. NS, not significant, $***P < 0.001$ (Wilcoxon rank sum test). $n \geq 200$ droplets for $N \geq 3$ biological replicates.

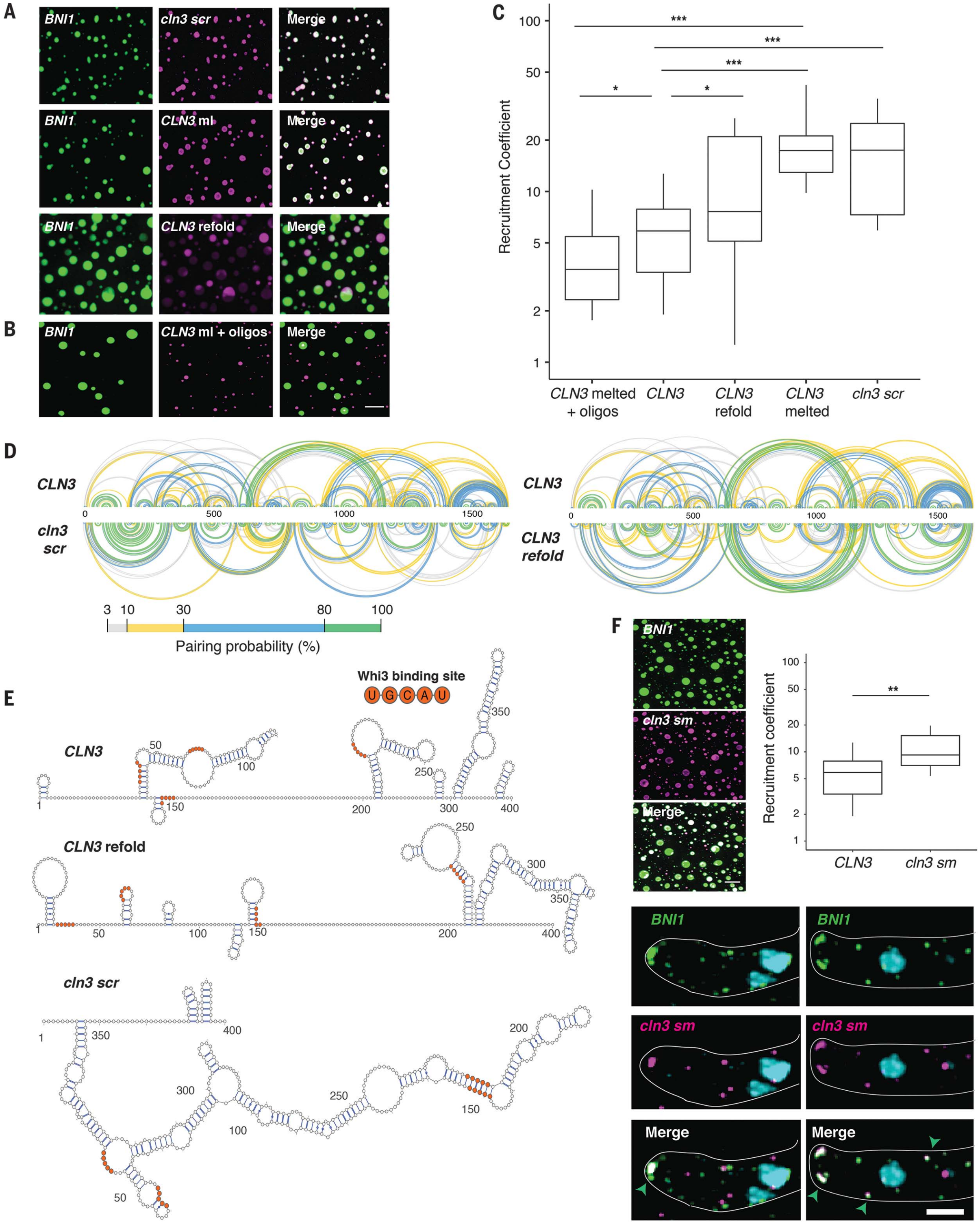


Fig. 3. RNA secondary structure determines specificity and identity of Whi3-*CLN3* droplets. (**A**) Fluorescence microscopy images showing the recruitment of scrambled (*cln3 scr*), melted (*CLN3* ml), and refolded *CLN3* (*CLN3* refold) mRNA (pink) into preformed Whi3-*BNI1* droplets (green). (**B**) Fluorescence microscopy images showing the loss of recruitment of *CLN3* ml when mixed with oligonucleotides targeting complementary sequences of *CLN3* to *BNI1*. Scale bar, 10 μm. (**C**) Quantification of (A) and (B). $^{*}P < 0.05$; $^{**}P < 0.01$; $^{***}P < 0.001$ (*t* test). $n \geq 500$ droplets for $N \geq 3$ biological replicates. (**D**) Base pairing probability from SHAPE-MaP of *CLN3*, *cln3 scr*, and *CLN3* refold show differences in the secondary structure in *CLN3*. Arcs connect base pairs and are colored by probability. (**E**) Secondary structure models from SHAPE-MaP for the first 400 nucleotides of *CLN3*, *CLN3* refold, and *cln3 scr*. Whi3 binding sites are in orange. (**F**) *CLN3* structure mutant (*cln3 sm*) mRNA is significantly recruited to Whi3-*BNI1* droplets in vitro and in vivo. $^{**}P < 0.01$ (*t* test). Green arrows denote sites of colocalization between *BNI1* mRNA (green) and *cln3 sm* mRNA (pink) by smFISH. Scale bar, 10 μm for in vitro, 2 μm for in vivo. $n \geq 500$ droplets for $N \geq 3$ biological replicates.

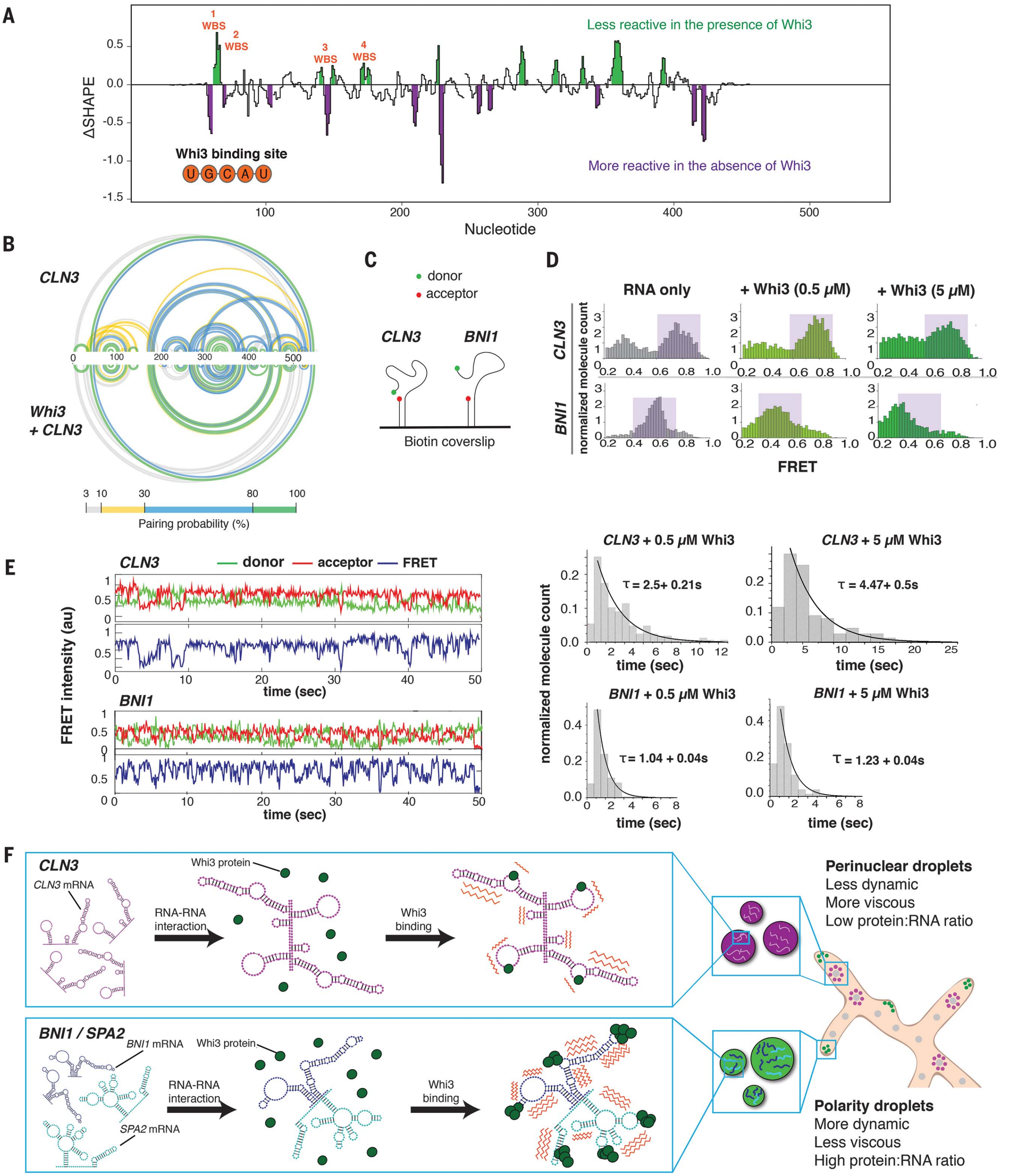


Fig. 4. Whi3 binding alters RNA behavior. (**A**) Differences in SHAPE reactivities (ΔSHAPE) were calculated by subtracting *CLN3* SHAPE reactivities from *CLN3* + Whi3 reactivities. Positive ΔSHAPE values indicate protection from modification in the presence of Whi3, and negative ΔSHAPE reports enhanced reactivity in the absence of Whi3 protein. (**B**) Base-pairing probability compared between *CLN3* and *CLN3* with Whi3 shows rearrangements in *CLN3* structure in the presence of Whi3. Arcs connect base-pairing sites and are colored by probability. (**C**) Schematic of smFRET experiment. (**D**) FRET histograms before (gray) and after (green) 0.5 or 5 μM Whi3 addition. Purple shaded regions denote high and mid FRET states for *CLN3* and *BNI1*, respectively. (**E**) Averaged cy3 (green), cy5 (red) intensities, and representative FRET traces (blue) obtained from smFRET experiments of *CLN3* and *BNI1* in the presence of 5 μM Whi3. Dwell-time analysis reveals slower FRET fluctuations for *CLN3* than *BNI1* in the presence of Whi3. (**F**) Proposed model in which RNA-RNA interactions derived from mRNA structure promote the selective uptake of distinct RNAs and protein constituents into droplets leading to distinct dynamics (orange zigzags) of different droplet complexes.

and fig. S10A) and revealed that protein binding causes structural rearrangements (Fig. 4B). We therefore hypothesize that Whi3 binding may have important contributions to structural rearrangements of target RNAs relevant to droplet identity.

To examine the consequence of Whi3 binding to RNA, we used smFRET (fluorescence resonance energy transfer) (Fig. 4C) to measure the conformational dynamics of *CLN3* and *BNI1* mRNAs with and without Whi3 (*16*). In the absence of protein, *CLN3* RNA showed high FRET values indicative of a compacted state, whereas *BNI1* RNA showed lower FRET values, indicating a less compact state (Fig. 4D, purple shaded regions). Upon addition of Whi3, *CLN3* FRET values decreased, indicating that more extended RNA conformations were induced, dependent on the ability of Whi3 to bind mRNA (fig. S10, B and C). In contrast, bound to Whi3, *BNI1* RNA showed a more substantial broadening of FRET values (Fig. 4D), indicating that Whi3-*BNI1* complexes are more dynamic. Dwell-time analysis revealed that Whi3-induced dynamics are three times faster for *BNI1* than *CLN3* (Fig. 4E). Different mRNAs thus react differentially in their intramolecular fluctuations to the presence of Whi3, providing an additional mode of RNA droplet regulation.

These FRET studies suggest that Whi3 binding alters the conformational dynamics of target RNAs. We speculate that these differential dynamics help maintain droplet identities established by RNA-RNA interactions. Once RNA-RNA interactions are formed, Whi3 binding may reduce the ability of the RNA to resample many alternate RNA structures to maintain the identity. Additionally, the slower fluctuations of *CLN3* bound to Whi3 may be one source of exclusion from the more rapidly fluctuating *BNI1*-Whi3 complexes in those droplets. Such dynamics may drive the droplet material properties reported previously (*3*) and serve as barriers to homogenization.

We show that mRNA structure defines the ability of an RNA to engage in homo- or heteromeric interactions and thus drives specificity in the composition of liquid droplet compartments (Fig. 4F). This mechanism is likely relevant for the sorting of specific RNAs to other RNA granules, such as stress and P granules, and P bodies (*17*, *18*). Future work will address the timing and location of how mRNA secondary structure influences selective uptake of cellular constituents into droplets. Protein binding to different RNAs can lead to varied dynamics of complexes, further distinguishing the physical properties of different droplets and promoting immiscibility of coexisting droplets. Given the large number of distinct, RNA-based condensates in the cell, these mechanisms are likely broadly relevant to explain how droplets achieve and maintain individuality.

REFERENCES AND NOTES

1. S. F. Banani, H. O. Lee, A. A. Hyman, M. K. Rosen, *Nat. Rev. Mol. Cell Biol.* **18**, 285–298 (2017).
2. Y. Shin, C. P. Brangwynne, *Science* **357**, eaaf4382 (2017).
3. H. Zhang *et al.*, *Mol. Cell* **60**, 220–230 (2015).
4. X. Zhang *et al.*, *PLOS Biol.* **15**, e2002183 (2017).
5. S. Elbaum-Garfinkle *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **112**, 7189–7194 (2015).
6. Y. Lin, D. S. W. Protter, M. K. Rosen, R. Parker, *Mol. Cell* **60**, 208–219 (2015).
7. R. S. Nash, T. Volpe, B. Futcher, *Genetics* **157**, 1469–1480 (2001).
8. N. Colomina, F. Ferrezuelo, E. Vergés, M. Aldea, E. Garí, *Cell Cycle* **8**, 1912–1920 (2009).
9. F. Caudron, Y. Barral, *Cell* **155**, 1244–1257 (2013).
10. G. Schlissel, M. K. Krzyzanowski, F. Caudron, Y. Barral, J. Rine, *Science* **355**, 1184–1187 (2017).
11. K. J. Holmes, D. M. Klass, E. L. Guiney, M. S. Cyert, *PLOS ONE* **8**, e84060 (2013).
12. C. Lee *et al.*, *Dev. Cell* **25**, 572–584 (2013).
13. C. Lee, P. Occhipinti, A. S. Gladfelter, *J. Cell Biol.* **208**, 533–544 (2015).
14. A. Jain, R. D. Vale, *Nature* **546**, 243–247 (2017).
15. M. J. Smola, G. M. Rice, S. Busan, N. A. Siegfried, K. M. Weeks, *Nat. Protoc.* **10**, 1643–1669 (2015).
16. Y. Kim, S. Myong, *Mol. Cell* **63**, 865–876 (2016).
17. T. Trcek *et al.*, *Nat. Commun.* **6**, 7962 (2015).
18. B. Van Treeck *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **115**, 2734–2739 (2018).

ACKNOWLEDGMENTS

We thank the Gladfelter, Weeks, and Laederach laboratories for critical discussions; E. Griffin, J. Moseley, D. Lew, M. Peifer, and H. Higgs for critically reading the manuscript; the HHMI HCIA at the Marine Biological Laboratory for intellectual community; and T. Straub for useful data analysis discussions. **Funding:** This work was supported by NIH GM R01-GM081506, the HHMI Faculty Scholars program, R35 GM122532, ACS 130845-RSG-17-114-01-RMC, NIH 1DP2 GM105453, and NIH R01 GM115631. **Author contributions:** E.M.L. and A.S.G. designed and performed experiments, analyzed data, prepared figures, and drafted the manuscript; Y.Q., A.G.N., and C.A.W. designed and performed experiments, analyzed data, and edited the manuscript; G.A.M. and C.M.T. performed experiments and analyzed data; T.M.G., J.A.S., and J.M.C. provided technical support and edited the manuscript; K.M.W. and S.M. designed experiments and edited the manuscript. **Competing interests:** K.M.W. is an advisor to and holds equity in Ribometrix, to which mutational profiling (MaP) technologies have been licensed. All other authors declare that they have no competing interests. **Data and materials availability:** All data are available upon request from E.M.L. or A.S.G.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/360/6391/922/suppl/DC1
Materials and Methods
Figs. S1 to S10
References (*19*–*22*)
Movie S1

12 December 2017; accepted 4 April 2018
Published online 12 April 2018
10.1126/science.aar7432


Exceptional
scientists wanted

2-Minute Next-Gen PCR

Introducing 2-minute PCR—NEXTGENPCR is redefining “fast.” The NEXTGENPCR device heats and cools samples instantly, losing no time getting samples to the desired temperature. Go from melting to annealing in less than one-tenth of a second. Complete PCR in as little as 2 minutes. Using one thermocycler, you can run more than 10 plates per hour. The math adds up quickly—10 plates per hour for 8 hours in a 384-well format means more than 30,000 data points in a single day. Process more samples per day for sequencing, genotyping, and other PCR applications by incorporating NEXTGENPCR into your laboratory.

Canon BioMedical
For info: 844-226-6624
www.canon-biomedical.com/gofaster

Antibody Labeling Kits

Mix-n-Stain antibody labeling kits are a revolutionary technology that dramatically simplifies the process of antibody labeling. The kits enable a researcher to covalently label ≤5 µg–100 µg of an antibody with one of 20 super bright and photostable CF dyes, biotin, or fluorescein isothiocyanate (FITC) in only 30 minutes, with 100% recovery of antibody after labeling. The prepared conjugates perform comparably to prelabeled antibodies from leading suppliers. Because the labeling is covalent, the conjugates are stable for long-term storage, and ideal for multicolor imaging. Each kit comes with everything you need to perform the conjugation. Simply mix your antibody with a vial of premeasured CF dye, biotin, or FITC in the buffer provided—a step taking less than 30 s. After waiting for 30 min, without a separation step, you will have an optimally labeled antibody conjugate ready for use in any immunofluorescence staining experiment.

Biotium
For info: 800-304-5357
www.biotium.com

Trypsin For Mass Spectrometry

MilliporeSigma’s Advanced Proteomics Grade SOLu-Trypsin is an exclusive, solution-stable enzyme for mass spectrometry. Designed to be stable in solution when refrigerated, it can be used immediately without preparation. Other forms of trypsin require thawing or reconstitution and must be discarded if not used immediately. SOLu-Trypsin allows excess product to be saved for future use, thus eliminating unnecessary waste and cost. It is formulated with a high-purity recombinant trypsin, free of chymotryptic activity, to ensure high-fidelity digestion, and is stable for short-term use at room temperature in an autosampler or on a liquid-handling robot.

MilliporeSigma
For info: 800-325-3010
www.sigma.com/solu-trypsin

Hematopoietic Progenitor Cell Culture Medium

Irvine Scientific’s PRIME-XV Hematopoietic Cell Basal Medium is a serum-free, xeno-free basal medium that specifically supports robust expansion of hematopoietic progenitors while maintaining their multipotency—the functional properties they need for cell-based therapies. Hematopoietic stem and progenitor cells (HSPCs) give rise to all other blood cells, and therefore have therapeutic potential for the treatment of cancers such as leukemia and lymphoma, and nonmalignant diseases such as sickle-cell anemia or immune-related disorders. With numbers of cells in the hundreds of millions required for therapy, the challenge has been obtaining sufficient numbers of HSPCs that maintain the progenitor cell properties required to differentiate and mature into the various hematopoietic lineage cells needed for therapies. PRIME-XV is manufactured under Current Good Manufacturing Practices using stringent raw-material qualification, and offers consistency and reliability while supporting the move from research toward clinical applications.

Irvine Scientific
For info: 800-577-6097
www.irvinesci.com

Bioreactor System

The BioFlo 120 is a bench-scale bioreactor/fermentor system perfectly suited for all levels of R&D. Flexibly designed to meet the wide-ranging needs of today’s scientists, the system features 24 interchangeable heat-blanketed and water-jacketed autoclavable vessels, along with BioBLU Single-Use Vessels ranging from 250 mL to 40 L working volume. Digital Mettler Toledo ISM (Intelligent Sensor Management) sensor technology is fully integrated, allowing the user to choose between pH, redox, dissolved oxygen (polarographic and/or optical), and CO_2 measurements without the need for additional equipment. A high-precision, expanded-range thermal mass flow controller is also available, which can provide low-flow mammalian cell culture through high-demand microbial processes on a single controller. Its ease of use and flexibility ensure that the BioFlo 120 will serve scientists for years to come—be it in academic, governmental, or industrial research settings.

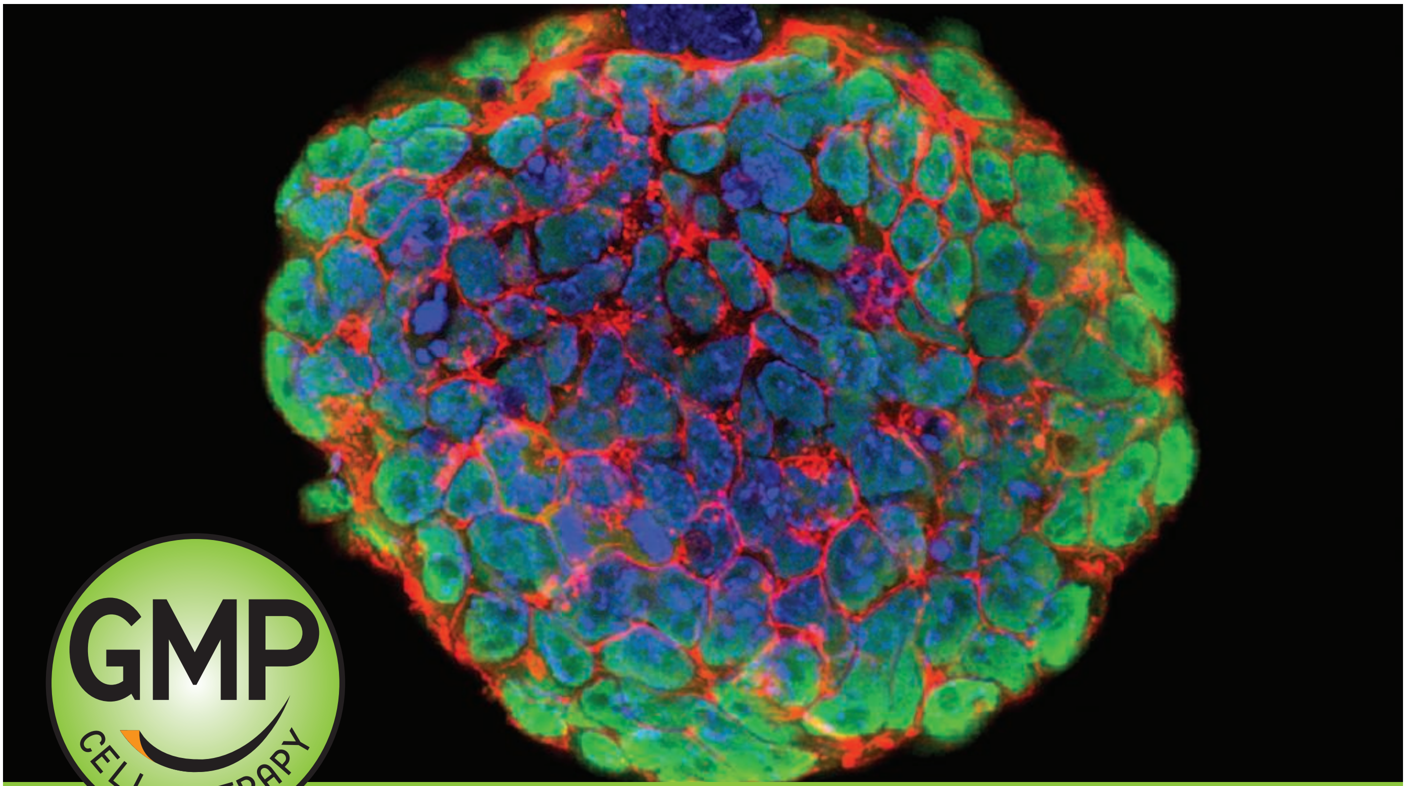
Eppendorf
For info: 800-645-3050
www.eppendorf.com/bioflo120

GC/Q-TOF System

The Agilent 7250 GC/Q-TOF system combines gas chromatography (GC) and quadrupole time-of-flight mass spectrometry (Q-TOF) with a low-energy electron ionization (EI) source that allows greater exploration of unknown chemical samples. The system is designed for laboratories doing food and environmental testing, life science research, forensics, and chemical analysis. It represents the only commercially available combination of a high-resolution, accurate-mass GC/MS and low-energy ionization source, enabling scientists to employ analytical techniques that were previously impractical, or in some cases impossible. The 7250 GC/Q-TOF system will now permit labs to quickly and easily identify volatile and semivolatile compounds that once required them to perform alternative techniques or longer sample preparations to detect and determine the compounds present in their samples. Furthermore, with low-energy EI, scientists will be better able to elucidate chemical structures.

Agilent Technologies
For info: 800-227-9770
www.agilent.com

Electronically submit your new product description or product literature information! Go to www.sciencemag.org/about/new-products-section for more information.



EPFL
ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

Faculty Position in Computational Environmental Sciences and Engineering

at the Ecole polytechnique fédérale de Lausanne (EPFL)

The EPFL School of Architecture, Civil and Environmental Engineering (ENAC) invites applications for a faculty position in its Institute of Environmental Engineering (IIE), either at the tenure track level (Assistant Professor) or the tenured level (Associate or Full Professor). The appointee will join the newly formed EPFL Centre for Changing Alpine and Polar Environments (CAPE), based in Sion, Switzerland, and contribute to research and teaching activities within IIE. This appointment is one of several CAPE professorships, and offers unrivalled collaboration opportunities at the local and European levels.

Analogously to the early days of computational biology, computational environmental sciences and engineering are now on the rise. The role of computation in understanding changing environments is pervasive, being central to, for example, process-based models, design of sensor networks, and big data analysis and visualization. The research vision of CAPE includes developing understanding of connections between models across scales, from global earth system simulation to small-scale ecosystem functioning. The appointee will investigate multiscale analysis of environmental systems found at high latitudes or high elevations. Research foci of interest include, but are not limited to, development of multiscale models linked with widespread environmental data, sensor networks, biogeochemical fluxes across scales, hydrological and ecological networks, and data-driven environmental system modelling.

We seek an outstanding individual who will lead an internationally recognized research program that extends and leverages the opportunities offered by CAPE/EPFL. The professor will be committed to excellence in research and in undergraduate and graduate level teaching, and will contribute to the teaching program in Environmental Engineering at EPFL, which views basic and translational research as the foundation for environmental adaption and engineering design.

With its main campus located in Lausanne and its developing antennae in neighbouring cantons in Switzerland, EPFL is a growing and well-funded institution fostering excellence and diversity. It is well equipped with experimental and computational infrastructure, and offers a fertile environment for research collaboration between different disciplines. The EPFL environment is multilingual and multicultural, with English serving as a common interface. EPFL offers internationally competitive start-up resources, salaries, and benefits. It is committed to increasing the diversity of its faculty, and strongly encourages women to apply.

The following documents are requested in PDF format: cover letter including a statement of motivation, curriculum vitae including explicit mention of career breaks, publications list, concise statements of research and teaching interests (3-5 pages) as well as the names and addresses, including emails, of at least three references for junior positions or five references for senior positions (may be contacted at a later stage). Applications should be uploaded to the EPFL recruitment web site:

https://facultyrecruiting.epfl.ch/position/10977282

Formal evaluation of the applications will begin on **September 1, 2018** and the search will continue until the position is filled.

Further enquiries should be made to:
Prof. D. Andrew Barry
Chair of the Search Committee
E-mail: searchenvironmental@epfl.ch

For additional information on EPFL:
http://www.epfl.ch, http://enac.epfl.ch; https://valais.epfl.ch/Home

EPFL is an equal opportunity employer and a family friendly university.

University of Pittsburgh
Graduate School of Public Health
Department of Epidemiology

ASSOCIATE/ASSISTANT PROFESSOR OF EPIDEMIOLOGY
Infectious Diseases

The Department of Epidemiology, Graduate School of Public Health (GSPH), University of Pittsburgh, Infectious Diseases Epidemiology Research Program is recruiting tenure stream faculty at the Associate or Assistant Professor levels with expertise in infectious diseases epidemiologic research. Opportunities are available for collaboration with the Pitt Center for Global Health, Pitt clinical research sites in Mozambique and South Africa, and Center for Vaccine Research. GSPH regularly ranks among the highest in NIH funding. The Department of Epidemiology has a strong research portfolio, collaborating locally with other institutions across the University, nationally, and internationally. Initial salary support, start-up funds, and laboratory space (if needed) will be provided.

Candidates should have a doctoral degree (MD and/or PhD) and at least 3 years postdoctoral experience with a demonstrated record of recent research accomplishments and sufficient evidence to establish an independent, funded research program. Salary and academic rank will be commensurate with experience.

Applications will be accepted until positions are filled. Send letter of intent, curriculum vitae, statement of research interests, and the names of three references by E-mail to: **Position #0124922, c/o D. Bushey, Department of Epidemiology, University of Pittsburgh, A528 Crabtree Hall, 130 DeSoto Street, Pittsburgh, PA 15261**; E-mail: **dlb22@pitt/edu**.

EEO/AA/M/F/Vets/Disabled.

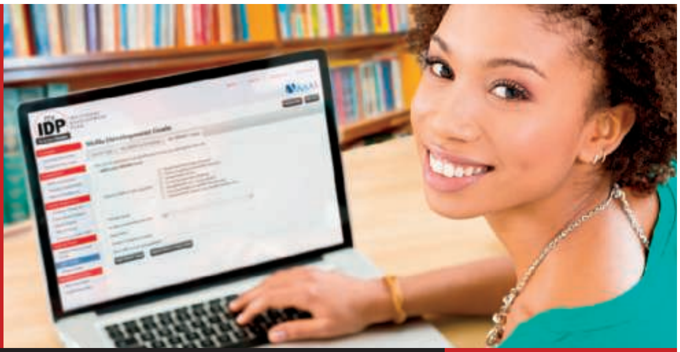
SCHOOL OF MEDICINE
CASE WESTERN RESERVE
UNIVERSITY

Open Rank Faculty Position (tenure-track/tenured), Dept. of Molecular Biology and Microbiology, Case Western Reserve University (CWRU), Part of the CWRU - Cleveland VAMC Center for Antimicrobial Resistance and Epidemiology (Case VA CARES). We are seeking applications for a 12-month tenure track or tenured position at the Assistant, Associate, or Full Professor level. Case VA CARES aims to strengthen and expand the research program targeted at understanding mechanisms of antibiotic resistance. We encourage applications from individuals with demonstrated experience in: molecular basis of antimicrobial action and resistance, alternative approaches to combating infections, animal models of infection with MDR organisms, as well as molecular epidemiology of MDR infections and bioinformatics.

Successful candidates will establish a vigorous research program, teach, and interact with a nationally-ranked team scientists interested in AMR and infectious diseases. A Ph.D. or M.D. degree, as well as at least 3 years of postdoctoral training are required. Candidates at the Assistant Professor level must have demonstrated a capacity for independent research and exhibit a high likelihood of attracting federal funding. At the Associate and Full Professor level, candidates should have a record of federal funding, an active research program, and evidence of a national/international reputation. Rank and salary will be commensurate with experience.

The University offers partner benefits and is responsive to the needs of dual-career couples. Applicants must email a letter of application, CV, brief statement of research goals and accomplishments, and 3 references as PDF files to: MBIO-VAMC Search Committee (**VAMBIO205@CASE.EDU**). Only complete applications will be considered.

In employment, as in education, Case Western Reserve University is committed to Equal Opportunity and Diversity. Women, veterans, members of underrepresented minority groups, and individuals with disabilities are encouraged to apply. Case Western Reserve University provides reasonable accommodations to applicants with disabilities. Applicants requiring a reasonable accommodation for any part of the application and hiring process should contact the Office of Inclusion, Diversity and Equal Opportunity at 216-368-8877 to request a reasonable accommodation. Determinations as to granting reasonable accommodations for any applicant will be made on a case-by-case basis.



ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

Faculty Position in Terrestrial Ecology

Joint Appointment between the Swiss Federal Institute for Forest Snow and Landscape Research (WSL) and the Ecole polytechnique fédérale de Lausanne (EPFL)

EPFL's School of Architecture, Civil and Environmental Engineering (ENAC) and the Swiss Federal Institute for Forest, Snow and Landscape Research (WSL) invite applications for either a tenure-track (Assistant) or tenured (Associate or Full) Professor in the Institute of Environmental Engineering. The appointee will lead the Swiss Romande WSL Site, and contribute to research and teaching activities within the EPFL Institute of Environmental Engineering.

The Institute of Environmental Engineering (IIE) in ENAC carries out basic and translational research spanning fundamental understanding of environmental systems and their resilience to design of adaption strategies. It covers a diverse portfolio in research, teaching and innovative technology development, including: Climate Change and Geochemical Cycles, Hydrology, Hydrodynamics, Precipitation, Water Quality, Bioremediation, Ecotoxicology, Air Quality, Renewable Energy, City and Landscape Development, Ecosystems, Ecology and Robotics. These research themes are underpinned by extensive facilities in, e.g., chemical, microbial and isotopic analysis. This professorship, along with the WSL Swiss Romande site, is located at EPFL's main campus at Lausanne. While IIE as a whole is mainly based in Lausanne, it is currently undergoing expansion at the EPFL Sion campus, with the creation of the new Centre for Changing Alpine and Polar Environments.

WSL is the national Swiss organization for basic and applied research into use and management of natural resources including forests, landscapes, and biodiversity. It focuses on use and protection of resources and the management of risks of natural hazards. Together with civic and governmental bodies, WSL seeks sustainable solutions to societally relevant questions.

The joint WSL/EPFL professor in terrestrial ecology will have acknowledged strengths in research related to forests, biodiversity and ecosystem services in dynamic environments. Areas of interest within these domains include, but are not limited to: functional biodiversity in different settings (urban to forested), responses and adaption of ecosystems to global changes, management of ecosystem services at local-to-global scales, plant-soil interactions, and ecosystem climate responses.

The WSL/EPFL professor will lead an internationally recognized research program that leverages the opportunities offered by WSL and EPFL. The professor will support WSL research and outreach activities, and enhance the group's integration into teaching and transdisciplinary research within ENAC. The appointee will promote excellence in research and in undergraduate and graduate level teaching.

The WSL/EPFL Professor of Terrestrial Ecology will be located at EPFL. EPFL is a growing and well-funded institution fostering excellence and diversity, with a highly international campus at an attractive location with excellent experimental and computational infrastructure. Teaching and research at EPFL covers essentially the entire palette of engineering and science, and offers a fertile environment for research collaboration between different disciplines. The EPFL environment is multilingual and multicultural, with English serving as a common interface. It is committed to increasing the diversity of its faculty, and strongly encourages women to apply.

The following documents are requested in PDF format: cover letter including a statement of motivation, curriculum vitae including explicit mention of career breaks, publications list, concise statements of research and teaching interests (3-5 pages) as well as the names and addresses, including emails, of at least three references for junior positions or five references for senior positions (may be contacted at a later stage).
Applications should be uploaded to the EPFL recruitment web site

https://facultyrecruiting.epfl.ch/position/10977284

Formal evaluation of the applications will begin on **August 15, 2018** and the search will continue until the position is filled.

Further enquiries should be made to:
Prof. D. Andrew Barry
Chair of the Search Committee
E-mail: SearchTerrestrialEco@epfl.ch

For additional information on WSL and EPFL:
http://wsl.ch, http://www.epfl.ch, http://enac.epfl.ch

EPFL is an equal opportunity employer and a family friendly university.

PIONEER INITIATIVE
Science
Careers
See all the latest jobs
Search Jobs
Two sides of the H-1B

ETH zürich

Professor of Solid State NMR Spectroscopy

→ The Department of Chemistry and Applied Biosciences (www.chab.ethz.ch) at ETH Zurich invites applications for the above-mentioned position. The professorship should have a research focus on NMR spectroscopy of the solid state with possible applications in either the life sciences or materials research. Candidates are expected to establish a research programme at the forefront of method development in their field. Collaboration with other groups at ETH in application work or method development is strongly encouraged and teaching in all areas of physical chemistry is expected. In general, at ETH Zurich undergraduate-level courses are taught in German or English and graduate level courses are taught in English.

→ **Please apply online: www.facultyaffairs.ethz.ch**

→ Applications should include a curriculum vitae, a list of publications, a statement of future research and teaching interests, and a description of the three most important achievements. The letter of application should be addressed **to the President of ETH Zurich, Prof. Dr. Lino Guzzella. The closing date for applications is 31 August 2018.** ETH Zurich is an equal opportunity and family friendly employer and is responsive to the needs of dual career couples. We specifically encourage women to apply.

Advance your career with expert advice from *Science* Careers.

Download Free Career Advice Booklets!
ScienceCareers.org/booklets

Featured Topics:

- Networking
- Industry or Academia
- Job Searching
- Non-Bench Careers
- And More

ScienceCareers
FROM THE JOURNAL SCIENCE AAAS

MPI

Max Planck Institute for Terrestrial Microbiology

The Max Planck Institute for Terrestrial Microbiology announces the position of

Department Director

The Max Planck Institute for Terrestrial Microbiology in Marburg, Germany invites expressions of interest from or nominations of outstanding scientists for the position of Director. A Max Planck Director pursues a bold long-term research agenda in his or her own Department, and leads the Institute jointly with the fellow Directors. We seek a director with a high international reputation and a proven track record in research areas that complement, strengthen or expand the research agenda of the Institute (for details, please see **http://www.mpi-marburg.mpg.de/**), preferably with a focus on interactions between microorganisms, on interactions of microorganisms with their environment or with their hosts. Chief criteria for a directorship are scientific and scholarly excellence, creativity, and an innovative and unique research program, rather than seniority or publication rankings.

We offer an excellent and international research environment, outstanding infrastructure, and long-term scientific funding. The working language is English and qualified individuals of all nationalities will be considered.

The Max Planck Society is deeply committed to the principles of gender equality, diversity, and inclusiveness and welcomes expressions of interest from all backgrounds. We actively encourage individuals with disabilities to submit an expression of interest. The Max Planck Society is aiming to increase the proportion of women among its scientific leadership. We therefore strongly encourage expressions of interests from and the nominations of women.

Both expressions of interest and nominations are welcome. Expressions of interest must consist of a letter of intent, a full curriculum vitae incl. a complete publication list, a brief summary of previous research accomplishments and a brief concept statement outlining the structure and long-term, visionary research agenda of the Department. Further information can be obtained from the managing director Prof. Dr. Victor Sourjik.

Expressions of interest and nominations will be held in strictest confidence and should be sent by e-mail as a single pdf file to **director@mpi-marburg.mpg.de**

Zertifikat seit 2011
audit berufundfamilie

Expressions of interest and nominations are due by **June 22nd 2018**.

Short-listed candidates will participate in a search symposium in Marburg in November 2018.

MAX-PLANCK-GESELLSCHAFT

By **Moamen Elmassry**

What I learned from teaching

When I started my Ph.D. program, 3 years ago, I was excited to conduct cutting-edge research. But I wasn't so sure about the teaching that I would have to do every semester to fund my education. I saw it as a waste of time and energy that I could otherwise spend in the lab. The stereotype of the teaching assistant (TA) rushing between classes and spending weekends grading while guzzling coffee did not help. I had no experience teaching, and the idea of being in front of a class made me anxious. But if you ask me now, I would say that teaching has been one of the most rewarding parts of my Ph.D. experience.

My first semester was tough. In addition to getting my research going, I was adjusting to living in a new country—I had just moved to the United States from my home country of Egypt. Adding teaching on top of that, with just a brief training workshop to prepare me, was not a recipe for success. I tried my best to help my students learn, but my inexperience was apparent.

I could have carried on as a mediocre teacher. But I recalled how some of my own teachers had inspired me over the years. I felt I owed my students the same—which, I realized, would require time and training. It was my responsibility to make that happen, even if it meant taking a little more time and focus away from my research.

ILLUSTRATION: ROBERT NEUBECKER

> *"Teaching has been one of the most rewarding parts of my Ph.D. experience."*

I started with the online Science Teaching Fellows Program from the American Society for Microbiology. I also began attending workshops and training sessions at my university's Teaching, Learning, and Professional Development Center (TLPDC). I loved being part of a supportive community and learned some valuable teaching strategies, which I experimented with in my second semester of teaching. For example, I introduced my students to epidemiology by asking them to write short stories about an epidemic spreading on campus, hoping to incorporate more creativity into their learning. This unconventional assignment surprised the students at first. But some of them got so into it that they wrote much more than the half page I had assigned. I loved seeing my students so engaged with an activity I had designed. In my end-of-semester evaluations, some students said that I was their favorite TA, and others asked me to write recommendation letters for them, which was both humbling and rewarding.

I was eager to continue my teacher training. I enthusiastically proposed to my adviser that I apply for a TLPDC program that requires 30 to 40 hours over the course of a semester. I was disappointed when he said I hadn't made enough progress on my dissertation research, but I also acknowledged that he was probably right. I found an alternate program that would still help me grow as an educator without taking as much of my time. And I later made enough progress on my research to participate in the original program—with my adviser's approval—the following year.

I am now focused on establishing a balance between my primary job—teaching, which is what I am paid for—and my primary goal: completing my doctoral degree and pursuing a career in academia. As I work to find this balance, I also remind myself—and my adviser and thesis committee members—that teaching has provided me with some unexpected benefits. Knowing that I have teaching commitments pushes me to conduct efficient, well-designed experiments. Answering undergraduate students' fundamental "why" questions helps keep me intellectually stimulated and forces me to think about science in new ways, which was useful during my Ph.D. candidacy exam. I also realized recently that my occasional bouts of near depression, triggered by failed experiments, hit only during summer and winter breaks, when I do not teach. If my research goes badly while I am teaching, at least I can see that my work is paying off for my students.

If you're viewing your teaching assignments with dread, I encourage you instead to approach them with an open mind and a willingness to learn. Maybe, like me, you'll find that teaching is full of unexpected rewards. ■

Moamen Elmassry is a Ph.D. candidate at Texas Tech University in Lubbock. Do you have an interesting career story? Send it to SciCareerEditor@aaas.org.

Published by AAAS