FDA documents reveal slipshod oversight of clinical trials p. 24

Freshwater conservation requires specific planning pp. 38 & 117

Thirst research wins neurobiology prize p. 45

Science

$15
2 OCTOBER 2020

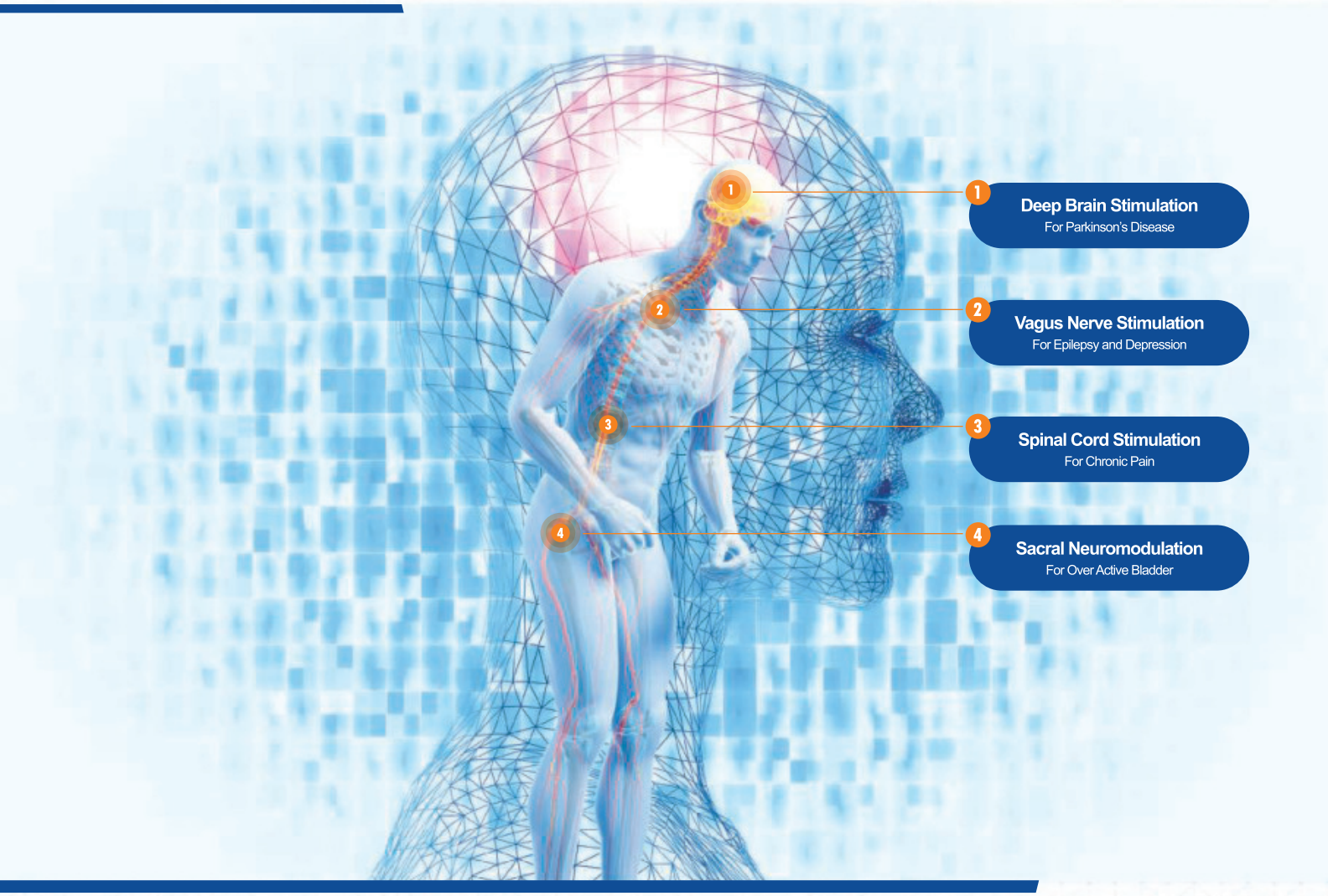
AAAS

SPECIAL ISSUE

NEURODEGENERATION

Understanding the brain's decline

PINS 品驰



Devoted to Neuromodulation for more than 20 years !

- **In 2000,** started to work on the DBS (Deep Brain Stimulation) initiative
- **In 2012,** set up the national research platform 'National Engineering Laboratory of Neuromodulation Technology'
- **In 2016,** established Science & PINS prize for Neuromodulation with Science

  PINS DBS products got CE marked
- **In 2017,** Professor Alim-Louis Benabid, founder of DBS therapy and winner of Lasker Medical Research Awards, visited PINS
- **In 2018,** PINS 'Remote programming system of implantable medical devices' won the gold medal award of Geneva International Invention Exhibition
- **In 2019,** PINS DBS project won the first prize of 2018 National Science and Technology Progress Award (China's top Science and Technology Progress Award)

Beijing PINS Medical Co., Ltd. focuses on the development and innovation of neuromodulation technologies, aiming to provide more comprehensive and professional solutions for patients suffering from functional neurological disorder.

www.pinsmedical.com/html/en
info@pinsmedical.com

Sethuraman Panchanathan
National Science Foundation

Ruha Benjamin
Princeton University

Anthony S. Fauci
National Institute of Allergy and Infectious Diseases

Mary L. Gray
Microsoft Research

Claire Fraser
AAAS President
University of Maryland School of Medicine

UNDERSTANDING DYNAMIC ECOSYSTEMS

AAAS | ANNUAL MEETING

The 2021 Annual Meeting will convene entirely online, February 8-11
with related pre-released materials available starting in late January.
Meeting registration will be available in November.

For more information, please visit:

aaas.org/meetings | #AAASmtg

AMERICAN ASSOCIATION FOR THE ADVANCEMENT OF SCIENCE

Eppendorf: Working for a better world for 75 years

Eppendorf is supporting essential businesses and institutions in the fight against the coronavirus. In the year of the company's 75th birthday, we are thus fulfilling the mission of our founders – to help improve people's living conditions – in a very special way.

"The mission of our company, as defined by its founders, is as current and relevant today as it was then," says Eva van Pelt, Co-CEO of Eppendorf AG. This commitment to improve people's living conditions drove a small group of technical experts led by the company's founders, Dr. Heinrich Netheler and Dr. Hans Hinz, toward the healthcare sector in the first weeks after World War II. In August 1945, they began repairing urgently needed but defective medical equipment and instruments belonging to University Medical Center Eppendorf in Hamburg. The group was so successful in that task that it soon received its first development orders for new types of equipment from the medical center's departments.

Eppendorf – an innovative pioneer in medical technology

The innovations developed by Eppendorf – such as the Stimulator, a device to stimulate muscles and nerves, or the Ophtalmochirurg, a kind of forerunner to eye lasers for the treatment of retinal detachment – can be regarded as the archetypes of modern medical technologies now used routinely all over the world. Eva van Pelt points to the important products developed by Eppendorf since its foundation and leaves no doubt that the company's innovative spirit and determination to advance technically will continue to be reflected in its products in the future. Nowadays, however, Eppendorf no longer operates in the field of medical technology, but as a globally successful manufacturer of laboratory equipment with around 4,000 employees all over the world.

"The mission of our company, as defined by its founders, is as current and relevant today as it was then."

Eva van Pelt, Co-CEO

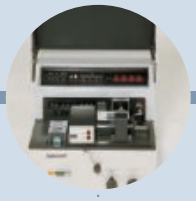
"In our work, we cooperate closely and trustfully with our partners and customers to address the issues of the future," adds Co-CEO Dr. Peter Fruhstorfer. He notes that Eppendorf is driven by the idea of enabling laboratories to work efficiently, sustainably and with high quality. "Specifically, our aim is to speed up laboratory processes while at the same time reducing potential errors. And we want that to go hand in hand with significant reductions in the workload of laboratory staff." This is a Herculean task, as Fruhstorfer points out, and one that requires digitalization, the best minds at Eppendorf and the special spirit that exists within the company.

Strong teamwork for quality and innovation

"The pronounced sense of community within the company is one of the keys to Eppendorf's long-standing success," Dr. Fruhstorfer explains. The well-being of every employee was always important to the founders of Eppendorf, and this attitude remains an integral part of the company's distinctive culture to this day. Mutual respect is just as important as the open-minded attitude toward others that can be encountered at Eppendorf throughout the world. In the eight decades of Eppendorf's presence on the market, these two qualities have been the basis for dialogue, the transfer of knowledge, exchanges of experience and global networking. All of these factors have contributed to the high reliability, good service and special quality of the company's products and thus to the efficient support provided by equipment, consumables and service that has long made Eppendorf indispensable for many laboratories.

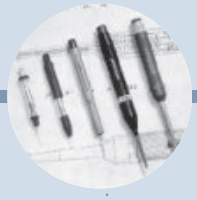
1945

Dr. Heinrich Netheler and Dr. Hans Hinz develop medical diagnostic products; in 1946 their company was renamed "Elektromedizinische Werkstätten GmbH."



1949

Eppendorf develops a photometer for clinical applications. Over time, it became a world standard for chemical and biochemical analyses.



1961

Eppendorf launches the first piston-stroke microliter pipette, heralding the age of precise and fast pipetting.

1963

A microliter system consisting of reaction vessels, mixers, centrifuges and pipettes is launched. It facilitates and improves laboratory work enormously.

1963

The "Eppi®" reaction vessel is launched on the market. It quickly became indispensable in medical and scientific laboratories worldwide.

Above
View into the clean room production: high-quality laboratory consumables made of high-tech plastics

Bottom left
Production halls and high-bay warehouse at the Eppendorf location Oldenburg in Holstein

Bottom right
Eppendorf centrifuges of the latest generation at the Leipzig location

Pioneer in accuracy, reliability and innovation

> In Liquid Handling, the product range extends from manual pipettes to electronic pipettes and comprehensive pipetting systems.

> Whether tubes, plates, pipette tips, stem cell cultureware or cuvettes: The unique product features of Eppendorf Consumables accelerate and simplify routine tasks and ensure greater reliability in every laboratory.

> Eppendorf is one of the world`s market leaders and full-spectrum supplier in the field of centrifugation: the product range extends from benchtop centrifuges, premium floor-standing centrifuges and ultracentrifuges to clinical and automated centrifuges.

High-tech developments together with users

In its Business Areas, Eppendorf focuses all its ideas and developments on the needs of the user. Customer requirements determine what innovations are developed, and these are constantly discussed and tested with users through agile working methods – a process that, by the way, was used in Eppendorf's early days, laying the foundation for the great market success of its equipment.

The result is state-of-the-art equipment technology with sophisticated ergonomics, ensuring that Eppendorf's products are always up to date. Ideally, they are even a step ahead of the times, like the piston-stroke pipette was when it was launched. Other examples are the Eppendorf centrifuges with their remarkable durability and Eppendorf's high-tech plastic consumables for use in laboratories.

State-of-the-art technology and the sophisticated ergonomics of the laboratory equipment ensure easy and reliable handling, allowing users to concentrate fully on their research.

"The pronounced sense of community within the company is one of the keys to Eppendorf's long-standing success."

Dr. Peter Fruhstorfer, Co-CEO

In this way, Eppendorf actively contributes to research activities around the world every day – and at the moment, to the fight against the Sars-CoV-2 virus – in keeping with the mission that the founders impressed on the company 75 years ago: to improve people's living conditions.

More information at:
eppendorf.com

1978

The first Multipette® with Combitips® is launched on the market and becomes a bestseller.

1996

Eppendorf develops the air and oil micro-injectors CellTram – for pressure control during manual micro-injection and dosing of liquids.

2003

The Eppendorf line of epMotion automated liquid handling systems is designed to help automate routine pipetting tasks to free up time.

2018

The 4 liter versatile Multipurpose Centrifuge 5910 R with universal rotor and adapter concept and the smallest footprint in the market expands the centrifuge range.

TOMORROW LAB SINCE 1945
75 YEARS

Today

Discover more product innovations online: eppendorf.com/75-years

NEVER
STOP
ADVANCING BIOPHARMACEUTICALS
With over 80 years' experience in healthcare, Fujifilm is committed to enhancing patient care and addressing unmet medical needs. As a world-leading contract biopharmaceutical manufacturer (CDMO), we're helping accelerate the development of new drugs with a complete range of services. Our state-of-the-art facilities deliver precise control of process conditions and unmatched productivity. Utilizing our cell biology and bioprocessing expertise, we provide high-quality cell culture media to our biopharma partners for the production of many of the world's leading drugs. At Fujifilm, we'll NEVER STOP pushing scientific boundaries to help develop novel therapies and life-changing treatments that create a healthier world for all.
FUJIFILM
Value from Innovation
Follow Fujifilm Life Sciences on
Image is an actual manufacturing facility of Fujifilm in Denmark.
FUJIFILM and Fujifilm Value from Innovation are trademarks of FUJIFILM Corporation.
©2020 FUJIFILM Corporation. All rights reserved.

CONTENTS

2 OCTOBER 2020 • VOLUME 370 • ISSUE 6512

SPECIAL SECTION

NEURODEGENERATION

INTRODUCTION

48 A cruel end to too many lives

REVIEWS

50 Glymphatic failure as a final common pathway to dementia
M. Nedergaard and S. A. Goldman

56 Beyond aggregation: Pathological phase transitions in neurodegenerative disease *C. Mathieu* et al.

61 Translating genetic risk of Alzheimer's disease into mechanistic insight and drug targets *A. Sierksma* et al.

66 Microglia modulate neurodegeneration in Alzheimer's and Parkinson's diseases *T. Bartels* et al.

ON THE COVER

Neurodegeneration in later life robs us of our abilities and our memories, slowly and inexorably, like leaves falling from a tree. A greater understanding of the pathophysiological underpinnings within the brain may provide crucial hints that will help us delay and perhaps even reverse symptoms. See page 48. *Illustration: Simon Prades*

SEE ALSO PERSPECTIVE p. 32

QQ群: 1074370165

NEWS

IN BRIEF

14 News at a glance

IN DEPTH

16 Europe builds 'digital twin' of Earth to hone climate forecasts
Ingesting more data than ever before, exascale model will simulate the impact of climate change on humans *By P. Voosen*

17 China's bold climate pledge earns praise—but is it feasible?
Plan to achieve carbon neutrality by 2060 would make China a global leader, but abandoning coal will be hard
By D. Normile

18 Record U.S. and Australian fires raise fears for many species
Scientists say fires likely wiped out some rare Australian organisms, and worry U.S. blazes now threaten more
By J. Pickrell and E. Pennisi
EDITORIAL p. 13

19 Stem cell studies probe origins of the placenta
Lab models of organ will track how it emerges—and what can go wrong
By K. Servick

20 The 'bat man' tackles COVID-19
After a career investigating why so many viruses come from bats, Linfa Wang eyes a new challenge *By K. Kupferschmidt*

22 A call for diagnostic tests to report viral load
Measure could help officials know who is most contagious *By R. F. Service*

FEATURES

24 Official inaction
A *Science* investigation shows that FDA oversight of clinical trials is lax, slow moving, and secretive—and that enforcement is declining *By C. Piller*

27 Disgraced researchers can still reap drug industry payouts
By C. Piller
PODCAST

INSIGHTS

LETTERS

30 NextGen Voices: Funding fix: Spend time

PERSPECTIVES

32 Shifts and drifts in prion science
Important questions remain unanswered since prions were discovered four decades ago *By A. Aguzzi and E. De Cecco*
NEURODEGENERATION SECTION p. 48

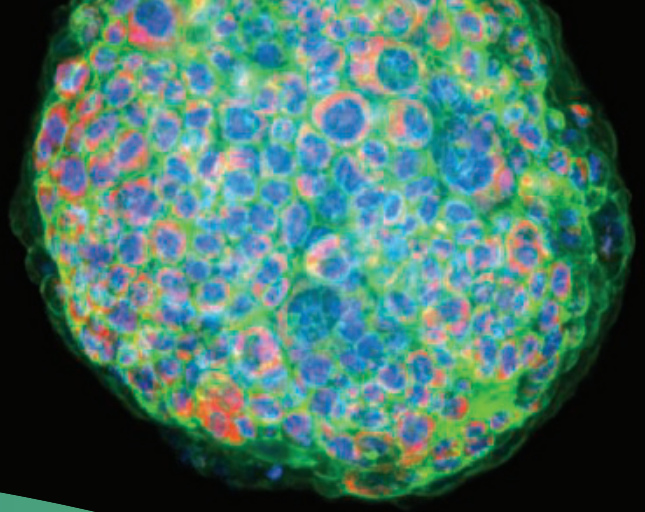
34 Mutational selection in normal urothelium
Mutations in normal tissue point to causes of DNA damage and set the stage for cancer
By S. G. Rozen
RESEARCH ARTICLES pp. 75 & 82

36 Coaxing stem cells to repair the spinal cord
Spinal cells in mice can be induced to generate protective oligodendrocytes after injury *By C. G. Becker and T. Becker*
RESEARCH ARTICLE p. 73

37 A rival to superalloys at high temperatures
Slip-pathway activation provides plasticity in a multiprincipal element alloy with high-temperature strength *By J. Cairney*
RESEARCH ARTICLE p. 95

ILLUSTRATION: SIMON PRADES

MOLECULAR DEVICES



CAPTURE the micro moments that matter

NEW Water Immersion Objectives

Gain the sensitivity to capture more phenotypic data at greater depths. Our high-content imagers and team are here to help your lab simplify imaging and analysis of complex, 3D assay workflows and fast-track discoveries.

Learn more at
moleculardevices.com/water

For Research Use Only. Not for use in diagnostic procedures.
© 2020 Molecular Devices, LLC. All Rights Reserved. The trademarks mentioned herein are the property of Molecular Devices, LLC or their respective owners.

38 A boost for freshwater conservation
Integrating freshwater and terrestrial conservation planning has high returns
By R. Abell and I. J. Harrison
REPORT p. 117

POLICY FORUM

40 How to fix the GDPR's frustration of global biomedical research
Sharing of data for research beyond the EU must improve *By J. Bovenberg* et al.

BOOKS *ET AL.*

43 Flawed research and its enduring repercussions
A journalist recounts how he exposed problems with a study linking vaccines and autism *By P. A. Offit*

44 Training tomorrow's scientists
Prioritizing STEM diversity, equity, and inclusion requires rethinking graduate education *By A. Huderson*

PRIZE ESSAY

45 The origins of thirst
Sensory signals arise throughout the body and converge in the brain to regulate drinking
By C. A. Zimmerman

PHOTO: CLAUS MEYER/MINDEN PICTURES

RESEARCH

IN BRIEF

70 From *Science* and other journals

RESEARCH ARTICLES

73 Neurodevelopment
A latent lineage potential in resident neural stem cells enables spinal cord repair
E. Llorens-Bobadilla et al.
RESEARCH ARTICLE SUMMARY; FOR FULL TEXT: DX.DOI.ORG/10.1126/SCIENCE.ABB8795
PERSPECTIVE p. 36

74 Solar cells
Vapor-assisted deposition of highly efficient, stable black-phase $FAPbI_3$ perovskite solar cells *H. Lu* et al.
RESEARCH ARTICLE SUMMARY; FOR FULL TEXT: DX.DOI.ORG/10.1126/SCIENCE.ABB8985

Mutation
75 Extensive heterogeneity in somatic mutation and selection in the human bladder *A. R. J. Lawson* et al.

82 Macroscopic somatic clonal expansion in morphologically normal human urothelium *R. Li* et al.
PERSPECTIVE p. 34

89 Coronavirus
Selective and cross-reactive SARS-CoV-2 T cell epitopes in unexposed humans
J. Mateus et al.

95 Metallurgy
Multiplicity of dislocation pathways in a refractory multiprincipal element alloy *F. Wang* et al.
PERSPECTIVE p. 37

101 Organic chemistry
A universal system for digitization and automatic execution of the chemical synthesis literature
S. H. M. Mehr et al.

REPORTS

108 Solar cells
Impact of strain relaxation on performance of α-formamidinium lead iodide perovskite solar cells *G. Kim* et al.

113 Developmental biology
An adhesion code ensures robust pattern formation during tissue morphogenesis
T. Y.-C. Tsai et al.

117 Conservation planning
Integrated terrestrial-freshwater planning doubles conservation of tropical aquatic species *C. G. Leal* et al.
PERSPECTIVE p. 38

121 Sex determination
The mouse *Sry* locus harbors a cryptic exon that is essential for male sex determination *S. Miyawaki* et al.

Electrocalorics
125 Giant temperature span in electrocaloric regenerator *A. Torelló* et al.

129 A high-performance solid-state electrocaloric cooling system
Y. Wang et al.

DEPARTMENTS

13 Editorial
Fire in our future
By William Wallace Covington and Stephen Pyne
NEWS STORY p. 18

138 Working Life
A leap of faith *By Rachel Mason*

Science Staff 10
New Products 134
Science Careers 135

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 © 2020 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): $2148; Foreign postage extra: Air assist delivery: $98. First class, airmail, student, and emeritus rates on request. Canadian rates with GST available upon request, GST #125488122. Publications Mail Agreement Number 1069624. **Printed in the U.S.A.**
Change of address: Allow 4 weeks, giving old and new addresses and 8-digit account number. **Postmaster:** Send change of address to AAAS, P.O. Box 96178, Washington, DC 20090–6178. **Single-copy sales:** $15 each plus shipping and handling available from backissues.sciencemag.org; bulk rate on request. **Authorization to reproduce** material for internal or personal use under circumstances not falling within the fair use provisions of the Copyright Act can be obtained through the Copyright Clearance Center (CCC), www.copyright.com. The identification code for *Science* is 0036-8075. *Science* is indexed in the *Reader's Guide to Periodical Literature* and in several specialized indexes.

Science AAAS

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

Executive Editor Monica M. Bradford

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

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

News Editor Tim Appenzeller

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

Creative Director Beth Rakouskas

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

Chief Executive Officer and Executive Publisher Sudip Parikh

Publisher, Science Family of Journals Bill Moran

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

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

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

DIRECTOR, PRODUCT & PUBLISHING DEVELOPMENT Chris Reid **DIRECTOR, BUSINESS STRATEGY AND PORTFOLIO MANAGEMENT** Sarah Whalen **ASSOCIATE DIRECTOR, PRODUCT MANAGMENT** Kris Bishop **PRODUCT DEVELOPMENT MANAGER** Scott Chernoff **PUBLISHING TECHNOLOGY MANAGER** Michael Di Natale **SR. PRODUCT ASSOCIATE** Robert Koepke **SPJ ASSOCIATE** Samantha Bruno Fuller

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

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

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

---

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

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

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

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

EDITORIAL
science_editors@aaas.org

NEWS
science_news@aaas.org

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

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

MEDIA CONTACTS
scipak@aaas.org

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

INSTITUTIONAL SALES AND SITE LICENSES
sciencemag.org/librarian

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

science_advertising@aaas.org

CLASSIFIED ADVERTISING
advertising.sciencemag.org/science-careers

advertise@sciencecareers.org

JOB POSTING CUSTOMER SERVICE
employers.sciencecareers.org

support@sciencecareers.org

MEMBERSHIP AND INDIVIDUAL SUBSCRIPTIONS
sciencemag.org/subscriptions

MEMBER BENEFITS
aaas.org/membercentral

AAAS BOARD OF DIRECTORS
CHAIR Steven Chu
PRESIDENT Claire M. Fraser
PRESIDENT-ELECT Susan G. Amara
TREASURER Carolyn N. Ainslie
CHIEF EXECUTIVE OFFICER
Sudip Parikh
BOARD Cynthia M. Beall
Rosina M. Bierbaum
Ann Bostrom
Stephen P.A. Fodor
S. James Gates, Jr.
Laura H. Greene
Kaye Husbands Fealing
Maria M. Klawe
Robert B. Millard
Alondra Nelson
William D. Provine

BOARD OF REVIEWING EDITORS (Statistics board members indicated with **(S)**)

Adriano Aguzzi, *U. Hospital Zürich*
Takuzo Aida, *U. of Tokyo*
Leslie Aiello, *Wenner-Gren Foundation*
Deji Akinwande, *UT Austin*
Judith Allen, *U. of Manchester*
Marcella Alsan, *Harvard U.*
Sebastian Amigorena, *Institut Curie*
James Analytis, *UC Berkeley*
Trevor Archer, *NIEHS, NIH*
Paola Arlotta, *Harvard U.*
Johan Auwerx, *EPFL*
David Awschalom, *U. of Chicago*
Clare Baker, *U. of Cambridge*
Nenad Ban, *ETH Zürich*
Franz Bauer, *Pontificia U. Católica de Chile*
Ray H. Baughman, *U. of Texas at Dallas*
Carlo Beenakker, *Leiden U.*
Yasmine Belkaid, *NIAID, NIH*
Philip Benfey, *Duke U.*
Gabriele Bergers, *VIB*
Kiros T. Berhane, *Columbia U.*
Bradley Bernstein, *Mass. General Hospital*
Joseph J. Berry, *NREL*
Alessandra Biffi, *Harvard Med. School*
Peer Bork, *EMBL*
Chris Bowler, *École Normale Supérieure*
Ian Boyd, *U. of St. Andrews*
Emily Brodsky, *UC Santa Cruz*
Ron Brookmeyer, *UCLA* **(S)**
Christian Büchel, *UKE Hamburg*
Dennis Burton, *Scripps Res.*
Carter Tribley Butts, *UC, Irvine*
György Buzsáki, *New York U. School of Med.*
Blanche Capel, *Duke U.*
Annmarie Carlton, *U. of California, Irvine*
Nick Chater, *U. of Warwick*
M. Keith Chen, *UCLA*
Zhijian Chen, *UT Southwestern Med. Ctr.*
Ib Chorkendorff, *Denmark TU*
James J. Collins, *MIT*
Robert Cook-Deegan, *Arizona State U.*
Alan Cowman, *Walter & Eliza Hall Inst.*
Carolyn Coyne, *U. of Pittsburgh*
Roberta Croce, *VU Amsterdam*
Ismaila Dabo, *Penn State U.*
Jeff L. Dangl, *U. of North Carolina*
Chiara Daraio, *Caltech*
Nicolas Dauphas, *U. of Chicago*
Christian Davenport, *U. of Michigan*
Frans de Waal, *Emory U.*
Claude Desplan, *New York U.*
Sandra Díaz, *U. Nacional de Córdoba*
Ulrike Diebold, *TU Wien*
Hong Ding, *Inst. of Physics, CAS*
Dennis Discher, *U. of Penn.*
Jennifer A. Doudna, *UC Berkeley*
Raissa M. D'Souza, *UC Davis*
Bruce Dunn, *UCLA*
William Dunphy, *Caltech*
Christopher Dye, *U. of Oxford*
Scott Edwards, *Harvard U.*
Todd Ehlers, *U. of Tübingen*
Jennifer Elisseeff, *Johns Hopkins U.*
Tim Elston, *U. of North Carolina*
Andrea Encalada, *U. San Francisco de Quito*
Nader Engheta, *U. of Penn.*
Karen Ersche, *U. of Cambridge*
Barry Everitt, *U. of Cambridge*
Vanessa Ezenwa, *U. of Georgia*
Michael Feuer, *The George Washington U.*
Toren Finkel, *U. of Pittsburgh Med. Ctr.*
Gwenn Flowers, *Simon Fraser U.*
Peter Fratzl, *Max Planck Inst. Potsdam*
Elaine Fuchs, *Rockefeller U.*
Eileen Furlong, *EMBL*
Jay Gallagher, *U. of Wisconsin*
Daniel Geschwind, *UCLA*
Karl-Heinz Glassmeier, *TU Braunschweig*
Ramon Gonzalez, *U. of South Florida*
Sandra González-Bailón, *U. of Penn*
Elizabeth Grove, *U. of Chicago*
Nicolas Gruber, *ETH Zürich*
Hua Guo, *U. of New Mexico*
Kip Guy, *U. of Kentucky College of Pharmacy*
Taekjip Ha, *Johns Hopkins U.*
Christian Haass, *Ludwig Maximilians U.*
Sharon Hammes-Schiffer, *Yale U.*
Wolf-Dietrich Hardt, *ETH Zürich*
Louise Harra, *U. College London*
Jian He, *Clemson U.*
Carl-Philipp Heisenberg, *IST Austria*
Ykä Helariutta, *U. of Cambridge*
Janet G. Hering, *Eawag*
Hans Hilgenkamp, *U. of Twente*
Kai-Uwe Hinrichs, *U. of Bremen*
Deirdre Hollingsworth, *U. of Oxford*
Lora Hooper, *UT Southwestern Med. Ctr.*
Fred Hughson, *Princeton U.*
Randall Hulet, *Rice U.*
Auke Ijspeert, *EPFL*
Akiko Iwasaki, *Yale U.*
Stephen Jackson, *USGS and U. of Arizona*
Kai Johnsson, *EPFL*
Peter Jonas, *IST Austria*
Matt Kaeberlein, *U. of Washington*
William Kaelin Jr., *Dana-Farber Cancer Inst.*
Daniel Kammen, *UC Berkeley*
V. Narry Kim, *Seoul Nat. U.*
Robert Kingston, *Harvard Med. School*
Nancy Knowlton, *Smithsonian Institution*
Etienne Koechlin, *École Normale Supérieure*
Alex L. Kolodkin, *Johns Hopkins U.*
Julija Krupic, *U. of Cambridge*
Thomas Langer, *Max Planck Inst. Cologne*
Mitchell A. Lazar, *U. of Penn.*
Wendell Lim, *UC San Francisco*
Jianguo Liu, *Michigan State U.*
Luis Liz-Marzán, *CIC biomaGUNE*
Omar Lizardo, *UCLA*
Jonathan Losos, *Washington U. in St. Louis*
Ke Lu, *Chinese Acad. of Sciences*
Christian Lüscher, *U. of Geneva*
Jean Lynch-Stieglitz, *Georgia Inst. of Tech.*
Fabienne Mackay, *QIMR Berghofer*
Anne Magurran, *U. of St. Andrews*
Asifa Majid, *U. of York*
Oscar Marín, *King's College London*
Charles Marshall, *UC Berkeley*
Christopher Marx, *U. of Idaho*
David Masopust, *U. of Minnesota*
Geraldine Masson, *CNRS*
C. Robertson McClung, *Dartmouth College*
Rodrigo Medellín, *U. Nacional Autónoma de México*
Graham Medley, *LSHTM*
Jane Memmott, *U. of Bristol*
C. Jessica Metcalf, *Princeton U.*
Baoxia Mi, *UC Berkeley*
Edward Miguel, *UC Berkeley*
Tom Misteli, *NCI, NIH*
Yasushi Miyashita, *U. of Tokyo*
Alison Motsinger-Reif, *NIEHS, NIH* **(S)**
Danielle Navarro, *U. of New South Wales*
Daniel Nettle, *Newcastle U.*
Daniel Neumark, *UC Berkeley*
Beatriz Noheda, *U. of Groningen*
Helga Nowotny, *Vienna Sci., Res. & Tech. Fund*
Rachel O'Reilly, *U. of Birmingham*
Harry Orr, *U. of Minnesota*
Pilar Ossorio, *U. of Wisconsin*
Andrew Oswald, *U. of Warwick*
Isabella Pagano, *Istituto Nazionale di Astrofisica*
Margaret Palmer, *U. of Maryland*
Elizabeth Levy Paluck, *Princeton U.*
Jane Parker, *Max Planck Inst. Cologne*
Giovanni Parmigiani, *Dana-Farber Cancer Inst.* **(S)**
Daniel Pauly, *U. of British Columbia*
Samuel Pfaff, *Salk Inst. for Biological Studies*
Julie Pfeiffer, *UT Southwestern Med. Ctr.*
Philip Phillips, *UIUC*
Matthieu Piel, *Institut Curie*
Kathrin Plath, *UCLA*
Martin Plenio, *Ulm U.*
Katherine Pollard, *UC San Francisco*
Elvira Poloczanska, *Alfred-Wegener-Inst.*
Julia Pongratz, *Ludwig Maximilians U.*
Philippe Poulin, *CNRS*
Jonathan Pritchard, *Stanford U.*
Félix A. Rey, *Institut Pasteur*
Trevor Robbins, *U. of Cambridge*
Joeri Rogelj, *Imperial College London*
Amy Rosenzweig, *Northwestern U.*
Mike Ryan, *UT Austin*
Shimon Sakaguchi, *Osaka U.*
Miquel Salmeron, *Lawrence Berkeley Nat. Lab*
Nitin Samarth, *Penn State U.*
Jürgen Sandkühler, *Med. U. of Vienna*
Erica Ollmann Saphire, *La Jolla Inst.*
Alexander Schier, *Harvard U.*
Wolfram Schlenker, *Columbia U.*
Susannah Scott, *UC Santa Barbara*
Rebecca Sear, *LSHTM*
Anuj Shah, *U. of Chicago*
Vladimir Shalaev, *Purdue U.*
Jie Shan, *Cornell U.*
Beth Shapiro, *UC Santa Cruz*
Jay Shendure, *U. of Washington*
Steve Sherwood, *U. of New South Wales*
Brian Shoichet, *UC San Francisco*
Robert Siliciano, *Johns Hopkins U. School of Med.*
Lucia Sivilotti, *U. College London*
Alison Smith, *John Innes Centre*
Richard Smith, *U. of North Carolina* **(S)**
Mark Smyth, *QIMR Berghofer*
Pam Soltis, *U. of Florida*
John Speakman, *U. of Aberdeen*
Tara Spires-Jones, *U. of Edinburgh*
Allan C. Spradling, *Carnegie Institution for Sci.*
V. S. Subrahmanian, *Dartmouth College*
Ira Tabas, *Columbia U.*
Sarah Teichmann, *Wellcome Sanger Inst.*
Rocio Titiunik, *Princeton U.*
Shubha Tole, *Tata Inst. of Fundamental Res.*
Wim van der Putten, *Netherlands Inst. of Ecology*
Reinhilde Veugelers, *KU Leuven*
Bert Vogelstein, *Johns Hopkins U.*
Kathleen Vohs, *U. of Minnesota*
David Wallach, *Weizmann Inst. of Sci.*
Jane-Ling Wang, *UC Davis* **(S)**
Jessica Ware, *Amer. Mus. of Natural Hist.*
David Waxman, *Fudan U.*
Jonathan Weissman, *UC San Francisco*
Chris Wikle, *U. of Missouri* **(S)**
Terrie Williams, *UC Santa Cruz*
Ian A. Wilson, *Scripps Res.* **(S)**
Yu Xie, *Princeton U.*
Jan Zaanen, *Leiden U.*

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.

novo nordisk fonden

International Research Leader Grants

Stage 1 application period for both calls
September 9, 2020 – October 21, 2020

Exceptional Grants for Exceptional Scientists
Two calls for applications within the areas of **biomedicine & biotechnology.** International Research Leader Grants from the Novo Nordisk Foundation are for outstanding scientists to establish and run their laboratories in Denmark.

Novo Nordisk Foundation
Laureate Research Grants

- Individual grants up to **DKK 50 million over 7 years (EUR ~6.7 million, USD ~7.4 million)**
- For principal investigators who have directed an independent research group for 7 or more years.
- Grant holders can apply for continued Laureate Research Grant funding, up to DKK 35 million over 7 additional years.

Novo Nordisk Foundation
Young Investigator Awards

- Individual grants up to **DKK 25 million over 7 years (EUR ~3.4 million, USD ~3.7 million)**
- For principal investigators who have directed an independent research group for less than 7 years.
- Award holders can apply for further funding from other Novo Nordisk Foundation grant programs.

To learn more, please visit: **novonordiskfoundation.com**

VERTEX®

THE SCIENCE *of* POSSIBILITY

At Vertex, science leads the way, but our people lead the science.

We believe that by bridging science, technology and business, we can transform our growing global biotechnology company in ways that will further our mission to create transformative medicines for people with serious diseases. We're hiring in ***Cell and Gene Therapy, Manufacturing, Clinical Development and Medical Affairs, Research, Quality*** and more.

Find open positions at
careers.vrtx.com

©2020 Vertex Pharmaceuticals Incorporated
Vertex and the Vertex triangle logo are registered trademarks of Vertex Pharmaceuticals Incorporated.

EDITORIAL

Fire in our future

It can seem like Earth itself is on fire. In places such as Australia and California for which fire is a natural feature, landscapes are burning at historic if not epic scales. In the Arctic and Greenland, where fire is rare, tundra is smoldering and melting permafrost. In Amazonia, Indonesia, and Mediterranean Europe, fires are interacting with the land clearing of rainforest, the draining of peatlands, and the abandonment of rural lands to create damaging, even lethal, conditions.

There is no single driver except humanity behind this outbreak. But increasingly, anthropogenic climate change is recognized as an enabler, performance enhancer, and globalizer. Fire seasons are lengthening, fire severity is escalating, and collateral damages are compounding.

Is this a "wicked" problem so entangled with scientific and social complexities that solutions are impossible? We think not.

We need to unbundle "fire" in all its shape-shifting avatars into manageable pieces. Some issues will have technical solutions—fires sparked by powerlines can be prevented. Some involve knotty ecological processes: Lands that have had fires removed can suffer an ecological fire deficit for which reinstating flame can be as complicated as restoring a vanished species. Most of the problems involve clashes of cultural values over how we get energy, organize our economy, and choose to live on the land. These will demand a political resolution.

Scales matter. Some reforms can be applied immediately and locally, as with protecting towns. Others will require decades of work across countries and regions. Restoring a suitable regimen of fire to tens of millions of hectares will be an arduous exercise in adaptive management. Confronting the effects of climate change will likely prove a century-long quest, but unless we reverse trends, they will overwhelm whatever type of management is implemented. We need to pursue all levels simultaneously.

Begin with ignition. Research shows that nationally, 97% of the fires that have threatened houses are started by people. There will always be accidental ignitions, and in the West and Florida, lightning kindles many fires. But prevention programs can reduce the risk to manageable levels.

Still, fires will escape. The power of fire, however, resides in its capacity to spread and inflict damages. Within limits, we can dampen fire intensities by modifying the landscapes that fire feeds upon, and we can harden communities to keep embers blown from the countryside from metastasizing into urban conflagrations. The strategies are the same as those used to contain urban fire. Concepts like the home ignition zone—the house and its immediate surroundings—identify points of vulnerability. Long-extant programs like Firewise, which also add concepts like defensible space, promote suites of tried-and-tested techniques to communities in nearly all kinds of environments.

In montane forests like the ponderosa pine of the Southwest, research shows that thinning and burning are effective methods to reduce fuel loads and allow surface fires to return. But many techniques are available, including prescribed grazing, the use of managed wildfire, and varieties of mechanical treatments like chipping and masticating. Most places will need a cocktail of treatments, appropriate to their local conditions.

> "There is no single driver except humanity..."

Smart treatments, done well, will enhance ecological integrity at the same time that they reduce hazardous fuels. Thinning, for example, resembles woody weeding and unlike logging removes the small stuff that powers fire. Moreover, fire is a biochemical process, not just a flaming woodchipper. Fire as fire matters biologically. Good fire can provide herd immunity against bad fire.

Yet all these interventions will be overpowered unless climate change is brought to heel. Paradoxically, as we ratchet down our binge-burning of fossil fuels, we'll have to ratchet up our burning of living landscapes to grant them the robustness they will need to survive the stresses to come.

Science can't do all the intellectual lifting. Fire is systemic: We need a systemic cultural response. We need art, new narratives and a poetry of flame, a revamping of liability laws to make controlled burning a default choice, a restoration of traditional knowledge to broaden techniques and purposes, a politics that can see the flames behind the smoke and engage with those who must live with its choices. In the end, science can advise; it can't decide.

But we need a solid empirical basis for the tough decisions heading our way. We need what science can do best, and the best of what science can do.

–William Wallace Covington and Stephen Pyne

William Wallace Covington is Emeritus Founding Executive Director of The Ecological Restoration Institute and Emeritus Professor of Forestry at Northern Arizona University, Flagstaff AZ, USA. w.wallace.covington@nau.edu

Stephen Pyne is an emeritus professor at Arizona State University, Tempe, AZ, USA, and a fire historian. stephen.pyne@asu.edu

10.1126/science.abe9780

NEWS

The global toll of COVID-19 this week passed 1 million deaths. A mourner cries near where a relative was laid to rest at the Cementerio General in La Paz, Bolivia. This summer, the cemetery saw as many as 90 burials per day and built three new pavilions and two mass graves.

PHOTO: GASTON BRITO/GETTY IMAGES

IN BRIEF

Edited by **Jeffrey Brainard**

U.S. is far from herd immunity

COVID-19 | Fewer than one in 10 Americans carried antibodies to the pandemic coronavirus in late July, according to a nationwide study of dialysis patients. One of the largest of its kind, the study indicates that the United States is a long way from reaching "herd immunity" to COVID-19, when a large portion of the population would become resistant to the virus that causes the disease, slowing the rate of new infections. Researchers reported last week in *The Lancet* that they found the virus in 8% of samples from leftover plasma of 28,503 people. Residents of neighborhoods that are predominantly Black or Hispanic or are densely populated were significantly more likely to have antibodies to the virus. The researchers also found regional differences: Thirty-three percent of New York state samples had antibodies, but only 3.8% of California samples did.

Trump wants shorter student visas

IMMIGRATION | The Trump administration last week proposed tightening visa policies for international students, a step it says is necessary to monitor their academic progress and protect national security. But many university officials believe the new rules, if adopted, would reinforce the message that foreign students aren't welcome. One change would impose a 4-year limit on their stay; a visa is now good for the duration of a student's academic program. And the limit would only be 2 years for students from 59 countries where more than 10% of visa holders violate the terms of their visit. The 4-year limit would apply to students from China and India, which combined are home to more than half the total number of international students in the United States. Any student wanting to remain to complete a degree would need to have "a compelling academic reason" or face "circumstances beyond their control." Even if finalized, however, the changes are unlikely to take effect if President Donald Trump loses the November election.

Genome gaps nearly filled

GENETICS | Despite the fanfare that greeted the first human genome sequence 17 years ago, researchers knew it was incomplete. For the sake of speed and because of limitations in technology, teams had skipped over regions of nearly identical DNA, which are difficult to map correctly. Last week, using newer technologies, a team of 70 researchers released a human cell line's genome sequence with almost all those difficult parts deciphered. It's a first attempt to sequence each chromosome from telomere to telomere, meaning from one end to the other. The work fills in about 135 million missing base pairs, bringing the total to 3.057 billion bases, the Telomere-to-Telomere consortium announced on Twitter and in the Genome Informatics Section of GitHub. Still missing are arrays of genes

coding for ribosomal RNA, an essential part of the cell's protein-building machinery. The consortium hopes to finish mapping remaining stretches within 1 year.

France to pair science, media

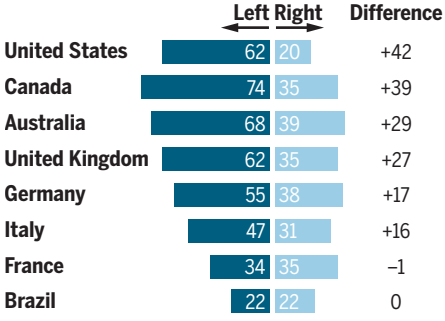
POLICY | France plans to bring scientists and journalists closer together in an effort to boost public access to reliable scientific information and combat misinformation. Parliament is expected to approve the project in coming weeks as part of a 10-year science plan. The science ministry said the plan is needed "at a time when French society is crossed by currents of irrationality and doubts about progress and knowledge." An early draft of the bill called for a science media center similar to those in the United Kingdom and Germany. But after critics said funders of a French center could distort its agenda and threaten journalistic independence, the bill was reworded. It now calls for a "series of actions or network."

HHMI mandates open access

PUBLISHING | The Howard Hughes Medical Institute (HHMI), one of the largest research philanthropies, said on 1 October it will begin to require its staff scientists to make research papers in which they played a leading role immediately free to read. The new policy will replace the current requirement that HHMI papers become open access within 12 months of publication. HHMI-supported scientists can comply with the rule, which takes effect in January 2022, by publishing in journals with open-access options or depositing a near-final, peer-reviewed version in a free archive. They can also use HHMI funds to publish in hybrid journals, which allow authors to pay for open-access publication but also charge subscriptions, but only if the journal is transitioning to immediate open access for all content. (For nonprofit publishers' journals, this provision takes effect in 2023.) HHMI also said this week it will join Coalition S, which includes funders mostly in Europe, in its campaign to flip journals to immediate open access. The new policy, which applies to papers on which an HHMI scientist is the first, last, or corresponding author, may divert papers from top-tier journals that lack open-access options, acknowledges biochemist Erin O'Shea, HHMI's president. In 2019, 13% of all papers in *Cell*, 5% in *Nature*, and 7% in *Science* would have been subject to the new requirement, the institute said. O'Shea says the open-access model helps accelerate science, so she hopes those subscription journals and others will find ways to help HHMI authors comply.

A partisan divide

Share of survey respondents, by political orientation, who trust scientists "a lot"



▲ U.S. leads in science trust gulf

PUBLIC OPINION | When it comes to trusting scientists, Americans are the most split ideologically, a survey of 20 countries has found. Some 62% of Americans who identified their politics as left wing said they held "a lot" of trust in scientists "to do what is right" for the public, the Pew Research Center reported on 29 September. But only 20% of those who identified with the right wing expressed that level of confidence. The polarization was smaller in other countries. Still, respondents in all countries had high regard for scientists, with 36% expressing "a lot of trust" in them, matching the military for most trusted and besting perceptions of government and business leaders and the news media.

U.S. census dispute is extended

DEMOGRAPHY | The political battle over when to stop tracking down Americans who haven't completed the 2020 U.S. census continued this week. After a court ruling last week overturned the Trump administration's deadline of 30 September, the Census Bureau said on 28 September it would finish the fieldwork by 5 October. Critics have worried ending the work too soon could cause a large undercount. In April, the Census Bureau said the COVID-19 pandemic had forced delays in deploying some 500,000 door-to-door enumerators. But in early August, the White House said the fieldwork must wrap up by the end of September. City and state officials and civil rights groups sued to block that plan, leading to last week's ruling.

Trade in Sri Lanka's critically endangered Knuckles pygmy lizard was banned last year.

CONSERVATION

Unregulated reptile trade threatens species

More than one-third of reptile species, or nearly 4000, have been offered for sale online in recent years, researchers at the Xishuangbanna Tropical Botanical Garden in China found when they searched 151 websites for ads for pet reptiles placed between 2000 and 2019. Some species for sale are endangered, reports the study, published on 29 September in *Nature Communications*. More than half of reptiles imported to the United States are captured from the wild, and the trade of most reptiles is not regulated, the study notes, even though researchers don't know enough about some 1500 species to tell whether they are endangered. The Convention on International Trade in Endangered Species of Wild Fauna and Flora (CITES) monitors the trade of just 9% of reptile species, typically those sold in larger volumes, such as crocodiles and pythons, whose skin is used in shoes, wallets, and other items. In all, CITES regulates trade of 856 reptile species and has banned sales for 98. The International Union for Conservation of Nature lists 1406 as species of concern.

CREDITS: (GRAPHIC) C. SMITH/*SCIENCE*; (DATA) PEW RESEARCH CENTER; (PHOTO) FRANK CANON

IN DEPTH

At 1-kilometer resolution, a European climate model (left) is nearly indistinguishable from reality (right).

IMAGES: (LEFT TO RIGHT) ECMWF; © EUMETSAT

CLIMATE CHANGE

Europe builds 'digital twin' of Earth to hone climate forecasts

Ingesting more data than ever before, exascale model will simulate the impact of climate change on humans

By **Paul Voosen**

The European Union is finalizing plans for an ambitious "digital twin" of planet Earth that would simulate the atmosphere, ocean, ice, and land with unrivaled precision, providing forecasts of floods, droughts, and fires from days to years in advance. Destination Earth, as the effort is called, won't stop there: It will also attempt to capture human behavior, enabling leaders to see the impacts of weather events and climate change on society and gauge the effects of different climate policies.

"It's a really bold mission, I like it a lot," says Ruby Leung, a climate scientist at the U.S. Department of Energy's (DOE's) Pacific Northwest National Laboratory. By rendering the planet's atmosphere in boxes only 1 kilometer across, a scale many times finer than existing climate models, Destination Earth can base its forecasts on far more detailed real-time data than ever before. The project, which will be described in detail in two workshops later this month, will start next year and run on one of the three supercomputers that Europe will deploy in Finland, Italy, and Spain.

Destination Earth rose out of the ashes of Extreme Earth, a proposal led by the European Centre for Medium-Range Weather Forecasts (ECMWF) for a billion-euro flagship research program. The European Union ultimately canceled the flagship program, but retained interest in the idea. Fears that Europe was falling behind China, Japan, and the United States in supercomputing led to the European High-Performance Computing Joint Undertaking, an €8 billion investment to lay the groundwork for eventual "exascale" machines capable of 1 billion billion calculations per second. The dormant Extreme Earth proposal offered a perfect use for such capacity. "This blows a soul into your digital infrastructure," says Peter Bauer, ECMWF's deputy director of research, who coordinated Extreme Earth and has been advising the European Union on the new program.

Typical climate models run at resolutions of 50 or 100 kilometers; even top ones like ECMWF's "European" model run at 9 kilometers. The new model's 1-kilometer resolution will enable it to directly render convection, the vertical transport of heat critical to the formation of clouds and storms, rather than relying on an algorithmic approximation. "I call it the third dimension of climate modeling," says Bjorn Stevens, a climate scientist at the Max Planck Institute for Meteorology. The model will also simulate the ocean in fine enough detail to capture the behavior of swirling eddies that are important movers of heat and carbon.

In Japan, pioneering runs of a 1-kilometer global climate model have shown that directly simulating storms and eddies leads to better short-term rainfall predictions. But it should also improve climate forecasts over periods of months and years. Recent work has shown climate models are not capturing predictable changes in wind patterns that drive swings in regional temperature and rainfall—probably because the models fail to reproduce storms and eddies (*Science*, 31 July, p. 490).

The high resolution will also enable Destination Earth to base its forecasts on more detailed data. Weather models suck in observations of temperature and pressure from satellites, weather stations, aircraft, and buoys to guide their simulations. But coarse grids mean the models can't assimilate measurements that don't average well or cover broad areas, such as fractures opening up in sea ice. Destination Earth will close this gap, says Sandrine Bony, a cloud scientist at the Pierre Simon Laplace Institute. "The scales that are resolved are closer to the scales that are measured."

The model will also incorporate real-time data charting atmospheric pollution, crop growth, forest fires, and other phenomena known to affect weather and climate, says Francisco Doblas-Reyes, an earth system scientist at the Barcelona Supercomputing Center. "If a volcano goes off tomorrow, that's important for the risk of tropical precipitation failure in a few months." And it will fold in data about society, such as energy use, traffic patterns, and human movements (traced by mobile phones).

The goal is to allow policymakers to directly gauge how climate change will impact society—and how society could alter the trajectory of climate change. For example, the model could predict how climate change

will affect agriculture and migration patterns in Brazil—and also how cuts in ethanol subsidies might limit deforestation in the Amazon. Currently, climate scientists extract regional results from global climate models and pass them to experts in agriculture or economics to understand effects on human behavior. Now, says Erin Coughlan de Perez, a climate hazard scientist at the Red Cross Red Crescent Climate Centre, modelers are "moving from just forecasting what weather will be, to what the weather will do."

Getting there won't be easy. Exascale supercomputers rely on both traditional computer chips as well as graphical processing units (GPUs), which are efficient at handling intensive calculations. GPUs are good for running model components in parallel and training artificial intelligence algorithms—two techniques Destination Earth will lean on to enhance performance. But old climate modeling code will have to be reworked. ECMWF has a head start: It is adapting its forecast model to a GPU-based environment, and last year tested it at 1-kilometer resolution for four simulated months on Summit, the U.S. supercomputer that was the world's fastest until a Japanese machine recently eclipsed it.

The massive amount of data generated by the model will be a problem of its own. When the Japanese team ran its 1-kilometer-scale experiment, it took half a year to extract something useful from a couple days of data, Doblas-Reyes says. "There's a bottleneck when we try to access the data and do something clever with it." A big part of Destination Earth will be solving this problem, designing ways to analyze model results in real time.

As an operational system, Destination Earth will likely run at several time scales, Bauer says. One will be near daily, perhaps targeting individual extreme weather events weeks or months in the future. Runs in the other mode—long-term predictions—would be less frequent: perhaps a decadelong prediction of the climate made every half-year or so. "If this works, it could be a template for other countries to follow," Bauer says.

The Europeans aren't alone in planning for exascale climate models. "We're heading in that direction as well, but we've yet to reach that level of effort," says Leung, who serves as chief scientist for DOE's earth system model.

Stevens says it's thrilling to be involved in a truly planetary-scale information system that can reveal not just the proverbial butterfly effect in weather and climate, but also how local human actions manifest globally. "That's the story of globalization. That's the story of the Anthropocene. And this is the scientific platform that will allow you to explore those." ■

CLIMATE CHANGE

China's bold climate pledge earns praise—but is it feasible?

Plan to achieve carbon neutrality by 2060 would make China a global leader, but abandoning coal will be hard

QQ群: 1074370165

By **Dennis Normile**

China's surprise pledge last week to cut its net carbon emissions to zero within 40 years has reignited hopes of limiting global climate change to tolerable levels. The country is the world's largest producer of carbon dioxide (CO_2), accounting for 28% of global emissions, and its move may inspire other countries to follow suit. But observers warn that China faces daunting challenges in reaching its goals. Kicking its coal habit will be particularly hard.

"We aim to have CO_2 emissions peak before 2030 and achieve carbon neutrality before 2060," Chinese President Xi Jinping told the United Nations General Assembly via a video link on 22 September. That's "a very significant and encouraging announcement," says Josep Canadell, an earth system scientist at Australia's Commonwealth Scientific and Industrial Research Organisation. He says the new targets "won't likely let us to stop at 1.5° Celsius [of global warming]," the preferred target set in the 2015 Paris agreement. "But below 2° might still be consistent with [Xi's] announcement." China's commitment also "ratchets up pressure on other major emitters" to set more ambitious targets "while further isolating the Trump administration in its climate myopia," Vance Wagner of Energy Foundation China wrote in a piece published online by the nonprofit China Dialogue.

China had previously said its CO_2 emissions would peak "around" 2030, a target most analysts considered within reach. But achieving carbon neutrality before 2060 will require drastically reducing the use of fossil fuels in transportation and electricity generation and offsetting any remaining emissions through carbon capture and storage or planting forests.

China has not yet revealed details of how it will do this. But a research group at Tsinghua University presented a $15 trillion, 30-year road map on 27 September that calls for ending the use of coal for electricity generation around 2050, dramatically increasing nuclear and renewable power generation, and relying on electricity for 80% of China's energy consumption by 2060.

Coal is both the biggest challenge and an opportunity. Last year, the carbon-heavy fuel accounted for about 58% of China's total energy consumption and 66% of its electricity generation. In coal-producing regions, coal is also used to heat buildings. Recent advances in renewable energy have made replacing coal easier than cutting oil use in transportation and emissions from farm fields and livestock. "The power sec-

PHOTO: XU CONGJUN/AP PHOTO

A coal-fired power plant in Jiangsu province. Coal accounted for 58% of China's energy consumption last year.

tor is the part of the energy system where zero emission technologies are the most mature and economically competitive," says Lauri Myllyvirta, an air pollution analyst at the Centre for Research on Energy and Clean Air in Helsinki. Zero-carbon electricity could make charging electric vehicles cleaner and supplant coal for heating.

But it will require a U-turn. A recent study by Myllyvirta and colleagues found that China's coal-fired generating capacity grew by about 40 gigawatts (GW) in 2019, to about 1050 GW. Another 100 GW is under construction and coal interests are lobbying for even more plants. "This is all despite significant overcapacity in the sector," with plants running at less than 50% of capacity and many coal-power companies losing money, the study said. Canadell says the building boom is the result of misplaced incentives to build coal plants and create construction jobs. He predicts many of the new plants will barely be used or become stranded assets that have to be written off.

A related challenge will be reforming the electricity market. Renewable energy is increasingly cost competitive with coal, says Li Shuo, a climate policy adviser to Greenpeace China. But regulators allocate operational time among electricity plants to match generation to demand, with little consideration of economic or environmental implications, Li says. The system overwhelmingly favors coal-fired generation, partly because it doesn't suffer from the variability of wind and solar power. The uncertain market access has already slowed investment in renewables, Li says. Given the power of coal and construction interests, the needed reforms will take considerable political will.

Expanding nuclear power presents challenges as well. The 2011 Fukushima Daiichi nuclear disaster in Japan sent ripples of concern through China, which mandated additional safety measures that made new plants more expensive. Public opposition is also growing. China has 48 nuclear power reactors in operation and 12 under construction, according to the World Nuclear Association. The government had aimed for 58 GW of nuclear capacity by this year but did not get beyond 52 GW.

China's Five-Year Plan for 2021–25, now being drafted, may contain concrete measures to help realize Xi's ambitious target. "China's interest in climate change has waned in recent years, due to the slowing down of economic growth and the U.S. withdrawal from the Paris agreement," says Zhang Junjie, an environmental economist at Duke Kunshan University. "The commitment on carbon neutrality reignited hopes for China's climate action." ■

Many koalas were killed by Australia's record wildfires, jeopardizing the survival of some populations.

PHOTO: PETER PARKS/AFP/GETTY IMAGES

CONSERVATION BIOLOGY

Record U.S. and Australian fires raise fears for many species

Scientists say fires likely wiped out some rare Australian organisms, and worry U.S. blazes now threaten more

By **John Pickrell** *and* **Elizabeth Pennisi**

For the past 3 months, arachnologist Jess Marsh has been searching for the Kangaroo Island assassin spider. Early this year, during the worst fire season ever recorded in Australia, a wildfire charred the spider's only known home on an island off the nation's south coast. Now, Marsh fears the tiny, rusty brown arachnid is another of the many Australian species that the blazes have put on a path to extinction: Countless hours of scouting haven't revealed a single survivor. "Its habitat is completely incinerated," says Marsh, who is affiliated with Charles Darwin University.

She isn't the only field biologist worried that the record wildfires around the globe are inflicting lasting damage on species and ecosystems. Even as Australia tallies the damage from its blazes, the worst fires in more than 70 years are burning in California, Oregon, and Washington; so far, they have consumed some 2 million hectares, killing at least 35 people. As in Australia, scientists fear the loss of habitat has threatened species with small populations or restricted ranges, and could potentially lead to permanent ecological changes if burned landscapes fail to rebound in a warming climate. "We are in uncharted territory here," says ecologist S. Mažeika Patricio Sullivan of Ohio State University, Columbus. "We just don't know how resilient species and ecosystems will be to wildfires of the magnitude, frequency, and intensity that we are currently experiencing in the U.S. West."

Australia's postfire experience offers cause for anxiety, researchers say. From September 2019 to March, more than 11 million hectares burned, mostly in the continent's southeastern forests, killing at least 34 people. More than 20% of the nation's total forest cover was lost, researchers at Western Sydney University reported in February. Even normally fire-proof rainforests and wetlands were scorched (*Science*, 20 December 2019, p. 1427). By one estimate, released early this year by the Australian government, 114 threatened plant and animal species lost 50% to 80% of their habitats; 327 species saw more than 10% of their ranges burn.

Those estimates, however, were based on satellite data, says John Woinarski, also at Charles Darwin University. To get better assessments, researchers have been trying to visit burned sites, an effort complicated by the COVID-19 pandemic.

In some cases, they've reported good news. There was grave concern for the en-

dangered Kangaroo Island glossy black cockatoo after 75% of its habitat burned. But observers have seen large flocks move to unburned areas of the island, says Karleah Berris of Natural Resources Kangaroo Island. And many of the birds appear to have bred and fledged young. "It seems they are coping with the reduction in food by [moving] to where the food is," she says. Researchers were also worried about the Kangaroo Island dunnart, a shrew-size carnivorous marsupial. Even before the fires, just 500 or so remained, and they lost 95% of their habitat to the flames. But automated cameras have revealed that at least some dunnarts survived, and managers moved quickly to build fences to protect the remaining animals from feral cats.

Other findings are more ominous. In New South Wales, fires killed about one-third of the state's koalas, a government inquiry found in July. It warned that the marsupial would be extinct in the state by 2050 if dramatic measures are not taken to conserve it. And in the state's Nightcap National Park, a survey found that fires destroyed 10% or more of the remaining stands of several critically endangered rainforest trees. Some species were down to fewer than 200 trees before the fires, says botanist Robert Kooyman of Macquarie University; they are now "certainly a few steps closer to extinction."

Such concerns have prompted scientists to ask Australia's government to expand its endangered species list. At least 41 vertebrates that were not endangered before the fires now face existential threats, Woinarski and others reported in July in *Nature Ecology & Evolution*. An additional 21, already tagged as threatened, might need greater protection. Marsh has recommended adding 16 invertebrate species found on Kangaroo Island to the list, including the assassin spider. "That species is really hanging in the balance," she says.

In the United States, researchers say it's too soon to know how many species the fires have put in jeopardy. But there are already worrying reports. In Washington, biologists estimate the fires have killed 50% of the state's endangered pygmy rabbits, which inhabit sagebrush flats that burned this year. They believe only about 50 of North America's smallest rabbit remain. Officials estimate the flames have also killed 30% to 70% of the state's sage grouse and sharp-tailed grouse, birds that also depend on sagebrush.

In California, the impact of fires in 2014 may offer a preview. After flames swept through habitat of the endangered spotted owl, many of the birds abandoned nesting sites, biologists Gavin Jones of the U.S. Forest Service's Rocky Mountain Research Station and M. Zachariah Peery from the University of Wisconsin (UW), Madison, found. In 2015, some 22% of nesting sites used by the birds in 2014 were not reoccupied and still are empty, Jones says, and this year's fires could add to the losses. Western fires also threaten the white-headed woodpecker, found only in pine forests in the Pacific Northwest and California, and the Grace's warbler, limited to pine and oak forests in the southwestern United States and northern Mexico, says wildlife biologist Vicki Saab, also at the Rocky Mountain Research Station.

Plants that have small ranges and are found in burned areas, such as the Coulter pine in California, might also face trouble, says Camille Stevens-Rumann, a fire ecologist at Colorado State University, Fort Collins. "California especially has a lot of endemic plant species that could be very much impacted," she says.

The longer term consequences for ecosystems are harder to predict, researchers say. In both Australia and the Western United States, many ecosystems are adapted to fire and even require it to thrive. "Many of the old-growth forests we know and love in the Pacific Northwest were born of large and severe fires centuries ago," says Brian Harvey, a wildfire ecologist at the University of Washington, Seattle. Fires can also help create a mosaic of habitats that support a wealth of species, he and others note.

But climate change adds to the uncertainty about how forests will respond this time. "The postfire climate is likely to be warmer and drier than when the parent trees established long ago," Harvey says, making it harder for ecosystems to recover, and boding more fire in the future. "Just a little more drought can lead to much bigger fires," says Monica Turner, a fire ecologist at UW who calls climate change "a threat multiplier."

Already, some ecosystems in North America that have had frequent or intense burns are not regenerating. In some places, such as the sagebrush ecosystem of the Great Basin west of the Sierra Nevada mountain range and forests in the Klamath Mountains along the California-Oregon border, invasive shrubs or grasses appear to have taken over. Because the invaders burn frequently, they appear to be preventing seedlings from maturing. In Australia, researchers have similar concerns. In the state of Victoria, forests of alpine ash, a towering eucalyptus tree found in moist regions, historically experienced fires less than once a century or so. Now, some forests have been hit by five fires in the past 20 years, and scientists fear some of the stately groves will disappear for good. ■

John Pickrell is a journalist in Sydney.

IMAGE: ANDY WALKER/MIDLAND FERTILITY SERVICES/SCIENCE SOURCE

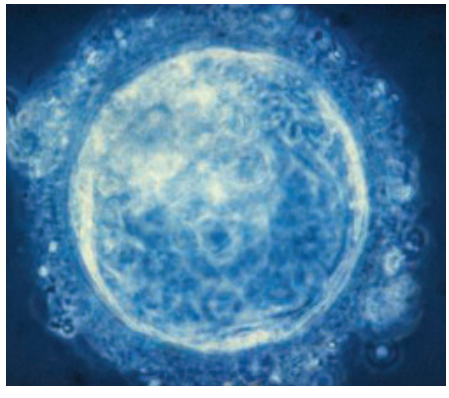
DEVELOPMENTAL BIOLOGY

Stem cell studies probe origins of the placenta

Lab models of organ will track how it emerges—and what can go wrong

By **Kelly Servick**

The placenta—a Frisbee-size hunk of tissue that chaperones a fetus in the uterus only to be tossed aside in the delivery room—has mysterious beginnings. The organ emerges from cells that develop alongside the embryo, and that have been difficult to grow in the lab. Now, researchers have devised a way to derive and observe early precursors of placental cells in a dish. They have found a method of "reprogramming" adult cells, reverting them



The outer layer of cells (thin ring above) in a several-days-old human embryo goes on to form the placenta.

to a primitive state, that can prompt them to become trophoblast stem cells (TSCs), which give rise to placental cells.

The method promises a window on how defects in placental development may lead to infertility, miscarriage, and preeclampsia, a dangerous complication of pregnancy. "It's like gaining a toehold on Mars," says reproductive biologist Susan Fisher at the University of California, San Francisco. "We know almost nothing about the early steps."

Those steps begin just days after a sperm and egg join. "The first decision in human life is to set aside the placental, supportive

cells," says Kathy Niakan, a developmental biologist at the Francis Crick Institute, whose team reported key molecular signals for that initial step in *Nature* last week. These cells go on to form the trophoblast, a multilayered ring that surrounds the embryo and helps it implant into the wall of the uterus. Some of these cells, TSCs, then give rise to cell types that will make up the bulk of the placenta, which enables mother and fetus to exchange nutrients and gases and helps protect the fetus from the mother's immune system.

Scientists have derived TSC-like cells from unused embryos created for in vitro fertilization (IVF) or from the placentas of terminated pregnancies, but both are limited resources. And in a dish, these cells have tended to mature and stop dividing. The same has been true of TSC-like cells created from cultured embryonic stem (ES) cells and from induced pluripotent stem (iPS) cells—mature cells reprogrammed to an ES-like state.

> **"It's like gaining a toehold on Mars."**
>
> **Susan Fisher**, University of California, San Francisco

But in 2017, Tohoku University stem cell biologist Takahiro Arima and colleagues described a broth of nutrients and other compounds that could make TSCs from IVF embryos or first trimester placentas thrive in a dish. "An enormous amount of work that was never possible before became possible," says William Pastor, a stem cell biologist at McGill University. This year, Pastor's group and two others showed this culture medium could also coax certain types of ES cells to become self-renewing TSCs.

To make TSCs that genetically match a patient, however, researchers want to be able to start from mature skin or blood cells. In the two new studies, teams led by stem cell biologists Jose Polo at Monash University and Laurent David at the University of Nantes found ways to convert adult skin cells into "induced" TSCs. Both teams had been studying how gene expression changes as mature cells are reprogrammed into iPS cells. They noticed that along the way, some expressed genetic signatures of so-called trophectoderm cells, which give rise to the trophoblast.

"That was very weird," Polo says, because a cell's decision to become trophectoderm happens so early in development—not anywhere along the expected path backward from skin cell to iPS cell. But by culturing the cells in the newly available medium, the researchers managed to push them to become TSCs.

In a 16 September *Nature* paper, Polo's team reported that these induced TSCs could develop into two major types of trophoblast cells and, like the cells surrounding an embryo, secrete human chorionic gonadotropin, a hormone whose signals are key to maintaining a pregnancy. David, a co-author on that paper, separately used gene expression data from human embryos to estimate that his own group's lab-derived TSCs are equivalent to those seen 8 to 10 days after fertilization, the team reported on 15 September in a preprint on bioRxiv.

It will be important to thoroughly compare these induced TSCs to placenta-derived and ES cell–derived TSCs, says Washington University in St. Louis stem cell biologist Thorold Theunissen, whose team recently derived TSCs from ES cells. That analysis should include comparing the chemical tags on DNA that influence cell function and sizing up how efficiently the cells differentiate into different types of specialized trophoblast cells.

Induced TSCs could now be used to study genetic defects that can end a pregnancy, says Soumen Paul, a stem cell biologist at the University of Kansas Medical Center. By making TSCs from cells from women with infertility and watching them develop in the lab, researchers could pinpoint how abnormal trophoblast cells prevent the embryo from implanting in the uterus or from developing normally once implanted.

Or TSCs could help root out causes of preeclampsia, in which a pregnant woman suddenly develops high blood pressure that sometimes can be relieved only by inducing an early delivery. Preeclampsia is thought to stem from a defect of the placenta, perhaps in the way it invades the uterine wall and interacts with the mother's blood vessels, Pastor says. Researchers should now be able to make TSCs from umbilical cord blood or from a baby's blood or skin cells to observe how placental precursor cells emerge and interact with uterine cells.

The new TSCs could also add realism to synthetic embryo models—stem cell–derived structures that mimic early human development in a lab dish. So far, they haven't included trophoblast or other such "extra-embryonic" cells, says Jianping Fu, a bioengineer developing such models at the University of Michigan, Ann Arbor. But signals from these cells are critical to normal embryo growth, he says. Adding them would take the models "to the next level."

Better approximations of real embryos will raise ethical concerns. The U.S. National Institutes of Health has not released formal guidelines, but Fu says the agency discouraged him from including trophoblast tissue in a recent grant application. But he thinks such experiments should proceed. "When you mix the cells together, allowing them to self-organize … they will do amazing things." ■

VOICES OF THE PANDEMIC

The 'bat man' tackles COVID-19

After a career investigating why so many viruses come from bats, Linfa Wang eyes a new challenge

By **Kai Kupferschmidt**

By pure chance, Linfa Wang, one of the world's foremost experts on emerging viruses, was in the Chinese city of Wuhan in January. The biologist was visiting collaborators at the Wuhan Institute of Virology (WIV) just as SARS-CoV-2 was starting to spread from the city to the rest of the world. Even among those experts there was little fear then. "I was mixing with all the lab people," Wang says. "We would go to a restaurant every night."

Only when he left on 18 January did he realize how serious the situation was. At the airport, staff checked his temperature three times before he could board his flight home to Singapore. Five days later, Wuhan, a city of 11 million people, was shut down. Wang later learned that a woman on his plane had carried the virus; luckily, he was not infected.

Wang, who heads the Emerging Infectious Diseases Program at Duke-NUS Medical School in Singapore, immediately got to work developing a new assay that can detect antibodies against SARS-CoV-2 in blood samples—an indication of prior infection. The tool could help untangle how the pandemic began. So far, the evidence is that the virus originated in bats, animals Wang has long argued are uniquely suited to harboring viruses that pose a danger to humans. Now, he hopes his assay can help trace the path of the virus to humans and pinpoint when and where it first spilled over.

The work is a natural next chapter for Wang, who has been tracking viruses from bats to humans for more than 2 decades. Marion Koopmans, a virologist at Erasmus Medical Center, credits him for essentially launching the field of bat immunology and developing the tools to pursue it. "He has made a heroic effort to establish a very chal-

Science's COVID-19 reporting is supported by the Pulitzer Center and the Heising-Simons Foundation.

lenging research line, which needed to start from scratch," she says.

As a child growing up in Shanghai during the Cultural Revolution, Wang would listen to Mao Zedong's speeches through a loudspeaker in kindergarten. "I was thinking: 'My God … how does his voice transfer from Beijing to Shanghai?'" Electrical engineering became his passion. But after getting into the prestigious East China Normal University, Wang was dismayed when the faculty assigned him to study biology. "I thought, 'I don't like plants, I don't like animals,'" he says. Going to a renowned university felt like going to heaven, he says, "but the wrong door of heaven, basically, because I went to a biology department."

Secretly listening to Voice of America, Wang eventually became so proficient at English that he was chosen for a scholarship to study abroad. He did a Ph.D. in molecular biology at the University of California, Davis, and later moved to Australia, where he studied infectious diseases in animals. His career took a turn when a new virus emerged in the leafy Brisbane suburb of Hendra in 1994, killing 14 horses and a trainer. Wang managed to sequence the virus, later named Hendra virus, and helped develop a vaccine for horses. The virus turned out to be transmitted by bats. A few years later Wang worked on another novel virus, Nipah virus, also from bats. Intrigued, Wang scoured the literature and found numerous other viruses linked to bats.

Then came severe acute respiratory syndrome (SARS). After the World Health Organization (WHO) declared the epidemic over in July 2003, it put together a mission of eight scientists, including Wang, to investigate the origins of the virus in China. Wang had a hunch bats could be the source, but the rest of the team was skeptical. At a meeting in Beijing, Wang met the head of WIV, who suggested he collaborate with a scientist at her institute: Shi Zhengli, who was then studying viruses in fish and shrimp. "She was the only virologist who believed me and was willing to collaborate with me," Wang says.

The two have since co-authored dozens of papers, including one in *Science* in 2005 that pinpointed horseshoe bats as a reservoir of SARS-like coronaviruses. They also like to team up in karaoke bars to sing classic Chinese ballads, says Peter Daszak, a researcher at the EcoHealth Alliance, a New York City nonprofit, and a longtime collaborator with Wang and Shi. "Linfa is an excellent singer and to see him and Shi Zhengli do a duet is very special."

Now, Wang hopes to home in on the origin of SARS-CoV-2—an effort that will likely require screening thousands of animals and humans for signs of a prior infection. The gold standard for doing that is called a virus neutralization assay, which combines human cells and live virus with a blood sample to see whether the sample contains antibodies that keep the virus from binding to the cells. But using live virus means working in a high-level biosafety lab—expensive and very slow work. An alternative called an enzyme-linked immunosorbent assay (ELISA) is much easier to handle, but a distinct version must be developed for every animal species. "You need to have a whole panel of ELISAs that are optimized for different bat species, and raccoon dogs, and civet cats, and pangolins, and God knows what," says Malik Peiris of the University of Hong Kong. "It's a never-ending business."

"I am now fascinated with bats [but] I am still not an animal fan."

Linfa Wang, Duke-NUS Medical School

Wang's new assay, published in July in *Nature Biotechnology* and now produced by Genscript Biotech, replaces the human cells and live SARS-CoV-2 virus of the gold standard assay with human and viral proteins, eliminating the need for a high-security lab. The sample is tested on a plate impregnated with angiotensin-converting enzyme 2 (ACE2), the human receptor protein that SARS-CoV-2 attaches to when it invades cells. Researchers then add a solution containing the fragment of the viral spike protein that can bind to ACE2. If the binding takes place, an enzyme turns the solution blue and then yellow. But when a sample contains antibodies against SARS-CoV-2, they prevent the binding, blocking the colorful reaction. Wang's assay works on a variety of species almost as well as the gold standard, says Peiris, who has been using it for several weeks in infected cats, dogs, and hamsters.

"This is an extremely interesting approach," says Isabella Eckerle, a virologist at the Centre for Emerging Viral Diseases at the University of Geneva. Eckerle and colleagues validated the test for WHO and published the result as a preprint in late September. "Especially for screening potential plasma donors or when looking for the animal reservoir it should be really useful."

Wang hopes to use the test to screen animals and people in Southeast Asia to identify "intermediate hosts"—species that may have picked up the virus from bats and transmitted it to people—and learn whether it crossed over into humans before the fateful outbreak in Wuhan.

The bigger question that drives his work is: why bats? Over the past decade he has started to piece together an evolutionary story as convoluted as his own path to bats. As the only flying mammals, bats expand huge amounts of energy. This eventually damages their DNA, and Wang contends that they have adapted, in part, by dampening immune responses to DNA damage. RNA viruses like SARS-CoV-2 can cause similar damage, so the upshot is that bats tolerate low levels of viruses in a kind of peaceful coexistence. "That's why they are such a good reservoir," Wang says.

Koopmans is not yet convinced by Wang's immune system argument—bat ecology may play a greater role, she says. For instance, bats often range over wide territories, potentially picking up a greater variety of viruses than other animals, and in many bat species millions of animals roost together, making it easier for viruses to spread. But she says that thanks to Wang's work, there's no doubt that bats are key viral reservoirs.

It's an ironic legacy for a student who studied biology despite disliking animals. "I am now fascinated with bats," he concedes. But, perhaps appropriately given what he has learned about emerging infections, he says: "I am still not an animal fan in the sense of keeping animals near me." ■

ILLUSTRATION: KATTY HUERTAS

COVID-19

A call for diagnostic tests to report viral load

Measure could help officials know who is most contagious

By **Robert F. Service**

Ever since the coronavirus pandemic began, battles have raged over testing: Which tests should be given, to whom, and how often? Now, epidemiologists and public health experts are opening a new debate. They say testing centers should report not just whether a person is positive, but also a number known as the cycle threshold (CT) value, which indicates how much virus an infected person harbors.

Advocates point to new research indicating that CT values could help doctors flag patients at high risk for serious disease. Recent findings also suggest the numbers could help officials determine who is infectious and should therefore be isolated and have their contacts tracked down. CT value is an imperfect measure, advocates concede. But whether to add it to test results "is one of the most pressing questions out there," says Michael Mina, a physician and epidemiologist at Harvard University's T.H. Chan School of Public Health.

Standard tests identify SARS-CoV-2 infections by isolating and amplifying viral RNA using a procedure known as the polymerase chain reaction (PCR), which relies on multiple cycles of amplification to produce a detectable amount of RNA. The CT value is the number of cycles necessary to spot the virus; PCR machines stop running at that point. If a positive signal isn't seen after 37 to 40 cycles, the test is negative. But samples that turn out positive can start out with vastly different amounts of virus, for which the CT value provides an inverse measure. A test that registers a positive result after 12 rounds, for a CT value of 12, starts out with more than 10 million times as much viral genetic material as a sample with a CT value of 35.

But the same sample can give different CT values on different testing machines, and different swabs from the same person can give different results. "The CT value isn't an absolute scale," says Marta Gaglia, a virologist at Tufts University. That makes many clinicians wary, Mina says. "Clinicians are cautious by nature," Mina says. "They say, 'If we can't rely on it, it's not reliable.'" In an August letter in *Clinical Infectious Diseases*, members of the College of American Pathologists urged caution in interpreting CT values.

Nevertheless, Mina, Gaglia, and others argue that knowing whether CT values are high or low can be highly informative. "Even with all the imperfections, knowing the viral load can be extremely powerful," Mina says.

Early studies showed that patients in the first days of infection have CT values below 30, and often below 20, indicating a high level of virus; as the body clears the coronavirus, CT values rise gradually. More recent studies have shown that a higher viral load can profoundly impact a person's contagiousness and reflect the severity of disease.

Positive coronavirus tests could reveal a person's infectiousness, too.

In a study published this week in *Clinical Infectious Diseases*, researchers led by Bernard La Scola, an infectious diseases expert at IHU-Méditerranée Infection, examined 3790 positive samples with known CT values to see whether they harbored viable virus, indicating the patients were likely infectious. La Scola and his colleagues found that 70% of samples with CT values of 25 or below could be cultured, compared with less than 3% of the cases with CT values above 35. "It's fair to say that having a higher viral load is associated with being more infectious," says Monica Gandhi, an infectious diseases specialist at the University of California, San Francisco.

Conversely, people often test positive for weeks or even months after they recover but have high CT values, suggesting the PCR has identified genetic material from noninfectious viral debris. Current guidelines from the Centers for Disease Control and Prevention and World Health Organization, which call for patients to isolate themselves for 10 days after onset of symptoms, recognize they are not likely to be infectious after that period. But Mina and others say the recent findings also suggest that a patient who has undergone multiple tests with high CT values is likely at the tail end of their infection and need not isolate themselves. He adds that contact tracers should triage their efforts based on CT values. "If 100 files land on my desk [as a contact tracer], I will prioritize the highest viral loads first, because they are the most infectious," Mina says.

Broad access to CT values could also help epidemiologists track outbreaks, Mina says. If researchers see many low CT values, they could conclude an outbreak is expanding. But if nearly all CT values are high, an outbreak is likely waning. "We have to stop thinking of people as positive or negative, and ask how positive?" Mina says.

CT values could also help clinicians flag patients most at risk for severe disease and death. A report in June from researchers at Weill Cornell Medicine found that among 678 hospitalized patients, 35% of those with a CT value of 25 or less died, compared with 17.6% with a CT value of 25 to 30 and 6.2% with a CT value above 30. In August, researchers in Brazil found that among 875 patients, those with a CT value of 25 or below were more likely to have severe disease or die.

Gandhi agrees that having access to CT values could help clinicians identify people at high risk for developing symptoms. Nevertheless, she and others note that a high viral load doesn't necessarily lead to disease; some 40% of people who contract SARS-CoV-2 stay healthy even though they have a similar amount of virus to patients who fall ill. "As a physician, having the CT value is not the only thing I will use" to diagnose and track patients, says Chanu Rhee, a hospital epidemiologist at Brigham and Women's Hospital. "But I do still find it helpful." ■

PHOTO: MARTIN BERNETTI/AFP/GETTY IMAGES

eppendorf & Science

PRIZE FOR NEURO BIOLOGY

Meet the Winner

Eppendorf & Science Prize for Neurobiology
Congratulations to Christopher Zimmerman on winning the 2020 Eppendorf & Science Prize for his work on the neural circuits that govern thirst and drinking behavior. Dr. Zimmerman discovered that sensory signals originating throughout the body come together within individual neurons in the brain to produce the sense of thirst. He demonstrated that this new class of body-to-brain signals predicts changes in hydration before they occur and, as a result, adjusts our level of thirst preemptively. Dr. Zimmerman's research has revealed fundamental principles of ingestive behavior and provided neural mechanisms to explain aspects of everyday human experience.

The annual US$25,000 Eppendorf & Science Prize for Neurobiology honors scientists, like Dr. Zimmerman, for their outstanding contributions to neurobiological research. Christopher Zimmerman is the 19th recipient of this international award. Due to coronavirus (COVID-19), the 2020 winner and finalists will be honored at the Prize Ceremony in November 2021 in Chicago.

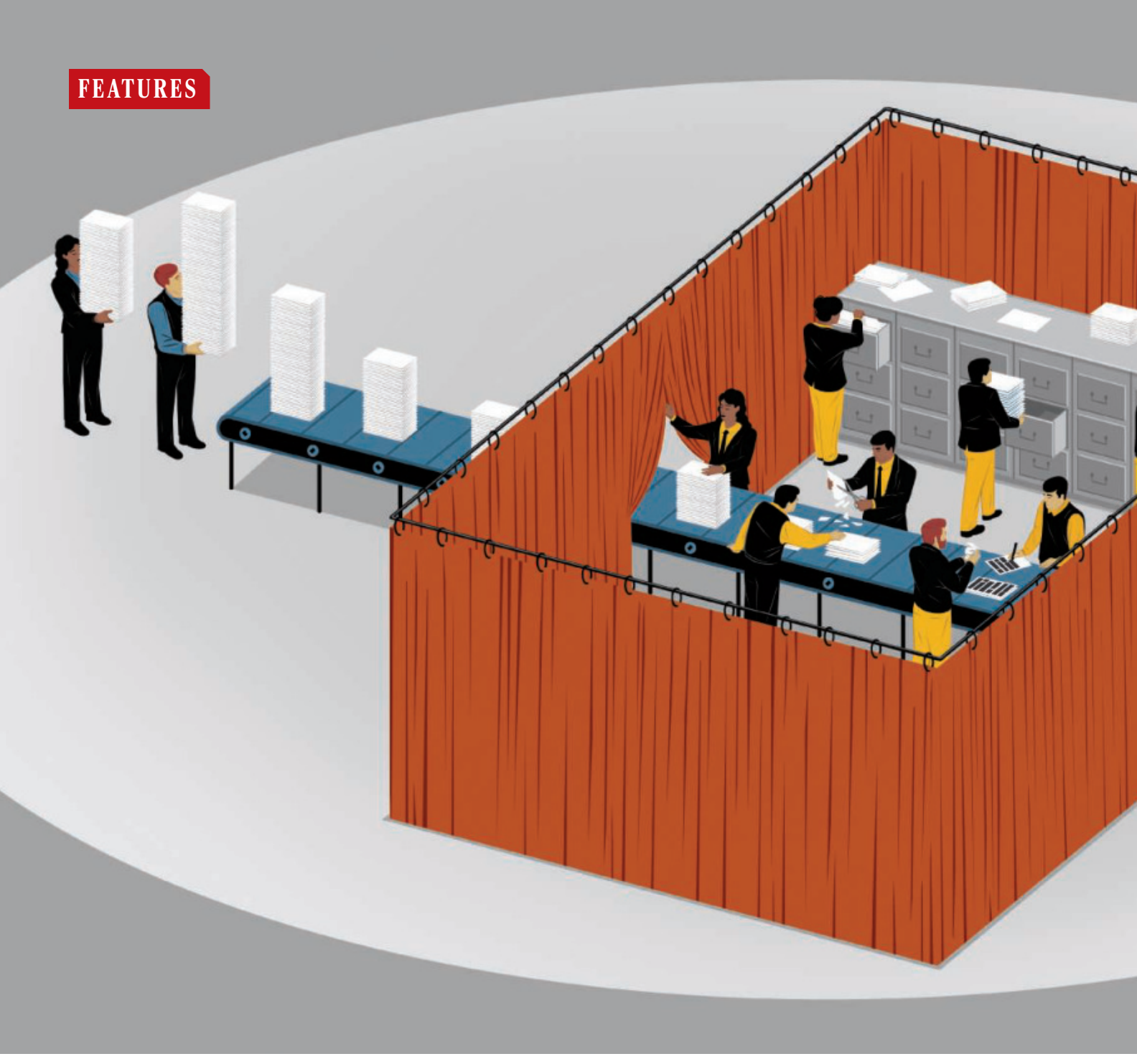
You could be next to win this prize.
If you are 35 years of age or younger and currently performing neurobiological research, you could be next to win the 2021 Prize.

Deadline for entries is June 15, 2021.

Learn more at: www.eppendorf.com/prize

eppendorf Science AAAS

AAAS® and Science® are registered trademarks of the American Association for the Advancement of Science, USA. Eppendorf® and the Eppendorf Brand Design are registered trademarks of Eppendorf AG Germany. All rights reserved, including graphics and images. Copyright © 2020 by Eppendorf AG. Photography: Sameer A. Khan.



OFFICIAL INACTION

A *Science* investigation shows that FDA oversight of clinical trials is lax, slow moving, and secretive–and that enforcement is declining

By **Charles Piller**

ILLUSTRATION: STEPHAN SCHMITZ

For nearly a decade, the Food and Drug Administration (FDA) cited osteopath Michael Harris for egregious errors in clinical trials he was overseeing. Agency inspectors found a litany of problems at Harris's private firm, Aspen Clinical Research in Orem, Utah, which had contracts to evaluate many drugs, including ones aiming to treat postoperative pain, pediatric schizophrenia, and migraines. FDA found there were serious lapses in obtaining informed consent from trial volunteers, unqualified staff made medical assessments, and Harris failed to properly report abnormal lab test results. He also did not disclose that trial participants were taking opioid, antidepressant, or antipsychotic drugs—which could have skewed results or posed safety concerns. The agency said Aspen's records were disorganized, contradictory, and sometimes backdated in a way that "begs the question of the authenticity and veracity of data collected."

Those "serious, ongoing deviations" might constitute "fraud, scientific misconduct," and "significant human subject protection violations," according to FDA documents obtained by *Science* through the Freedom of Information Act (FOIA). Inspectors told Harris he could be subject to fines, permanent disqualification from clinical research in the United States, and legal prosecution. Repeat problems and a raft of new ones emerged during inspections in 2014, 2015, and 2019. Each time, in responses to FDA, Harris admitted some transgressions, strenuously disputed others, and promised to improve.

Through all that, FDA never formally sanctioned Harris or pursued other penalties. The agency never made public the alleged offenses or told trial participants they might have been put at risk. Nor did it tell companies sponsoring some of the trials that their data might have been compromised. (The documents *Science* obtained were heavily redacted, making it impossible to know which trials were in doubt and, thus, which volunteers might have been harmed or endangered.) Meanwhile, pharmaceutical and medical device companies continued to contract with Aspen. Since 2011, they have paid the firm millions of dollars for work on at least 65 trials, and Aspen is now recruiting people for nine new trials on Alzheimer's disease, autism, depression, and other serious disorders.

Harris declined to comment on the FDA reports and the agency refused to discuss its dealings with him. Wayne Croft, Harris's business partner at Aspen, also would not reply to questions, but he describes FDA's treatment of Harris as bureaucratic harassment that drove away business. "Anytime he hears about the FDA," Croft says, Harris experiences "post-traumatic stress."

Scrutiny of the agency's clinical trial oversight for the past 11 years suggests its interactions with Harris and Aspen are commonplace. FDA's enforcement of clinical research regulations is often light-handed, slow-moving, and secretive, *Science* has found in an investigation that included a review of some 1600 agency inspection and enforcement documents for other trials that FDA said violated rules and law. Almost all were acquired via FOIA requests, including many made by FDAzilla, a commercial service that tracks the agency.

The period examined covered former President Barack Obama's administration and the first 3 years of President Donald Trump's term. Clear corrections of inspector-reported dangerous or unlawful clinical trial practices were the exception, even amid signs that trial participants were harmed and that data underpinning evidence-based medicine were corrupted. On the rare occasions when FDA formally warned researchers of findings that they had broken the law, the agency often neglected to ensure that fixes occurred,

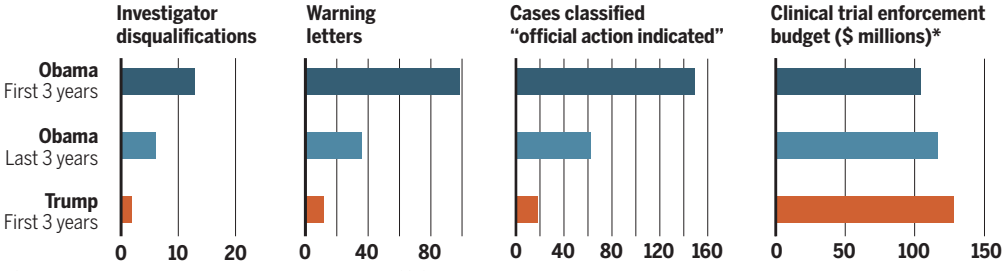
Science found. Moreover, the agency frequently closed cases on the basis of unverified claims by those accused.

The apparent neglect appears to be worsening. The agency issued 99 "warning letters" for serious clinical trial transgressions during Obama's first 3 years in office, 36 in his last 3 years, and just 12 during the first 3 years under Trump. "It certainly looks like FDA is enforcing clinical trial requirements much less frequently, which is troubling for protecting subjects' welfare and ensuring the validity of data for our medical products," says Patricia Zettler, a lawyer who reviewed *Science*'s findings. She worked for FDA from 2009 to 2012, rising to associate chief counsel.

Jill Fisher, a social scientist at the University of North Carolina, Chapel Hill, who studies clinical trials, is also troubled that when an FDA inspection exposes apparent mistakes, recklessness, or fraud in a trial, neither the agency nor the scientists running it are obliged to notify participants. She calls that failure to inform "a travesty." At the least, Fisher says, volunteers in a trial should be told when the integrity of a research center running the study or its investigator is in question.

A watchdog loses its bite

Despite a steady or rising enforcement budget, the Food and Drug Administration has wielded its power to enforce laws meant to safeguard volunteers for clinical trials—and ensure the integrity of testing data—less and less frequently since the beginning of the Obama administration. Under President Donald Trump, the number of enforcement actions has plummeted.



*Combined fiscal year budgets for the periods shown, constant 2019 dollars

LIGHT ENFORCEMENT

Although best known for overseeing food safety and evaluating experimental drugs, vaccines, and medical devices for marketing approval, FDA performs many other tasks with its $5.7 billion annual budget (for fiscal year 2019). Among them is overseeing most clinical research in the United States and some conducted elsewhere for the U.S. market—including many high-stakes vaccine and drug trials for COVID-19. The agency now deploys 102 inspectors to respond to whistleblower complaints of dangerous operations, conduct routine visits to trial sites, and review records of those sites or the institutional review boards (IRBs) that oversee trials locally. Over the period *Science* examined, FDA conducted roughly 6700 inspections of clinical researchers or IRBs.

FDA inspections can result in a range of responses, from "no action" declarations; to "official action indicated" (OAI) reports, which require a violator to clean up serious transgressions; to warning letters, which threaten further regulatory action unless corrections are made promptly. In extreme cases, the agency can even disqualify a scientist from clinical research.

But *Science* found that FDA rarely levels sanctions. The agency almost always rules that no action is warranted or requests voluntary corrections, according to documentation on all inspections that drew an OAI designation or more severe enforcement action. Since 2009, FDA has used the OAI designation to direct only 291 of the inspected researchers or institutions—about 4%—to correct serious, illegal, and potentially dangerous clinical trial problems. Furthermore, FDA officials sometimes downgraded a problem originally classified by inspectors as OAI—in which compliance must be confirmed by a follow-up inspection or other actions—to "voluntary action."

Holly Fernandez Lynch, a lawyer and bioethicist at the University of Pennsylvania (UPenn), says the agency prefers voluntary action. "When people fail to comply, it's hard to know what to do about it, short of completely ending a trial or all research at that institution," which might deprive patients of a chance at an effective treatment in trials, Lynch notes.

Under Trump, the number of FDA inspections per year has increased significantly, yet enforcement actions have nosedived (see graphic, above). About 6% of FDA inspections were classified as OAI during the Obama administration, although the proportion began to decline in his second term. During Trump's first 3 years, however, OAI reports fell to less than 1% of the total. Even voluntary action designations declined sharply under Trump, whereas "no action" inspections spiked. The agency disqualified an average of three investigators per year under Obama but just two total during Trump's first 3 years.

FDA did not dispute *Science*'s figures on OAIs and warning letters—but the agency said Trump-era policies were not a factor. "The number of warning letters can ebb and flow," the agency wrote in an email.

For extreme problems in a clinical trial, when patients or data reliability are at immediate and serious risk, FDA can use warning letters to immediately halt the trial or restrict an IRB from approving new trials. The agency did so seven times during Obama's first term and not once since Trump took office.

LITTLE FOLLOW-UP

Among the 291 OAI cases *Science* identified, only 71 resulted in a clear regulatory endpoint—such as disqualification, a "closeout letter" certifying that corrections were completed, or another formal statement. For the remaining 220, no clear outcomes could be found in public documents or data banks, leaving trial participants and others in the dark. FDA can also deem violations "not correctable," ending its enforcement with neither compliance nor any form of disclosure. From outside the agency, it's impossible to know how often that occurs.

In addition to searching public documents and FDA databases and filing FOIA requests for key documents in nearly all 291 cases, *Science* filed additional FOIA requests in 28 of those cases for every essential document, including any that described a resolution. After about 10 months, FDA released documents describing how 16 cases were resolved. Most were based on written promises from researchers. In one, the researcher pledged to never again conduct clinical studies and then violated that pledge, according to records posted by drug companies that hired him for clinical trials—with no apparent FDA response.

In an email, FDA defended its approach, noting its "strategy to focus inspectional resources on higher-risk facilities." FDA tries to address clinical trial problems early in the process, it also noted, so "an issue that might warrant a warning letter could be resolved before the problem rises to that level." (Agency officials would not agree to an interview. Spokespeople instead emailed selective responses to written questions.)

Yet FDA sometimes issues closeout letters even when nothing material has changed. In 2016, the agency sent a warning letter to the

CREDITS: (GRAPHIC) N. DESAI/*SCIENCE*; (DATA) FOOD AND DRUG ADMINISTRATION

ILLUSTRATION: STEPHAN SCHMITZ

Disgraced researchers can still reap drug industry payouts

When the Food and Drug Administration (FDA) wields its ultimate penalty, disqualification, against clinical researchers who it determines have violated the law, falsified data, or committed grave errors or misconduct, they can no longer run human trials in the United States. But that doesn't always sever their financially rewarding relationships with big pharma.

In 2008, FDA filed a public notice that it had disqualified Texas urologist James Vestal after its inspectors discovered egregious problems in clinical trials he had led of an experimental hormonal treatment for advanced prostate cancer. The agency said Vestal admitted to fabricating medical exams, faking signatures, enrolling ineligible patients, and other actions that "exposed [his] subjects to unnecessary risks."

Yet a *Science* examination of corporate compensation disclosures from the federal Open Payments database showed that from 2013 to 2019, 27 drugmakers—including heavyweights Bayer, the Johnson & Johnson subsidiary Janssen, and Sanofi—paid Vestal about $422,000, including $340,000 for consulting and teaching. (The system only began to record pharmaceutical compensation to physicians in 2013.) Vestal did not respond to requests for comment.

In 2005, FDA disqualified Philadelphia dermatologist Harold Farber just before he pleaded guilty to two federal crimes: illegal possession of anabolic steroids and (in his role as sole owner of a private practice) "knowingly and willfully" making a "false, fictitious, or fraudulent statement or representation" to the sponsor of a clinical trial. Federal prosecutors found that during the clinical trial, for an experimental cream to treat Sun-damaged skin, Farber delegated patient exams that he claimed to have conducted himself to an assistant. He was sentenced to probation and medical supervision of his practice and paid about $220,000 in fines and restitution, mostly to the sponsor of the skin cream trial. And he was barred from future earnings for conducting trials.

But Farber has since recouped those losses in other ways, thanks to the pharmaceutical industry. From 2013 to 2019 alone, Open Payments data show, he took in about $665,000 from 45 medical research companies—including Actavis, Allergan, Bayer, and Genentech—for consulting and teaching. Pfizer paid him the most, $90,000. He received $169,000 in other fees, food, and travel perks. Asked whether he informed the companies about his FDA disqualification and legal history, Farber declined to comment.

Science asked 33 drug companies why they paid Vestal or Farber to teach or consult after FDA expelled them as clinical investigators. The 22 that responded all said the two remain qualified physicians who were not hired for clinical trials. None would say whether they knew about the FDA sanctions, or Farber's criminal conviction, when they made the payments.

Among the 42 other physicians FDA disqualified since 2005, data from 2013 to 2019 show five others took in about $10,000 to $36,000 in drug company payments after their bans. One was Miami internist Farid Marquez, barred from research in 2015 after FDA found what it said were phony documents in his work for drug giants Boehringer Ingelheim and Eli Lilly. But those companies seem to have no hard feelings. Although Marquez no longer gets their research money, Boehringer Ingelheim has treated him to at least 168 meals since his disqualification; and Lilly, 17. Marquez declined to comment.

"The most favorable sort of interpretation is that the companies' internal vetting process is deeply broken," says Vinay Prasad, a hematologist-oncologist at the University of California, San Francisco, who studies how drug industry funding influences research and medical practice. "The more pessimistic interpretation is that they turned a blind eye because the investigators were not disqualified for something that hurts the companies' bottom lines." —*C.P.*

medical director of a Texas research firm. He had enrolled trial participants who suffered from kidney impairments, an agency inspection reported, and then overdosed them with an experimental drug that could have further damaged their kidneys. The physician responded with a plan to prevent future overdoses, but FDA replied that it was inadequate. Yet on the same day as that reply, the agency sent a letter closing the case.

"The agency should follow through and complete actions—so they don't initiate a process and never resolve it," says Joshua Sharfstein, vice dean at the Johns Hopkins Bloomberg School of Public Health and a top FDA official under Obama. "And it's important for the agency to share information about its processes, so people can understand what happened."

When FDA does follow up on the problems it identifies, the process can be slow. Agency documents and databases covering the 291 OAI cases show dozens of cases that took 10 months to 14 years to be escalated to a warning letter. Some delays can be "egregious," Lynch says.

In 2015, the agency issued an OAI after it inspected a Pennsylvania hospital IRB that oversees more than 100 trials annually. The agency's findings read like a functional collapse: records glitches, including failure to keep a list of IRB members and their qualifications or record attendance at meetings; approval of trials without a quorum; sloppy reviews of trials; a lack of required written policies on how to determine the risk of devices being tested on patients; neglecting to promptly inform FDA or the hospital's own leaders about trial problems; and poor or no training for IRB members. According to FDA documents, the agency waited 4 years to follow up. In a second inspection, in 2019, many of the same problems were found, resulting in a warning letter. Only then did the IRB finally resolve the deficiencies.

A St. Louis IRB received citations about similar problems in a 2012 warning letter, which noted that pervasive lapses at the IRB were also found in 2003 and in 1998. That IRB finally addressed the issues after the warning, according to an FDA closeout letter.

That sparse, slow enforcement puts participants and the research process itself at risk, critics say. Given the extremity of violations that lead to little or no punishment, Fisher asks: "What would it take for the FDA to take serious action to ensure that clinical research is being conducted with the most integrity possible?"

The secrecy of the process also means trial participants are kept unaware of possible harm, violating "a basic human right," says Vinay Prasad, a hematologist-oncologist at the University of California, San Francisco, who studies how drug funding affects research. Steven Joffe, a medical ethicist at UPenn, agrees, citing an obligation to be transparent to the people who have volunteered their bodies for clinical research: "We owe that back to them."

"SYSTEMIC PROBLEM" IGNORED

A clinical trial scandal that erupted in Minnesota in 2018 illustrates the potential public health hazards posed by FDA's lax enforcement. That year, the consumer advocacy group Public Citizen and more than 60 clinicians and medical ethicists alleged that a large Minneapolis community hospital operated by Hennepin Healthcare had violated informed consent requirements in trials comparing ketamine, a powerful sedative and anesthetic, with other drugs for people who came to the emergency room with agitation. The critics said the emergency department doctors did not obtain informed consent from trial participants, even though ketamine was much more likely to cause serious breathing problems and movement disorders. Many patients who received ketamine developed such symptoms.

> **"It certainly looks like FDA is enforcing clinical trial requirements much less frequently, which is troubling."**
>
> **Patricia Zettler**, former FDA lawyer

After Public Citizen's allegations, FDA inspectors confirmed the lapses in informed consent. They concluded that Hennepin's IRB had waived the requirement to obtain consent and did not institute additional safeguards "to protect the rights and welfare of those subjects." And inspectors said researchers running the trials neglected to properly report serious adverse events. Nevertheless, the agency did not classify the inspection as OAI or send a warning letter. Hennepin rejected most of the FDA findings, although the health system agreed to some procedural changes and suspended one trial for review.

FDA already knew of alleged problems in Hennepin's ketamine research at the same hospital before the Public Citizen allegations, according to new documents obtained by *Science*. In 2014, the agency inspected a Hennepin trial led by Brian Driver, a staff physician, and James Miner, then and now chief of emergency medicine at the hospital where the later trials took place. The trial compared ketamine with another drug used to sedate patients who required breathing tubes. As in the later trials, the agency detailed failures to properly obtain informed consent from patients.

A 2016 follow-up review by FDA of the ketamine trial detailed failures by Driver and Miner to properly report 13 patient deaths to Hennepin's IRB—a required practice to verify that the trial itself does not pose safety hazards. In a response to FDA, Driver acknowledged his reporting error in three of the deaths but did not address the other 10 deaths, and FDA documented no further information on them. In an email to *Science*, Driver wrote that FDA had never discussed those additional deaths with him, but he acknowledged failing to report the deaths of the 13 FDA initially flagged, as well as 10 others in the trial that it did not document. Driver attributed all 23 deaths to acute illnesses or injuries unrelated to the study drugs, including sepsis, stroke, and heart attack.

Driver became a member of the hospital's IRB in 2016 and was on it when the group approved or reviewed the later ketamine trials flagged by Public Citizen. *Science* showed the FDA documentation of his earlier ketamine trial to Michael Carome, who directs health research at Public Citizen. "We have a group of physicians, all in the emergency department, over a period of several years … not complying with FDA regulations regarding human subjects protection," says Carome, a physician himself. The clinical trial missteps exposed "a pattern of sloppiness and carelessness" by Hennepin researchers and their IRB, he adds, raising broad doubts about its oversight of human subjects research.

The IRB chair, a hospital spokesperson, Miner, and Driver all declined to comment about the later clinical trial problems FDA identified. In a public statement, Hennepin defended the trials while apologizing for "not involving the community in a proactive way" before the studies. Hennepin said it would increase training for clinical researchers and improve study monitoring and public disclosure. But the health system refused to answer questions such as why—after years of FDA inspections cited serious problems with ketamine trials—the IRB and emergency department launched new trials with no special precautions to protect patients.

Hennepin's ketamine controversy also raises questions about FDA oversight. Despite the problems it identified as early as 2014, FDA never issued a warning letter or otherwise demanded that Hennepin researchers rectify trial procedures. The agency imposed no restrictions on the hospital's clinical research or IRB operations. FDA never mandated that trial participants who suffered adverse effects—or their loved ones—be advised about inspection findings. FDA declined to comment on Hennepin's

clinical trials—or on any specific cases described in the *Science* investigation.

"FDA is just dropping the ball here," says Carome, a former research regulation executive at the Department of Health and Human Services. "I think the institution should have gotten a very strong warning letter for the [latter] ketamine trials, and that letter perhaps should have noted the prior history with Dr. Driver, to signal a systemic problem."

By not coming down hard on IRBs and the scientists running clinical trials, FDA lets those parties avoid possible negligence lawsuits or other legal liability if patients are harmed, Prasad says. He adds that the lack of enforcement also lets the drug companies sponsoring the trials off the hook. "FDA is a regulatory agency charged with protecting the public's best interests. But at times it behaves like an attorney working on behalf of the companies."

SLOW-MOTION DISQUALIFICATIONS

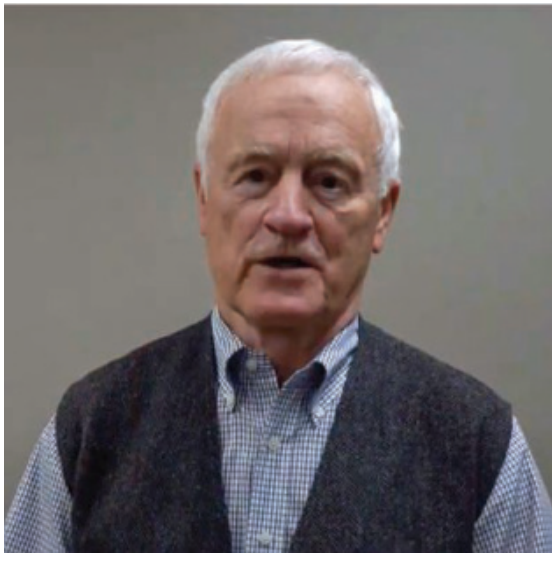
For the most extreme violations of rules governing human subjects research, FDA can brandish the threat of a publicly announced disqualification—effectively the death penalty for a clinical research career. But how the agency has used that sanction captures many flaws of FDA's clinical trial oversight.

Since the beginning of Obama's presidency, the agency has disqualified just 24 clinical trial investigators, mostly for deliberate and repeated falsification of data or other information such as patient referrals for a trial. And although disqualification bars a scientist from receiving any experimental drug or device or conducting any trial involving FDA-regulated drugs or other substances, it can sting less than it sounds. Drug companies often continue to lavish money and prestige on disgraced investigators (see sidebar, p. 27).

Investigators under a cloud also can continue to run lucrative trials while FDA considers whether to disqualify them, a process that takes, on average, more than 3 years. Scientists or IRBs facing FDA accusations, including those that ultimately lead to disqualification, have collectively conducted or overseen thousands of trials during the past decade, *Science* found in a review of public records. Take Oregon allergist-immunologist James Baker, one of only two disqualifications by Trump's FDA.

Between 2009 and 2017, FDA inspected Baker-led trials on four occasions and each time found violations. Time after time, Baker promised to do better, according to agency documents. Again and again, he broke his promises, FDA documents conclude. The agency, however, waited until 2018 to disqualify him.

Baker's abuses were not mere technicalities. FDA said they "placed unnecessary risks to human subjects and jeopardized the integrity of data." Inspectors found informed consent breakdowns involving children and adults—some of whom were injected with experimental allergy or immunology drugs or biologics despite taking other drugs, such as cough medicine and antibiotics, that could interact with them. He failed to properly document or follow up on a child's emergency department visit for severe abdominal pain. People were enrolled against eligibility rules—with some participating in more than one trial, for example, or having conditions that should have disqualified them. FDA said Baker kept substandard case histories, overdosed patients on experimental medicines, and repeatedly conducted trials without IRB approval. Baker conceded to agency inspectors that one of his employees, who had no relevant medical credentials, prescribed drugs to trial volunteers. She worked on 30 of his trials.



James Baker is one of just two scientists disqualified from U.S. clinical research by FDA during President Donald Trump's first 3 years.

FDA finally disqualified Baker after he was caught backdating many clinical trial reports beginning in 2013. The records indicated he performed patient exams in Oregon when he was actually in Las Vegas, Hawaii, and elsewhere attending medical meetings and, in one instance, a wedding. (FDA inspectors had earlier caught him similarly falsifying records in 2015, when he admitted he had "gone fishing" on days he purportedly saw trial patients.)

FDA documents show Baker conducted at least 115 clinical trials from 2005 until 2017, often testing drugs for which companies later sought agency approval. Yet his disqualification stemmed from FDA's review of just a few of those trials. Not long before his disqualification, the agency asked Baker whether any data or patient records had also been doctored in his other completed trials. Baker said he lacked the capacity to review those studies. As far as *Science* could ascertain, FDA did not investigate the other trials.

Baker declined to say whether he had informed participants in his many trials that he had been disqualified. *Science* also contacted multiple companies that sponsored Baker's trials during the period for which FDA found falsified information. Asked whether they told trial volunteers about the government's findings, none responded.

Baker, who declined to answer most questions submitted by *Science*, contended in an email that his patients suffered no adverse events. But he also added: "I agree with [the] way the FDA controls the research world and regret running afoul with those studies."

Had FDA acted earlier, Baker's clinical trial venture—the Baker Allergy, Asthma and Dermatology Research Center—might have lost a lot of business. The French drug giant Sanofi and other companies paid him at least $1.2 million for trials conducted between the time FDA detected his falsified exams in 2016 through 2019, according to the federal Open Payments system, which tracks pharmaceutical industry payments to doctors.

Yet the work kept flowing from Sanofi to his research business even after he was disqualified. Sanofi assigned responsibility for Baker's trials to another doctor at his center, which Baker operated with his dermatologist wife in the same building as their medical practice.

Likewise, the disqualification leaves Baker free to operate his other company, Xtract Solutions, a suburban Portland, Oregon, purveyor of allergy and immunology technology to hospitals and physicians. In 2014, FDA issued a warning letter to Baker and Xtract for allegedly selling "adulterated devices" and failing to establish and maintain quality-control procedures for its syringe-filling systems—meant to speed up vaccinations during pandemics.

Xtract eventually dropped its injection device. But the company still sells software to guide doctors and hospitals on a subject Baker has ample experience, if not always success, with: FDA regulatory compliance. ■

With reporting by Meagan Weiland. This story was supported by the *Science* Fund for Investigative Journalism. For more on how the story was reported, see http://scim.ag/FDA_Methods.

PHOTO: CLOUDSHINE MEDIA, "BAKER ALLERGY, ASTHMA & DERMATOLOGY (COVID-19 SAFETY)," 2 APRIL 2020/YOUTUBE.COM

INSIGHTS

LETTERS

NEXTGEN VOICES: ASK A PEER MENTOR

Funding fix: Spend time

As coronavirus disease 2019 (COVID-19) restrictions isolate scientists in their homes, close labs, and cancel conferences, some researchers are overwhelmed by professional and domestic responsibilities, whereas others, like the advice-seeker below, have a surfeit of time but a dearth of funding. Here, peer mentors offer advice to this researcher. Do you need advice for the COVID-19 era? See the box on page 31 to submit your own questions. Follow NextGen Voices on Twitter with hashtag #NextGenSci. Read previous NextGen Voices survey results at https://science.sciencemag.org/collection/nextgen-voices. **—Jennifer Sills**

Dear NextGen VOICES peer mentors,

I am a young researcher with research proposals but no funding. My grant applications are still pending. My contract requires me to publish, but I can't conduct experiments or produce reliable results without funding. With restrictions in place for COVID-19, I have limited access to my lab. How can I use this time most effectively to ensure that I can stay in academia long-term?

Sincerely, Funding Fix

Hone your skills

How can you turn the time created by this pandemic into a blessing? Faced with similar challenges, I have tried to upgrade my skills to work remotely, manage effectively, network extensively, and expand my knowledge through webinars, online conferences, and seminars. I suggest using this time to increase your knowledge by reading extensively in your research area. Write reviews, book chapters, or a short piece analyzing old data. Find collaborators who can give you access to a lab that is operational. Apply for funding that is open to young researchers or that requires collaboration with national or international partners. Attend interesting webinars, and never miss an opportunity to present your work to the research community.

Charu Lata
Council of Scientific and Industrial Research–National Institute of Science Communication and Information Resources, New Delhi, Delhi, 110067, India. Email: charulata@niscair.res.in

Have you considered using publicly available data as a source for research? I suggest that you stay in your area of research and ask some new questions. In this era of genomics, the wealth of publicly available databases allows you to conduct research at home. In the medical field, medical records and big data discovery are considered an essential part of the health infrastructure and represent a valuable resource for translational research.

Ruty Mehrian-Shai
Pediatric Hemato-oncology, Sheba Medical Center, Ramat Gan, 52621, Israel.
Email: ruty.shai@sheba.health.gov.il

Are you sure that spending all your time doing experiments in the lab is the best way to obtain academic achievements? In my experience, analyzing experimental results, organizing figures, and rethinking scientific ideas at regular intervals are more important to improving quality of research. Consider using this time to better understand your current results. Prepare research figures that help you tell the scientific story of your findings. Review the literature with an eye toward adjusting the aims and methods of your ongoing research. Practice communicating your work through oral and poster presentations in virtual meetings. These activities will prepare you to use your time more productively once you return to the lab.

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

ILLUSTRATION: ROBERT NEUBECKER

Connect and collaborate

Have you reached out to your departmental peers and administrators for direction? As a faculty member in a non–tenure-track position, I have found that it's helpful to form a peer group to brainstorm ideas. It might also be useful to ask your school's administration about their expectations for the current pandemic situation. Perhaps they will be flexible about your contractual obligations.

Naga Rama Kothapalli
Department of Chemistry and Biochemistry, University of Oklahoma, Norman, OK 73019, USA.
Email: ramakn@ou.edu

Have you thought about how much time you have been devoting to conducting lab work, writing manuscripts, securing grants, managing students, and reading? Many young scientists are pressured to focus on publications and funding to the exclusion of other important aspects of a career in academia. This time presents an opportunity to find a balance. Consider helping the graduate students and postdoctoral students in your lab. Spending some time bonding with these researchers, professionally and personally, will help you support them. Help them contribute by asking them to think in innovative ways. Fostering new networks will help you excel in the post-COVID world.

Garima Singh
Fleming Fellowships (Antimicrobial Resistance), South Asia, Delhi, India.
Email: singhg20@gmail.com

Step back and rethink

How can you use this time in non-traditional ways? Science teaches us to adapt when new information arrives or situations change. I have found that using my creativity has helped inspire my best science and service. I suggest that you find ways to be innovative by making art, learning about a new topic or perspective, volunteering for peer review, doing outreach, educating others, or taking part in citizen science.

Daniel Ari Friedman
Department of Entomology, University of California, Davis, CA 95616, USA and Remotor Consulting Group, Davis, CA 95616, USA.
Keybase: @docxology

What is the one experiment that you are most excited to complete? I have noticed from my work-from-home experience that it is good to stop and rethink the research from time to time. By prioritizing different experiments, goals become clearer. You might realize that you can avoid doing tangential experiments that are not necessary to support your primary hypothesis. This introspection will make your time in the lab more efficient when you return.

Felix Man-Him Cheung
School of Biomedical Sciences, University of Hong Kong, Hong Kong, China.
Email: felixcheung@connect.hku.hk

How can you use this time to protect your mental health? We always complain about not having enough time to read and learn everything we need. Now that many of us have plenty of time, it is important not to waste this gift. Remember to save some time for leisure. Read a romance novel, play a game, or listen to music. When you go back to the lab, you need not only the knowledge to complete your experiments but also the mental health to continue your work.

Wagner Eduardo Richter
Department of Chemical Engineering, Technological Federal University–Paraná, Paraná, Brazil.Email: richter@utfpr.edu.br

Explore a new field

Have you considered reading current literature outside of your primary area of expertise? Waiting for funding outcomes necessary to commence data collection is stressful, but it does provide opportunities to spend time in other ways. Reading widely enabled me to bring novel perspectives to my research areas, which led to my highest-funded and most rewarding projects. Although reading broadly will not result in immediate publications, it can help differentiate you from your peers, which will ultimately help you stay in academia long-term.

Samuel Nathan Kirshner
School of Information Systems and Technology Management, University of New South Wales, Sydney, NSW 2052, Australia.
Email: s.kirshner@unsw.edu.au

Would your papers, science, and current datasets benefit from a shift in paradigm? Early in my career, I found that working with oral historians exposed me to a new way of seeing the archaeological landscape. It made my science more accessible to a general audience, brought community support for my work, and even helped pass legislation protecting archaeological remains in the countries where I recorded stories. I encourage you to keep an open mind and be flexible. Taking a step away from the lab could create an opportunity to examine your experimental paradigm, expand your literature searches to include fields outside your own, and seek different perspectives by opening dialogue with members of another field.

Felicia Beardsley
Department of Anthropology, University of La Verne, La Verne, CA 91750, USA.
Email: fbeardsley@laverne.edu

Have you thought about taking a break from research? Moving away from academia for a short period helped me acquire new skills and gain momentum later in my career. I took a break during the third year of my Ph.D. program to do an internship in industry, where I started to look at problems with a new perspective. Later, I incorporated some of that experience into my thesis, my postdoc work, and even my tenure track position. I suggest that you take a break and look for "virtual" internship opportunities. You may acquire new skills that you can apply to your research later. You might even realize that there are many other options for you to apply your knowledge and skills, which will allow you to make better career decisions.

Xiao-Yu Wu
Department of Mechanical and Mechatronics Engineering, University of Waterloo, Waterloo, ON N2L 3G1, Canada.
Email: xiaoyu.wu@uwaterloo.ca

10.1126/science.abe6015

NEXTGEN VOICES: CALL FOR QUESTIONS

Need advice? Ask your peers!

Do your COVID-19 experiences differ from this young scientist, who faces too much time and not enough funding? Are you affected by increased family duties, cancelled travel, health uncertainties, or limited career options? Are you facing unique challenges that others have overlooked? *Science* would like to support you by asking readers to provide peer mentoring advice. Do you have a question that you would like your peers to address? Send it to *Science* at the link below!

To submit, go to www.sciencemag.org/nextgen-voices-covid-19-questions
Please submit by 30 October. If your question is selected, *Science* will post it anonymously and ask young scientists to respond with advice to be published in a later issue.

1 Scrapie is a highly infectious disease of sheep that is caused by neurotoxic prion protein aggregation.

PHOTO: WAYNE HUTCHINSON/MINDEN PICTURES

LANDMARK: NEURODEGENERATION

Shifts and drifts in prion science

Important questions remain unanswered since prions were discovered four decades ago

By **Adriano Aguzzi** *and* **Elena De Cecco**

Paradigm shifts are drivers of scientific progress, yet the shifters of the paradigms often experience scorn rather than immediate applause. That was the fate of Stanley Prusiner's 1982 paper claiming—to the initial amusement of his colleagues—that scrapie, a degenerative disease that affects the central nervous system of sheep, is caused by "proteinaceous infectious particles," which he called prions (*1*). Prusiner's intuition, which earned him the 1997 Nobel Prize, is influencing our approach to an ever-expanding variety of seemingly unrelated diseases and physiological processes, and its implications reverberate to the present day.

The two decades preceding Prusiner's paper had witnessed the immense success of molecular biology, including the cracking of the genetic code; the elucidation of DNA replication, transcription, and translation; and the cloning of genes. These discoveries prompted Francis Crick to conceptualize the "central dogma": Information flows unidirectionally from DNA to proteins. But although religious dogmas may be eternal, the shelf life of scientific dogmas is inevitably limited. Prusiner postulated that prions carry on their replicative cycle without the participation of nucleic acids. This hypothesis, reminiscent of John Griffith's 1967 suggestion of the existence of self-replicating proteins (*2*), had the potential to explain the prodigious resistance of the scrapie agent to DNA-damaging radiation.

Daniel Carleton Gajdusek, who won a Nobel Prize for showing that Kuru was a human disease transmitted by cannibalism in Papua New Guinea, proposed in 1959 that the neurodegenerative disorders Kuru, scrapie, and Creutzfeldt-Jakob disease (CJD) are caused by "slow viruses." Indeed, prions behave similarly to neurotropic viruses in many surprising ways, including the colonization of extraneural organs followed by neuroinvasion of the brain through peripheral nerves (*3*). Yet, Prusiner purified the agent and found it to be smaller than a virus: No informational nucleic acid would fit into it. Over time, the group of prion diseases grew to include other human (fatal familial insomnia) and animal (bovine spongiform encephalopathy, also called mad cow disease, and chronic wasting disease) disorders, but no causative virus has been identified and their prion etiology is now well accepted.

But prions did not contradict Crick's central dogma after all. Charles Weissmann, refusing to believe that a protein could exist without its respective gene, discovered in hamsters the gene encoding the cellular prion protein (PrP^{C}), whose misfolding yields tightly packed aggregates called PrP^{Sc}. It is generally believed that prion replication occurs when coalesced PrP^{Sc} is broken down into smaller species. Those species then accrue further PrP^{Sc}, in a process akin to the growth of crystals, and eventually break again, perpetuating their replicative cycle. Infectious prion seeds then move to neighboring cells and wreak havoc in the central nervous system by inducing vacuolation ("spongiosis") within neurons (see the figure).

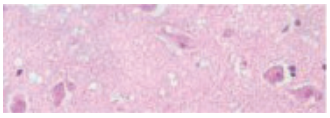
Does this mean that PrP^{Sc} is the prion? Weissmann's discovery in 1993 that prion protein (*Prnp*)–ablated mice are resistant to scrapie (*4*) was designed to disprove the protein-only hypothesis but failed to do so. However, *Prnp* deletion in mice also fell short of proving the prion hypothesis. If PrP^{C} were the receptor of an imaginary "scrapie virus," its ablation may also render mice resistant to scrapie. More direct evidence for Prusiner's ideas emerged in 2001 from Claudio Soto's landmark experiment: Repeated cycles of PrP^{Sc} fragmentation, when followed by addition of PrP^{C} and aggregate regeneration, can multiply prions ad libitum (*5*). These findings strengthen the hypothesis that the transfer of structural information can occur horizontally between proteins.

More recently, the prion concept has been applied, sometimes overenthusiastically, to virtually all diseases characterized by progressive deposition of aggregated proteins in the central nervous system, whether infectious or not—and even to physiological processes such as memory formation (*6*). α-Synuclein aggregates can self-propagate

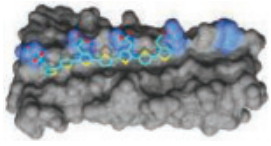
Institute of Neuropathology, University of Zürich, Rämistrasse 100, CH-8091 Zürich, Switzerland. Email: adriano.aguzzi@usz.ch

Three centuries of prion science

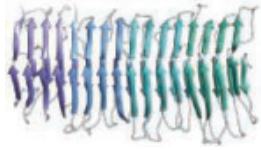
The timeline shows key prion-related discoveries. In 1982, Prusiner suggested that the prion protein (PrP) is the infectious cause of spongiform encephalopathies, including Kuru, scrapie, and Creutzfeldt-Jakob disease (CJD). These insights have had implications for many neurodegenerative diseases involving prionoids, but many questions still remain unanswered.



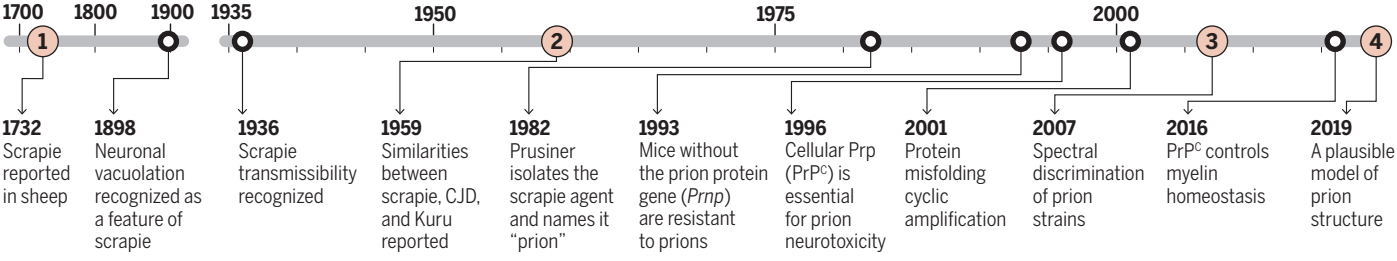
2 CJD brain with vacuolation (spongiosis)



3 An antiprion molecule targets prion aggregates



4 Structure of PrP aggregate (PrP^{Sc})



in the brains of Parkinson's disease patients (*7*), in cultured cells, and in mice (*8*). This implies that α-synuclein is a de facto prion and that its handling demands high biosafety standards. Similar arguments were made for tau and amyloid-β (Aβ) aggregates, the major hallmarks of Alzheimer's disease (*9*). However, prions caused many epidemics, whereas infectiousness has not been conclusively demonstrated for other protein aggregates—and specifically not through oral transmission. Protein aggregates that were not shown to be serially transmissible across multiple generations of hosts are better regarded as "prionoids," even if they share molecular mechanisms of amplification with bona fide prions in vitro.

As predicted by Prusiner in the closing lines of his paper, the "prion revolution" boosted research in the field of neurodegeneration by providing an intellectual framework that might explain many aspects of Alzheimer's disease, Parkinson's disease, and many other neurodegenerative diseases featuring protein aggregates. Although cellular PrP^{C} is now known to be crucial for the maintenance of peripheral myelin (*10*), our understanding of prions has essentially stagnated for more than a decade and may now be lagging behind that of prionoids. What is really known about prions, after almost 40 years since Prusiner's discovery?

One crucial obstacle to advancing prion research is the lack of high-resolution structures of PrP^{Sc} owing to its insolubility, its noncrystalline aggregational state, and the persistent difficulties in preparing high-purity infectious material de novo from recombinant protein (*11*). This raises the possibility that infectious aggregates may constitute a sparsely populated conformational variant within such preparations. If so, most material aggregated in vitro may be noninfectious and may not be informative of the structure of the prion or of its replicative mechanism.

Of all the models that have been proposed so far, the most plausible suggests that the prion consists of fibrils arranged as four-rung β-solenoids (*12*) stacked either head-to-tail or head-to-head. Cryo–electron microscopy of purified glycosylphosphatidylinositol (GPI)–anchorless prion fibrils (*13*) supports this model, thus providing the first high-magnification images of infectious prions, albeit the resolution does not suffice to determine the precise arrangement of the monomers within the fibrils. These structures are quite different from those of tau, α-synuclein, and Aβ and also differ from recombinant PrP fibrils—all of which are arranged in long fibers with no cavity. Hence, PrP^{Sc} has distinctive structural characteristics, but it is unknown whether and how these peculiarities relate to their frightening infectivity.

The link between the generation of PrP aggregates and their neurotoxicity is also unclear. A large body of evidence (*14*) indicates that PrP^{C} is necessary for toxicity, perhaps because extracellular PrP^{Sc} oligomers dock to PrP^{C} on the surface of diverse cell types. Another aspect specific to prion infections pertains to the peculiar morphology of the damage that it wreaks on the brain. Of all aggregation-prone proteins, prions are the only ones that cause extensive intraneuronal vacuolation (spongiosis), the severity of which increases during disease progression. This phenomenon is as much intriguing as it is mysterious. To date, almost nothing is known about the cellular and molecular pathologies underlying vacuole formation; yet its ubiquity in all known prion diseases suggests that vacuolation is a prime driver of toxicity—and therefore also a target for therapeutic interventions.

High-resolution three-dimensional structures of prions are also required to solve the long-standing question of prion strains, which share the same PrP sequence and yet cause distinct diseases (e.g., "hyper" and "drowsy" phenotypes in minks), the traits of which are maintained over successive rounds of infection. Viral strains are defined by specific polymorphisms in their respective genomes, and the existence of strains in prion diseases was long thought to be incontrovertible evidence for the involvement of nucleic acids. However, after four decades of failed attempts to isolate any scrapie-specific genomes, strains are now thought to be caused by different PrP^{Sc} conformations that can be distinguished with conformer-sensitive fluorescent polythiophenes.

Embarrassingly for the prion field, no definitive structural evidence for these presumptions has come forward, and the "strainness" of bona fide infectious prions is still diagnosed using imperfect surrogate biomarkers such as differential resistance to disaggregation and proteolysis. By contrast, conformational heterogeneity was reported to correlate with distinct clinical phenotypes in some prionoid pathologies, although the stability of different conformations in serial transmission experiments is not yet fully established.

But how stable are prion strains across generations? RNA viruses achieve maximal fitness by creating quasispecies, clouds of variants in precarious equilibrium between adaptive mutagenesis and error catastrophe. Notably, prions can also engender quasispecies whose monoclonal constituents can be isolated from cultured cells by applying various kinds of selective pressure (*15*).

GRAPHIC: MELISSA THOMAS BAUM/*SCIENCE*; (IMAGES, LEFT TO RIGHT) K. FRONTZEK; U. S. HERMANN *ET AL.*, *SCI. TRANSL. MED.* **7**, 299RA123 (2015); G. SPAGNOLLI. ADAPTED FROM (*12*)

The structural mechanisms underlying this phenomenon are unknown and may involve conformational selection of distinct PrP^{Sc} species. The conformational selection model predicates the coexistence of multiple conformers within a single infected organism, some of which may replicate more efficiently in their host under certain environmental circumstances. The incubation time of prion infections can vary immensely between different strains, and the delay in the onset of the pathology might reflect the time needed for such selection to occur.

PrP^{Sc} conformer heterogeneity may also underlie the barriers that control interspecies prion transmission, the strength of which is variable and depends both on host factors and on prion strains. Although prion propagation from cows to humans results in variant CJD, sheep prions appear to be largely innocuous to humans. This species barrier relies both on the structural diversity of the PrP^{Sc} contained in the inoculum and the PrP^{C} of the host, which cannot always interact with the misfolded conformer efficiently.

The ideas promulgated by Prusiner have undergone a marked metamorphosis. Templated nucleation of protein aggregates is now known to underlie not only diseases but also many physiological processes, some of which bear little resemblance to the original set of diseases that attracted Prusiner's attention. Notably, the structural predictions of the prion model were verified for several prionoids but not for prions. As such, many of the questions raised by Prusiner in 1982—prion structure, mechanism of replication, and drivers of toxicity—are still open. Based on historical evidence, addressing these questions in the prion arena may, once again, provide answers that will also apply to more prevalent neurodegenerative diseases. ■

REFERENCES AND NOTES

1. S. B. Prusiner, *Science* **216**, 136 (1982).
2. J. S. Griffith, *Nature* **215**, 1043 (1967).
3. M. Prinz *et al.*, *Nature* **425**, 957 (2003).
4. H. Büeler *et al.*, *Cell* **73**, 1339 (1993).
5. G. P. Saborio, B. Permanne, C. Soto, *Nature* **411**, 810 (2001).
6. R. Hervas *et al.*, *Science* **367**, 1230 (2020).
7. J.-Y. Li *et al.*, *Nat. Med.* **14**, 501 (2008).
8. K. C. Luk *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **106**, 20051 (2009).
9. F. Clavaguera *et al.*, *Nat. Cell Biol.* **11**, 909 (2009).
10. A. Küffer *et al.*, *Nature* **536**, 464 (2016).
11. G. Legname *et al.*, *Science* **305**, 673 (2004).
12. G. Spagnolli *et al.*, *PLOS Pathog.* **15**, e1007864 (2019).
13. E. Vázquez-Fernández *et al.*, *PLOS Pathog.* **12**, e1005835 (2016).
14. S. Brandner *et al.*, *Nature* **379**, 339 (1996).
15. J. Li, S. Browning, S. P. Mahal, A. M. Oelschlegel, C. Weissmann, *Science* **327**, 869 (2010).

ACKNOWLEDGMENTS

The authors thank K. Frontzek for the micrograph of a CJD brain and G. Spagnolli for the PrP^{Sc} molecular model. The Aguzzi lab is supported by the Nomis Foundation and the Swiss National Science Foundation.

10.1126/science.abb8577

CANCER

Mutational selection in normal urothelium

Mutations in normal tissue point to causes of DNA damage and set the stage for cancer

By **Steven G. Rozen**

Cells have elaborate machinery to preserve the integrity of their genomes, which nevertheless relentlessly gather new mutations over time. Recent technical advances have enabled high-resolution delineation of these accumulated mutations and their spatial organization in tissues (*1*–*8*). It is now possible to deduce which mutations in which genes allowed cells to outcompete their neighbors and colonize nearby regions of normal tissue (clonal expansion). One can also sometimes ascertain which endogenous mutational processes or external mutagens caused these somatic mutations. These recent findings have profound implications for understanding aging and the early stages of cancer initiation. On pages 75 and 82 of this issue, Lawson *et al.* (*9*) and Li *et al.* (*10*), respectively, delve into the somatic mutations lurking in the normal urothelium—the lining of the bladder and ureter—and relate them to cancers in these tissues.

Although the findings of the two studies are broadly consistent, substantial methodological differences account for some divergence. Lawson *et al.* studied the urothelium of the bladder, primarily in organ donors from Europe, whereas Li *et al.* studied the nonmalignant urothelium of both the bladder and the ureter in cancer patients from China. Additionally, Lawson *et al.* studied an average of ~100 tiny (~0.01 mm^2) urothelial samples each from 15 organ donors and 5 cancer patients; by contrast, Li *et al.* studied an average of 1.3 large (~2 mm^2) samples of normal tissue from 120 cancer patients. The studies also differed in their sequencing approaches. Lawson *et al.* sequenced a mix of whole genomes, whole exomes, and a targeted panel of 321 cancer-associated genes, whereas Li *et al.* sequenced whole exomes.

Despite the methodological differences, both studies analyzed somatic mutations in normal urothelium to show that clonal expansion occurred. Both studies found that a selective advantage (indicated by high prevalence of a clone) usually stemmed from mutations in genes encoding proteins involved in histone modification and chromatin remodeling. Within these genes, truncating, and presumably inactivating, mutations were often strongly selected. The affected genes prominently included *KMT2D* (histone-lysine *N*-methyltransferase 2D) and *KDM6A* (lysine-specific demethylase 6A). These chromatin remodeling genes are also mutated in many cancers of the urothelium, suggesting that the nonmalignant expanded clones driven by these genes regularly, but not frequently, become malignant. This contrasts with the esophagus, in which some genes that often drive clonal expansion in the normal tissue only rarely act as drivers in cancers (*3*, *6*).

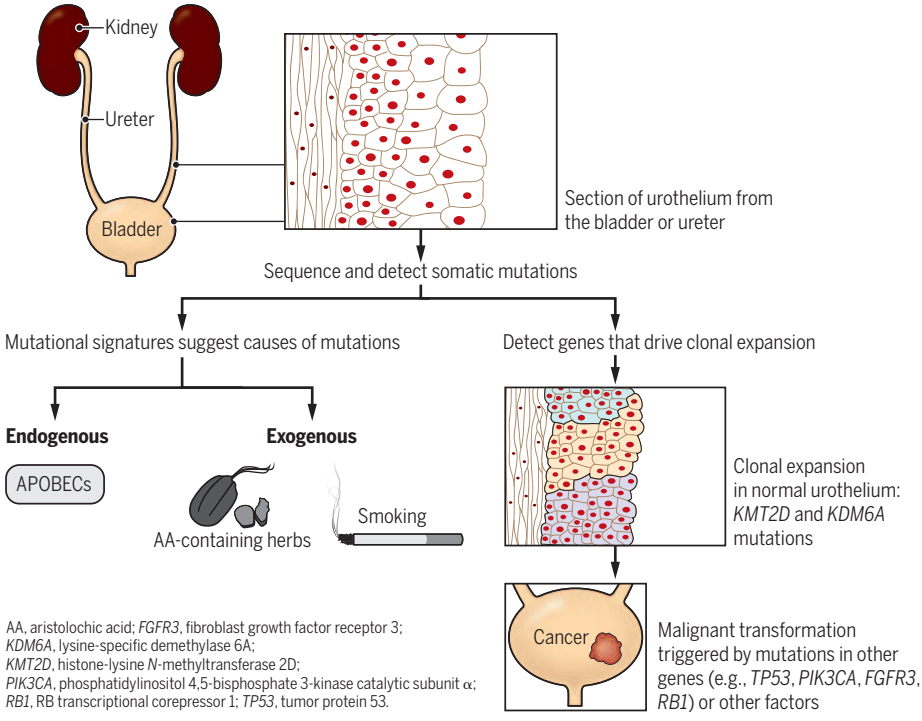
By examining many samples per individual, Lawson *et al.* found that, in some individuals, the urothelium showed strong selective preference for mutations in particular genes. For example, one person had 35 distinct *KDM6A* mutations distributed over multiple clones and 2 different *ARID1A* (AT-rich interactive domain-containing protein 1A, also involved in chromatin remodeling) mutations, whereas another person had 4 different *KDM6A* mutations and 20 different *ARID1A* mutations distributed over multiple clones. It was not possible to make such observations with the study design of Li *et al.*, but with a larger sample of individuals, Li *et al.* may have been better able to assess the prevalence of mutations in different genes.

Both studies found that, with the exception of the chromatin remodeling genes, most other genes that are commonly mutated in urothelial cancers were rarely mutated in normal urothelium. These include the well-established cancer driver genes *PIK3CA* (phosphatidylinositol 4,5-bisphosphate 3-kinase catalytic subunit α), *FGFR3* (fibroblast growth factor receptor 3), and *RB1* (RB transcriptional corepressor 1). Thus, mutations in these driver genes might be later events that finally trigger malignant transformation of cells harboring mutations that drive clonal expansion (see the figure). Both studies also found few large-scale copy number alterations in the normal urothelium, in contrast to the abundant large-scale

Centre for Computational Biology and Programme in Cancer and Stem Cell Biology, Duke–National University of Singapore (NUS) Medical School, 169857 Singapore. Email: steve.rozen@duke-nus.edu.sg

Mutations in the bladder and ureter linings

Cells in the normal urothelium that lines the ureter and bladder accumulate mutations caused by endogenous mutagenic processes (e.g., APOBEC cytidine deaminases) or by exogenous mutagens (e.g., AA or tobacco smoke). Some mutations (e.g., in the chromatin modifying genes *KMT2D* and *KDM6A*) confer competitive advantages that drive cells to colonize larger regions of the urothelium. Additional mutations (e.g., in the genes *TP53*, *PIK3CA*, *FGFR3*, or *RB1*) and perhaps other changes are needed to trigger malignant transformation.



AA, aristolochic acid; *FGFR3*, fibroblast growth factor receptor 3; *KDM6A*, lysine-specific demethylase 6A; *KMT2D*, histone-lysine *N*-methyltransferase 2D; *PIK3CA*, phosphatidylinositol 4,5-bisphosphate 3-kinase catalytic subunit α; *RB1*, RB transcriptional corepressor 1; *TP53*, tumor protein 53.

GRAPHIC: MELISSA THOMAS BAUM/*SCIENCE*

copy number alterations found in urothelial cancers. Lawson *et al.* detected copy number alterations in only 28% of exomes, whereas the average bladder cancer has 200 per exome. Thus, genome instability occurs late in the malignant transformation of the urothelium, a pattern also seen in skin, esophageal, and colon transformation (*1, 3, 5*).

Slowing the accumulation of mutations in normal tissue might slow aging and reduce the risk of cancer, and therefore it is important to understand the causes of these mutations. The two studies analyzed their respective sequencing data for "mutational signatures" that might point to these causes. Mutational signatures are the patterns of single base mutations within distinct sequences of preceding and following bases that can distinguish various mutational processes. Both studies identified the mutational signature caused by APOBEC cytidine deaminase activity in more than half of the individuals studied, although it was not found in every sample from those individuals. This signature is almost always found in bladder cancer and often in cancer of the ureter; it is also found in many other cancer types (*11*). The two studies show that increased APOBEC activity seems to accompany malignant transformation in urothelial cancer. The mechanisms responsible for this activation, and whether it is a cause or an effect of transformation, are unknown.

Lawson *et al.* sequenced the entire genomes of a subset of the urothelial samples, which enabled them to uncover previously unknown mutational signatures. One of these signatures correlated with tobacco smoking. This may solve the puzzle that no specific smoking-associated mutational signature has been identified in bladder cancer, even though tobacco smoking is a well-established risk factor for bladder cancer and causes a highly recognizable signature in lung cancer. The smoking signature in lung cancer bears no resemblance to the smoking-associated signature identified by Lawson *et al.*, and the mechanisms by which tobacco might cause it are unknown. Li *et al.* did not detect this signature, possibly because they analyzed whole exomes rather than whole genomes.

A prominent difference between the mutational signatures detected in the two studies is the presence of the signature caused by aristolochic acid (AA) in more than half of the ureter samples and about a third of the bladder samples examined by Li *et al.* AA is a kidney toxin and carcinogen that occurs naturally in some herbs used as medicine. AA exposure is widespread in East Asia, including China (the location of patients studied by Li *et al.*), but very limited in Europe (the location of individuals studied by Lawson *et al.*) (*11–13*). AA mutagenesis was previously detected in a few normal urothelial samples (*13*, *14*), and Li *et al.* now show that AA mutagenesis is widespread in normal urothelium. This suggests that it may be possible to reliably assess the AA mutational signature in cells or DNA shed from normal urothelium and, using this signature, noninvasively assess previous AA exposure (*13*). A noninvasive test for AA exposure would offer substantial benefits to research into the epidemiology of AA-associated disease and to secondary prevention of cancer and kidney failure in AA-exposed individuals.

An unexpected finding was the strong difference in driver mutation preferences between individuals, with, for example, one person having multiple independent mutations in *KDM6A* and few in *ARID1A*, and another person with the opposite pattern. It will be interesting to see how general this phenomenon is and whether driver preferences can predict cancer risk or, very speculatively, suggest prophylactic therapies. Each study explored mutations in a different dimension. One studied one or two samples from a broad swath of individuals, and the other was a deep study of many samples from a few individuals. To understand the implications of differences in preferences of driver gene mutations between individuals, a study that is simultaneously broad and deep is required. Nevertheless, the shared message of Li *et al.* and Lawson *et al.* is that histologically normal urothelium contains many clones that are only a few steps away from turning malignant but that rarely do. Studies using mouse models of the development of urothelial cancers, similar to a recent study of the development of esophageal cancer (*15*), would further delineate the mechanisms by which mutations in the normal human urothelium sometimes give rise to cancer. ■

REFERENCES AND NOTES

1. I. Martincorena *et al.*, *Science* **348**, 880 (2015).
2. M. A. Lodato *et al.*, *Science* **359**, 555 (2018).
3. I. Martincorena *et al.*, *Science* **362**, 911 (2018).
4. S. F. Brunner *et al.*, *Nature* **574**, 538 (2019).
5. H. Lee-Six *et al.*, *Nature* **574**, 532 (2019).
6. A. Yokoyama *et al.*, *Nature* **565**, 312 (2019).
7. L. Moore *et al.*, *Nature* **580**, 640 (2020).
8. K. Yoshida *et al.*, *Nature* **578**, 266 (2020).
9. A. R. J. Lawson *et al.*, *Science* **370**, 75 (2020).
10. R. Li *et al.*, *Science* **370**, 82 (2020).
11. L. B. Alexandrov *et al.*, *Nature* **578**, 94 (2020).
12. A. W. T. Ng *et al.*, *Sci. Transl. Med.* **9**, eaan6446 (2017).
13. H. Lu *et al.*, *Theranostics* **10**, 4323 (2020).
14. Y. Du *et al.*, *Eur. Urol.* **71**, 841 (2017).
15. B. Colom *et al.*, *Nat. Genet.* **52**, 604 (2020).

ACKNOWLEDGMENTS

This work was supported by Singapore National Medical Research Council grant MOH-000032/MOH-CIRG18may-0004 and by the Singapore Ministry of Health via the Duke-NUS Signature Research Programmes. I thank A. Boot for comments.

10.1126/science.abe0955

REGENERATION

Coaxing stem cells to repair the spinal cord

Spinal cells in mice can be induced to generate protective oligodendrocytes after injury

By **Catherina G. Becker** *and* **Thomas Becker**

Spinal cord injuries destroy neurons, axonal processes, and oligodendrocytes that provide insulation and protection of axons by means of membrane wrappings, called myelin sheaths. None of these cellular structures are efficiently replaced after injury. This can lead to lifelong disabilities, including paralysis. Endogenous stem cells exist in the spinal cord, but after injury they produce mainly astrocytic scar tissue, no neurons, and very few oligodendrocytes (*1*). On page 73 of this issue, Llorens-Bobadilla *et al.* (*2*) show that by overexpressing a single factor in spinal stem cells, they can boost post-injury production of oligodendrocytes in mice. This leads to better remyelination of remaining axons that lost their myelin and to improved axonal impulse conduction in vivo. The study raises hope that endogenous stem cells in the injured spinal cord can be recruited to generate neural cell types in a more balanced way after injury to promote recovery of function.

One of the problems with spinal cord injury is secondary cell death around an injury site that leads to the loss of not only neurons, but also oligodendrocytes and the myelin sheaths they produce. This in turn causes denuding of spared axons, which, bereft of their insulation and trophic support, function inefficiently and may ultimately degenerate. Llorens-Bobadilla *et al.* analyzed mouse ependymal cells (stem cells in the lining of the spinal cord central canal) to find that chromatin regions with binding motifs for the oligodendrocyte-determining transcription factors OLIG2 (oligodendrocyte transcription factor 2) and SOX10 were accessible even though the transcription factors were not expressed. This suggested a latent capacity of ependymal cells to generate oligodendrocytes. Indeed, inducing overexpression of OLIG2 in ependymal cells in vivo strongly increased the accessibility of OLIG2 binding sites and the production of oligodendrocytes from these cells after spinal injury.

Neither promoter accessibility nor oligodendrocyte production increased without an injury, indicating that factors in addition to OLIG2 expression are necessary to realize the latent potential of ependymal cells for oligodendrocyte production. Injury induces ependymal cells to proliferate, which changes gene accessibility. Moreover, spinal injury induces inflammation (*3*) and attendant signaling molecules that may influence gene expression in ependymal cells, but these factors are largely unknown.

The observed increase in oligodendrocyte numbers was substantial. High numbers of cells may have been reached because the progeny of ependymal cells comprises proliferative oligodendrocyte progenitors that can be considered transit-amplifying cells for oligodendrocytes. Even though oligodendrocyte production was boosted from less than 1% to more than one-third of ependymal progeny, the astrocytic scar, which also consists of cells derived from ependymal cells, was not depleted. The astrocytic scar has protective functions in wound healing (*4*), so it is important that any therapeutic approach does not compromise scar tissue. Together, these findings indicate that ependymal cells in the spinal cord can be reprogrammed for oligodendrocyte production after injury in mice.

Full differentiation of oligodendrocytes is a prerequisite for efficient impulse propagation in axons. One problem with naturally occurring remyelination is that new myelin sheaths are usually thinner than those that were originally present, compromising conduction velocity. It will be interesting to find out whether new myelin displays "full thickness" and is comparable to myelin that has arisen during development. It will also be interesting to determine whether new myelin persists for longer than the 3 months reported in this study to indicate permanent repair.

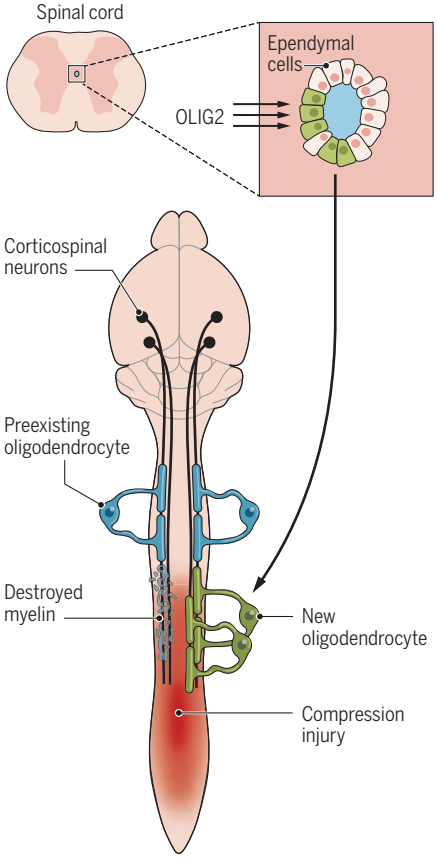
Llorens-Bobadilla *et al.* found that induced myelin improved conductance velocity in a mouse model of spinal cord injury. The model is a contusion injury, which resembles the physical impact associated with injuries in humans. Conductance velocity in spared axons above the injury site was improved in OLIG2-overexpressing animals compared to injured controls, ostensibly because of improved myelination. Below the injury, no such effects were observed, likely because of the scarcity of spared axons. Consequently, recovery of function was not better in OLIG2-overexpressing animals.

To bring about functional recovery, combinations with other treatments would be needed (*5*) because although oligodendrocytes can support axons, remyelination alone is insufficient to boost recovery (*6*). For example, axon growth can be enhanced by modulating intrinsic axon growth propensity (*7*) or the inhibitory environment (*8*). Moreover, transplanted neural stem cells can form new neurons that integrate and improve neuronal impulse conduction over an injury site (*9*). Electrical stimulation of spared fibers may also lead to more efficient myelination (*10*). In all approaches, newly grown axons could benefit from additional myelination capacity from induced oligodendrocytes.

To realize the potential of ependymal cells in human spinal cord injuries, it will be nec-

Induced protection

Chromatin in adult ependymal cells in the mouse spinal cord is accessible to the transcription factor OLIG2 (oligodendrocyte transcription factor 2). After spinal cord injury, myelin is lost near the injury site and can be lost from spared axons. Overexpression of OLIG2 in ependymal cells in mice reprograms these stem cells to produce more oligodendrocytes, which form new myelin near the injury and improve axon conduction.



GRAPHIC: KELLIE HOLOSKI/*SCIENCE*

University of Edinburgh, Edinburgh, Scotland, UK. Email: catherina.becker@ed.ac.uk; thomas.becker@ed.ac.uk

essary to determine whether similar stem cells exist near the human spinal cord central canal in sufficient numbers (*11*). The lumen of the central canal progressively disappears during childhood, but cells with stem cell potential can be isolated from the human spinal cord (*12*). It will be necessary to determine how similar these cells are to mouse ependymal cells and whether their chromatin is similarly poised to generate oligodendrocytes.

From a therapeutic viewpoint, efficient means of inducing gene expression are needed. Viral delivery systems are under development that evade immune detection and can be switched off—an important safety feature to avoid unwanted proliferation of cells (*13*). Such strategies could be used to treat any demyelination after spinal injury or in demyelinating diseases, such as multiple sclerosis.

Ependymal cells in the mouse spinal cord have a stem cell potential that sets them apart from other spinal cell types. For example, Llorens-Bobadilla *et al.* found that astrocytes did not produce oligodendrocytes after forced OLIG2 expression. What else could ependymal cells do? In rats, stem cells derived from whole spinal cord, including ependymal cells, formed neurons when transplanted into the hippocampus, a neurogenic region of the brain (*14*). In anamniotes (salamanders and fishes), spinal ependymal progenitor cells generate neurons after injury in situ (*12*). This indicates a potential for neurogenesis in spinal stem cells across vertebrates. It might be informative to determine the nature of the gene regulatory programs activated by anamniotes to produce neurons from ependymal cells in the spinal cord and whether these could be activated by gene therapy in nonregenerating mammals to contribute to repair after injury. ■

REFERENCES AND NOTES

1. K. Meletis *et al.*, *PLOS Biol.* **6**, e182 (2008).
2. E. Llorens-Bobadilla *et al.*, *Science* **370**, eabb8795 (2020).
3. A. D. Greenhalgh, S. David, F. C. Bennett, *Nat. Rev. Neurosci.* **21**, 139 (2020).
4. M. A. Anderson *et al.*, *Nature* **532**, 195 (2016).
5. G. Courtine, M. V. Sofroniew, *Nat. Med.* **25**, 898 (2019).
6. G. J. Duncan *et al.*, *Nat. Commun.* **9**, 3066 (2018).
7. J. Ruschel *et al.*, *Science* **348**, 347 (2015).
8. B. T. Lang *et al.*, *Nature* **518**, 404 (2015).
9. S. Ceto, K. J. Sekiguchi, Y. Takashima, A. Nimmerjahn, M. H. Tuszynski, *Cell Stem Cell* **27**, 430 (2020).
10. S. Mensch *et al.*, *Nat. Neurosci.* **18**, 628 (2015).
11. D. Garcia-Ovejero *et al.*, *Brain* **138**, 1583 (2015).
12. C. G. Becker, T. Becker, J.-P. Hugnot, *Prog. Neurobiol.* **170**, 67 (2018).
13. E. R. Burnside *et al.*, *Brain* **141**, 2362 (2018).
14. L. S. Shihabuddin, P. J. Horner, J. Ray, F. H. Gage, *J. Neurosci.* **20**, 8727 (2000).

ACKNOWLEDGMENTS

Supported by the EU Cofund ERANET NEURON consortium NEURONICHE with contributions from MRC (MR/R001049/1), Spinal Research, and Wings for Life (C.G.B.) and by BBSRC project grant BB/R003742/1 (T.B.).

10.1126/science.abe1661

GRAPHIC: KELLIE HOLOSKI/*SCIENCE*

METALLURGY

A rival to superalloys at high temperatures

Slip-pathway activation provides plasticity in a multiprincipal element alloy with high-temperature strength

By **Julie Cairney**

Although conventional alloys are based mainly on one element, recent design efforts have focused on multiprincipal element alloys (MPEAs) that contain substantial quantities of several elements. Success with this approach requires a robust understanding of the mechanistic origin of MPEA properties. On page 95 of this issue, Wang *et al.* (*1*) report the deformation behavior in a promising body-centered cubic (bcc) MPEA with good room-temperature plasticity and high strength at the temperatures at which conventional alloys would soften. They observed multiplanar, multicharacter dislocation slip that was not expected in bcc systems. This property is attributed to variations in the glide resistance along the core of dislocations, created by the atomic-scale fluctuations in composition that are characteristic of MPEAs. This mechanism explains the plasticity and could be used to guide the design of new MPEA candidates for high-temperature applications in aerospace and power generation.

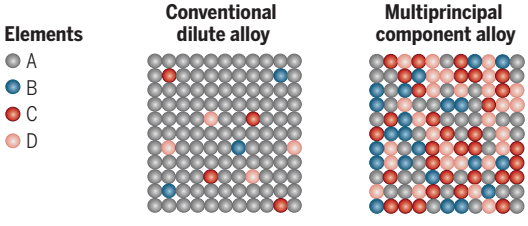
For centuries, alloy design has involved taking a base metal and adding small amounts of other elements to improve the properties. For example, adding carbon to iron enhances its strength, creating steel, and adding yet another element, nickel, improves its corrosion response, creating stainless steel. Sophisticated superalloys have complex compositions that provide high performance near their melting point, but they are still based on a primary element, usually nickel, cobalt, or iron.

To expand the alloy design space, more recent efforts have shifted toward the development of alloys that contain substantial quantities of three or more elements (see the figure). These materials are variously referred to as MPEAs, complex concentrated alloys, or high-entropy alloys (a subset of MPEAs). Some of these new alloys display unprecedented combinations of strength, ductility, high-temperature performance, or functional properties (*2*, *3*). Each new composition can result in the formation of different phases and microstructures within the alloy, which can in turn be altered by mechanical deformation and heat treatment. With such an enormous range of possibilities, traditional trial-and-error approaches are ineffective, and researchers are turning to computational and combinatorial approaches to predict the elemental combinations that could lead to alloys with desirable properties (*4*, *5*). However, for these approaches to be successful, it is critical that the alloy design process is guided by an understanding of the origins of the specific properties that are desired.

Refractory MPEA alloys, with their excellent high-temperature strength, show great promise. They are composed of combinations of three or more of the elements chromium (Cr), molybdenum (Mo), niobium (Nb), vanadium (V), tantalum (Ta), tungsten (W), hafnium (Hf), titanium (Ti), or zirconium (Zr) at nearly equal concentrations. The alloys usually have a bcc crystal structure and, for many of the most desirable combinations, are single-phase solid solutions. However, unlike dilute bcc alloys, many of these alloys show excellent retention of strength up to 1900 K (*6*). Aerospace, petrochemical, and power-

A more equal union

The schematics represent the atomic structure of conventional and multifunctional alloys. Conventional alloys are based mainly on a single element, compared with multicomponent alloys that contain several elements in similar proportions.



School of Aerospace, Mechanical and Mechatronic Engineering, University of Sydney, Sydney, Australia. Email: julie.cairney@sydney.edu.au

generation industries all require tough components that are exposed to high temperatures. Superalloys, the best available option, have an operational limit of around 1100 K (*2*). This material constraint affects the efficiency of potential new technologies for power generation by limiting operating temperatures (*7*). It is also a serious issue in aerospace applications. Presently, aircraft parts that are exposed to hot environments in engines, or components that will be exposed to the high temperatures caused by hypersonic travel, require complex and expensive ceramic thermal barrier coatings to withstand the service environment (*8*).

MoNbTaW, MoNbTaVW (*9*), HfNbTaTiZr (*10*), and HfMoNbTiZr (*11*) are all examples of MPEAs that display excellent high-temperature strength (*1*). However, many of these systems have limited room-temperature ductility, which is characteristic of bcc alloys. The lack of ductility in conventional bcc systems is related to the mobility of defects (dislocations), but dislocations could behave differently in bcc MPEAs because of local variations in composition along their core (*2*, *6*). Local compositional fluctuations are intrinsic to MPEAs, in which the elements that surround each individual atom vary (see the figure, right). The high-temperature strength in MPEA alloys has been attributed to solid-solution strengthening by regions of concentrated solute, the mobility of certain dislocation structures, or both effects (*12*). Understanding the details of the dislocation structure and motion is crucial for a mechanism-guided search for the best refractory bcc alloys across the immense range of possible compositions.

Wang *et al.* present a new MPEA MoNbTi alloy with good room-temperature strength but considerably lower density than many of the other refractory MPEA options. Density is important for transport applications, especially in rotating parts, where lower density increases the allowable service temperature by decreasing the stress caused by self-loading. The alloy displays homogeneous plasticity in microscale tension tests performed at room temperature in a scanning electron microscope. To understand the deformation process, Wang *et al.* used a sophisticated experimental setup that combines microscale mechanical testing and advanced microscopy. They used a focused ion beam (FIB) to prepare cross sections underneath nano-indents and performed in situ deformation on a single-crystal specimen, again prepared with a FIB. The tensile axis was aligned with the [001] direction so that the four 1/2⟨111⟩–type Burgers vectors are equally stressed. They observed unexpected nonscrew dislocation structures. Multiple slip systems appear to be operative in addition to those expected for a bcc alloy at room temperature. Wang *et al.* attribute this multiplanar, multicharacter dislocation slip to variations in the glide resistance for dislocations caused by the atomic-scale chemical fluctuations in composition along the core of the dislocation. Atomistic simulations show that the plane that has the lowest stress required for the movement of dislocations can vary in this system depending on the local atomic configuration.

The implication of this finding is that there are additional pathways for dislocation slip, which is desirable for plasticity and toughness. This observation explains the plasticity and supports an explanation for high-temperature strength based around the slip mechanism instead of solid-solution strengthening. Deformation was studied at room temperature, and future work at high temperature may reveal more details about the active slip systems.

Activation of additional slip pathways as a design goal will require renewed effort to understand how atoms are arranged along the core of the dislocation. Despite suggestions of local chemical order in MPEAs, experimental verification has been ambiguous (*3*). In Wang *et al.*'s work, atom probe tomography suggested that the atoms were randomly distributed. A smaller number of species in this alloy, as compared with many other MPEAs, has reduced the complexity of analysis. Further atom probe work could be carried out on carefully chosen model MPEAs to minimize overlapping peaks in data. Last, if bcc MPEA alloys are truly to rival superalloys for high-temperature use, consideration must be given to factors beyond the strength, ductility, and toughness. This includes oxidation resistance, creep strength, fatigue strength, and routes for manufacture, offering directions for future research. ■

REFERENCES AND NOTES

1. F. Wang *et al.*, *Science* **370**, 95 (2020).
2. D. B. Miracle, O. N. Senkov, *Acta Mater.* **122**, 448 (2017).
3. E. P. George, D. Raabe, R. O. Ritchie, *Nat. Rev. Mater.* **4**, 515 (2019).
4. A. Abu-Odeh *et al.*, *Acta Mater.* **152**, 41 (2018).
5. T. Borkar *et al.*, *Acta Mater.* **116**, 63 (2016).
6. O. N. Senkov, S. Gorsse, D. B. Miracle, *Acta Mater.* **175**, 394 (2019).
7. Y. Ahn *et al.*, *Nucl. Eng. Technol.* **47**, 647 (2015).
8. B. Liu *et al.*, *J. Mater. Sci. Technol.* **35**, 833 (2019).
9. O. N. Senkov, G. B. Wilks, J. M. Scott, D. B. Miracle, *Intermetallics* **19**, 698 (2011).
10. O. N. Senkov *et al.*, *J. Mater. Sci.* **47**, 4062 (2012).
11. N. N. Guo *et al.*, *Mater. Des.* **81**, 87 (2015).
12. F. Maresca, W. A. Curtin, *Acta Mater.* **182**, 235 (2020).

ACKNOWLEDGMENTS

The author thanks M. Griffith for feedback on the manuscript and acknowledges funding from the Australian Research Council Future Fellowship (FT180100232).

10.1126/science.abd6587

CONSERVATION

A boost for freshwater conservation

Integrating freshwater and terrestrial conservation planning has high returns

By **Robin Abell** and **Ian J. Harrison**

Systematic conservation planning—a data-driven process for prioritizing biodiversity conservation resources—has been strongly biased over the past two decades toward terrestrial and marine species and systems (*1*). Freshwater ecosystems, which are among the most threatened on Earth, have received less attention. Wetland extent is estimated to have declined globally by nearly 70% since 1900 and, on average, freshwater vertebrate populations declined by 84% between 1970 and 2016 (*2*). There is an urgent need for prioritizing resources toward freshwater conservation. On page 117 of this issue, Leal *et al.* (*3*) show that such prioritizations need not be a zero-sum game: Integrated cross-realm conservation planning can, for a negligible reduction in terrestrial benefits, increase freshwater benefits up to 600%.

These results are important because, at this moment, the global conservation community is setting targets for the next 30 years. The Convention on Biological Diversity (CBD) is defining its Post-2020 Global Biodiversity Framework, to replace the expiring Aichi Biodiversity Targets. There is an opportunity with the renewed CBD framework to create a policy environment and commitments designed to "bend the curve" of freshwater biodiversity loss (*4*).

Leal *et al.*'s findings, from their studies in Pará, Brazil, are of particular relevance to two targets in the draft framework, which will be negotiated during the first quarter of 2021 and finalized at the 15th meeting of the Conference of the Parties to the CBD, scheduled for May 2021. One target aims for a proportion of global land and sea areas to be under spatial planning, as a precursor to protecting and restoring natural ecosystems. The second aims for a propor-

Conservation International, Arlington, VA, USA.
Email: rabell@conservation.org

The Amazonas Lowlands ecoregion contains a rich and diverse fauna. The main types of vegetation include seasonally flooded forests. There are strong ecological links between the aquatic and terrestrial habitats. Rivers and forests supply multiple ecosystem services to the communities living in the region.

PHOTO: CLAUS MEYER/MINDEN PICTURES

tion of the planet to be conserved using a combination of conventional protected areas and other effective area-based conservation measures (*5*). The formulation of these targets may yet change, but the underlying intent will likely remain: We need to protect more of our planet than we currently do, and we need to use spatial planning to guide that protection.

Historically, protected areas have been designed and managed first for terrestrial ecosystems and their species, with freshwater considered as an afterthought, if at all. For example, at least 1249 large dams are located in protected areas, with the processes of "downgrading, downsizing, and degazettement" having legalized dam construction within many of these areas (*6*). A study of continental Africa (*7*) showed that the proportion of freshwater species whose ranges are substantially covered by protected areas and Ramsar sites (wetlands designated to be of international importance under the Ramsar Convention) is much smaller than for birds or mammals. Globally, about 70% of river reaches (by length) have no protected areas in their upstream catchments, rendering them partly conserved at best (*8*). These and other shortfalls can be traced back in part to a lack of freshwater-focused conservation planning and investment (*1*).

At the same time, as Leal *et al.* rightly observe, conventional protected areas are no silver bullet for conserving freshwater biodiversity. Instead, they are part of a toolbox of options that should also embrace a growing set of strategic adaptive management approaches for freshwater protection. Among these options is area-based protection of systems beyond protected areas, including lands with collective tenure rights for indigenous peoples and local communities.

Recommendations already exist for integrating freshwater conservation priorities into larger landscape strategies, in particular through entire watershed management plans (*9*, *10*). The International Union for Conservation of Nature's approach for identifying freshwater Key Biodiversity Areas (KBAs), which then serve as spatial conservation planning inputs to critical site networks, is one applied example [e.g., Lake Victoria, East Africa (*11*)]. These holistic approaches consider not only aquatic habitats of importance but also the landscapes in which these habitats are embedded, which strongly regulate flow, sediment, and nutrient regimes in rivers, lakes, and wetlands. Freshwater KBAs and similar efforts acknowledge the impossibility of securing freshwater species and systems without addressing the lands draining to them. Leal *et al.*'s study shows not only high reward for achieving freshwater conservation targets from such an integrated approach but also low risk to achieving paired terrestrial targets.

The good news from Leal *et al.*'s study—namely that when we conserve freshwater species and habitats, we receive multiple biodiversity cobenefits—also applies to the delivery of ecosystem services. Two-thirds of today's human population lives downstream from protected areas, and more than one-quarter of water provisions supplied by the world's protected areas are exposed to low levels of threat (*12*). There is a win-win from protecting important source water areas for human communities; an assessment of the likely source catchments of 4000 cities supplying water to up to 1.7 billion people showed that 85% of the area of these catchments overlaps with high biodiversity value freshwater ecoregions (*13*).

Achieving this wealth of benefits will require strong policy that recognizes the connections between terrestrial and freshwater systems and that treats those systems as equal in importance. And this strong policy must be matched with the resources to enact it. At present, freshwater systems receive a tiny percentage of overall environmental funding; for example, only 3.2% of the environmental funding provided by European foundations, and about 8% provided by North American foundations, goes to freshwater conservation in those regions or elsewhere (*14*). These funding challenges are likely to get worse, in the face of new economic crises and humanitarian needs in a world shaped by coronavirus disease 2019 and other potential pandemics. There are always trade-offs associated with any prioritization—in spatial planning, policy, and resource allocation—but the study by Leal *et al.*, combined with complementary analyses of the multiple benefits of healthy fresh waters and their watersheds, suggests that those trade-offs may be more acceptable than we think. ■

REFERENCES AND NOTES

1. S. Linke, E. Turak, M. G. Asmyhr, G. Hose, *Aquat. Conserv.* **29**, 1063 (2019).
2. C. Baker *et al.*, "A deep dive into freshwater: Living Planet Report 2020" (World Wildlife Fund, 2020).
3. C. G. Leal *et al.*, *Science* **370**, 117 (2020).
4. D. Tickner *et al.*, *Bioscience* **70**, 330 (2020).
5. CBD, Updated Zero Draft of the Post-2020 Global Biodiversity Framework; www.cbd.int/article/zero-draft-update-august-2020.
6. M. L. Thieme *et al.*, *Conserv. Lett.* **13**, e12719 (2020).
7. W. R. T. Darwall *et al.*, *Conserv. Lett.* **4**, 474 (2011).
8. R. Abell, B. Lehner, M. Thieme, S. Linke, *Conserv. Lett.* **10**, 384 (2017).
9. R. Abell, J. D. Allan, B. Lehner, *Biol. Conserv.* **134**, 48 (2007).
10. R. Flitcroft *et al.*, in *Freshwater Ecosystems in Protected Areas: Conservation and Management*, C. M. Finlayson, A. H. Arthington, J. Pittock, Eds. (Routledge, 2018), pp. 190–203.
11. C. A. Sayer, J. A. Carr, W. R. T. Darwall, *Fish. Manag. Ecol.* **26**, 435 (2019).
12. I. J. Harrison *et al.*, *Aquat. Conserv.* **26**, 103 (2016).
13. R. Abell *et al.*, *Aquat. Conserv.* **29**, 1022 (2019).
14. Synchronicity Earth Insight, "Freshwater"; www.synchronicityearth.org/wp-content/uploads/2018/02/Synchronicity-Earth-Freshwater-Insight.pdf.

10.1126/science.abe3887

POLICY FORUM

DATA SHARING

How to fix the GDPR's frustration of global biomedical research

Sharing of data for research beyond the EU must improve

By **Jasper Bovenberg[1], David Peloquin[2], Barbara Bierer[3], Mark Barnes[2,4], Bartha Maria Knoppers[5]**

Since the advent of the European Union (EU) General Data Protection Regulation (GDPR) in 2018, the biomedical research community has struggled to share data with colleagues and consortia outside the EU, as the GDPR limits international transfers of personal data. A July 2020 ruling of the Court of Justice of the European Union (CJEU) reinforced obstacles to sharing, and even data transfer to enable essential research into coronavirus disease 2019 (COVID-19) has been restricted in a recent Guidance of the European Data Protection Board (EDPB). We acknowledge the valid concerns that gave rise to the GDPR, but we are concerned that the GDPR's limitations on data transfers will hamper science globally in general and biomedical science in particular (see the text box) (*1*)—even though one stated objective of the GDPR is that processing of personal data should serve humankind, and even though the GDPR explicitly acknowledges that the right to the protection of personal data is not absolute and must be considered in relation to its function in society and be balanced against other fundamental rights. We examine whether there is room under the GDPR for EU biomedical researchers to share data from the EU with the rest of the world to facilitate biomedical research. We then propose solutions for consideration by either the EU legislature, the EU Commission, or the EDPB in its planned Guidance on the processing of health data for scientific research. Finally, we urge the EDPB to revisit its recent Guidance on COVID-19 research.

Concerns that gave rise to the GDPR include that data subjects be informed of use of their personal data and be afforded appropriate rights with respect to the use of their data, and that data users be required to follow certain standards in processing those data. But balancing these concerns against the concerns over research should be informed by the generally scientific research–friendly approach of the GDPR. Current interpretations of the GDPR fail to recognize how research uses of personal data differ from other uses, particularly because data used for research purposes are often pseudonymized, used to derive generalizable knowledge that can benefit society, and can be used in this way without identification of, or perceptible harm to, data subjects. Thus, the balance between privacy of the individual and the benefit to society in the research context is different than in other contexts, such as many commercial contexts in which data are used to construct a profile of an individual to permit targeted advertising with demonstrable impact on the individual.

GLOBAL SHARING OF RESEARCH DATA

The rationale behind the GDPR's limitations on transfers of data outside the EU is simple: When personal data are transferred to non-EU countries, the level of protection ensured in the EU should not be undermined. The limitations aim to ensure that the "GDPR travels with the data." Several routes for valid transfer of research data have been proposed, which we discuss below.

Data may be transferred on the basis of "an adequacy decision." This means that the European Commission has decided that the third country or international organization in question ensures an "adequate level of protection." Such a transnational data transfer does not require any specific authorization. However, to date, adequacy decisions are in place for only a limited number of countries worldwide: Andorra, Argentina, Canada (commercial organizations), Israel, Japan, New Zealand, Switzerland, Uruguay, and the self-governing dependencies of the Isle of Man, Guernsey, Jersey, and the Faroe Islands. The adequacy decision that was in place for the United States, the EU-U.S. "Privacy Shield" framework, was available only to for-profit organizations and today can no longer be used, as it has been invalidated by the recent decision of the CJEU (*2*).

Standard contractual clauses, which bind data transferees to comply with certain data protection standards when they receive and process personal data, are commonly used for cross-border transfer in the commercial context, but they pose particular difficulties for transfers to certain types of data recipients, including governmental agencies such as the U.S. National Institutes of Health or universities outside the EU. Such entities are often barred by their own national laws from agreeing to certain terms required to be included in the standard contractual clauses, including those specifying auditing of data systems by a foreign entity and submission to the jurisdiction of foreign courts (*3*). Many research entities that are arms of sovereign governments either lack authorization to waive their sovereign immunity or have a long-standing policy not to waive such immunity. Moreover, because the EU data transferors are often private universities or research institutes and transferees are governmental or parastatal entities, the individually negotiated interstate transfer agreements contemplated by the GDPR for transfers between two public bodies are not routinely available as an alternative to the standard contractual clauses (*4*).

Although the CJEU has upheld the validity of at least one set of the standard contractual clauses to permit cross-border data transfer, it has also ruled that a data exporter and the recipient of personal data using the clauses are required to verify, prior to any transfer, whether the level of protection required by EU law is respected in the importing country (*2*). It also made clear that recipients outside the EU must return any received data or destroy them "in their entirety" when their domestic laws no longer allow them to comply with the EU clauses (*2*). The verification must consider, as regards any access by public authorities of the importing country to the personal data transferred, the relevant aspects of the legal system of the importing country (*2*). Such an assessment on a case-by-case basis (and its monitoring on an ongoing basis) will probably be beyond the capabilities of most, if not all, EU researchers and their institutions. In essence, this requires resource-limited private parties to undertake the adequacy assessment process that would typically be done by the European Commission.

Even if researchers would somehow be able to complete such an assessment (and to monitor it on a going-forward basis), the standard contractual clauses present complications for multi-party research collaborations, when the recipient organization needs to share the data with other organizations

[1]Legal Pathways Life Sciences Law, Haarlem, Netherlands. [2]Ropes & Gray LLP, Boston, MA, USA. [3]Multi-Regional Clinical Trials Center of Harvard University and Brigham and Women's Hospital, Cambridge, MA, USA. [4]Yale Law School, New Haven, CT, USA. [5]Centre of Genomics and Policy, Faculty of Medicine, McGill University, Montreal, Quebec, Canada. Email: jabovenberg@xs4all.nl

in their own country or in another "third country" in order to complete the research. Unfortunately, the standard contractual clauses are not clear regarding such "onward transfers" and the mechanisms they offer; unambiguous consent by the data subject to the onward transfer, or adherence to the clauses by the onward transferee(s), are often not viable options.

Although the GDPR provides that entities may enter into bespoke clauses that are tailored to the circumstances, such bespoke clauses must be approved by the national competent supervisory authority (*5*). Yet in many EU jurisdictions, the lack of guidance from the EDPB on the requirements for bespoke clauses means that the competent authorities have not yet established a process for the review of bespoke clauses [e.g., (*6*)].

For prospective research, such as interventional clinical trials, in which data subjects provide informed consent at the time they enroll in the study, researchers have often relied on the explicit consent of the data subject as the means to legitimize data transfer. However, under the GDPR, this "transfer consent" is subject to a number of requirements and limitations. The researcher must inform data subjects about the possible risk that their personal data will be transferred to a country for which there is no adequacy decision or appropriate safeguards. Pursuant to Guidance from the EDPB, invoking data subjects' consent as a basis for transfer is limited to occasional and "not repetitive" transfers. Consent therefore is not a viable option for research consortia, data repositories, and legacy collections that store data for the global research community. Also, the general GDPR requirements for a valid consent continue to apply, including that it must be free, informed, specific and explicit, and subject to immediate withdrawal by the data subject at any time (withdrawal of consent should be as easy as giving consent). Upon withdrawal of consent, processing must be stopped, unless there is another legal basis to continue.

In addition, since the advent of the GDPR, ethics committees have asked researchers to provide a detailed list of all countries that will receive data collected as part of the study (*7*). Yet at the outset of a data-intensive research study, it is usually not possible to know all of the countries to which data may be sent, given the large number of collaborating parties and service providers involved in multinational collaborative studies. Moreover, with respect to data gathered in an interventional clinical trial of a medicinal product, EDPB Guidance disfavors the use of consent as a legal basis and condition for the use of the data for research processing, as it does not believe that such consent can be freely given, and this logic might also be extended to consent asked for transfer of data out of the EU.

Recently, in at least one long-standing research collaboration involving the NIH, an EU research institute agreed to permit the transfer of genetic data from Finland to the United States on the basis that the transfer is necessary "for important reasons of public interest" (*8*). This recognized GDPR derogation to the prohibition on cross-border transfer of personal data requires that the "public interest … shall be recognized in Union law or in the law of the Member State to which the controller is subject" (*9*). Some examples of when this public interest provision may be relied upon include international data exchange between competition authorities, tax or customs administrations, financial supervisory authorities, services competent for social security matters, or for public health, such as tracing for contagious diseases and/or elimination of doping in sport (*10*). The EDPB, however, fails to provide clear guidelines and only complicates matters by stating that "the derogation only applies when it can also be deduced from EU law or the law of the Member State to which the controller is subject, that such data transfers are possible for important public interest purposes including in the spirit of reciprocity for international cooperation" (*7*). The EDPB claims that this transfer mechanism, as an exception to the requirement of an adequacy decision or appropriate safeguards, although not expressly limited to "occasional" or "not repetitive" transfers, must be interpreted restrictively.

In sum, the inability to find a suitable mechanism amid the above legal bases for transfer has stymied research collaborations between the EU and the rest of the world, resulting in the cessation or harmful delays of critical data flows (see the text box) (*1*).

ILLUSTRATION: MIXMAGIC/ISTOCK.COM

SHARING DATA FOR COVID-19 RESEARCH

Sadly, the EDPB Guidelines published on 21 April 2020 regarding data for COVID-19 research (*11*) lack both any sense of urgency and any consideration of the public good, and fail to take into account other fundamental rights, societal interests, and scientific considerations. They stress that consent must be "specific," the "derogations and limitations in relation to the protection of data used in research must apply only in so far as is strictly necessary," and that "the current COVID-19 outbreak does not suspend or restrict the possibility of data subjects to exercise their rights." The Guidelines go so far as to state that "storage periods" (timelines) must be set for COVID data. This is an inexplicable limitation, and arguably even violates the GDPR's exemption to the storage limitation for data processed for scientific purposes. Pandemic researchers need access to past, present, and future collections of human biospecimens and associated personal data, to anticipate new waves of infections and any new mutations. Guidance on the "compatibility presumption" for research, announced in January 2019, is simply being postponed, leaving the research community in the dark.

Luckily, the EDPB does consider COVID research as qualifying as an "important public interest" to allow international transfer of data. However, the EDPB also notes that this derogation may only be justified for "initial transfers," as a "temporary measure," and that repetitive transfers as part of a long-lasting research project cannot proceed under this derogation. Thus, even in the important context of COVID-19 research when the ability to transfer personal data across international borders for research purposes is urgently needed, the EDPB minimizes the ability of the research community to rely on derogations that are included in the very text of the GDPR.

BEYOND EUROPE IN THE PUBLIC INTEREST

The above limitations for transfer of research data outside the EU appear to be at odds with the generally research-friendly intent of the GDPR. During its drafting and negotiation, the European Parliament and the European Council paid extensive attention to research issues, which has resulted in several provisions that may facilitate the processing of personal data for scientific research. Processing data for research is deemed compatible with the initial purpose of data collection, and data may be stored longer if for research purposes. The GDPR explicitly allows data subjects to give general consent rather than specific consent for processing

Examples of biomedical research frustrated by the GDPR

Inefficient distributed analysis of international data

The International Genomics of Alzheimer's Consortium and the U.S.-based Alzheimer's Disease Sequencing Project based at the University of Pennsylvania have been unable to pool personal data on a single server because EU investigators believe that the GDPR prevents them from sharing the European personal data with U.S.-based researchers. This creates a scientifically compromised, inefficient, and more expensive distributed analysis of international Alzheimer's disease data because investigators must run identical analyses on segregated pools of data in different locations. This distributed analysis model both slows research and limits the scope of research projects in which they can engage.

Protections in place, but struggling to identify a transfer mechanism

European research centers used to send de-identified human genetic data to the Imputation Server hosted by the University of Michigan. The server has been certified by an outside auditor for conformance with recognized information technology security and privacy standards [National Institute of Standards and Technology (NIST)]. Measures are in place to secure physical security of the location, space, and equipment and for identification and authentication (logging in). Users upload their private data, which is not accessed by server administrators. Once imputation is complete, the results are encrypted and uploaded files are deleted. Server administrators do not have access to users' private encryption passwords. Measures are also in place for encryption of data during storage and transmission. Server administrators cannot access completed imputation data. Despite the measures and protections in place, EU centers are now unable to send their data for imputation to the Michigan Imputation Server, as they struggle to identify a viable transfer mechanism under the GDPR.

for research purposes. The GDPR also allows for an exception to the notice requirement when providing notice proves impossible or would involve a disproportionate effort—in particular, processing for scientific research purposes. The GDPR further exempts from the right of erasure personal data processed for scientific research purposes if erasure is likely to render impossible or seriously impair the achievement of the objectives of the processing. Furthermore, the GDPR explicitly provides for an exemption to the right to object when personal data are processed for scientific research purposes, and permits member states to enact derogations from various data subject rights in the research context. Notably, all exemptions are subject to appropriate safeguards for the rights and freedoms of the data subject, such as technical and organizational measures, including pseudonymization.

The common rationale behind these exceptions and exemptions is the notion that scientific research is a "public interest" and the notion that the GDPR should facilitate processing of personal data in the public interest. Likewise, the conduct of science has been acknowledged by EDPB as a legitimate interest. Missing, however, from the GDPR list of research-friendly provisions is an appreciation of the international dimensions of research and, consequently, a corresponding appropriate provision to enable scientific research data transfers across the globe. As the conduct of science is a global affair, research-friendly provisions for the sharing of data for science beyond the borders of Europe should be part and parcel of the GDPR. The GDPR legislature has failed to take this crucial aspect of sharing data with scientific collaborators around the globe into account when drafting the research provisions.

REFORM AND GUIDANCE

We suggest a number of solutions, in the form of GDPR reform per se, dialogue between the Commission and the EDPB and the Commission's global counterparts, or as part of the Guidance planned by the EDPB on the processing of health data for the purpose of scientific research. First, we recommend that the GDPR transfer mechanisms be expanded by adding processing necessary for scientific research as an express public interest, subject to appropriate safeguards, such as pseudonymization (coding), data protection by design and default, and the requirements of notice and choice [e.g., (*12*)]. This basis for global sharing of research data should also extend to onward transfers. Second, it should be clarified that pseudonymized data should not be considered personal data in the hands of an entity that does not possess the key needed to re-identify such data, as was understood by many researchers and institutions under the law preceding the GDPR (*1*, *13*, *14*). Third, as part of its ongoing modernization of the standard contractual clauses, the EU Commission should adopt specific standard contractual clauses for scientific biomedical research. These clauses should reflect the specific context, purposes, and practices of such transfers—for example, review of sharing or access requests by independent Data Access Committees. Fourth, the EDPB should (i) issue guidance for the approval by the competent supervisory authorities of bespoke clauses for specific research studies and (ii) issue guidance identifying when data processing for scientific research, if carried out outside of the EU by a non-EU entity, would fall under GDPR standards. Finally, with respect to COVID-19 research, we recommend that the EDPB revisit its Guidance on processing of health data for scientific research, to reaffirm the validity of broad consent and to clarify that the exemption for transfers of research data for important reasons of public interest is not restricted to time-limited, occasional, and nonrepetitive transfers with respect to COVID-19 research.

We believe that our recommendations can help to redress the unfortunate consequences created by the existing GDPR approach to international transfers of research data and will enable the biomedical research community to share data beyond the EU for scientific research, while ensuring a high level of protection for data subjects. ■

REFERENCES AND NOTES

1. R. Eiss, *Nature* **584**, 498 (2020).
2. CJEU Case C-311/18, 16 July 2020 ("Schrems II"), specifically paragraphs 104–105, 135–143, and 203.
3. European Commission, Standard Contractual Clauses; https://ec.europa.eu/info/law/law-topic/data-protection/international-dimension-data-protection/standard-contractual-clauses-scc_en.
4. GDPR, Articles 46(2)(a), 46(3)(b); see also EDPB, *Guidelines 2/2020 on Articles 46(2)(a) and 46(3)(b) of Regulation 2016/679 for Transfers of Personal Data Between EEA and non-EEA Public Authorities and Bodies* version 1 (18 January 2020).
5. GDPR, Article 46(3).
6. United Kingdom, Information Commissioner's Office, *Guide to the GDPR: International Transfers*; https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/international-transfers/.
7. GDPR, Recital 42, EDPB 2/2018 Guidelines on Derogations of Article 49 of Regulation 2016/679 (adopted 25 May 2018).
8. GDPR, Article 49(1)(d).
9. GDPR, Article 49(4).
10. GDPR, Recital 112.
11. EDPB, *Guidelines 03/2020 on the Processing of Data Concerning Health for the Purpose of Scientific Research in the Context of the COVID-19 Outbreak* (21 April 2020).
12. PHG Foundation of the University of Cambridge, *The GDPR and Genomic Data: The Impact of the GDPR and DPA 2018 on Genomic Healthcare and Research* (May 2020); www.phgfoundation.org/documents/gdpr-and-genomic-data-report.pdf.
13. United Kingdom, Information Commissioner's Office, *Anonymisation: Managing Data Protection Risk: Code of Practice* (November 2012); https://ico.org.uk/media/1061/anonymisation-code.pdf.
14. CJEU Case C-582/14 of 19 October 2016 ("Breyer").

ACKNOWLEDGMENTS

J.B., D.P., and M.B. provide legal counsel to the biomedical research community, inter alia on issues of data protection and data transfers. B.M.K. received funding from Genome Canada/Genome Quebec and under EU-CIHR grant agreements No. 825903 euCanSHare and No. 160202 EUCANCan.

10.1126/science.abd2499

Andrew Wakefield addresses the media outside the General Medical Council on 28 January 2010.

PHOTO: SHAUN CURRY/AFP/GETTY IMAGES

PUBLIC HEALTH

Flawed research and its enduring repercussions

A journalist recounts how he exposed problems with a study linking vaccines and autism

By **Paul A. Offit**

On 26 February 1998, the Royal Free Hospital in London held a press conference to announce that a study conducted by one of the hospital's clinicians would be published in *The Lancet*, one of the world's oldest and most prestigious medical journals. Sitting at the front of the room was the senior author, Andrew Wakefield, who explained that the combination measles-mumps-rubella (MMR) vaccine could cause developmental delays, including autism. Wakefield argued that the MMR vaccine suppressed the immune system in some children, freeing the measles vaccine virus to damage the intestine, which allowed encephalopathic proteins to enter the circulation, cross the blood-brain barrier, and destroy brain cells. He called for MMR vaccinations to cease until more research could be conducted.

The reviewer is director of the Vaccine Education Center and an attending physician in the Division of Infectious Diseases at Children's Hospital of Philadelphia, Philadelphia, PA 19104, USA. Email: offit@email.chop.edu

Wakefield became an international hero. A biopic starring British actor Hugh Bonneville portrayed him as a courageous man willing to speak truth to power. In the United States, Wakefield testified before the Congressional Committee on Government Reform and appeared on *60 Minutes* with Ed Bradley. Then along came Brian Deer, an investigative reporter working for *The Sunday Times*. Deer would become the first to expose the clinician's undisclosed financial associations and unearth troubling problems with the *Lancet* paper. In *The Doctor Who Fooled the World*, Deer recounts in vivid detail how he came to learn that Wakefield and his study were not what they appeared to be.

The Doctor Who Fooled the World
Brian Deer
Johns Hopkins University Press, 2020. 408 pp.

Deer reveals that children admitted to the Royal Free Hospital with developmental delays in the 1990s were subjected to a regime known as the Wakefield protocol, which entailed magnetic resonance imaging, electroencephalography, spinal taps, abdominal x-rays, blood tests, and intestinal biopsies—procedures often not indicated by the children's symptoms. "The Royal Free would become the Mecca, or Lourdes," he writes, "for the desperately questing families of developmentally challenged children."

Wakefield, we learn, had received £435,643 (the equivalent of $846,000 today) to conduct studies that would help build a legal case against MMR vaccine producers 2 years prior to the *Lancet* publication. And although he reported that the children in his study were referred to his hospital through routine channels, many came from an antivaccine group called JABS and the lawyer preparing to sue vaccine makers. In June 1997, further undermining the sentiment the physician would convey at the 1998 press conference ("It's a moral issue for me"), Wakefield submitted a patent for a product that claimed to treat so-called "autistic enterocolitis," rid the body of harmful toxins, and immunize safely against measles.

Deer reveals that Wakefield also misrepresented clinical, biological, and molecular data. Although the 12 children in his paper were described as having suffered developmental problems within 2 weeks of vaccination, for example, Deer discovered that some had begun displaying symptoms before receiving the MMR vaccine whereas others did not begin exhibiting symptoms until months afterward. Moreover, there were instances in which normal intestinal biopsy specimens were mischaracterized as colitis, and Wakefield's claim that the measles vaccine virus genome was present in intestinal epithelial cells of children with autistic enterocolitis was inconsistent and irreproducible.

As a consequence of these and other revelations, *The Lancet* retracted the paper, and Wakefield lost his license to practice medicine. Subsequent studies have shown that children who receive the MMR vaccine are at no greater risk of developmental delays than those who do not receive it. Nonetheless, the damage was done. The Wakefield study helped to accelerate the antivaccination movement that has imperiled children and led to the resurgence of once-controlled diseases.

Although many people think they know this now-infamous story, it is likely they are unaware of all its dramatic details. Curious lay readers and vaccine experts alike are sure to learn something worthwhile from Deer's well-chronicled account. ■

10.1126/science.abb3587

EDUCATION

Training tomorrow's scientists

Prioritizing STEM diversity, equity, and inclusion requires rethinking graduate education

By **Ashley Huderson**

How do you engage with an institution that was not designed with women or people of color in mind? How do you dismantle years of discrimination and unequal treatment? When addressing such questions, I am reminded of a saying my mother liked to use. When asked "How do you eat an elephant?" she would reply, "One bite at a time."

This approach works in many contexts but is especially applicable to the challenges faced by academic institutions as science, technology, engineering, and mathematics (STEM) fields aim to increase diversity and become more inclusive and equitable. Although the importance of these elements to the scientific enterprise is implied, many STEM educators hesitate to participate in discussions surrounding diversity, equity, and inclusion because they think they lack the knowledge and/or training needed to engage intelligently or because they fear that their efforts will not result in change. However, as institutions begin to prioritize diversifying STEM, faculty must consider how their teaching and mentoring approaches affect these issues. In *Equity in Science*, Julie Posselt demonstrates how we can take steps to mitigate systemic discrimination in STEM education.

Posselt examines existing equity, diversity, and inclusion efforts across a number of graduate STEM fields, including astronomy, physics, chemistry, geology, and psychology. Her case studies—which include an ethnographic assessment of a geology field course; a comparative analysis of the trajectories of high-diversity STEM graduates; and an overview of the barriers, lessons learned, and design of several STEM Ph.D. programs—reveal key ways that privilege and power operate in scientific organizations and have created a culture of exclusion and sameness. The book closes with targeted recommendations for how individuals, departments, and scholarly societies can create systemic and sustainable change.

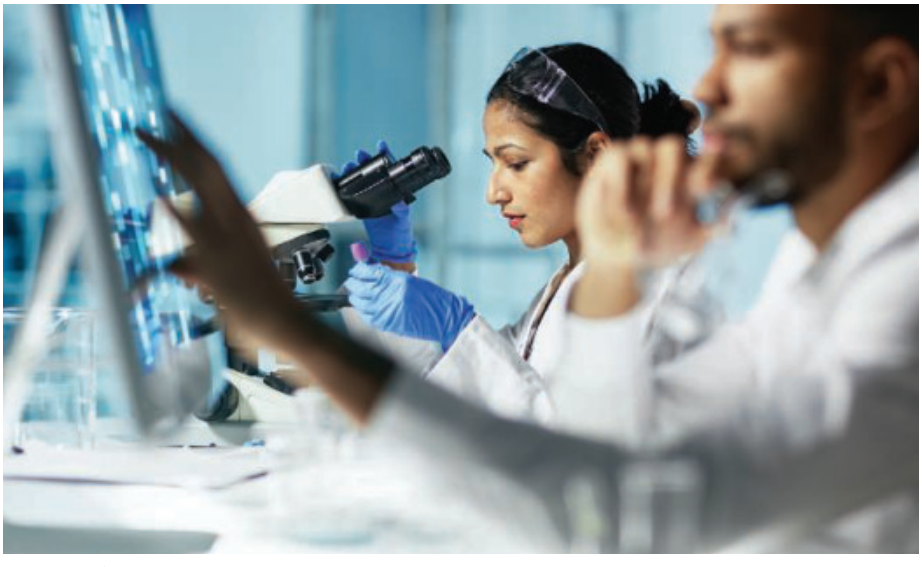
So how does Posselt suggest we dismantle our "elephant"? What does that first bite look like? She argues that we must revisit and reorganize the practices and priorities that have been socialized throughout STEM culture, and she encourages readers to reimagine the community's boundaries of what constitutes good science and to rethink the practices and the qualities we associate with good scientists. The only way to ensure that science is equitable, she argues, is to restructure scientific culture through a lens that respects and shows cultural differences and encourages those in nonmarginalized positions to acknowledge their power and privilege and the benefits that are conferred to certain groups. We must also center the voices, needs, and stories of people from marginalized groups.

Equity in Science
Julie R. Posselt
Stanford University Press, 2020. 240 pp.

Posselt argues that we must design STEM graduate programs with a diversity, equity, and inclusion lens that dismantles traditional community norms and values, including assumptions about scholastic ability, admission requirements, and curriculum structure. She lays out recommendations for the retention and recruitment of traditionally marginalized groups in graduate programs. These recommendations include downplaying or eliminating Graduate Record Examination scores as admission criteria; creating "bridge" programs that create a clear and intentional pathway through the STEM pipeline at various critical junctions; providing faculty mentors who share students' identities; and tracking program-level data, disaggregated by race or ethnicity and gender. She even provides recommendations for improving equity and inclusion within existing scientific collaborations, including ways to manage impervious and wayward colleagues. Here, she advises how to assess a collaborator's willingness to change and discusses how to overcome different types of resistance.

Posselt argues that advancing the movement for diversity, equity, and inclusion in science requires more effective collaboration across boundaries that typically separate scholars. She highlights how these collaborations tend to lie at the intersection of diverse identities, including gender, race, economic status, and discipline. Although its conclusions and recommendations are not exactly novel, the book succeeds in illustrating the depth to which diversity, equity, and inclusion are lacking at every level of STEM culture. ■

The reviewer is at the American Society of Mechanical Engineers, Washington, DC 20036, USA; the Department of Biology, University of the District of Columbia, Washington, DC 20008, USA; and STEM Innovation Consulting, Washington, DC 20018, USA.
Email: ahuderson@steminnovationconsulting.com

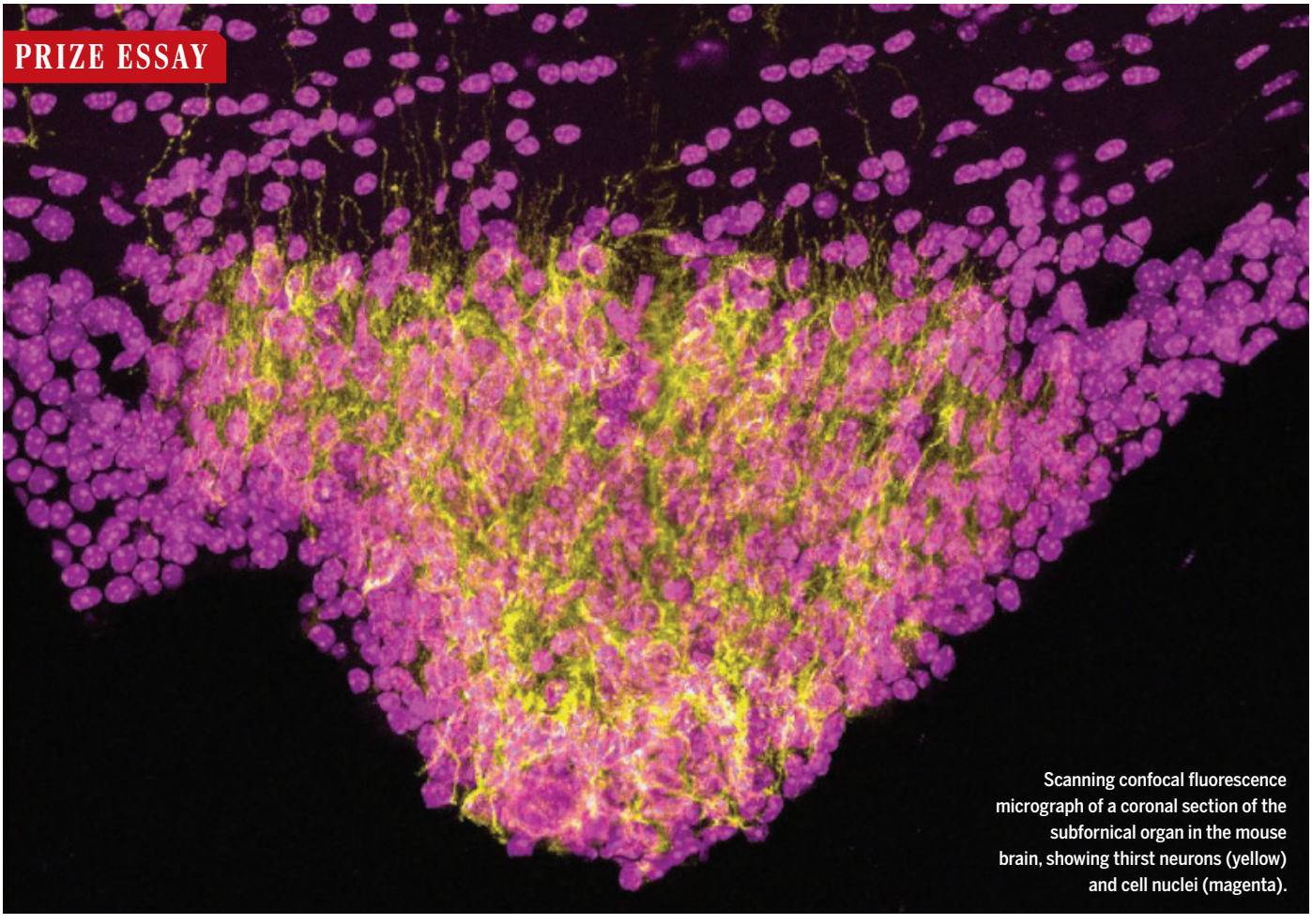


To increase STEM workforce diversity, we must dismantle norms that discriminate against marginalized groups.

PHOTO: ISTOCK.COM/POBA

10.1126/science.abd9804

PRIZE ESSAY



Scanning confocal fluorescence micrograph of a coronal section of the subfornical organ in the mouse brain, showing thirst neurons (yellow) and cell nuclei (magenta).

PHOTO: CHRISTOPHER ZIMMERMAN

NEUROBIOLOGY

The origins of thirst

Sensory signals arise throughout the body and converge in the brain to regulate drinking

By **Christopher A. Zimmerman**

We experience thirst every day, but where does this sensation come from? In the 1950s, Bengt Andersson proposed a tantalizing answer: Our brains might contain an "osmosensor" (*1*) that governs thirst, which consists of a group of cells that sense when we are dehydrated by directly monitoring the osmolarity of the blood. In a series of pioneering experiments, Andersson systematically infused salt into the brains of goats in an attempt to locate this osmosensor (*2*, *3*). He ultimately discovered a small area within the hypothalamus where even minute amounts of salt triggered immediate, voracious drinking. Subsequent studies established that Andersson's osmosensor encompasses the subfornical organ (SFO), a brain region that is distinctively suited to detecting blood osmolarity because it lies outside the blood-brain barrier (*4*).

The osmosensor model is powerful because it explains how dehydration generates thirst, but it has a crucial shortcoming: Drinking behavior is regulated on a fast, moment-by-moment basis that cannot be explained by slow changes in blood osmolarity. Consider that drinking immediately satiates thirst, even though the water imbibed is not absorbed for many minutes (*5*, *6*), and that eating stimulates prandial drinking long before the ingested food enters the bloodstream (*7*, *8*). How does the brain bridge these disparate time scales to dynamically adjust our sense of thirst?

I reasoned that we might gain new insight into this longstanding question by recording the activity of thirst-promoting neurons in living animals. My colleagues and I thus began by genetically labeling the SFO neurons that comprise Andersson's osmosensor and confirming that these cells are essential for dehydration-induced drinking (*9*). We then set out to observe the neural dynamics underlying thirst in behaving mice (*10*, *11*).

THIRST NEURONS ARE MORE THAN SIMPLE DEHYDRATION SENSORS

If SFO neurons are genuine osmosensors, then we would expect them to simply encode an animal's dehydration level. Consistent with this idea, our initial fiber photometry recordings demonstrated that these neurons are dose-dependently activated by increases in blood osmolarity (*10*).

It was therefore surprising to discover that SFO neurons are also rapidly regulated

Princeton Neuroscience Institute, Princeton University, Princeton, NJ, USA.
Email: czimmerman@princeton.edu

eppendorf & Science PRIZE FOR NEURO BIOLOGY

during eating and drinking, well in advance of any impact food and drink might have on the blood (*10*). For example, their activity decreases every time a mouse licks from a water bottle and increases with every bite of food. This counterintuitive finding indicated that SFO neurons—long viewed as merely passive sensors of dehydration—must receive a second class of signals that operate on the fast time scale of behavior.

LAYERS OF SIGNALS ARISE FROM THE DIGESTIVE TRACT DURING INGESTION

To pinpoint the origin of these signals, we traced the flow of water through the digestive tract of the mouse. We found that fluid detection in the mouth triggers a near-instantaneous inhibitory signal that closely tracks the volume ingested (*10*). Temperature sensing contributes to this process—SFO neurons are most efficiently inhibited by drinking cold water, a phenomenon that could be reproduced through isolated oral cooling. This may explain why we experience cold drinks as especially thirst-quenching and pleasurable (*12*, *13*).

Using an intragastric infusion paradigm, we next discovered that the osmolarity of ingested fluids is precisely measured in the gastrointestinal tract and then rapidly transmitted to the brain by the vagus nerve (*11*). This gut-to-brain osmolarity signal sustains the inhibition of SFO neurons produced by oral volume signals and satiates thirst if pure water is drunk. By contrast, detection of hypertonic fluids in the gut causes SFO activity to rebound to the "thirsty" state. Thus, drinking generates layers of signals that enable thirst neurons to predict how ingested fluids will affect hydration in the future and then adjust drinking preemptively. This simple model explains how drinking can rapidly quench thirst yet also be properly calibrated to match an animal's level of dehydration (*5*, *6*).

Does the body notify the thirst system about other behaviors that affect hydration? We found that eating triggers additional signals that activate SFO neurons in anticipation of food absorption (*10*). This activation drives prandial drinking or, if water is unavailable, suppresses further feeding. This suggests a neural basis for the widespread coordination of eating and drinking (*7*, *8*).

To test the causal role of the body-to-brain signals identified by our recording experiments, we used optogenetics to precisely manipulate each of them during behavior. This allowed us to confirm that these signals are necessary for thirst satiation, prandial thirst, and dehydration-induced anorexia (*10*, *11*), and thus account for most normal drinking behavior.

GRAND PRIZE WINNER

Christopher Zimmerman

Christopher Zimmerman received his undergraduate degrees from the University of Pittsburgh and a Ph.D. from the University of California, San Francisco. His thesis research focused on the neural mechanisms that govern thirst and drinking behavior. Zimmerman is currently a postdoctoral fellow at the Princeton Neuroscience Institute, where he continues to study the neural processes underlying motivated behaviors.

FINALIST

Tara LeGates

Tara LeGates received her B.S. in Biopsychology from Rider University and a Ph.D. from Johns Hopkins University. She completed a postdoctoral fellowship at the University of Maryland School of Medicine, where she established the importance of the strength and plasticity of hippocampus-nucleus accumbens synapses and reward behavior. LeGates is now an assistant professor at the University of Maryland, Baltimore County (UMBC). Her lab studies how neuronal circuits integrate information to regulate behavior and their alterations in psychiatric disorders. **www.sciencemag.org/content/370/6512/46.1**

FINALIST

Riccardo Beltramo

Riccardo Beltramo received his undergraduate degree from the University of Turin and a Ph.D. from the Italian Institute of Technology. After his doctoral training, Beltramo joined the Howard Hughes Medical Institute at the University of California, San Diego and the University of California, San Francisco, where he is completing his postdoctoral work. He studies sensory perception in the mouse visual system, focusing on understanding how cortical and subcortical neural circuits process visual information to drive behavior. **www.sciencemag.org/content/370/6512/46.2**

SIGNALS CONVERGE ONTO INDIVIDUAL NEURONS TO DYNAMICALLY ADJUST THIRST

The discovery of diverse inputs to SFO neurons raises the fundamental question of how signals are processed by the individual cells that comprise the thirst system. Do they flow in segregated "streams" or do they interact? To answer this question, we used microendoscopic imaging to track the activity of single neurons during dehydration, drinking, and intragastric infusion (*11*). This revealed a simple processing logic: The signals arising from the mouth, gut, and blood converge onto the same individual thirst neurons, thereby enabling every cell to continuously integrate information about current hydration status with the predicted consequences of ongoing ingestion.

In a parallel series of experiments, we showed that downstream brain regions use this integrated representation to coordinate the various components of the body's response to dehydration, including not only drinking but also cardiovascular adjustments, hormone secretion, and changes to emotional valence (*11*, *14*).

CONCLUSIONS

Thirst is governed by a sensory system, analogous to vision or hearing. Unlike these exterosensory systems, however, the neural dynamics underlying thirst were previously unknown. Our recordings revealed that thirst is regulated by layers of signals that arise throughout the body and converge onto individual neurons in the forebrain. This convergence occurs at the first node in the thirst system—the SFO—and generates a real-time estimate of the body's need for water that downstream nodes use to dynamically adjust drinking, valence, and cardiovascular physiology (*10*, *11*, *14*). Our findings reveal fundamental principles that govern ingestive behavior (*15*, *16*) and provide neural mechanisms that can potentially explain long-enigmatic elements of everyday human experience, including the speed of thirst satiation, the prevalence of drinking during meals, and the thirst-quenching power of oral cooling. ■

REFERENCES AND NOTES

1. E. B. Verney, *Proc. R. Soc. London Ser. B* **135**, 25 (1947).
2. B. Andersson, *Acta Physiol. Scand.* **28**, 188 (1953).
3. B. Andersson, S. M. McCann, *Acta Physiol. Scand.* **33**, 333 (1955).
4. M. J. McKinley *et al.*, *The Sensory Circumventricular Organs of the Mammalian Brain* (Springer, 2003).
5. R. T. Bellows, *Am. J. Physiol.* **125**, 87 (1938).
6. B. J. Rolls *et al.*, *Am. J. Physiol. Regul. Integr. Comp. Physiol.* **239**, R476 (1980).
7. T. J. Fitzsimons, J. Le Magnen, *J. Comp. Physiol. Psychol.* **67**, 273 (1969).
8. F. Bellisle, J. Le Magnen, *Physiol. Behav.* **27**, 649 (1981).
9. C. A. Zimmerman, D. E. Leib, Z. A. Knight, *Nat. Rev. Neurosci.* **18**, 459 (2017).
10. C. A. Zimmerman *et al.*, *Nature* **537**, 680 (2016).
11. C. A. Zimmerman *et al.*, *Nature* **568**, 98 (2019).
12. G. Kapatos, R. M. Gold, *Science* **176**, 685 (1972).
13. D. Boulze, P. Montastruc, M. Cabanac, *Physiol. Behav.* **30**, 97 (1983).
14. D. E. Leib *et al.*, *Neuron* **96**, 1272 (2017).
15. B. B. Lowell, *N. Engl. J. Med.* **380**, 459 (2019).
16. C. A. Zimmerman, Z. A. Knight, *Curr. Opin. Neurobiol.* **64**, 79 (2020).

10.1126/science.abe1479

PHOTOS: (TOP TO BOTTOM) SAMEER A. KHAN; MELISSA PENLEY CORMIER/RESEARCHGRAPHICS/UMBC; VICKY PAPAVASILEIOU

Where Science Gets Social.

AAAS.ORG/ COMMUNITY

AAAS' Member Community is a one-stop destination for scientists and STEM enthusiasts alike. It's "Where Science Gets Social": a community where facts matter, ideas are big and there's always a reason to come hang out, share, discuss and explore.

Member COMMUNITY
AAAS

The Most ADVANCED Patch Clamp System Available Today

Visit
www.dpatch.com
to learn more

SUTTER INSTRUMENT

+1.415.883.0128
info@sutter.com
www.sutter.com

MilliporeSigma

CRISPR-enabled drugs for one?

Whatever projects you're working on, from exploring mechanisms of disease to creating personalized therapies, you can count on our innovative Sigma-Aldrich® portfolio to drive your work forward. With cutting-edge products and deep technical expertise in genome engineering, we're here to help you make the next great breakthrough.

Unleash the impossible

To find out more, visit:
SigmaAldrich.com/UnleashTheImpossible

MilliporeSigma, the vibrant M and Sigma-Aldrich are trademarks of Merck KGaA, Darmstadt, Germany or its affiliates. All other trademarks are the property of their respective owners. Detailed information on trademarks is available via publicly accessible resources.

© 2020 Merck KGaA, Darmstadt, Germany and/or its affiliates. All Rights Reserved.

The Life Science Business of Merck KGaA, Darmstadt, Germany operates as MilliporeSigma in the US and Canada.

Sigma-Aldrich®
Lab & Production Materials

NEURODEGENERATION

A cruel end to too many lives

By **Stella Hurtley** *and* **Gemma Alderton**

Neurodegeneration is an all-too-common set of symptoms seen in people as they reach life's later stages. Most famously, neurodegeneration can lead to memory loss, robbing people of their quality of life and their essential personality.

Much fundamental research has led to a greater understanding of the pathophysiological processes involved in neurodegeneration and has revealed that accumulation of misfolded, toxic proteins is frequently a cause of neurodegeneration. This understanding is being refined toward the view that a common underlying principle of neurodegenerative misfolded proteins involves pathological phase transitions.

Many different disorders can cause neurodegeneration, but by far the most prevalent is Alzheimer's disease. Much is already known about the genetic underpinnings of susceptibility to Alzheimer's disease, but there is still a lot to learn about how to translate this knowledge into therapies.

Disrupted sleep often occurs before or in the early stages of neurogenerative disease. An important function of sleep is to allow waste products—including misfolded proteins associated with neurodegeneration—to be cleared from the brain via the glymphatic system.

Microglia, the major tissue-resident immune cells of the brain, directly modulate neuronal function in development and in neurodegenerative diseases such as Alzheimer's and Parkinson's diseases.

Elucidating the biological foundations of neurodegenerative disease will be key to the development of targeted therapies to slow or prevent the onset of dementia. In such late-onset diseases, delaying symptoms by even a few years can make an enormous difference to the quality of life of patients, as well as their families and caregivers.

INSIDE

REVIEWS

Glymphatic failure as a final common pathway to dementia p. 50

Beyond aggregation: Pathological phase transitions in neurodegenerative disease p. 56

Translating genetic risk of Alzheimer's disease into mechanistic insight and drug targets p. 61

Microglia modulate neurodegeneration in Alzheimer's and Parkinson's diseases p. 66

RELATED ITEM

PERSPECTIVE p. 32

ILLUSTRATION: SIMON PRADES

Neurodegeneration can stop people from being able to perform the most basic tasks in life. However, possibly the most difficult thing is the gradual loss of their most precious memories of loved ones.

REVIEW

Glymphatic failure as a final common pathway to dementia

Maiken Nedergaard[1,2]* and Steven A. Goldman[1,2]*

Sleep is evolutionarily conserved across all species, and impaired sleep is a common trait of the diseased brain. Sleep quality decreases as we age, and disruption of the regular sleep architecture is a frequent antecedent to the onset of dementia in neurodegenerative diseases. The glymphatic system, which clears the brain of protein waste products, is mostly active during sleep. Yet the glymphatic system degrades with age, suggesting a causal relationship between sleep disturbance and symptomatic progression in the neurodegenerative dementias. The ties that bind sleep, aging, glymphatic clearance, and protein aggregation have shed new light on the pathogenesis of a broad range of neurodegenerative diseases, for which glymphatic failure may constitute a therapeutically targetable final common pathway.

Little can replace the rejuvenating feeling of a good night's sleep. Our mood and affect, as well as our ability to attend, focus, and problem-solve, are all directly linked to how well we sleep. The benefits of sleep are cumulative; they are not restricted to the morning hours or even to a given day. Good sleepers live longer, weigh less, have a reduced incidence of psychiatric disorders, and remain cognitively intact longer (*1*–*4*).

Why do we sleep?

The idea that our brains rest during sleep to preserve energy was both posited and rejected in the 1950s, when electroencephalographic (EEG) recordings of brain activity made it clear that rapid eye movement (REM) sleep, which comprises ~20% of normal sleep, is linked to cortex-wide neuronal activation (*5*, *6*). Indeed, energy consumption declines by only 15% in the remaining non-REM (NREM) periods of sleep. Borbély proposed 40 years ago that the sleep-wake cycle is determined by the interaction of two processes: a circadian oscillator, which cycles with the solar day, and a homeostatic drive for sleep (*7*). A key element in that model is that a sleep deficit (i.e., sleep deprivation) causes a quantifiable "pressure to go to sleep." Subsequent NREM sleep is both longer and deeper than normal, and the antecedent sleep loss can be identified post hoc by an increase in EEG slow-wave activity during recovery sleep (*8*). Slow-wave activity is characterized by a wave of synchronous local neural firing that typically begins in the frontal cortex and propagates posteriorly, occurring roughly every second during NREM sleep (*9*). One of the predictions of the Borbély model is that daytime sleep is lighter, because it is not aligned with the circadian clock, and hence fails to fulfill the homeostatic function of sleep. This prediction has been supported by numerous studies of night-shift workers, who as a group are predisposed to stress, obesity, cognitive deficits, and an elevated risk of neurodegenerative diseases (*10*–*13*). One of the most prominent current models of sleep posits that the purpose of sleep is to restore synaptic homeostasis (*14*). The synaptic homeostasis hypothesis of sleep is based on the observations that wakefulness is associated with the sustained potentiation of excitatory transmission, as well as with the structural expansion of postsynaptic dendritic spines (*15*, *16*). The larger size of spines during wakefulness increases their postsynaptic currents and thereby strengthens excitatory transmission. This model is supported by the observation that sleep deprivation is linked to an increased risk of seizures in predisposed individuals (*17*). It is only during subsequent recovery sleep that excitatory transmission tone and spine volume fall, each returning to its sleep-associated baseline (*18*).

Recent studies in mice have offered molecular insights into the synaptic homeostasis hypothesis by mapping the impact of the sleep-wake cycle on synaptic gene expression (*19*, *20*). These studies showed that genes involved in synaptic signaling were predominantly transcribed before the mice woke up, whereas transcripts of genes involved in metabolism rose a few hours before their expected bedtime. Thus, the circadian clock dictates the transcription of genes in anticipation of the tasks appropriate for the time of day. Similarly, translation of mRNAs into proteins largely followed transcription, so that proteins involved in synaptic signaling were produced during wakefulness, whereas those with a role in metabolism were translated during sleep. Surprisingly, when the mice were kept awake longer than normal, the translation of proteins involved in synaptic signaling continued during sleep deprivation, concurrently with suppressed production of proteins associated with metabolism (*19*, *20*). Thus, the behavioral state, rather than the circadian clock, controls synaptic protein production. Under continued wakefulness, proteins involved in synaptic signaling are continuously produced, whereas proteins needed for restorative metabolic processes are not translated. Thus, extended wakefulness is associated with a dysregulation of translation that enables the sustained potentiation of excitatory transmission; this supports a critical homeostatic role of sleep that cannot occur in the awake state. It is intriguing to speculate that the depth of recovery sleep, detected as slow-wave activity, controls the translation of proteins needed to restore metabolic homeostasis.

The glymphatic and lymphatic systems

A fundamental tenet of brain homeostasis is that protein clearance must approximate protein synthesis. Is removal of protein waste also controlled by the sleep-wake cycle? Until 2012 it was believed that the brain, singular among organs, was recycling all of its own protein waste (*21*). Only a small number of proteins were known to be transported across the blood-brain barrier, and these did not include most of the primary proteins made or shed by brain cells (*22*). In the absence of lymphatic vessels or any overt pathways for fluid export, it was unclear how protein waste might exit the mature brain parenchyma. The default conclusion was that the classical cellular protein degradation pathways—autophagy and ubiquitination—must be responsible for all central nervous system (CNS) protein recycling (*23*).

This supposition, that the brain must recycle its own waste, was questioned after the discovery of the glymphatic system (*24*). The glymphatic system is a highly organized cerebrospinal fluid (CSF) transport system that shares several key functions, including the export of excess interstitial fluid and proteins, with the lymphatic vessels of peripheral tissues (Fig. 1A). Indeed, both the brain's CSF and peripheral lymph are drained together into the venous system, from which protein waste is removed and recycled by the liver (*25*). Yet brain tissue itself lacks histologically distinct lymphatic vessels. Rather, fluid clearance from the brain proceeds via the glymphatic pathway, a structurally distinct system of fluid transport that uses the perivascular spaces created by the vascular endfeet of astrocytes (*26*). The endfeet surround arteries, capillaries, and veins, serving as a second wall that covers the entire cerebral vascular bed. The perivascular spaces are open, fluid-filled tunnels that offer little resistance to flow. This is in sharp contrast to the disorientingly crowded and compact architecture of adult brain tissue, the neuropil, through which interstitial fluid flow is necessarily slow and restricted—akin to a marsh, flowing to the glymphatic system's creeks and then rivers (*27*). The glymphatic system's perivascular tunnels are directly connected to the

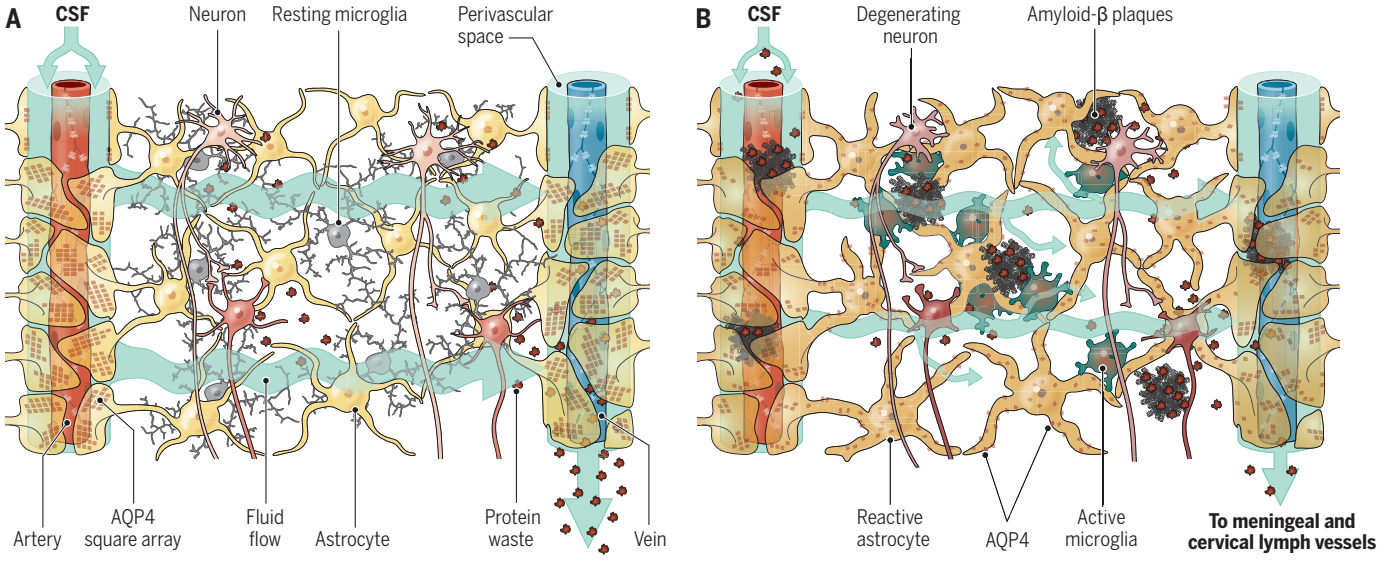
[1]Center for Translational Neuromedicine, Faculty of Health and Medical Sciences, University of Copenhagen, 2200 Copenhagen, Denmark. [2]Center for Translational Neuromedicine, University of Rochester Medical Center, Rochester, NY 14642, USA.
*Corresponding author. Email: nedergaard@sund.ku.dk (M.N.); goldman@sund.ku.dk (S.A.G.)

subarachnoid spaces surrounding the brain, from which CSF is rapidly driven into deep regions of the brain by the cardiac rhythm–linked pulsations of the arterial wall (*28*). The vascular endfeet of astrocytes, a primary subtype of glial cells, surround the perivascular spaces and can be regarded as open gates for fluid influx into the neuropil. The astrocytic endfeet are connected by gap junctions, and almost 50% of their plasma membrane facing the vessel wall is occupied by square arrays composed of the water channel protein aquaporin-4 (AQP4) (*29*). Deletion of AQP4 channels in mice reduces both the influx of CSF tracers and the efflux of solutes from the neuropil (*24*, *30*, *31*). Given this pathway's functional similarities to the peripheral lymphatic system, we termed this astrocyte-regulated mechanism of brain fluid transport the "glymphatic (glial-lymphatic) system."

Notably, fluid transport through the glymphatic system is directionally polarized, with influx along penetrating arteries, fluid entry into the neuropil supported by AQP4, and efflux along the perivenous spaces, as well as along the cranial and spinal nerves (*24*, *32*–*34*). In addition to its vectorial nature, glymphatic clearance is temporally regulated, and cyclically so, whereby fluid transport is enabled by sleep and suppressed during wakefulness. Brain fluid transport initiates and proceeds during NREM sleep, and CSF tracer influx correlates with the prevalence of EEG slow-wave activity (*35*, *36*). Fluid flow through the glymphatic system is thus inextricably linked with sleep, to the extent that flow appears to stop with the onset of wakefulness. In this regard, slow-wave activity predominates in the early hours of sleep and is a direct measure of sleep pressure, increasing with antecedent sleep deprivation (*8*). As such, waste removal is likely most efficient in the early hours of sleep and especially during recovery sleep after prolonged wakefulness (*37*). Yet it is easy to imagine why the awake state might be incompatible with active parenchymal fluid flow. Wakefulness relies on the precision of synaptic transmission in both time and space. Active flow might be expected to increase glutamate spillover during synaptic activity, resulting in bystander activation of local synapses and hence a loss of both the temporal and spatial fidelity of synaptic transmission. A recent analysis showed that glymphatic flow is also regulated by circadian rhythmicity, such that fluid transport peaks during the sleep phase of diurnal activity and falls during the active phase, independent of the light cycle. This rhythm is supported by the temporally regulated localization of AQP4 via the dystrophin-associated complex, providing a dynamic link to the molecular circadian clock (*38*).

A functionally integrated unit

Upon discovery and characterization of the glymphatic system, it quickly became apparent that glymphatic efflux pathways needed to be more comprehensively defined. Then came the reports that classical lymphatic vessels draining brain interstitial CSF might also be identified in the dura, the fibrous external layer of the meningeal membranes (*39*, *40*). The meningeal lymphatic vessels are separated from CSF by the arachnoid membrane, an internal meningeal layer whose cells constitute a tight fluid barrier by virtue of their dense expression of tight junctions, identified by their expression of claudin-11 (*41*). Yet the glymphatic and meningeal lymphatic systems are clearly connected: CSF tracers can exit the CNS via the meningeal lymphatic vessels, particularly by way of the lymph vessels of the ventral aspect of the brain draining to the cervical lymph nodes (*39*, *40*, *42*). CSF exit from the CNS by way of the meningeal lymph vessels, as well as via both cranial and spinal nerve roots, is rapid; contrast agents can be detected



CREDIT: D. XUE; ADAPTED BY KELLIE HOLOSKI/*SCIENCE*

Fig. 1. The brain glymphatic system is a highly organized fluid transport system. (**A**) Vascular endfeet of astrocytes create the perivascular spaces through which CSF enters the brain and pervades its interstitium. CSF enters these perivascular spaces from the subarachnoid space and is propelled by arterial pulsatility deep into the brain, from where CSF enters the neuropil, facilitated by the dense astrocytic expression of the water channel AQP4, which is arrayed in nanoclusters within the endfeet. CSF mixes with fluid in the extracellular space and leaves the brain via the perivenous spaces, as well as along cranial and spinal nerves. Interstitial solutes, including protein waste, are then carried through the glymphatic system and exported from the CNS via meningeal and cervical lymphatic vessels. (**B**) Amyloid-β plaque formation is associated with an inflammatory response, including reactive micro- and astrogliosis with dispersal of AQP4 nanoclusters. Age-related decline in CSF production, decrease in perivascular AQP4 polarization, gliosis, and plaque formation all impede directional glymphatic flow and thereby impair waste clearance. Notably, vascular amyloidosis might be initiated by several mechanisms. Amyloid-β might be taken up from the CSF by vascular smooth muscle cells expressing the low-density lipoprotein receptor-related protein 1 (LRP1) (*111*). Alternatively, amyloid deposition might be initiated by the backflow of extracellular fluid containing amyloid-β into the periarterial space from the neuropil, rather than proceeding to the perivenous spaces, because of an increase in hydrostatic pressure on the venous side or an inflammation-associated loss of AQP4 localization to astrocytic endfeet.

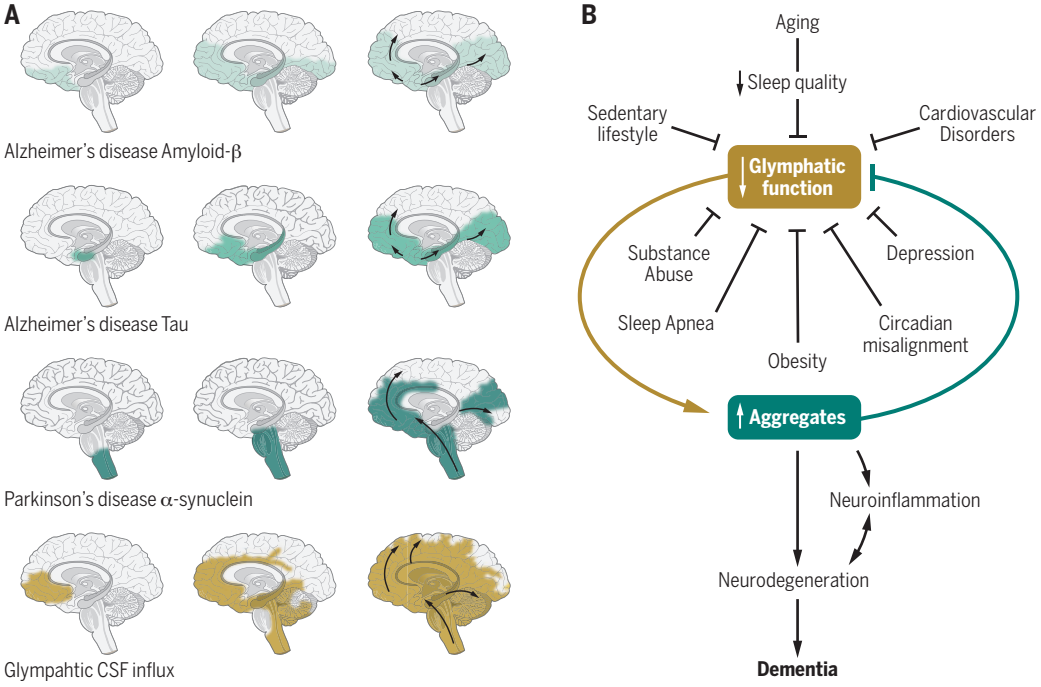


Fig. 2. Prion-like spread of protein aggregates and proposed role of glymphatic transport. (**A**) Seeding and prion-like spread of protein aggregates (amyloid-β and tau) in Alzheimer's disease and of α-synuclein in Parkinson's disease, relative to the distribution of glymphatic influx of a CSF tracer after intrathecal delivery (*67*). Prion-like spread of protein aggregates includes an extracellular component and, hence, the possibility that the seeds are transported by the glymphatic system. (**B**) In this model, the glymphatic system resides at the intersection of a broad scope of disorders, which share an association with diminished brain fluid clearance. Normal aging is also linked to a sharp decline in sleep quality and decreased glymphatic flow. In turn, the stagnation of glymphatic flow, and hence that of extracellular proteins, contributes to protein aggregation, with misfolding and seeding, leading to local inflammation, neuronal loss, and ultimately dementia.

in the deep cervical lymph nodes within minutes after CSF delivery (*42*–*45*). Nonetheless, proteins and tracers can circulate back into the brain along the periarterial spaces, which suggests that our understanding of flow vectors in the CNS is incomplete. More work is needed to comprehensively account for all of the paths by which extracellular fluid and its solutes are cleared from the adult brain (*46*).

Regardless of its precise efflux pathways, CSF ultimately drains into the cervical lymphatic vasculature, by which it returns to the venous system. In a mouse model of Alzheimer's disease (AD), amyloid-β was present in high concentrations in the cervical and axillary lymph nodes, at levels analogous to those in the brain, and yet was either undetectable or barely so in the spleen and other peripheral tissues (*47*). A large proportion of brain waste proteins and metabolites might then be expected to pass through and be cleared by the cervical lymphatics. Lymphatic vessels undergo atrophy in aging (*48*, *49*); thus, lymphatic drainage of CSF may pose a checkpoint—and with aging, a bottleneck—for brain protein clearance. In this regard, overexpression of vascular endothelial growth factor C induced sprouting of the meningeal lymphatic vessels and slowed cognitive decline in a mouse model of AD (*50*). Conversely, both ultraviolet photoablation of meningeal lymphatic vessels and mechanical ligation of cervical lymphatics aggravated amyloid plaque formation in the same mouse models of AD (*50*, *51*). Therefore, the glymphatic and lymphatic systems are intimately connected, both structurally and functionally, such that interference with fluid transport at any segment or node risks upstream fluid stasis and, hence, the aggregation of proteins otherwise destined for clearance.

Why do proteins aggregate?

Aging is also associated with a steep fall in glymphatic flow in the brains of both rodents and humans. CSF inflow of larger tracers is reduced by up to 85% in aged wild-type mice, whereas contrast agent clearance in human brain tissue was inversely correlated to age in all individuals studied (*50*, *52*–*54*). The decrease in glymphatic flow in old mice is partly mediated by mislocalization of AQP4 away from the vascular wall (*52*) and by possible atrophy of meningeal lymphatic vessels (*42*). In addition to age-related decreases in brain fluid transport, glymphatic CSF influx and CSF clearance are each reduced in early stages of amyloid-β deposition in the APP/PS1 model of AD compared with in littermate controls, and CSF clearance continues to decline as the amyloid burden increases (Fig. 1B). Infusion of amyloid-β into CSF acutely reduced glymphatic activity in wild-type mice, suggesting a direct toxic effect (*50*, *55*).

The suppressive effects of both age and amyloid-β overexpression on glymphatic flow can be extended to other experimental rodent models of neurodegeneration: Both traumatic brain injury and Parkinson's disease are similarly linked to a sustained reduction of glymphatic fluid transport (*56*–*58*). Notably, most of these age-related primary neurodegenerative diseases involve disorders of protein processing and aggregation. The hallmark features of these proteinopathies are the fibrillary aggregates of misfolded or hyperphosphorylated proteins (*59*). The protein aggregates can range in size from oligomers to large fibrillary structures. These aggregation-prone proteins include amyloid-β in AD; phosphorylated tau in frontotemporal dementia (FTD), chronic traumatic encephalopathy, and AD; α-synuclein in Parkinson's disease, Lewy body disease, and the multisystem atrophies; mutant huntingtin in Huntington's disease; and TAR DNA-binding protein 43 (TDP-43) in amyotrophic lateral sclerosis and FTD (*60*). Although the specific protein species differ in the different neurodegenerative disorders, in most cases their protein aggregates are formed in part by the interactions of intermolecular β-sheet–rich strands. Once a seed is formed, the aggregates attract monomers of the same protein, as well as other proteins, which may be preferentially bound and entrapped (*60*).

To understand why aging predisposes organisms to these proteinopathies, we need to consider those conditions that favor nucleation, the growth of protein aggregates, and their subsequent seeding to neighboring cells. Protein self-assembly and aggregation depend on a number of factors, among which are structure, concentration, ionic strength, and local pH, as well as their interactions with nucleating interfaces, such as phospholipid membranes (*61*, *62*). Ex vivo aggregation can be induced by simply mixing hydrophobic nanoparticles into an aqueous solution that contains proteins (*63*). A lack of fluid flow (stagnation) or its opposite (shear stress) can also promote aggregation (*64*, *65*), which can occur at a distance from the protein source—for example, along the cerebral vasculature (Fig. 1) (*66*). Depending on the protein, each of these factors, alone

CREDIT: D. XUE; ADAPTED BY KELLIE HOLOSKI/*SCIENCE*

or in combination, can lead to self-aggregation with the formation of stable β-sheet–rich strands. Reduced glymphatic clearance might then be predicted to increase the risk of protein aggregation, given the combination of locally stagnant fluid flow and elevated extracellular concentration of the protein of interest.

Spread of protein aggregates

The recent discovery that specific misfolded and aggregated proteins can propagate and spread in a prion-like fashion has sparked considerable interest (*67*). It has been generally posited that seeding occurs across regions that are synaptically connected (*68*). However, the evidence for synaptic spread is largely based on post hoc analysis of anatomic networks; it remains unclear how synaptic relationships by themselves can mediate seeding. The arguments for synaptic spread are somewhat weakened by the fact that aggregate spread happens in both antero- and retrograde directions across regions that are anatomical neighbors (*68*). An alternative hypothesis is that aggregates simply spread via the extracellular spaces and that the age-dependent reduction in glymphatic flow, with its attendant fluid stagnation, raises the local protein concentration to a level that favors aggregation. In support of this hypothesis, the suppression of glymphatic flow by deletion of AQP4 water channels sharply increased both amyloid-β plaque formation and cognitive deficits in a mouse model of AD (*69*). Similarly, in humans, efflux of CSF containing amyloid-β and phosphorylated tau is reduced in patients with AD compared with age-matched controls. The suppression of CSF clearance in AD is so substantial that it can possibly serve as a biomarker (*70*).

What do we know about the spread of protein aggregates on a macroscopic scale? In AD, amyloid-β deposition typically first occurs in the basal portions of the frontal, temporal, and occipital lobes. Later, the plaques spread to include the hippocampus and posterior parietal cortex, initially sparing both the motor and sensory cortices. These latter regions are first recruited in the final stages of the disease, along with subcortical gray matter regions. Yet the cognitive decline of AD patients correlates more closely with the later-occurring tauopathy and microglial activation than with the earlier amyloid-β plaque formation (*71*, *72*). In the initial stages of AD, phosphorylated tau deposits in the entorhinal cortex, followed by the hippocampus and dorsal thalamus, whereas the neocortex becomes involved later. In Parkinson's disease and Lewy body disease, α-synuclein aggregates initially spread through the brainstem and olfactory bulb, followed by limbic structures, and only then to the neocortex (Fig. 2A). In each of these cases, the aggregates initially deposit at the ventral base of the forebrain and midbrain and then extend rostrally and dorsally to the cortex.

How does this pattern of spread compare to glymphatic CSF inflow (Fig. 2A) (*67*, *73*)? Neuroimaging studies have shown that intrathecally delivered contrast agents are first propelled into the brain along the large cerebral arteries, entering the mediobasal frontal lobe and cingulate cortex along the anterior cerebral artery, the insula via the middle cerebral artery, and the limbic structures (including the hippocampus and entorhinal cortex) via the posterior circulation. The contrast agent remains trapped in the same regions for prolonged periods of time, especially if an underlying pathology is present (*74*, *75*). The accumulation of low–molecular weight CSF contrast agents (<1 kDa) supports the idea that much larger proteins also get trapped in the tortuous extracellular spaces of deep brain regions.

Although the conditions by which pathogenic proteins may become entrapped and aggregate in glymphatic channels remain unclear, the geographic spread of aggregates in AD and Parkinson's disease clearly mirrors the pattern of glymphatic inflow in the human brain, as mapped by magnetic resonance imaging. Indeed, the geographic pattern of macroscopic aggregate formation closely resembles that of entrapped CSF contrast agents during restriction of glymphatic flow in those brains (Fig. 2B). On that basis, we propose that trapping of aggregation-prone proteins in the extracellular space, rather than synaptic connectivity, is responsible for the patterns of protein spread in at least some proteinopathies. As such, the regional variations in the path of seeding across the different types of neurodegenerative diseases may reflect region- and patient-specific variability in the rates of neuronal production of amyloid-β, tau, and α-synuclein.

Notably, although proteins associated with neurodegenerative diseases may normally be either intracellular or extracellular in nature, all are present in the extracellular space. Efforts to sample CSF and extracellular fluid have shown that amyloid-β, tau, and α-synuclein are present outside the cytosol. These proteins all lack N-terminal signal sequences, so unconventional mechanisms must be responsible for their release (*76*). In each of these cases, it is unclear whether oligomers or the larger protein aggregates constitute the principal neurotoxic species (*60*). Although no consensus has been reached, several studies have highlighted the critical role of oligomers as directly toxic and as a nidus for macromolecular aggregation. Immune therapies have attempted to clear the extracellular space and CSF of amyloid-β in AD patients. The failure of such clinical trials may reflect the relatively late initiation of treatment or that the antibody load was not sufficient to clear enough amyloid-β to yield clinical benefit. Alternatively, it is possible that the underlying model of direct, aggregation-associated neurotoxicity is fundamentally incorrect, in AD as well as more broadly (*77*).

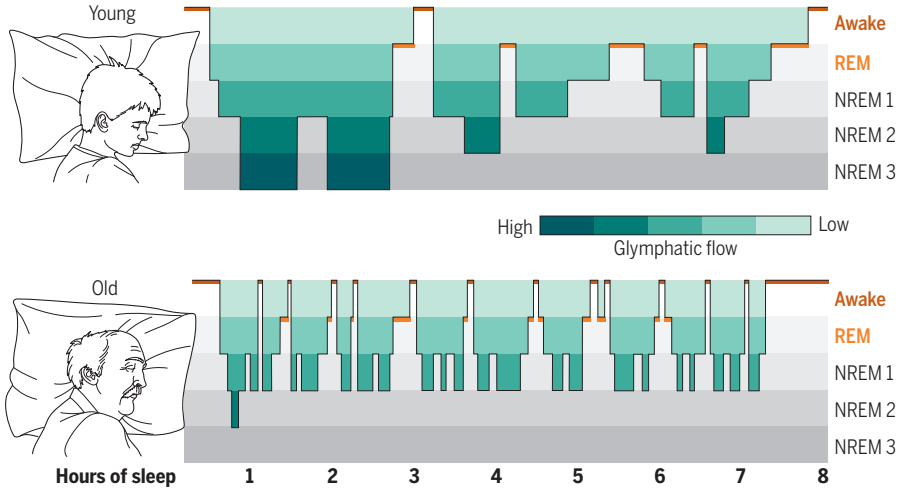


Fig. 3. Sleep architecture in young and old individuals. Hypnograms are constructed from EEG recordings and display the cyclic transitions between sleep stages. The two schematic hypnograms illustrate the sleep architecture of young and old individuals who transition spontaneously between the awake state, REM sleep, and NREM (stages 1 to 3) sleep. Stage 1 NREM sleep is light sleep, whereas stage 3 NREM sleep is the deepest sleep stage and is characterized by slow-wave EEG activity. For young people, deep (stage 3) NREM sleep dominates in the early phases of sleep, whereas REM sleep is more frequent in the later phases. Sleep spindles are most frequent in stage 2 NREM sleep. By contrast, for people older than 60 years of age, sleep is often interrupted by short awake episodes, and older individuals do not typically enter stage 3 NREM sleep. Total sleep time decreases by 10 min for each decade of life (*79*). Green shading indicates the proposed efficacy of glymphatic clearance on the basis of data collected in rodents (*35*, *36*). The lack of stage 3 NREM sleep, the frequent interruptions of stage 1 and 2 NREM sleep, and the shorter total sleep time all serve to decrease glymphatic activity in aging. Critically, a number of disorders and conditions can suppress glymphatic function during NREM sleep, further exacerbating the effects of glymphatic dysfunction in neurodegenerative disease.

CREDIT: D. XUE; ADAPTED BY KELLIE HOLOSKI/*SCIENCE*

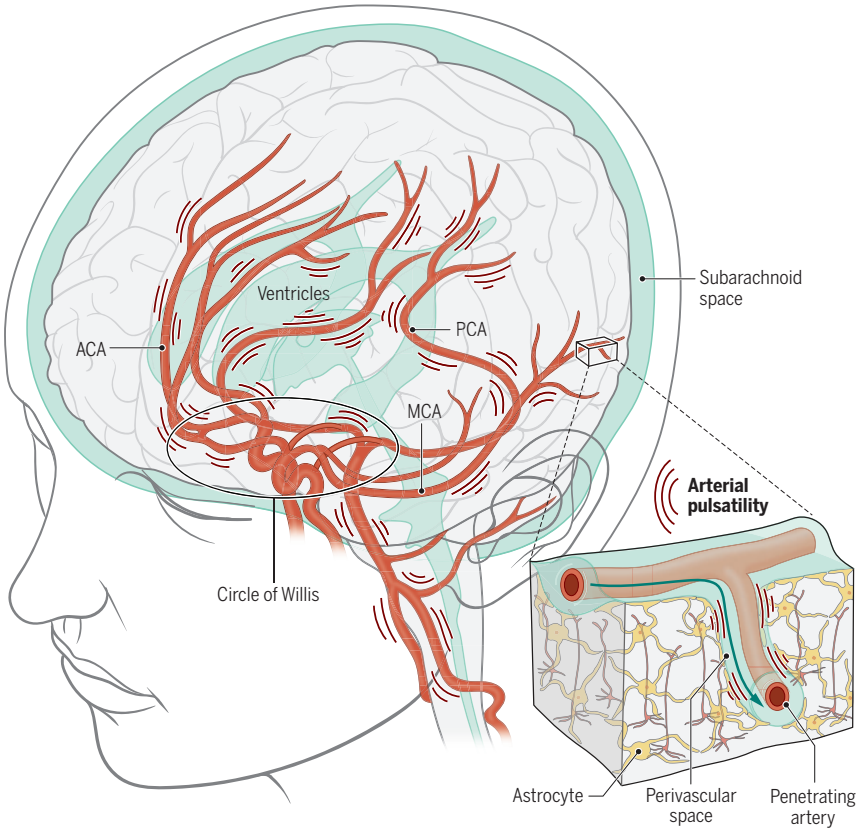


Fig. 4. Arterial pulsatility propels fluid flow in the brain. The brain receives 20 to 25% of a person's cardiac output but constitutes only ~2% of total body weight. The large-caliber arteries of the circle of Willis are positioned in the CSF-containing basal cisterns below the ventral surface of the brain. Arterial pulsatility provides the motive force for CSF transit into the perivascular spaces surrounding the major arteries, whereas respiration and slow vasomotion contribute to sustaining its flow (*112*). The anterior (ACA), middle (MCA), and posterior (PCA) arteries transport CSF to the penetrating arteries (inset), from which CSF is then driven into the neuropil via the still-contiguous perivascular spaces. Cardiovascular diseases associated with reduced cardiac output, such as left heart failure and atrial arrhythmias, reduce arterial wall pulsatility, resulting in diminished CSF flow. In addition, thickening of the arterial wall in SVD, hypertension, and diabetes reduces arterial wall compliance and, hence, pulsatility. Each of these fundamentally cardiovascular disorders serves to attenuate glymphatic flow, providing a potential causal link between these vascular etiologies and AD (*113*).

Sleep, aging, neurodegeneration, and the glymphatic system

The most substantial risk factor for developing protein aggregation, as for developing dementia, is age (*78*). With the glymphatic system in mind, it is notable that sleep quality decreases as a function of normal aging. Insomnia is more frequent with increasing age, and total sleep duration becomes shorter and more interrupted. Perhaps more critically, older individuals rarely enter deep NREM (stage 3) sleep. Most NREM sleep in people older than 60 years of age is light, consisting of the more superficial stages 1 and 2 (*79*) (Fig. 3). Thus, the aged brain spends less time in NREM sleep, potentially causing a catastrophic decline in clearance of brain waste, as the efficacy of glymphatic fluid transport correlates directly with the prevalence of slow-wave activity (*36*). The age-related impairment in sleep quality may thus be causally related to the increased incidence and accelerated course of neurodegenerative disease in older people, whose disrupted sleep architecture may sharply diminish the clearance of brain fluid and its attendant export of protein waste, thus leading to the stagnant interstitial flow that favors aggregate formation.

In addition to the deterioration of sleep architecture in aging, the neurodegenerative diseases—including AD, Parkinson's disease, Huntington's disease, the multisystem atrophies, and the FTDs—are all associated with sleep disturbances (*80*). The best characterized among these are the sleep pathologies associated with Parkinson's disease, in which REM sleep disturbances often precede the onset of motor symptoms by several years or even decades (*80*, *81*). Future work should define whether sleep disturbances that preceded the clinical diagnosis contribute to aggregate seeding and whether sleep disturbances during disease progression accelerate aggregate spread. It would seem axiomatic that a stronger focus on age-related impairment of sleep quality should benefit the aging population.

AQP4 polymorphisms

The polarized expression of AQP4 in the vascular endfeet of astrocytes facilitates glymphatic fluid transport and amyloid-β export in rodents (*24*, *30*) (Fig. 1). In humans, genetic variation in *AQP4* affects both sleep and amyloid-β burden (*82*). A recent study established a link between AQP4, sleep, and the effects of prolonged wakefulness on cognitive function. The study demonstrated that a common single-nucleotide polymorphism (SNP) of *AQP4* was linked to changes in slow-wave activity during NREM sleep that were mirrored by changes in daytime sleepiness as well as in altered reaction times during extended wakefulness (*83*). Yet AQP4 SNPs have also been associated with the rate of cognitive decline in longitudinally followed cohorts of AD patients (*84*). Patients with two specific *AQP4* SNPs exhibited slower cognitive decline after AD diagnosis, whereas cognitive decline progressed more rapidly in individuals with two other *AQP4* SNPs (*85*). Structurally, the integrity of perivascular AQP4 localization was found to degrade with AD, whereas it was preserved in patients older than 85 years of age who remained cognitively intact (*84*). Similarly, the expression of a cluster of transcripts encoding proteins associated with astrocytic endfeet predicted lower amounts of cortical phosphorylated tau in humans (*86*). Indeed, a recent study reported that deletion of *Aqp4* accelerated amyloid plaque formation in a mouse model of AD (*69*). Thus, although AQP4 is expressed only in astrocytes, and not in amyloid-producing neurons, considerable evidence indicates that AQP4 modulates sleep architecture, tolerance to sleep deprivation, amyloid-β accumulation, and the progression of AD. Targeting the brain's waste removal system may thus be an attractive approach for alleviating the waste burden of the proteinopathies because aggregation-prone proteins are removed by bulk flow, without the requirement for specific transporters.

Links to cardiovascular disease

Neurodegenerative diseases are not the only cause of dementia. It has been known for decades that poor cardiovascular health negatively affects cognitive abilities (*87*, *88*), whereas cardiovascular fitness positively correlates with cognition in young adults (*89*) and preserves cognitive performance in aging individuals (*90*). Why is a healthy heart so important for higher brain function? It has been shown that glymphatic function is suppressed in hypertensive rats (*91*, *92*). It is also well established that sleep quality is compromised in cardiovascular diseases (*93*), perhaps providing a link to impaired glymphatic clearance and subsequent protein aggregation and dementia (*94*).

We also propose that a healthy cardiovascular system, besides its role in delivering energy metabolites to the brain, plays a hitherto

CREDIT: D. XUE; ADAPTED BY KELLIE HOLOSKI/*SCIENCE*

unappreciated role in the clearance of neurotoxic waste from the brain. In particular, we have found that the brain's fluid transport system is designed to take advantage of cardiac pulsatility to drive CSF transport in the neuropil (*28*). The ejection pressure of blood from the left ventricle is partly absorbed by the elastic arterial wall of the aorta. As the ejected blood transits the arteries, it enlarges the arterial diameter as its pulse wave propagates downstream (*28*). About 20 to 25% of the total ejected blood volume enters the CNS via the paired internal carotid and posterior cerebral arteries. Pulsatility in these large-caliber arteries constantly transmits pressure waves along the axis of the major vessels, as well as through the soft brain tissue (Fig. 4). The motion of the brain is locally supplemented by the pulsatility of the penetrating arteries, as they enter the brain from the CSF-filled subarachnoid space, thereby driving CSF into the neuropil along the periarterial spaces (*24*). It should not be surprising that heart diseases associated with reduced cardiac output, including congestive heart failure and atrial dysrhythmias (*95*), are also associated with diminished glymphatic flow, because the pulsatility of the cerebral arteries and hence the driving forces within the glymphatic system are reduced. Indeed, the cognitive decline frequently noted in patients with a low cardiac ejection fraction, often attributed to low cerebral perfusion, may also reflect poor glymphatic flow and incomplete waste clearance, as well as a consequent predisposition to aggregate formation and still-slower glymphatic flow (*95*).

Small vessel disease (SVD) is a vascular disorder that targets the small cerebral vessels, in which penetrating arterioles undergo progressive thickening of their walls (*96*). Deterioration of the vascular bed may occur alone or in combination with other pathologies (*97*), leading to progressive demyelination and loss of white matter (*98*). SVD is common in patients with hypertension, many of whom are concurrently diabetic or smokers, and it progresses silently for years before dementia is clinically evident (*99*). Hypertension induces hypertrophy of vascular smooth muscle cells, with a stiffening of the arterial wall that dampens arterial wall pulsatility and compliance, thus reducing convective perivascular flow (*94*, *100*). The stiffening of the perivascular glycocalyx of diabetic patients has a similar effect (*101*), and the two disorders are in frequent combination as the incidence of obesity, a predisposing factor and comorbidity to both, increases worldwide. SVD is linked to glymphatic dysfunction in experimental models (*91*) and may potentiate the progression of neurodegenerative dementias in the same patients at risk for SVD-associated vascular dementia. It is not surprising, then, that the clinical distinctions between AD and the vascular dementias are often blurred by their frequent co-association (*102*).

Outlook

Fundamentally, the studies discussed here highlight the benefits of a good night's sleep. Sleep is an evolutionarily conserved mechanism that serves multiple purposes, with benefits to the homeostatic support of the cardiovascular system, immune system, and memory (*103*–*105*). Yet the most fundamental incentive for the brain to sleep lies in its own self-preservation: Only the sleeping brain is capable of efficiently clearing the waste products generated during active wakefulness. Amyloid-β, tau, and α-synuclein are all present in the brain extracellular fluid and CSF at higher concentrations during wakefulness than during sleep, and sleep deprivation further increases these levels (*106*–*108*). Indeed, positron emission tomography imaging has shown that a single night of sleep deprivation resulted in a significant increase in amyloid-β burden in the hippocampus and thalamus (*109*). Humans need sleep to clear proteins from the brain extracellular space, or these proteins will aggregate, impede fluid flow, and potentiate further fibril polymerization. Together with local inflammation, this process may be expected to progressively suppress glymphatic flow in the most affected regions.

Overall, these observations suggest a causal linkage between the sleep-wake cycle and its regulation of fluid flow via the glymphatic system, and thereby the modulation of the balance between protein clearance and aggregation. As such, the observations suggest a basis for the increased incidence of protein aggregation–related disorders that occur with aging, the appearance of which tracks age-related declines in both vascular health and glymphatic patency. The neurodegenerative dementias may thus be viewed as the products of a final common pathway that integrates the dysfunction of any and all of these closely interdependent upstream mechanisms (Fig. 3). These various processes are linked in their regulation by the brain's glymphatic system, the directed regulation of which may, in turn, present new therapeutic opportunities for the disease-modifying treatment of patients with these disorders (*75*). In particular, the development of small-molecule agonists of glymphatic efflux might present opportunities to slow disease progression in the aggregation disorders, just as the optimization of cardiovascular health might be expected to delay disease onset. These systems are intimately connected such that modulation of glymphatic flow, and hence protein clearance from the brain, will ultimately require a deeper understanding of the dependence of both glymphatic and lymphatic flow on intracardiac pressures.

Recent advances in neuroimaging have provided multiple approaches to map the human glymphatic system and to assess its functional competence in the context of disease, as well as the effects thereof on sleep-dependent glymphatic cycling (*72*, *73*, *108*, *110*). The diagnostic neuroimaging of glymphatic function via such "glymphograms" may provide both a means to predict the risk of developing proteinopathies and an approach by which to evaluate the efficacy of glymphatic flow–directed treatments as they are developed. Until then, the most assured means of preserving effective glymphatic clearance is to get a good night's sleep.

REFERENCES AND NOTES

1. B. A. Mander *et al*., *Nat. Neurosci.* **16**, 357–364 (2013).
2. K. Spiegelhalder, W. Regen, S. Nanovska, C. Baglioni, D. Riemann, *Curr. Psychiatry Rep.* **15**, 364 (2013).
3. A. P. Spira, L. P. Chen-Edinboro, M. N. Wu, K. Yaffe, *Curr. Opin. Psychiatry* **27**, 478–483 (2014).
4. L. Leigh, I. L. Hudson, J. E. Byles, *J. Sleep Res.* **24**, 648–657 (2015).
5. E. Aserinsky, N. Kleitman, *Science* **118**, 273–274 (1953).
6. J. A. Hobson, *Nature* **437**, 1254–1256 (2005).
7. A. A. Borbély, *Hum. Neurobiol.* **1**, 195–204 (1982).
8. A. V. Rodriguez *et al*., *J. Neurosci.* **36**, 12436–12447 (2016).
9. M. Massimini, R. Huber, F. Ferrarelli, S. Hill, G. Tononi, *J. Neurosci.* **24**, 6862–6870 (2004).
10. S. M. Rajaratnam, M. E. Howard, R. R. Grunstein, *Med. J. Aust.* **199**, S11–S15 (2013).
11. G. Kecklund, J. Axelsson, *BMJ* **355**, i5210 (2016).
12. O. Itani *et al*., *Sleep Med.* **39**, 87–94 (2017).
13. K. Bokenberger *et al*., *Eur. J. Epidemiol.* **33**, 977–987 (2018).
14. G. Tononi, C. Cirelli, *Neuron* **81**, 12–34 (2014).
15. V. V. Vyazovskiy, C. Cirelli, M. Pfister-Genskow, U. Faraguna, G. Tononi, *Nat. Neurosci.* **11**, 200–208 (2008).
16. R. Huber *et al*., *Cereb. Cortex* **23**, 332–338 (2013).
17. J. Logothetis, I. Milonas, S. Bostantzopoulou, *Eur. Neurol.* **25** (suppl. 2), 134–140 (1986).
18. L. de Vivo *et al*., *Science* **355**, 507–510 (2017).
19. S. B. Noya *et al*., *Science* **366**, eaav2642 (2019).
20. F. Brüning *et al*., *Science* **366**, eaav3617 (2019).
21. D. C. Rubinsztein, *Nature* **443**, 780–786 (2006).
22. P. Ballabh, A. Braun, M. Nedergaard, *Neurobiol. Dis.* **16**, 1–13 (2004).
23. T. A. Thibaudeau, R. T. Anderson, D. M. Smith, *Nat. Commun.* **9**, 1097 (2018).
24. J. J. Iliff *et al*., *Sci. Transl. Med.* **4**, 147ra111 (2012).
25. M. Nedergaard, *Science* **340**, 1529–1530 (2013).
26. M. Simard, G. Arcuino, T. Takano, Q. S. Liu, M. Nedergaard, *J. Neurosci.* **23**, 9254–9262 (2003).
27. J. Tithof, D. H. Kelley, H. Mestre, M. Nedergaard, J. H. Thomas, *Fluids Barriers CNS* **16**, 19 (2019).
28. H. Mestre *et al*., *Nat. Commun.* **9**, 4878 (2018).
29. E. A. Nagelhus, O. P. Ottersen, *Physiol. Rev.* **93**, 1543–1562 (2013).
30. H. Mestre *et al*., *eLife* **7**, e40070 (2018).
31. H. Mestre *et al*., *Science* **367**, eaax7171 (2020).
32. M. Braun, J. J. Iliff, *Int. Rev. Neurobiol.* **154**, 413–436 (2020).
33. X. Wang *et al*., *Sci. Transl. Med.* **12**, eaaw3210 (2020).
34. H. Mestre, Y. Mori, M. Nedergaard, *Trends Neurosci.* **43**, 458–466 (2020).
35. L. Xie *et al*., *Science* **342**, 373–377 (2013).
36. L. M. Hablitz *et al*., *Sci. Adv.* **5**, eaav5447 (2019).
37. L. Hauglund, C. Pavan, M. Nedergaard, *Curr. Opin. Physiol.* **15**, 1–6 (2020).
38. L. Hablitz *et al*., *Nat. Commun.* **11**, 4411 (2020).
39. A. Louveau *et al*., *Nature* **523**, 337–341 (2015).
40. A. Aspelund *et al*., *J. Exp. Med.* **212**, 991–999 (2015).
41. C. B. Brøchner, C. B. Holst, K. Møllgård, *Front. Neurosci.* **9**, 75 (2015).
42. J. H. Ahn *et al*., *Nature* **572**, 62–66 (2019).
43. Q. Ma *et al*., *Acta Neuropathol.* **137**, 151–165 (2019).
44. Q. Ma, Y. Decker, A. Müller, B. V. Ineichen, S. T. Proulx, *J. Exp. Med.* **216**, 2492–2502 (2019).
45. L. Jacob *et al*., *Nat. Commun.* **10**, 4594 (2019).
46. A. Louveau *et al*., *J. Clin. Invest.* **127**, 3210–3219 (2017).
47. M. Pappolla *et al*., *Neurobiol. Dis.* **71**, 215–219 (2014).
48. J. P. Scallan, S. D. Zawieja, J. A. Castorena-Gonzalez, M. J. Davis, *J. Physiol.* **594**, 5749–5768 (2016).
49. B. Jakic, D. Kerjaschki, G. Wick, *Gerontology* 10.1159/000508459 (2020).
50. S. Da Mesquita *et al*., *Nature* **560**, 185–191 (2018).
51. L. Wang *et al*., *Brain Pathol.* **29**, 176–192 (2019).
52. B. T. Kress *et al*., *Ann. Neurol.* **76**, 845–861 (2014).
53. H. Benveniste *et al*., *Gerontology* **65**, 106–119 (2019).
54. Y. Zhou *et al*., *Ann. Neurol.* **87**, 357–369 (2020).
55. W. Peng *et al*., *Neurobiol. Dis.* **93**, 215–225 (2016).
56. J. J. Iliff *et al*., *J. Neurosci.* **34**, 16180–16193 (2014).
57. W. Zou *et al*., *Transl. Neurodegener.* **8**, 7 (2019).
58. S. Sundaram *et al*., *Neurosci. Biobehav. Rev.* **103**, 305–315 (2019).

59. C. Soto, *Nat. Rev. Neurosci.* **4**, 49–60 (2003).
60. C. Soto, S. Pritzkow, *Nat. Neurosci.* **21**, 1332–1340 (2018).
61. K. A. Burke, E. A. Yates, J. Legleiter, *Front. Neurol.* **4**, 17 (2013).
62. E. A. Yates *et al.*, *J. Mol. Biol.* **425**, 1915–1933 (2013).
63. M. R. G. Kopp, U. Capasso Palmiero, P. Arosio, *Mol. Pharm.* **17**, 909–918 (2020).
64. C. N. Trumbore, *J. Alzheimers Dis.* **54**, 457–470 (2016).
65. S. J. Roeters *et al.*, *Sci. Rep.* **7**, 41051 (2017).
66. R. O. Weller, *J. Neuropathol. Exp. Neurol.* **57**, 885–894 (1998).
67. M. Jucker, L. C. Walker, *Nature* **501**, 45–51 (2013).
68. M. X. Henderson *et al.*, *Nat. Neurosci.* **22**, 1248–1257 (2019).
69. Z. Xu *et al.*, *Mol. Neurodegener.* **10**, 58 (2015).
70. M. J. de Leon *et al.*, *J. Nucl. Med.* **58**, 1471–1476 (2017).
71. S. Hong *et al.*, *Science* **352**, 712–716 (2016).
72. T. Terada *et al.*, *J. Neurol.* **266**, 2186–2196 (2019).
73. P. Brundin, R. Melki, R. Kopito, *Nat. Rev. Mol. Cell Biol.* **11**, 301–307 (2010).
74. G. Ringstad *et al.*, *JCI Insight* **3**, e121537 (2018).
75. G. Ringstad, S. A. S. Vatnehol, P. K. Eide, *Brain* **140**, 2691–2705 (2017).
76. A. A. Davis, C. E. G. Leyns, D. M. Holtzman, *Annu. Rev. Cell Dev. Biol.* **34**, 545–568 (2018).
77. D. S. Knopman, *Neurology* **90**, 447–448 (2018).
78. S. Hoyer, *Ann. N. Y. Acad. Sci.* **719**, 248–256 (1994).
79. H. P. Landolt, A. A. Borbély, *Clin. Neurophysiol.* **112**, 369–377 (2001).
80. R. K. Malhotra, *Sleep Med. Clin.* **13**, 63–70 (2018).
81. R. Malkani, H. Attarian, *Curr. Sleep Med. Rep.* **1**, 81–90 (2015).
82. S. R. Rainey-Smith *et al.*, *Transl. Psychiatry* **8**, 47 (2018).
83. S. M. Ulv Larsen *et al.*, *PLOS Biol.* **18**, e3000623 (2020).
84. D. M. Zeppenfeld *et al.*, *JAMA Neurol.* **74**, 91–99 (2017).
85. K. G. Burfeind *et al.*, *Alzheimers Dement.* **3**, 348–359 (2017).
86. M. J. Simon *et al.*, *Sci. Rep.* **8**, 12389 (2018).
87. D. Knopman *et al.*, *Neurology* **56**, 42–48 (2001).
88. E. C. Leritz, R. E. McGlinchey, I. Kellison, J. L. Rudolph, W. P. Milberg, *Curr. Cardiovasc. Risk Rep.* **5**, 407–412 (2011).
89. M. A. I. Aberg *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **106**, 20906–20911 (2009).
90. C. R. Wendell *et al.*, *J. Gerontol. A* **69**, 455–462 (2014).
91. K. N. Mortensen *et al.*, *J. Neurosci.* **39**, 6365–6377 (2019).
92. Y. Xue *et al.*, *Brain Res. Bull.* **161**, 78–83 (2020).
93. C. Gonzaga, A. Bertolami, M. Bertolami, C. Amodeo, D. Calhoun, *J. Hum. Hypertens.* **29**, 705–712 (2015).
94. H. Mestre, S. Kostrikov, R. I. Mehta, M. Nedergaard, *Clin. Sci.* **131**, 2257–2274 (2017).
95. R. L. Vogels *et al.*, *J. Am. Geriatr. Soc.* **55**, 1764–1770 (2007).
96. J. M. Wardlaw, C. Smith, M. Dichgans, *Lancet Neurol.* **18**, 684–696 (2019).
97. C. Iadecola, *Neuron* **80**, 844–866 (2013).
98. B. V. Zlokovic, *Nat. Rev. Neurosci.* **12**, 723–738 (2011).
99. J. M. Wardlaw, C. Smith, M. Dichgans, *Lancet Neurol.* **12**, 483–497 (2013).
100. J. M. Wardlaw *et al.*, *Nat. Rev. Neurol.* **16**, 137–153 (2020).
101. Q. Jiang *et al.*, *J. Cereb. Blood Flow Metab.* **37**, 1326–1337 (2017).
102. M. D. Sweeney *et al.*, *Alzheimers Dement.* **15**, 158–167 (2019).
103. H. Toda, J. A. Williams, M. Gulledge, A. Sehgal, *Science* **363**, 509–515 (2019).
104. C. S. McAlpine *et al.*, *Nature* **566**, 383–387 (2019).
105. N. Niethard, A. Burgalossi, J. Born, *Front. Neural Circuits* **11**, 65 (2017).
106. R. J. Bateman *et al.*, *Nat. Med.* **12**, 856–861 (2006).
107. J. E. Kang *et al.*, *Science* **326**, 1005–1007 (2009).
108. J. K. Holth *et al.*, *Science* **363**, 880–884 (2019).
109. E. Shokri-Kojori *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **115**, 4483–4488 (2018).
110. N. E. Fultz *et al.*, *Science* **366**, 628–631 (2019).
111. T. Kanekiyo, C. C. Liu, M. Shinohara, J. Li, G. Bu, *J. Neurosci.* **32**, 16458–16465 (2012).
112. V. Kiviniemi *et al.*, *J. Cereb. Blood Flow Metab.* **36**, 1033–1045 (2016).
113. S. Rius-Pérez, A. M. Tormos, S. Pérez, R. Taléns-Visconti, *Neurologia* **33**, 112–120 (2018).

ACKNOWLEDGMENTS

We thank D. Xue for assistance with illustrations and C. Cirelli and N. Beschorner for discussions. **Funding:** The authors are funded by the European Research Council under the European Union's Horizon 2020 research and innovation program (742112), the Lundbeck and Novo Nordisk foundations, the Dr. Miriam and Sheldon Adelson Medical Research Foundation, Foundation Leducq, the National Institute of Neurological Diseases and Stroke and the National Institute on Aging, and the U.S. Army Research Office MURI program, grant W911NF1910280. S.A.G. is additionally supported by Oscine Corp. and Sana Biotechnology. **Competing interests:** The authors declare no relevant competing interests.

10.1126/science.abb8739

REVIEW

Beyond aggregation: Pathological phase transitions in neurodegenerative disease

Cécile Mathieu[1], Rohit V. Pappu[2,3], J. Paul Taylor[1]*

Over the past decade, phase transitions have emerged as a fundamental mechanism of cellular organization. In parallel, a wealth of evidence has accrued indicating that aberrations in phase transitions are early events in the pathogenesis of several neurodegenerative diseases. We review the key evidence of defects at multiple levels, from phase transition of individual proteins to the dynamic behavior of complex, multicomponent condensates in neurodegeneration. We also highlight two concepts, dynamical arrest and heterotypic buffering, that are key to understanding how pathological phase transitions relate to pleiotropic defects in cellular functions and the accrual of proteinaceous deposits at end-stage disease. These insights not only illuminate disease etiology but also are likely to guide the development of therapeutic interventions to restore homeostasis.

Proteinaceous deposits in neuronal tissues have long been recognized as a hallmark of late-onset neurodegenerative diseases. Over the past 20 years, the dominant paradigm to relate protein deposits to cellular demise has centered on the concept of aggregation. Mechanistically, protein deposits are thought to impose neomorphic, toxic gains of function intrinsic to the deposited proteins, and much recent research has been centered on the physicochemical nature and aggregation state of the presumed toxic species (e.g., whether they are monomers, oligomers, or fibrils). A wealth of evidence has accrued in recent years leading to an evolution of the aggregation paradigm into a deeper and more direct understanding of how these protein deposits arise and relate to cellular dysfunction and death—specifically, via pathological phase transitions. Among this additional evidence is a deeper understanding of how cells are organized by phase transitions and how they govern vital biological processes. In particular, the concept of pathological phase transitions has arisen from increasing recognition that disease-associated proteins participate in these phase transitions, and from genetic insights demonstrating that disease-causing mutations in these same proteins promote this process. These advances have implications not only for understanding gain and/or loss of function of proteins associated with disease but also for the biological function of the condensates (Box 1) to which they belong. Here, we review these lines of evidence and synthesize these correlations into a framework that seeks to draw a direct line from genetics and biophysics to molecular mechanisms of disease and pathology. This framework is important in defining disease etiology and is also likely to guide the development of therapeutic interventions to restore homeostasis.

"Pathological fibrils that are found in late-stage neurodegeneration form via liquid-to-solid phase transitions..."

Disease proteins undergo phase transition via homotypic interactions

A phase transition (Box 1) is a sharp change to one or more physical properties of a physicochemical system. In macromolecular solutions, the relevant physical properties are termed symmetries (Box 1). A disordered system has high symmetry in that measurable properties of the system are invariant to molecular translations, rotations, vibrations, density fluctuations, and changes in conformation. Thus, in general terms, phase transitions represent the breaking of one or more symmetries. Phase transitions abound in nature and are particularly important in cell biology.

A system comprising macromolecules plus solvent can undergo a particular type of phase transition known as phase separation (Box 1). Beyond a system-specific threshold concentration, the macromolecular solution can separate into two coexistent phases: a dense phase enriched in macromolecules, and a dilute phase that is deficient in macromolecules. Of particular relevance in biological systems are liquid-liquid phase separation (LLPS) and liquid-to-solid phase transitions. In LLPS, macromolecules separate from solution to form a dense liquid phase. Pathological fibrils that are found in late-stage neurodegeneration form via liquid-to-solid phase transitions,

[1]Howard Hughes Medical Institute, Department of Cell and Molecular Biology, St. Jude Children's Research Hospital, Memphis, TN 38105, USA. [2]Department of Biomedical Engineering, Washington University, St. Louis, MO 63130, USA. [3]Center for Science and Engineering of Living Systems, Washington University, St. Louis, MO 63130, USA.
***Corresponding author. Email: jpaul.taylor@stjude.org**

where the precursor liquid phase can be either the dilute solution of macromolecules or the dense phase formed via LLPS.

Many proteins associated with neurodegeneration, including tau, α-synuclein, TDP-43, hnRNPA1, TIA1, and FUS, are capable of fibril formation in vitro (*1*–*7*). Although this process is often referred to as aggregation, phase transitions and aggregation are distinct phenomena: Whereas fibril formation is mediated by homotypic interactions and governed by the principles of phase transitions, aggregation refers to the sticking of molecules to one another, unconstrained by concentration thresholds or accompanied by symmetry breaking. This distinction between pathological phase transition and aggregation is more than academic because it is informative with respect to how such fibrils may arise in a pathological context, the cellular processes that may be disturbed by pathological phase transitions, and how such pathology may be reversed.

Beginning in 2015, it was appreciated that many neurodegenerative disease-related proteins not only assemble into fibrillar solids but also undergo LLPS to form liquid droplets. This phenomenon was first illustrated for hnRNPA1 (*8*, *9*), TDP-43 (*8*), and FUS (*10*, *11*). These results indicated that many disease proteins exhibit distinct concentration thresholds that correspond to the onset of two types of phase transitions: one threshold for LLPS, and a higher threshold for liquid-to-solid phase transition (*8*, *10*). It was also shown that the liquid-to-solid phase transition can be enhanced within the liquid phase (*8*, *10*). Similar observations have been made for tau (*12*), α-synuclein (*13*), huntingtin (*14*), and TIA1 (*7*). These observations have highlighted two distinct routes to fibril formation (Fig. 1). Fibril formation may be initiated by a combination of primary and secondary nucleation (*15*) in dilute solution. Alternatively, fibril formation may occur via liquid-to-solid phase transition within the dense liquid phase. In the latter case, the condensed liquid state that arises from LLPS facilitates fibril formation by concentrating proteins and enabling the crossing of the threshold concentration for liquid-to-solid phase transition. These routes are not mutually exclusive and might differ for different proteins and different contexts (e.g., in vitro versus in living cells).

Neurodegenerative disease–causing mutations in the low-complexity domains (LCDs) of hnRNPA1 (*6*), hnRNPA2B1 (*6*), TDP-43 (*16*), FUS (*10*, *11*), and TIA1 (*7*) are known to reduce the concentration threshold for LLPS. These mutations can also reduce the threshold for liquid-to-solid phase transitions within dense liquid phases, giving rise to pathological phase transitions. Likewise, it was recently shown that disease-causing mutations in α-synuclein also alter the threshold for liquid-to-solid phase transition (*13*).

The biophysical properties of proteins can be regulated by local changes in the cellular milieu (e.g., pH) or chemical modifications. Indeed, a number of posttranslational modifications associated with disease in these proteins also reduce the threshold for phase transitions driven by homotypic interactions that promote fibril formation (*12*, *13*, *17*).

Neurodegenerative disease proteins are constituents of complex condensates

A common feature of purified disease proteins is their ability to undergo phase transitions driven by homotypic interactions, and such phase transitions are promoted by disease-causing mutations. However, the situation is far more complex in living cells. Moreover, the degeneration of neurons cannot be explained solely by fibril formation by a disease protein. Rather, understanding the pathogenesis of neurodegeneration requires consideration of underlying cellular processes that are corrupted over time. Notably, many proteins associated with neurodegeneration reside primarily within cellular bodies known as biomolecular condensates that assemble via phase separation and encompass hundreds of distinct protein and nucleic acid components (*13*, *18*–*22*) (Fig. 2). Biomolecular condensates are distinct from simple liquid droplets or solids formed via phase transitions mediated by homotypic interactions of specific proteins. Instead, they form and dissolve via reversible phase transitions of multicomponent systems that are controlled by dynamic networks of homotypic and heterotypic interactions (*23*).

In cells, biomolecular condensation provides spatial and temporal control over cellular components and biochemical reactions (*24*). A plethora of condensates are found in cells spanning a vast range of sizes and compositions. For example, the central channels of nuclear pores are small condensates composed of a few different biomolecules, whereas ribonucleoprotein (RNP) granules are large, complex

Box 1. Glossary of terms.

Condensate: A generic term used to refer to membraneless cellular structures that concentrate biomolecules. These structures can form through reversible phase transitions.

Phase transition: An abrupt, highly cooperative change to order parameters caused by the breaking of symmetries that in turn leads to a change in the state of matter.

Symmetry: The invariance of a physical system to operations such as translations of molecules along, or rotations about, defined axes. Disorder is characterized by a state of high symmetry; disorder-to-order transitions occur by the breaking of specific symmetries.

Phase separation: A type of phase transition in which a system separates into one or more coexisting phases. In a binary mixture, the coexisting phases are dense and dilute phases. If the dense and dilute phases that result from phase separation are liquids, then the transition is referred to as liquid-liquid phase separation (LLPS).

Percolation: Multivalent macromolecules behave like associative polymers that form reversible, noncovalent cross-links. When the number of cross-links crosses a threshold known as the percolation threshold, a majority of the molecules are incorporated into a system-spanning (percolated) network.

Material properties: With respect to biomolecular condensates, this refers to features such as viscosity, elasticity, and surface tension of the dense phase. These features are manifestations of elastic and dissipative moduli that are governed by the extent of physical cross-linking and the time scales for making and breaking cross-links within condensates. These material properties influence the spatial organization and diffusion of macromolecules within the dense phase, as well as selective permeability to molecules entering the condensate. These material properties are tightly regulated and directly linked to condensate function.

Dynamical arrest: The situation in which making and breaking of cross-links within a condensate becomes so inefficient that dynamism is lost. A network of physical cross-links can trap macromolecules in arrested phases characterized by irregular (aspherical) morphologies and immobile molecules. When cross-links are made and broken efficiently, the resulting condensate can have liquid properties. Excessive cross-linking or reduced efficiency in the breaking of cross-links—such as accompanies many disease-causing mutations—results in arrested dynamism, altered material properties, and impaired function.

Heterotypic buffering: The ability of heterotypic interactions to buffer against the deleterious effects of homotypic interactions that can drive pathological liquid-to-solid phase transitions. This term also refers to the positive effects of heterotypic interactions that suppress dynamical arrest.

condensates containing hundreds of distinct biomolecules that function as discrete membraneless organelles. Indeed, a cell may be viewed as a complex, dynamic network of condensates that are in constant communication through exchange of materials. Biomolecular condensates provide advantages over membrane-mediated compartments in that they concentrate macromolecules in space while enabling rapid exchange of constituents with the surrounding intracellular milieu (*25*). Moreover, many condensates can assemble or disassemble rapidly in response to changes in cellular state. The functional consequences of this dynamic organization include positive and negative regulation of biochemical processes. For example, condensation of multiple enzymes in a common pathway can promote "substrate channeling" wherein the intermediary metabolic product of one enzyme is passed directly to another enzyme without its release into solution, thereby increasing overall efficiency of the pathway (*26*). Such a mechanism may underlie the regulation of multiple glycolytic enzymes at neuronal synapses in an activity-dependent manner (*27*). Condensate formation via phase separation may also have the opposite effect, wherein sequestering one or more constituents in the dense phase may negatively regulate biological activities in the dilute phase (*28*). Condensates can also orchestrate the assembly of complex higher-order structures, such as ribosome subunit assembly in the nucleolus (*29*).

Dynamism in complex biomolecular condensates

The dynamic behavior of condensates reflects the nature of the interactions that underlie their assembly: weak, transient interactions among multivalent biomolecules that form noncovalent cross-links of varying strengths and durations. Above a system-specific threshold, the system becomes populated with sufficient cross-links to form a system-spanning network that holds the condensate together, a phenomenon known as percolation (Box 1). Thus, unlike phase transitions that are driven purely by homotypic interactions, condensate formation requires the crossing of a collective threshold defined by condensate-specific networks of homotypic and heterotypic interactions (*23*). The macromolecular partners engaging in physical cross-links will evolve dynamically, and if such cross-links are made and broken efficiently, the condensate can be highly dynamic and exhibit liquid properties (*30*). The extent of networking and the time scales for making and breaking cross-links contribute directly to the material properties (Box 1) of a condensate. These include properties such as viscosity, elasticity, and surface tension. These material properties are tightly regulated because they influence the spatial organization and diffusion of macromolecules within the dense phase, selective permeability to molecules entering the dense phase, exchange of constituents with the light phase, and ultimately condensate function.

Many proteins associated with neurodegenerative diseases reside within distinct biomolecular condensates. For example, disease-related RNA-binding proteins such as TDP-43, FUS, hnRNPA1, hnRNPA2B1, and TIA1 are constituents of multiple types of RNP condensates that control the fate of RNA molecules as they transit through processes such as splicing, nuclear export, trafficking in the cytoplasm, translation, and degradation (*31*). Disease-causing mutations alter the balance of homotypic and heterotypic interactions in these RNP assemblies, thereby changing their material properties, even in the absence of pathological liquid-to-solid phase transitions (*32*). Indeed, a recurrent observation is that disease-causing mutations lead to dynamical arrest (Box 1) of RNP granules that can impair functions with adverse consequences for RNA metabolism (*6*, *7*, *10*–*13*, *33*). At the same time, concentration of proteins in the dense liquid phase increases the probability of a liquid-to-solid phase transition—particularly in proteins harboring prion-like LCDs.

However, the rarity of this pathological event points to the presence of mechanisms that hold such pathological transitions in check. The primary factor suppressing potentially deleterious excess homotypic interactions is the collective effect of functional networks of heterotypic interactions within condensates, which we term heterotypic buffering (Box 1). Additional checks, most notably the activities of chaperones, are in place to reverse any excess homotypic interactions that escape heterotypic buffering. The concept of heterotypic buffering is particularly useful as a framework to understand sporadic neurodegenerative disease, which culminates in the same pathology as disease arising from rare genetic mutations. According to this view, a variety of insults may intersect at a common point, collectively altering the dynamic network of condensate interactions in such a way as to impair heterotypic buffering, leading to pathological liquid-to-solid transitions and/or dynamical arrest (Fig. 3).

Defining the relationship between pathological phase transitions and disease

Understanding the process of neurodegeneration requires consideration not only of how end-stage proteinaceous deposits arise, but how specific cellular processes are corrupted over time to give rise to neuronal dysfunction and death. Such disturbances can be considered from two perspectives: first, a mode of toxicity in which pathological consequences arise directly from unchecked

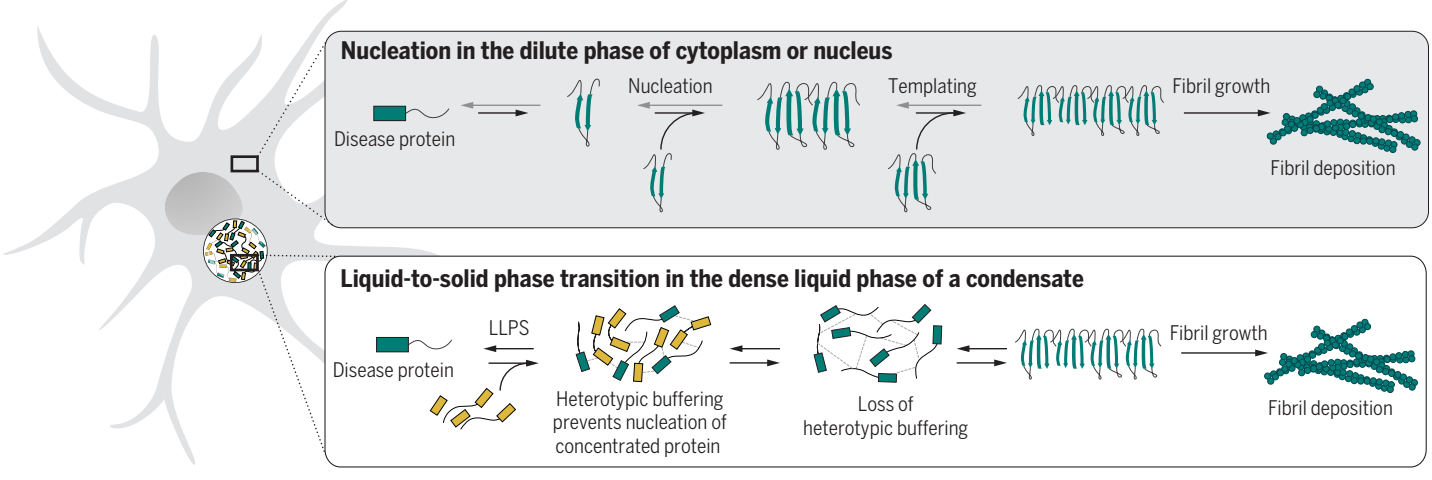


Fig. 1. Two nonexclusive routes lead to fibril formation in neurodegenerative diseases. Fibril formation may be initiated by a primary and secondary nucleation in dilute solution, with subsequent growth through templating of additional units. Alternatively, fibril formation may occur via a liquid-to-solid phase transition within the dense liquid phase. In the condensed liquid state, fibril formation is facilitated by concentrating proteins and bringing them closer to the threshold for liquid-to-solid phase transition. These routes are not mutually exclusive and may be influenced by context.

homotypic phase transition leading to fibril formation; and second, a mode of toxicity in which pathological phase transitions have broader functional consequences by altering key properties of biomolecular condensates, thereby impairing the ability of condensates to regulate biological activities.

With regard to the first mode of toxicity, the accumulation of disease proteins in fibrillar deposits may directly impair cellular function. One possible mechanism is that pathological protein deposits create a sink that depletes cells of that particular protein. For example, accumulation of fibrillar TDP-43 or FUS pathology in the cytoplasm is accompanied by gradual depletion from nuclei (*34*), and this nuclear depletion may lead to a partial loss of function (*35*, *36*). Pathological proteinaceous deposits may also sequester additional factors leading to their functional depletion. For example, it has been suggested that recruitment of chaperones to such pathology may deplete the capacity of protein quality control mechanisms, with widespread implications (*37*).

Alternatively, pathological phase transitions may exert broad cellular toxicity by influencing the network of interactions that define the nature and function of biomolecular condensates. A prominent example is amyotrophic lateral sclerosis with frontotemporal dementia (ALS-FTD), which arises from pathological variants in at least eight different RNA-binding proteins, including TDP-43 (*38*), FUS (*39*, *40*), hnRNPA1 (*6*), hnRNPA2 (*6*), TIA1 (*7*, *41*), matrin 3 (*42*), ataxin 2 (*43*), and annexin A11 (*44*). All of these proteins reside within biomolecular condensates that are distributed throughout the nucleus and cytoplasm of cells and govern many aspects of RNA metabolism. Disease-associated mutations in these RNA-binding proteins alter the material properties of their native condensates, and it is therefore unsurprising that ALS-FTD is associated with widespread disturbance of RNA metabolism (*45*). It remains to be determined whether disease-causing mutations in other condensate-resident proteins (e.g., tau and α-synuclein) also impair functions of specific condensates. Beyond impairing function, perturbation of condensate material properties can simultaneously enhance the driving forces for additional symmetry-breaking operations, notably liquid-to-solid phase transition. As a result, dynamically arrested condensates can also become crucibles driving the formation of fibrillar deposits that arise from homotypic interactions, as described above (Fig. 1). Thus, disturbance of the material properties of condensates can drive pathology through two consequences that are inextricably linked: impairment of native condensate function and production of proteinaceous pathology.

Remarkably, disease-causing mutations in a variety of proteins that are not typically thought of as constituents of condensates can also cause dynamical arrest. These include proteins involved in the maintenance and clearance of condensates, such as valosin-containing protein (VCP) (*46*–*48*), UBQLN2 (*49*, *50*), and OPTN (*51*, *52*). Furthermore, pathological polydipeptides arising from expanded *C9ORF72* produce widespread disturbances in biomolecular condensate function. Specifically, arginine-containing polydipeptides (polyGR and polyPR) become concentrated within biomolecular condensates and alter their material properties through extensive interactions with LCDs (*53*). For example, *C9ORF72*-related polydipeptides impair the central channel of the nuclear pore (*54*), disturb the dynamics of stress granules and RNA transport granules (*53*, *55*), and impair the dynamics and material properties of nucleoli, resulting in reduced ribosome biogenesis (*53*, *56*, *57*) and a decrease in the ability of the nucleolus to buffer against nuclear protein misfolding (*58*).

The mechanisms described above focus on pathological phase transitions of proteins. However, we note that RNA molecules are also well suited to driving phase transitions. Indeed, pathological expansion of RNA repeats, such as those observed in myotonic dystrophy types 1 and 2 and *C9ORF72*-related ALS-FTD, is marked by pathological RNA phase transition resulting in RNA foci (*59*). These RNA-driven pathological phase transitions can also initiate toxicity via mechanisms that are remarkably similar to those observed with protein deposits, including sequestration of RNAs and RNA-binding proteins (*60*).

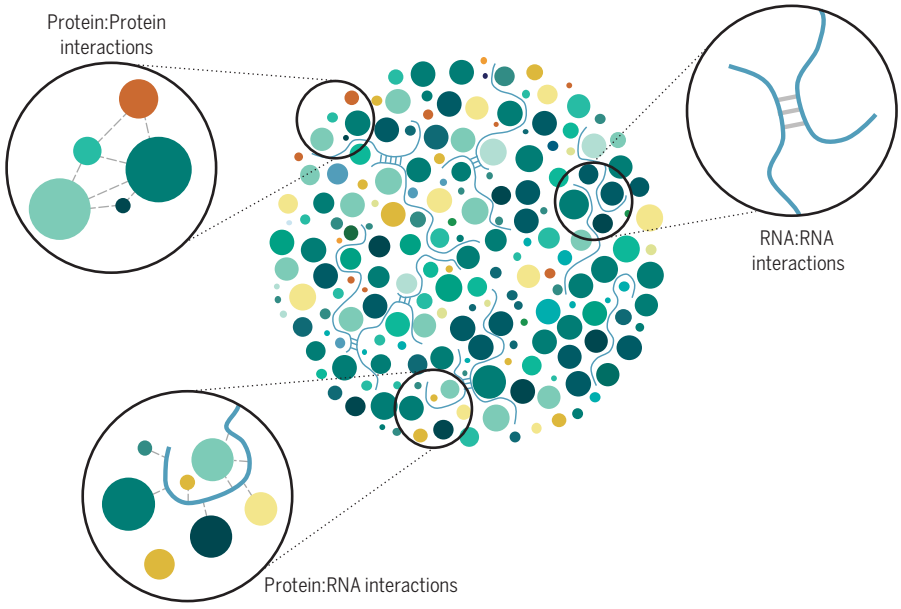


Fig. 2. Condensates arise through a network of heterotypic and homotypic interactions. Condensates form through phase separation of multiple types of macromolecules. In RNP granules, for example, multivalent proteins and RNA molecules participate in a network of homotypic and heterotypic interactions that collectively determine the concentration threshold for LLPS and the material properties of the resulting condensate. The material properties of biomolecular condensates, such as viscosity, elasticity, and surface tension of the dense phase, are governed by the extent of physical cross-linking and the time scales for making and breaking cross-links within condensates. These material properties influence the spatial organization and diffusion of macromolecules within the dense phase, as well as selective permeability to molecules entering the condensate. These material properties are tightly regulated and directly linked to condensate function.

Perspectives

In light of the past decade's worth of research, the commonly held perspective that neurodegeneration is caused by aggregation of misfolded, toxic proteins is evolving toward the view that pathological phase transitions represent a common principle underlying neurodegeneration. This insight is important because it focuses our attention squarely on the dynamic cellular condensates that are assembled from these proteins. Pathological phase transitions of disease proteins, irrespective of which route they take to fibril formation, are inextricably linked to the functions of the condensates in which they reside. According to this view, the primary manifestations of cellular dysfunction in the context of disease are twofold: (i) altered material properties due to dynamical arrest of condensates, and (ii) pathological liquid-to-solid transitions. Accordingly, reversing these defects should be the objective of therapeutic intervention.

As described above, the percolation threshold and material properties of condensates are defined by the dynamic network of homotypic and heterotypic interactions that define them. Indeed, it is now evident that manipulation of individual constituents is sufficient to alter the

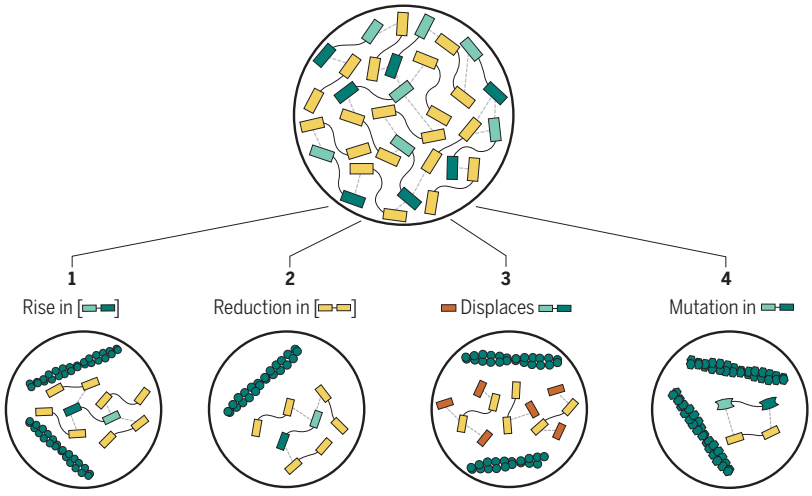


Fig. 3. Disruption of the physiologically relevant interplay between homotypic and heterotypic interaction can lead to pathological phase transitions. Within multicomponent condensates, homotypic interactions are typically buffered by an abundance of heterotypic interactions. This buffering guards against pathological homotypic interactions that may give rise to liquid-to-solid phase transition and fibril formation. In this simplified example, two types of macromolecules are illustrated, each depicted with two interacting domains (rectangles) connected by a spacer region (black line). Green macromolecules can form homotypic interactions with other green macromolecules or heterotypic interactions with yellow macromolecules. Yellow macromolecules can form heterotypic interactions with green macromolecules. Under normal conditions (top), the concentration of yellow macromolecules is such that heterotypic interactions buffer any homotypic interactions. Four scenarios are shown in which this buffering is disrupted: (1 and 2) imbalances in the relative concentrations of condensate constituents, (3) competition for heterotypic interactions by the presence of an additional molecule, and (4) pathological mutations that favor homotypic interactions. In each case, excessive homotypic interactions can lead to the deposition of solid-like structures within condensates if they escape protein quality control mechanisms.

percolation threshold and material properties of specific condensates. Thus, one promising approach to restoring material properties to condensates is to target individual constituents that are deemed druggable by conventional criteria, with the desired effect of shifting the percolation threshold and material properties of a specific condensate in the desired direction. For example, in the setting of ALS, there is evidence of altered material properties and functions of a variety of RNP granules, such as stress granules and RNA transport granules. It has been proposed that depletion of ataxin 2 ameliorates neurodegeneration in a mouse model of ALS through precisely this mechanism, resulting in decreased TDP-43 pathology (*61*). Another attractive target to restore normal material properties in this way may be G3BP, which serves as a more central node in the assembly of these condensates. Indeed, knockdown or inhibition of G3BP enhances local translation in neuronal processes, protects against axonal injury, and promotes axonal regeneration (*62*).

Beyond targeting the network of condensate constituents to influence material properties, an alternative approach might be to target key pathological phase transitions themselves, such as assembly of pathological TDP-43 or FUS fibrils. In this strategy, one might exploit the chaperones that specifically target these assemblies. Specifically, it was recently shown that nuclear import receptors (karyopherins) function as chaperones for and can reverse pathological phase transitions of their clients. Indeed, the activity of KapB2 reverses pathological phase transitions by proteins harboring the cognate PY-NLS (i.e., FUS, hnRNPA1, hnRNPA2), whereas the activity of KapB1 together with importin-α reverses pathological phase transitions of TDP-43 in vitro and in vivo (*63*). Thus, strategies to augment the activity of these chaperones may be beneficial in a disease setting. Related to this, VCP serves as an important segregase in the dismantling of stress granules and perhaps other condensates. Indeed, it was recently shown that a small molecule agonist of ULK1/2 kinases, which phosphorylate and activate VCP, accelerates stress granule disassembly (*64*). It is likely that our improved understanding of the causes of dynamical arrest, pathological liquid-to-solid transitions, and the loss of heterotypic buffering will pave the way for targeting functional restoration and regulation of condensates as potent therapeutic strategies.

REFERENCES AND NOTES

1. M. Kato *et al.*, *Cell* **149**, 753–767 (2012).
2. Q. Cao, D. R. Boyer, M. R. Sawaya, P. Ge, D. S. Eisenberg, *Nat. Struct. Mol. Biol.* **26**, 619–627 (2019).
3. B. Li *et al.*, *Nat. Commun.* **9**, 3609 (2018).
4. M. Goedert, D. S. Eisenberg, R. A. Crowther, *Annu. Rev. Neurosci.* **40**, 189–210 (2017).
5. D. T. Murray *et al.*, *Cell* **171**, 615–627.e16 (2017).
6. H. J. Kim *et al.*, *Nature* **495**, 467–473 (2013).
7. I. R. Mackenzie *et al.*, *Neuron* **95**, 808–816.e9 (2017).
8. A. Molliex *et al.*, *Cell* **163**, 123–133 (2015).
9. Y. Lin, D. S. Protter, M. K. Rosen, R. Parker, *Mol. Cell* **60**, 208–219 (2015).
10. A. Patel *et al.*, *Cell* **162**, 1066–1077 (2015).
11. T. Murakami *et al.*, *Neuron* **88**, 678–690 (2015).
12. S. Wegmann *et al.*, *EMBO J.* **37**, e98049 (2018).
13. S. Ray *et al.*, *Nat. Chem.* **12**, 705–716 (2020).
14. J. Yang, X. Yang, *Front. Genet.* **11**, 754 (2020).
15. A. K. Buell *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **111**, 7671–7676 (2014).
16. A. E. Conicella, G. H. Zerze, J. Mittal, N. L. Fawzi, *Structure* **24**, 1537–1549 (2016).
17. S. Qamar *et al.*, *Cell* **173**, 720–734.e15 (2018).
18. D. S. W. Protter, R. Parker, *Trends Cell Biol.* **26**, 668–679 (2016).
19. Y. R. Li, O. D. King, J. Shorter, A. D. Gitler, *J. Cell Biol.* **201**, 361–372 (2013).
20. N. H. Alami *et al.*, *Neuron* **81**, 536–543 (2014).
21. R. Tan *et al.*, *Nat. Cell Biol.* **21**, 1078–1085 (2019).
22. J. Wang *et al.*, *Cell* **174**, 688–699.e16 (2018).
23. P. Yang *et al.*, *Cell* **181**, 325–345.e28 (2020).
24. S. F. Banani, H. O. Lee, A. A. Hyman, M. K. Rosen, *Nat. Rev. Mol. Cell Biol.* **18**, 285–298 (2017).
25. Y. Shin, C. P. Brangwynne, *Science* **357**, eaaf4382 (2017).
26. L. J. Sweetlove, A. R. Fernie, *Nat. Commun.* **9**, 2136 (2018).
27. S. Jang *et al.*, *Neuron* **90**, 278–291 (2016).
28. F. Wippich *et al.*, *Cell* **152**, 791–805 (2013).
29. D. M. Mitrea *et al.*, *Nat. Commun.* **9**, 842 (2018).
30. T. S. Harmon, A. S. Holehouse, M. K. Rosen, R. V. Pappu, *eLife* **6**, e30294 (2017).
31. S. L. McKnight, *Trends Biochem. Sci.* **44**, 899–901 (2019).
32. S. Boeynaems *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **116**, 7889–7898 (2019).
33. C. Fallini, G. J. Bassell, W. Rossoll, *Hum. Mol. Genet.* **21**, 3703–3718 (2012).
34. J. P. Taylor, R. H. Brown Jr., D. W. Cleveland, *Nature* **539**, 197–206 (2016).
35. J. P. Ling, O. Pletnikova, J. C. Troncoso, P. C. Wong, *Science* **349**, 650–655 (2015).
36. C. Lagier-Tourenne *et al.*, *Nat. Neurosci.* **15**, 1488–1497 (2012).
37. H. Olzscha *et al.*, *Cell* **144**, 67–78 (2011).
38. J. Sreedharan *et al.*, *Science* **319**, 1668–1672 (2008).
39. T. J. Kwiatkowski Jr. *et al.*, *Science* **323**, 1205–1208 (2009).
40. C. Vance *et al.*, *Science* **323**, 1208–1211 (2009).
41. Y. Lee *et al.*, *J. Clin. Invest.* **128**, 1164–1177 (2018).
42. A. Mensch *et al.*, *Exp. Neurol.* **306**, 222–231 (2018).
43. A. C. Elden *et al.*, *Nature* **466**, 1069–1075 (2010).
44. Y. C. Liao *et al.*, *Cell* **179**, 147–164.e20 (2019).
45. N. B. Nedelsky, J. P. Taylor, *Nat. Rev. Neurol.* **15**, 272–286 (2019).
46. J. O. Johnson *et al.*, *Neuron* **68**, 857–864 (2010).
47. G. D. Watts *et al.*, *Nat. Genet.* **36**, 377–381 (2004).
48. J. R. Buchan, R. M. Kolaitis, J. P. Taylor, R. Parker, *Cell* **153**, 1461–1474 (2013).
49. H. X. Deng *et al.*, *Nature* **477**, 211–215 (2011).
50. E. J. Alexander *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **115**, E11485–E11494 (2018).
51. H. Maruyama *et al.*, *Nature* **465**, 223–226 (2010).
52. C. Pottier *et al.*, *Acta Neuropathol.* **130**, 77–92 (2015).
53. K. H. Lee *et al.*, *Cell* **167**, 774–788.e17 (2016).
54. K. Y. Shi *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **114**, E1111–E1117 (2017).
55. S. Boeynaems *et al.*, *Mol. Cell* **65**, 1044–1055.e5 (2017).
56. A. R. Haeusler *et al.*, *Nature* **507**, 195–200 (2014).
57. I. Kwon *et al.*, *Science* **345**, 1139–1145 (2014).
58. F. Frottin *et al.*, *Science* **365**, 342–347 (2019).
59. A. Jain, R. D. Vale, *Nature* **546**, 243–247 (2017).
60. L. J. Sznajder, M. S. Swanson, *Int. J. Mol. Sci.* **20**, 3365 (2019).
61. L. A. Becker *et al.*, *Nature* **544**, 367–371 (2017).
62. P. K. Sahoo *et al.*, *Nat. Commun.* **9**, 3358 (2018).
63. L. Guo *et al.*, *Cell* **173**, 677–692.e20 (2018).
64. B. Wang *et al.*, *Mol. Cell* **74**, 742–757.e8 (2019).

ACKNOWLEDGMENTS

We thank N. Nedelsky for editorial assistance and M. White for assistance with Fig. 2. J.P.T. acknowledges helpful interactions with R. Parker, M. Rosen, T. Mittag, and B. Seeley. R.V.P. acknowledges helpful interactions with J.-M. Choi, F. Dar, M. Farag, A. Holehouse, and K. Ruff. **Funding:** Supported by HHMI (C.M. and J.P.T.), NIH grants R35NS097974 (J.P.T.) and 5R01NS056114 (R.V.P.), NSF grant MCB1614766 (R.V.P.), and the St. Jude Children's Research Hospital Research Collaborative on Membraneless Organelles (J.P.T. and R.V.P.). The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH. **Competing interests:** J.P.T. is a consultant for Nido Biosciences and Faze Medicines. R.V.P. is a member of the scientific advisory board of DewpointX. This work was not funded or influenced in any way by these affiliations.

10.1126/science.abb8032

REVIEW

Translating genetic risk of Alzheimer's disease into mechanistic insight and drug targets

Annerieke Sierksma[1,2], Valentina Escott-Price[3,4]*, Bart De Strooper[1,2,5]*

To provide better prevention and treatment, we need to understand the environmental and genetic risks of Alzheimer's disease (AD). However, the definition of AD has been confounded with dementia in many studies. Thus, overinterpretation of genetic findings with regard to mechanisms and drug targets may explain, in part, controversies in the field. Here, we analyze the different forms of genetic risk of AD and how these can be used to model disease. We stress the importance of studying gene variants in the right cell types and in the right pathological context. The lack of mechanistic understanding of genetic variation has become the major bottleneck in the search for new drug targets for AD.

The number of people worldwide suffering from dementia already slightly exceeds the number of people with cancer and is poised to increase even more over the next decades. Dementia is, however, a container term for the end symptoms of a wide variety of brain diseases, including Alzheimer's disease (AD). AD is a slowly progressing disorder characterized by specific protein accumulations in the brain. Clinical dementia manifests only late, which confounds many case-control studies using this criterion as a proxy for AD. Prodromal AD patients become excluded, and not-yet-cognitively-altered AD cases are mixed with controls. About 30% of clinically diagnosed patients have no neuropathological or biomarker characteristics of AD (*1*), and 56% of cases defined as AD present with common comorbidities such as Lewy body disease, vascular pathology, or hippocampal sclerosis (*2*). Unfortunately, the advice of the National Institute of Aging and the Alzheimer Association Research Framework to define AD as a biological construct (*3*) is not yet widely adopted.

A recent comprehensive overview estimated that 35% of lifetime risk of dementia is modifiable, including factors such as education, vascular aspects, hearing loss, social deprivation, and so on (*4*). The Framingham Heart Study confirms the modifiable nature of dementia risk, with decreasing incidence of dementia over the past few decades (*5*). However, although this trend was highly statistically significant for dementia overall and for vascular dementia in particular, it was not significant for AD alone (*5*). Given the high heritability of AD (*6*, *7*), studying genetic risk seems a more fruitful way forward to identify molecular mechanisms of disease. The central question is whether there is one central route to AD, and therefore one "type" of AD, or whether various pathogenic mechanisms exist that converge on the defining amyloid plaque and tangle pathology.

The heritability of AD

Heritability, formally defined as the proportion of phenotypic variance that is due to genetic factors, can be used as a population-based measure for the risk of disease (see the glossary in Box 1). Importantly, the inheritance of genetic risk variants does not necessarily imply disease, and not all individuals with AD carry the same risk variants.

The best-studied risk (or, better, causal) genetic variants in AD are the fully penetrant mutations in the genes encoding amyloid precursor protein (*APP*) and presenilin 1 and 2 (*PSEN1/2*). They affect the processing of the amyloid β (Aβ) peptide, indicating that Aβ peptide aggregation is an upstream event in the pathogenesis of AD (*8*). These mutations were identified in families with a Mendelian, dominantly inherited form of AD (*8*) that manifests clinically as early-onset dementia (onset before <65 years). Estimates for the heritability of early-onset AD are very high, ranging between 0.92 and 1 (*7*). Even in this smaller group (<10% of total AD patients), *APP* and *PSEN1/2* mutations explain only about 10% of these early-onset cases (*7*). The remaining heritability is explained by *APP* duplications; by an increasing number of rare variants in genes encoding, for example, the sortilin-related receptor (*SORL1*), triggering receptor expressed on myeloid cells 2 (*TREM2*), and ATP binding cassette subfamily A member 7 (*ABCA7*) (*9*); and, finally, by not-yet-identified, but likely recessive, mutations (*7*). An example of a recessive AD mutation is Ala^{673}→Val (A673V) in *APP* (*10*).

In the large group of patients in which dementia manifests after age 65, the heredity is also large, estimated between 0.58 and 0.78 with rather large 95% confidential intervals, [0.19 to 0.87] and [0.67 to 0.88], respectively (*6*). This is high compared with other late-onset diseases. The genetic architecture underlying AD at >65 years old is far from fully charted (see Fig. 1). Apolipoprotein ε4 allele (*APOE4*) is the only common high-risk genetic variant [odds ratio (OR) = 3.32] (*11*, *12*). Genome-wide association studies (GWASs) have further identified many common genetic variants with low risk (OR = 1.1 to 1.2), of which 40 have genome-wide significance (*12*–*14*). Exome chip analyses have additionally yielded rare variants in the very same genes—that is, *SORL1*, *TREM2*, and *ABCA7*—that strongly increase risk of early-onset AD (*9*, *15*). Variants that protect have also been discovered: the *APOE2* allele (OR = 0.6) (*16*), rare mutations in phosphatidylinositol-specific phospholipase C-gamma 2 (*PLCG2*) (OR = 0.68) (*15*), and the Icelandic mutation Ala^{673}→Thr (A673T) in *APP* (OR = 0.19) (*10*). The genetics of early- and late-onset disease suggest that AD should be considered a continuum.

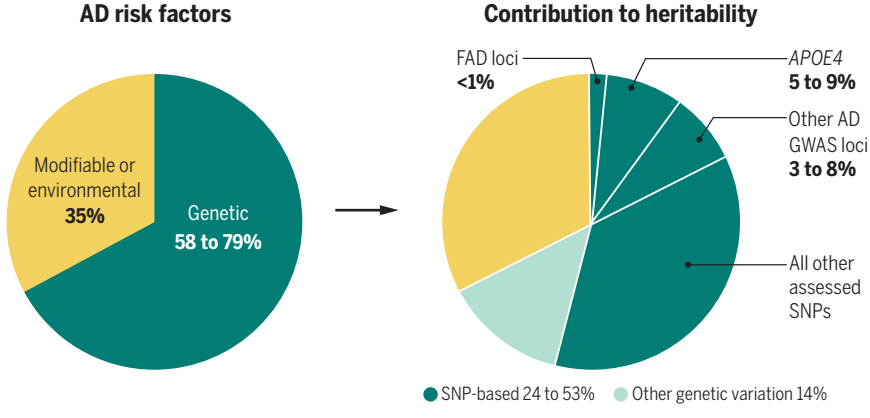


Fig. 1. Risk factors and heritability for AD. Whereas 35% of lifetime risk for AD is composed of modifiable or environmental risk factors, 58 to 79% of AD risk is genetic. The genetics of AD can be broken down into SNP-based heritability and other types of genetic variation, including rare variants, structural and copy-number variation, duplications, SNP×SNP interaction, dominance, and so on. FAD, familial AD.

[1]VIB Center for Brain & Disease Research, Leuven, Belgium. [2]Laboratory for the Research of Neurodegenerative Diseases, Department of Neurosciences, Leuven Brain Institute (LBI), KU Leuven (University of Leuven), Leuven, Belgium. [3]Medical Research Council Centre for Neuropsychiatric Genetics and Genomics, Cardiff University, Cardiff, UK. [4]UK Dementia Research Institute, Cardiff University, Cardiff, UK. [5]UK Dementia Research Institute, University College London, London, UK.
*Corresponding author. Email. escottpricev@cardiff.ac.uk (V.E.-P.); b.strooper@ukdri.ucl.ac.uk (B.D.S.)

As indicated by the broad confidence intervals, the AD heritability estimates remain imprecise, as with many polygenetic disorders (*17*). Increasing efforts to create larger datasets for GWASs, to directly sequence full genomes, and to develop new data analysis methodologies are under way to tackle the "missing heritability" in AD.

From heritability to mechanisms of disease

Translating genetic information into disease mechanisms is anything but trivial. It took 20 years to understand that AD-causing mutations destabilize presenilin, leading to premature release of long Aβ peptides (*18*). Similar efforts to understand how rare mutations in the open reading frames of *TREM2*, *PLCG2*, *SORL1*, and *ABCA7* affect protein function will be needed. In addition, most available genetic information for AD remains imprecise. The causal variant is known for only 40% of the identified GWAS loci (*14*), the effect of these variants is only known in a minority of cases, and literally thousands of these variants contribute to the heritability of the phenotype.

The question is what the core genes are, that is, which genes execute a direct effect on the disease process. Unfortunately, more than 70% of variants that determine phenotypic variation are in "peripheral" genes. Such genes have only indirect effects on expression or posttranslational modification of core gene products and, as such, are not very informative for the molecular mechanisms driving the phenotype (*19*). The individual "trans" effects of these peripheral genes are small (*19*), but, because there are many, they underlie a large part of the heritability (*20*, *21*). Even using the p value of $p < 5 \times 10^{-8}$ for genome-wide significance to prioritize gene loci [which now comprise 40 loci in AD (*14*)] does not provide certainty of finding "core disease-pathway genes" (*19*). The frustrating conclusion is that the bulk of the heredity in AD likely only indirectly points to key biological pathways of disease.

One group of peripheral genes—that is, master regulator genes—is nevertheless of particular interest. These genes—encoding, for example, transcription factors, chromatin modifiers, regulatory RNA, or enzymes—regulate the expression or function of several disease core genes. For example, the AD risk locus Spi-1 proto-oncogene (*Spi1*) codes for the transcription factor Pu.1, which regulates many microglia genes, pointing to a role for inflammation in AD (*22*). Such master regulators are usually under strong evolutionary constraint and so not easily detected in GWASs (*19*).

One could try to investigate how peripheral genes affect the expression of core genes. A prerequisite is to understand in which cells these peripheral genes exert their effect, and hence, single-cell analyses of gene expression in brain cells is crucial (*23*). Such trans-expression quantitative trait loci (trans-eQTL) mapping, however, needs huge datasets and is only readily available for peripheral blood cells. Another possibility is to focus on gene variants with large effects on heredity. *APOE4* is the only example in AD. Finally, one could ignore the quantitative contributions of genes to heredity and focus on rare variants, which are likely more central to the disease mechanisms because of their large effect sizes. A potentially fruitful avenue of research is to investigate how the common variants that define heredity regulate these rare variant genes.

The complexity of the *APOE* locus

The three major isoforms of APOE (ε2, ε3, and ε4) are defined by two single-nucleotide polymorphisms (SNPs; rs429358 and rs7412) within exon 4 of the gene (*11*). The ε4 allele (frequency 0.14 in the Caucasian population) provides a 3-fold increased risk of AD, which increases to 14-fold in ε4 homozygotes (*11*, *16*). Conversely, the ε2 allele (frequency 0.08) confers a 1.7-fold decreased risk. This risk is more pronounced in women than in men and is strongly dependent on ethnic background, that is, the ε4 effect is much smaller in the African-American and Hispanic populations (*16*). This illustrates the importance of multiethnic genetic studies when studying the heritability of AD.

The *APOE* locus is highly complex, spanning almost 2 Mb and covering more than 70 genes. Despite being in low linkage disequilibrium with the APOE SNPs for ε2 and ε4, there are many other SNPs in this large locus that show significant association to AD. This might point to other AD risk genes in this locus. Several of these SNPs, however, likely affect expression of APOE. Understanding this will be of tremendous value because it would clarify whether and under what conditions up- or down-regulation of this multifunctional protein could affect the risk of AD.

Under physiological conditions, APOE is mainly expressed by astrocytes, but microglia exposed to Aβ plaques highly up-regulate APOE. It will be critical to unravel how microglial function is affected by different APOE isoforms and how this contributes to disease. Knock out of the gene eliminates the AD-induced inflammatory response in mice (*24*).

Although very relevant to AD, the role of APOE in brain inflammation remains poorly understood. APOE obviously plays a crucial role in cholesterol transport and lipid homeostasis, but it also plays a role in Aβ aggregation, clearance, and cellular uptake and also affects, through less well-understood molecular pathways, synapse number and function, blood-brain barrier integrity, and TAU-mediated neurodegeneration (*16*, *24*, *25*). It is important to decipher which roles of APOE are directly relevant to AD because the variety of functional effects of APOE deficiency in different cell types and in different tissues suggests that the *APOE* gene is a master peripheral regulator in the disease. Not all affected pathways are necessarily relevant to AD. Directly modulating APOE to protect against AD is likely to have a variety of effects, and the outcome of such treatments will need careful monitoring.

Causal, high-risk, and protective variants are involved in APP processing and in microglial function

Evidence of genotype-phenotype dose-responses in an allelic series strongly argues for a core gene function. Such gene-dosage effects are observed

Box 1. Glossary of terms.

Heritability: The proportion of phenotypic variance that is due to genetic factors.

Missing heritability: The difference between the genetic heritability observed in families and the estimated heritability of identified genetic variants in the population.

Core gene: A mutation in this gene will directly affect disease.

Peripheral gene: A mutation in this gene will only indirectly affect disease, most likely through a trans-regulatory effect on core genes.

Core disease-pathway genes: Genes directly affecting pathways that determine disease onset.

Master regulatory gene: A peripheral gene that regulates the expression or function of several core genes in the disease. Examples include transcription factors, regulatory RNAs or enzymes, or chromatin modifiers.

Genotype-phenotype dose-response: Several alleles of a gene affect disease risk, possibly to different degrees—for example, common and rare, or loss- and gain-of-function variants. Either multiple alleles can affect the same gene or causal alleles are present in different genes that cooperate within the same disease pathway. An example is the aβ pathway, where mutations in APP, the presenilins, and α-secretase affect the same pathway and both protective and risk variants have been identified.

Polygenic risk score: A single genetic score indicating a person's risk of developing a trait. Calculated by summing the number of risk alleles present and multiplying this by their effect size, that is, the weight of disease risk.

Linkage disequilibrium: The observation that specific alleles at a particular genomic locus or region are more often co-inherited within the population than is expected by chance.

with genes involved in Aβ generation. Next to the fully penetrant *APP* and *PSEN* mutations, *APP* gene duplications and triplications, including Down syndrome, cause AD (*8*, *10*). A recessive (A673V) and a protective (A673T) allele (*10*) affect the propensity of Aβ to aggregate. A673T also lowers β-secretase processing of APP (*26*). A common allele (rs2154481) in the *APP* locus lowers risk (OR = 0.95) although, counter intuitively, slightly increases APP expression (*14*). Finally, variants in the gene loci of the α-secretases, *ADAM17* and *ADAM10* (*27*), all demonstrate that APP itself and the enzymes processing it to Aβ carry risk and even cause AD.

SORL1 provides another example of an allelic series with increasing risk of AD. *SORL1* encodes the sorting-related receptor with A-type repeats SORLA (also called LR11) involved in retromer-related endosomal traffic. SorlA contains functional domains that can bind monomeric Aβ or APP. Several of the deleterious variants affect those domains (*28*). SorlA lowers Aβ production by redirecting APP to the cell membrane and trans-Golgi network and Aβ to lysosomes in neurons (see Fig. 2) (*28*). Notably, *SORL1* expression is 20-fold higher in human than in mouse microglia, warranting further characterization of the impact of *SORL1* deficiency on microglia functions (*29*). The different *SORL1* variants aggregate into categories with increasing risk burden (*30*), from OR = 1.21 for missense variants up to OR = 16.73 for protein-truncating variants (*9*). These ORs are comparable to heterozygous (OR = 3.2) and homozygous (OR = 14.9) *APOE4* carriers. Variants are present in 2% of AD, compared with <1% for *APP* and *PSEN1* mutations (*30*).

Common and rare *ABCA7* variants provide a third allelic series. ABCA7 promotes the efflux of phospholipids out of cells. Protein-truncating (OR = 2.6) and missense mutations (OR = 1.8) are associated with AD (*31*). In addition, a tandem repeat in intron 18, ranging from 300 base pairs to more than 10 kb, provides relative high risk of AD (OR = 4.5). It remains unclear how loss of function of ABCA7 increases risk of AD, although in mice it causes higher Aβ plaque burden related to impaired Aβ phagocytosis in macrophages and microglia (*32*). Loss of ABCA7, because of its role in lipid transport, might have broad effects on cell physiology (see Fig. 2). The relatively low OR suggests that it is not directly causally involved, and its broad function suggests that it acts as a master regulator peripheral gene in AD.

In addition to core and master regulator genes affecting Aβ processing, strong genetic evidence implicates microglia in AD. Many common variants associated with risk of AD occur in genes that are expressed in microglia [see table 1 in (*33*)]. Rare missense mutations in the open reading frames of *TREM2*, *PLCG2*,

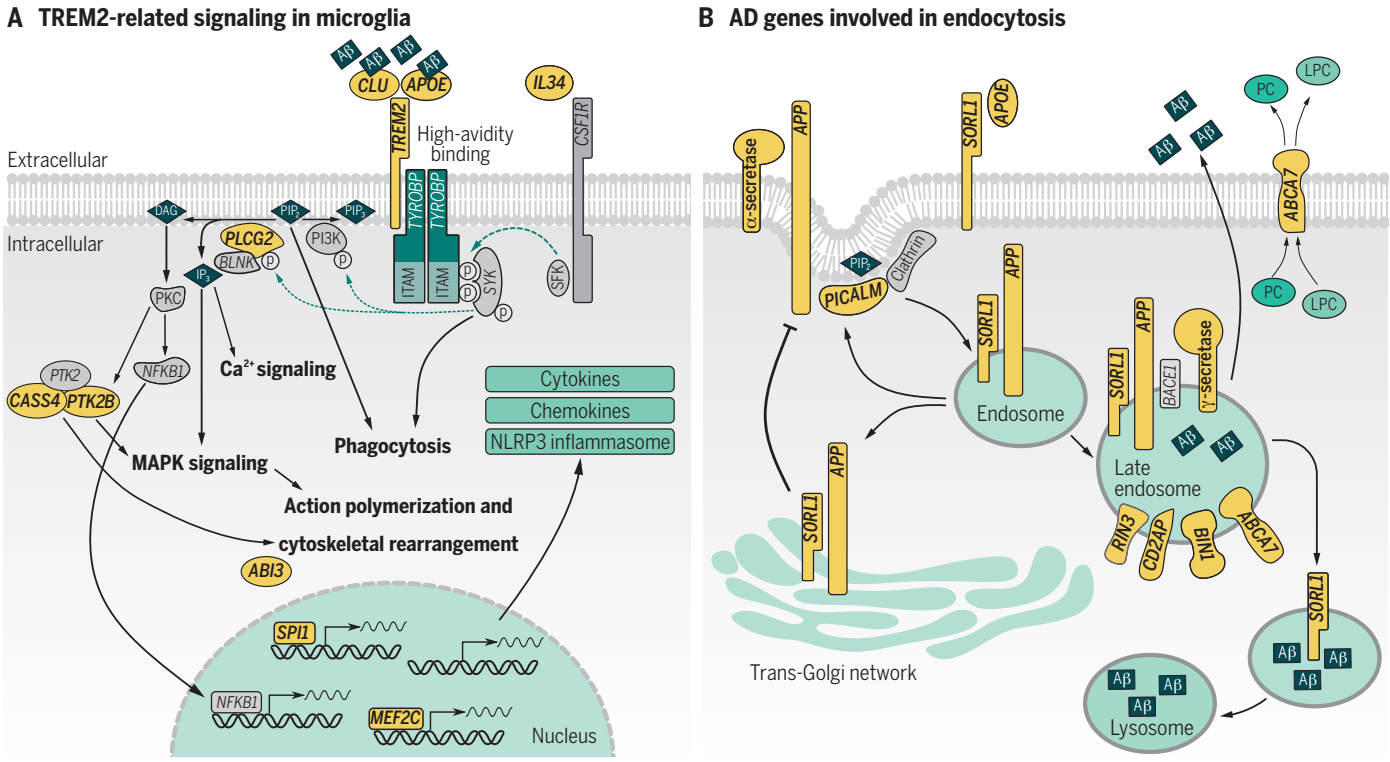


Fig. 2. Emerging signaling pathways in AD. Proteins encoded by *TREM2*, *SORL1* and *ABCA7* interact with proteins encoded by other genetic risk genes for AD (risk genes highlighted in yellow), affecting microglial function and APP processing. (**A**) AD pathway in microglia. TREM2 can bind Aβ that may need to be lipidated by APOE or APOJ [also known as clusterin (*CLU*)] and associates with TYRO protein tyrosine kinase–binding protein (*TYROBP*) to constitute intracellular signaling through its ITAM. The ITAM domain undergoes double phosphorylation by the SRC family kinases (SFK) to allow binding of spleen tyrosine kinase (*SYK*). SYK can phosphorylate phosphoinositide 3-kinase (PI3K) and PLCγ2 (encoded by *PLCG2*). Activation of these proteins ultimately leads to calcium and mitogen-activated protein kinase (MAPK) signaling and nuclear factor κβ (*NFKB1*) transcription. Protein kinase C (PKC) can also activate proline-rich tyrosine kinase 2 (*PTK2B*), which can activate MAPK signaling, but also associates with Cas scaffold protein family member 4 (*CASS4*) and protein tyrosine kinase 2 (*PTK2*) to affect actin polymerization, as does ABI3. Overall, these signaling pathways affect cytoskeletal rearrangements associated with microglial motility and increase phagocytosis. DAG, diacylglycerol; PIP_3, phosphatidylinositol 3,4,5-trisphosphate; P, phosphate. (**B**) Endocytosis and AD genes. SORLA (encoded by *SORL1*) can transfer APP to the trans-Golgi network and late endosomes where it undergoes amyloidogenic processing to Aβ. SorlA can also directly bind Aβ and facilitate its degradation in the lysosomes. Although ABCA7 is involved in cellular lipid homeostasis—for example, regulating the efflux of lysophosphatidyl choline (LPC) and phosphatidyl choline (PC)—ABCA7 can also affect amyloidogenic proteolysis by affecting beta-site APP cleaving enzyme 1 (*BACE1*) expression levels. Several other AD risk genes involved in endocytic pathway are indicated in yellow.

and *ABI3* (Abelson interactor family protein 3), which are genes mainly or exclusively expressed in myeloid cells, also point in that direction (*15*). Many of the risk genes of AD become up-regulated in microglia when exposed to Aβ but less so in TAU pathology, as evidenced when one Aβ model and one TAU model were directly compared (*34*). Thus, a large part of the genetic risk of AD, as opposed to genetic cause of AD, seems to converge into the microglial response to amyloid plaques.

One of the best studied genes in this series is *TREM2*. TREM2 is a receptor for anionic ligands, including phospholipids, lipopolysaccharide, and DNA (*35*). In mouse models of AD, Trem2 is required for the transit of microglia from homeostatic to activated cell states in response to amyloid plaques. Interestingly, these microglia strongly up-regulate APOE expression (*36*). *Trem2* deficiency leads to more diffuse plaques with greater neuritic damage and less recruitment of microglia to amyloid plaques (*36*). The rare Arg^{47}→His (R47H) and more common R62H variants of *TREM2* (*15*) alter its stability, affect phagocytic capacity, and impair TREM2 affinity for APOE, clusterin (ApoJ), low-density lipoproteins, and Aβ (*33*). The impact of other more common variants on TREM2 function remains unclear. Because the R47H and R62H mutations cause partial loss of function of TREM2, and because Trem2 deficiency seems to aggravate amyloid plaque pathology in mice, most drug development efforts are focused on enhancing TREM2 function (*37*). However, enhancing microglia activity might be a double-edged sword, with opposite effects on Aβ and TAU pathology (*24*). A big question that also remains is to what extent observations in mice can be extrapolated to the human pathophysiology: The cellular reactions around amyloid plaques are much more complex in human than in the available mouse models [see, for instance, (*38*, *39*)]. That several additional AD-associated variants have been observed in genes that act downstream of TREM2 nevertheless underlines the importance of TREM2-signaling in AD (see Fig. 2).

One example is the rare protective Pro^{522}→Arg (P522R) variant in the *PLCG2* gene with moderate effect size (OR = 0.57 to 0.68) (*15*). This mutation increases the activity of the microglial signaling enzyme phospholipase C-gamma 2 (PLCγ2) downstream of TREM2 (see Fig. 2). The variant is overrepresented in a cohort of cognitively healthy centenarians and, anecdotally, provides full protection to *APOE4* in a more than 100-year-old homozygous carrier (*40*). PLCγ2 becomes phosphorylated on stimulation and affects phagocytosis, migration, and chemokine and cytokine release (*41*). Structurally, the P522R variant modifies an autoinhibitory domain of PLCγ2, leading to greater phosphatidylinositol 4,5-bisphosphate (PIP_2) conversion and increased cellular calcium release (*41*). The P522R variant enhances Aβ endocytosis, suggesting that this protective variant may facilitate microglial clearance of Aβ.

In conclusion, the genetics of AD provide strong evidence for a major pathway centered on Aβ generation, aggregation, and clearance that operates in early- and late-onset disease. The genetics also strongly implicate microglia responses to amyloid plaques in AD. Assuming that these responses are directed by the genetic-risk profile of the patient, one would predict that some patients are protected from the damage caused by amyloid plaques because of their advantageous microglia (*34*). Major questions for the field are what aspects of the microglia response on amyloid plaques are beneficial or detrimental, how genetic risk affects this balance, and whether this contributes to TAU pathology. Drug development will have to move cautiously, taking into account this fine yin and yang of the cellular response in AD (*23*).

Leveraging polygenic risk

A large proportion of the genetic risk of AD is explained by common variation in the genome and is captured by SNPs in GWASs (*21*). Such single variants on their own do not predict an individual's risk of AD but can be combined in a polygenic risk score (PRS). PRS is a "genetic score" defined as the sum of the number of SNP risk alleles that an individual carries, weighted by their contribution to the disease risk (effect size).

Most investigators currently use a partial AD PRS calculated with the lead SNP in the 40 canonical GWAS genome loci mentioned before (*12*–*14*, *42*). However, a more complete PRS calculation includes the thousands of other SNPs in loci that are associated with risk of AD but did not reach the threshold for genome-wide significant association ($p < 5 \times 10^{-8}$). Such calculation improves the prediction accuracy of AD, something also observed with psychiatric and other complex disorders (*43*). Indeed, the prediction accuracy of AD using the complete PRS is high, with an area under the receiver-operator curve (AUC) of 75% in clinical and 84% in pathologically confirmed samples (*21*, *44*).

Using only the canonical GWAS loci biases the score to the effect of the *APOE* region (*21*). If all genetic risk of AD is used as proposed for the complete PRS, the bulk of associated SNPs of small effect sizes will eventually outperform the effect size of the *APOE* locus alone. Accordingly, the predictive accuracy of complete PRS in pathologically confirmed *APOE3* homozygotes is high, with AUC > 80 (*45*). To date, the PRS approach has mostly been assessed in European populations, owing to a lack of multiethnic GWAS data.

The field is currently struggling to translate the concept of PRS into meaningful functional hypotheses. An interesting recent development is to include only SNPs associated with genes from putative disease-specific pathways—for instance, APP metabolism, lipid metabolism, and endocytosis—to generate pathway-specific PRSs (*27*). However, the definition of these disease pathways is based mostly on the different functional categories defined by gene ontology (*46*). This is problematic (*46*) because there is little expert scrutiny, inclusion thresholds are low, and almost all AD genes are implicated in more than one pathway (*12*, *47*, *48*). It turns out that the AD predictability using such categories is low (*47*). Finally, it is important to mention that the PRS is currently designed as a linear combination of SNP effect sizes without accounting for nonlinear effects, also known as epistasis or SNP×SNP interaction. Biologically, it is very unlikely that genetic risk of AD is the simple additive sum of the individual SNP risks.

From polygenic risk to mechanisms of disease and drug targets: Cellular state and disease context matter

Drugs developed against targets supported by genetic evidence have a better chance to become approved (*49*). However, in the AD field, the causal SNPs are in many cases unknown or assigned to the wrong gene. There are still large gaps in understanding how SNPs affect the functional genomic architecture. Information on the effects of drugs on eQTL are often not in the public domain, making it difficult to link experimental drugs to candidate targets. Overall, the single most important limiting factor in the translation of knowledge from genetics to drugs is, however, the lack of good models for AD (see Fig. 3).

Assessing the functional impact of noncoding risk variants is challenging and starts with the question of whether a particular SNP is functional or is only in linkage disequilibrium with the real functional SNP. Risk mechanisms will only manifest in disease-relevant conditions. Thus, cell type and experimental context really matter when analyzing the functional consequences of SNPs (*23*). Finally, SNPs are frequently assigned to the genes to which they are closest in the linear DNA sequence representation of the genome. However, chromatin has a complex three-dimensional (3D) structure, and enhancers or suppressors can exert their effects on the expression of genes that are remote from their location (*50*). Recent work has indeed shown that many causal variants affect enhancers that are highly specific to brain region, cell type, and cell state (*22*, *50*, *51*). It was noted that mainly myeloid and microglial enhancer regions, and not the promoter regions, are significantly enriched for AD-associated variants (*50*–*52*). An elegant knockout experiment underscored this conclusion. Nott *et al*. deleted in human induced pluripotent stem cells (iPSCs) the *BIN1* (bridging

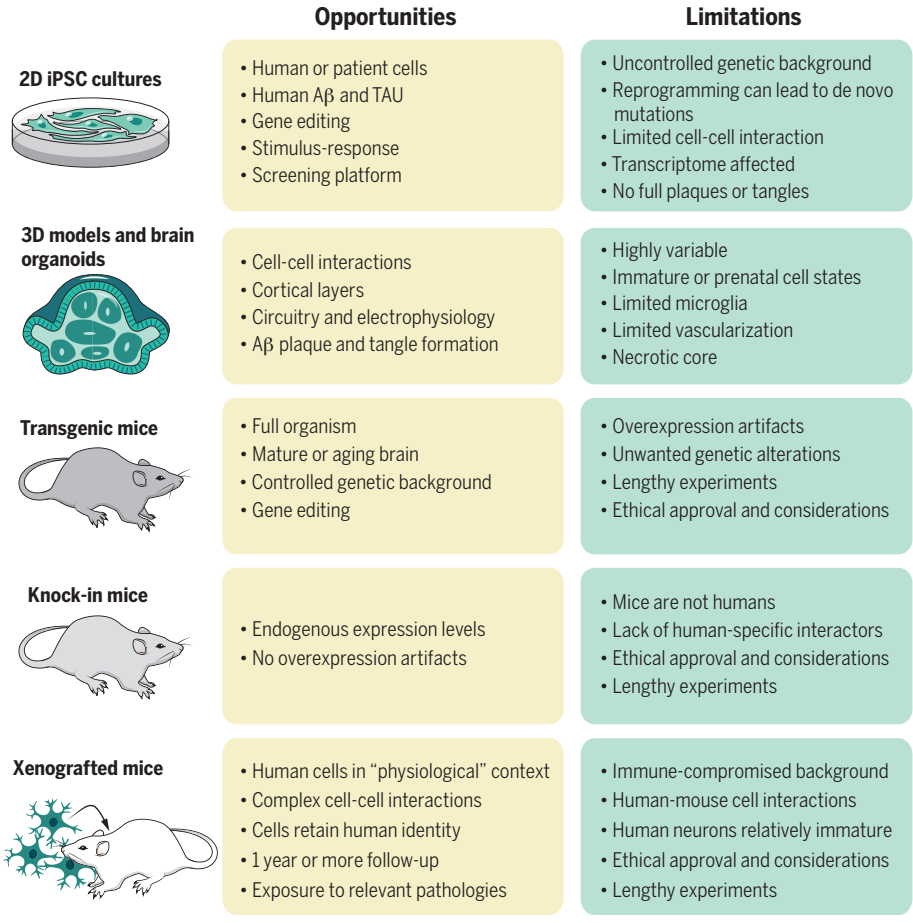


Fig. 3. The opportunities and limitations of commonly used models in AD research.

integrator-1 and amphiphysin-2) enhancer, which carries one of the higher AD risk variants, rs6733839 (OR = 1.2) (*12*, *50*). When the authors differentiated these cells into microglia, astroglia, and neurons, expression of BIN1 was only affected in microglia (*50*).

The AD field really struggles to generate good models that reproduce all features of disease (see Fig. 3). Double and even triple transgenic mice overexpressing human TAU, PSEN, and APP, all with familial AD or frontotemporal degeneration mutations, are needed to obtain amyloid plaques and tangles, and it remains a tantalizing question to what extent cellular phenotypes induced in these mice mimic the situation in human. Sixty-five million years of evolutionary divergence cannot be ignored when modeling a human polygenic disease. With respect to human-specific cell biology, research on human iPSCs has taken off, including in vitro 3D (*53*) and organoid cultures (*54*). All are promising, but each approach comes with its own limitations (see Fig. 3). For instance, the 3D in vitro cultures provide a very artificial conformation to grow the cells and use high overexpression of the *APP* gene in the neurons to obtain AD phenotypes (*53*). Human organoid cultures are promising, but their usefulness to study nondevelopmental disorders remains debated (*54*).

The xenograft, or chimeric, mouse model approach, in which human iPSC-derived brain cells are transplanted into the mouse brain (*55*–*57*), provides an interesting alternative combining several advantages. The rodent brain functions as a superior "physiological" 3D matrix for human cells compared with other more artificial environments. Human neurons (*55*), microglia (*56*, *57*), and astroglia have been grown in rodent brains for more than 1 year and reproduce many human features. Although the rodent brain background and the immune suppression are confounders in these experiments, microglia cells, even after exposure to a cell culture environment, fully regain their identity when returned to the central nervous system and transcriptionally closely resemble freshly isolated human microglia from surgical samples (*57*).

In theory, human iPSCs and their derived models can be used to functionally evaluate the impact of PRS-defined risk in different cell types and AD-relevant contexts. Obviously, the genomic variants captured from different patients will be different, but, because the pathological phenotype of AD patients is very similar, it is assumed that the cellular pathology converges and that shared pathways leading to disease may be identified. Once a critical mass of PRS-defined iPSCs has been analyzed, one can also envision eQTL and regulatory landscape analyses to define how specific AD-associated variants may exert their effects. This can subsequently refine the list of SNPs, including only core variants driving AD pathogenesis. Ultimately, such functional insights will lead to better and more relevant PRSs that will be used for diagnostics, stratification of patients for clinical trials, and personalized medicine based on genetic profile.

Conclusions

The genetic component in AD risk is surprisingly large for a late-onset disorder. Tremendous progress has been made to map this genetic landscape, but now it becomes critically dependent on a better definition of AD and the underlying mechanisms of disease. "More," with respect to cases, is never going to replace "quality," and deeper clinical phenotyping and biomarkers are needed to better interpret the role of genetic variation in specific aspects of the AD phenotype.

While working further along those lines and also from a therapeutic development perspective, it is crucial to take into account the long preclinical phase of AD (*23*). At the functional level, we need to get away from the classical molecular biology paradigms of one gene, one function, one drug target. Gene variants affect gene function in specific genetic backgrounds (mice are not humans), in specific cell types, in specific cell states, and in specific stages of the disease. In silico predictions and simple cell biology experiments, although tempting because of the high throughput, can be very misleading and can jeopardize a whole drug development campaign. Finding drugs for a complicated multifactorial disease like AD requires deep knowledge of the mechanisms that are targeted. The full mapping of the cellular phase of AD is now a priority for the field (*23*).

One should, however, acknowledge the tremendous progress made in AD research. We can now build further on the many hints coming from genetic work over the past decade to generate more sophisticated models that will better represent specific mechanisms underlying AD. This thinking will open many opportunities for drug development, and better stratification of patients will accelerate the road from concept to clinic.

REFERENCES AND NOTES

1. T. G. Beach, S. E. Monsell, L. E. Phillips, W. Kukull, *J. Neuropathol. Exp. Neurol.* **71**, 266–273 (2012).
2. G. D. Rabinovici *et al.*, *Alzheimers Dement.* **3**, 83–91 (2016).
3. C. R. Jack Jr. *et al.*, *Alzheimers Dement.* **14**, 535–562 (2018).
4. G. Livingston *et al.*, *Lancet* **390**, 2673–2734 (2017).
5. C. L. Satizabal *et al.*, *N. Engl. J. Med.* **374**, 523–532 (2016).
6. M. Gatz *et al.*, *Arch. Gen. Psychiatry* **63**, 168–174 (2006).

7. T. S. Wingo, J. J. Lah, A. I. Levey, D. J. Cutler, *Arch. Neurol.* **69**, 59–64 (2012).
8. R. E. Tanzi, L. Bertram, *Cell* **120**, 545–555 (2005).
9. C. Bellenguez *et al.*, *Neurobiol. Aging* **59**, 220.e1–220.e9 (2017).
10. J. Tcw, A. M. Goate, *Cold Spring Harb. Perspect. Med.* **7**, a024539 (2017).
11. E. H. Corder *et al.*, *Science* **261**, 921–923 (1993).
12. B. W. Kunkle *et al.*, *Nat. Genet.* **51**, 414–430 (2019).
13. I. E. Jansen *et al.*, *Nat. Genet.* **51**, 404–413 (2019).
14. I. de Rojas *et al.*, medRxiv 19012021 [Preprint]. 17 January 2020; https://doi.org/10.1101/19012021.
15. R. Sims *et al.*, *Nat. Genet.* **49**, 1373–1384 (2017).
16. M. E. Belloy, V. Napolioni, M. D. Greicius, *Neuron* **101**, 820–838 (2019).
17. A. J. Mayhew, D. Meyre, *Curr. Genomics* **18**, 332–340 (2017).
18. M. Szaruga *et al.*, *Cell* **170**, 443–456.e14 (2017).
19. X. Liu, Y. I. Li, J. K. Pritchard, *Cell* **177**, 1022–1034.e6 (2019).
20. R. S. Desikan *et al.*, *PLOS Med.* **14**, e1002258 (2017).
21. V. Escott-Price *et al.*, *Brain* **138**, 3673–3684 (2015).
22. K. L. Huang *et al.*, *Nat. Neurosci.* **20**, 1052–1061 (2017).
23. B. De Strooper, E. Karran, *Cell* **164**, 603–615 (2016).
24. J. M. Long, D. M. Holtzman, *Cell* **179**, 312–339 (2019).
25. A. Montagne *et al.*, *Nature* **581**, 71–76 (2020).
26. I. Benilova *et al.*, *J. Biol. Chem.* **289**, 30977–30989 (2014).
27. R. Sims, M. Hill, J. Williams, *Nat. Neurosci.* **23**, 311–322 (2020).
28. D. Campion, C. Charbonnier, G. Nicolas, *Acta Neuropathol.* **138**, 173–186 (2019).
29. D. Gosselin *et al.*, *Science* **356**, eaal3222 (2017).
30. H. Holstege *et al.*, *Eur. J. Hum. Genet.* **25**, 973–981 (2017).
31. A. De Roeck, C. Van Broeckhoven, K. Sleegers, *Acta Neuropathol.* **138**, 201–220 (2019).
32. T. Aikawa *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **116**, 23790–23796 (2019).
33. D. V. Hansen, J. E. Hanson, M. Sheng, *J. Cell Biol.* **217**, 459–472 (2018).
34. A. Sierksma *et al.*, *EMBO Mol. Med.* **12**, e10606 (2020).
35. T. K. Ulland, M. Colonna, *Nat. Rev. Neurol.* **14**, 667–675 (2018).
36. S. Parhizkar *et al.*, *Nat. Neurosci.* **22**, 191–204 (2019).
37. K. Schlepckow *et al.*, *EMBO Mol. Med.* **12**, e11227 (2020).
38. Y. Zhou *et al.*, *Nat. Med.* **26**, 131–142 (2020).
39. W.-T. Chen *et al.*, *Cell* **182**, 976–991.e19 (2020).
40. S. J. van der Lee *et al.*, *Acta Neuropathol.* **138**, 237–250 (2019).
41. L. Magno *et al.*, *Alzheimers Res. Ther.* **11**, 16 (2019).
42. R. E. Marioni *et al.*, *Transl. Psychiatry* **8**, 99 (2018).
43. International Schizophrenia Consortium, *Nature* **460**, 748–752 (2009).
44. V. Escott-Price, A. J. Myers, M. Huentelman, J. Hardy, *Ann. Neurol.* **82**, 311–314 (2017).
45. V. Escott-Price, A. Myers, M. Huentelman, M. Shoai, J. Hardy, *J. Prev. Alzheimers Dis.* **6**, 16–19 (2019).
46. F. Koopmans *et al.*, *Neuron* **103**, 217–234.e4 (2019).
47. E. Bellou *et al.*, *Neurobiol. Aging* **93**, 69–77 (2020).
48. G. Leonenko *et al.*, *Ann. Neurol.* **86**, 427–435 (2019).
49. E. A. King, J. W. Davis, J. F. Degner, *PLOS Genet.* **15**, e1008489 (2019).
50. A. Nott *et al.*, *Science* **366**, 1134–1139 (2019).
51. G. Novikova *et al.*, bioRxiv 694281 [Preprint]. 12 August 2019; https://doi.org/10.1101/694281.
52. E. Gjoneska *et al.*, *Nature* **518**, 365–369 (2015).
53. J. Park *et al.*, *Nat. Neurosci.* **21**, 941–951 (2018).
54. S. Wray, *Semin. Cell Dev. Biol.* 10.1016/j.semcdb.2020.05.012 (2020).
55. I. Espuny-Camacho *et al.*, *Neuron* **93**, 1066–1081.e8 (2017).
56. J. Hasselmann *et al.*, *Neuron* **103**, 1016–1033.e10 (2019).
57. R. Mancuso *et al.*, *Nat. Neurosci.* **22**, 2111–2116 (2019).

ACKNOWLEDGMENTS

Funding: We thank the Dementia Research Institute [UKDRI supported by the Medical Research Council (MRC), Alzheimer's Research UK, and Alzheimer's Society], MRC Centre for Neuropsychiatric Genetics and Genomics (MR/L010305/1), Welsh Government, Joint Programming for Neurodegeneration (JPND), VIB and KU Leuven (Methusalem grant), the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant no. ERC-834682 CELLPHASE_AD), the "Fonds voor Wetenschappelijk Onderzoek," the "Geneeskundige Stichting Koningin Elisabeth," Opening the Future campaign of the Leuven Universitair Fonds, VLAIO ICON Personalized Medicine grant (grant no. HBC.2019.2523 "PRISMA"), the Belgian Alzheimer Research Foundation, and the Alzheimer's Association USA. B.D.S. is a holder of the Bax-Vanluffelen Chair for Alzheimer's disease. **Competing interests:** B.D.S. is an ad hoc consultant for various companies but has no direct financial interest in the current manuscript.

10.1126/science.abb8575

REVIEW

Microglia modulate neurodegeneration in Alzheimer's and Parkinson's diseases

Tim Bartels, Sebastiaan De Schepper, Soyon Hong*

Dementia is a rapidly rising global health crisis that silently disables families and ends lives and livelihoods around the world. To date, however, no early biomarkers or effective therapies exist. It is now clear that brain microglia are more than mere bystanders or amyloid phagocytes; they can act as governors of neuronal function and homeostasis in the adult brain. Here, we highlight the fundamental role of microglia as tissue-resident macrophages in neuronal health. Then, we suggest how chronic impairment in microglia-neuron cross-talk may secure the permanence of the failure of synaptic and neuronal function and health in Alzheimer's and Parkinson's diseases. Understanding how to assess and modulate microglia-neuron interactions critical for brain health will be key to developing effective therapies for dementia.

It is becoming increasingly clear that amyloids are not necessarily the smoking gun of neuronal dysfunction and cognitive decline in neurodegenerative diseases. This is displayed in centenarians who have both apparently good cognitive health and brains populated with amyloids (*1*). Epidemiological data also point to the concept of cognitive reserve, where certain individuals appear more resilient to pathological changes in their brains (*2*). Hence, a challenge in neurodegeneration is to understand how certain aging brains successfully maintain proper neuronal function despite chronic amyloid build-up, whereas others do not. Several genetic studies have suggested that microglia, the primary tissue-resident macrophages of the brain, may be key in determining this success (*3*, *4*). Emerging data in developing, adult, and diseased brains collectively suggest that microglia are critical to neuronal homeostasis and health. These observations raise the question of whether, and which, microglia-neuron interactions may be impaired in Alzheimer's disease (AD) and Parkinson's disease (PD) to confer neurodegeneration. Insight into this question will enable the development of methods to assess and modulate microglia-neuron interactions in the aging brain and allow for a desperately needed expansion of focus from clearing amyloids alone to monitoring neuronal health in biomarker and target engagement efforts.

Here, we propose several pathways by which microglia may contribute to neuronal dysfunction in AD and PD. In AD, we focus specifically on complement-mediated synaptic loss and suggest that there are lipid-centric mechanisms in microglia-neuron cross-talk at the synapse. In PD, we discuss the potential roles of tissue-resident macrophages in the brain and gut in modulating amyloid spreading and toxicity, including lysosomal degradation pathways.

UK Dementia Research Institute, Institute of Neurology, University College London, London WC1E 6BT, UK.
***Corresponding author. Email: soyon.hong@ucl.ac.uk**

Microglia are central to neuronal function and health

Genomic and proteomic tools indicate that microglia, akin to other tissue-resident macrophages, are functionally diverse depending on context—i.e., brain region, age, health, and metabolic needs. [Region-specific microglial heterogeneity and their interdependence on neuronal microenvironment have recently been reviewed (*5*).] Microglia constantly survey their local milieu for signals of danger and injury, including pathogens, disease stimuli, and apoptotic neurons. In addition to their immune functions, microglia crucially support brain development—for example, by sculpting neuronal synapses in the developing brain. In the adult brain, microglia perform multiple functions, including monitoring changes in neuronal activity, modulating learning and memory, and acting as local phagocytes and damage sensors in the brain parenchyma (*6*–*12*).

Many of these microglia-neuron interactions are mediated by cell-cell signaling pathways, including purinergic signaling, cytokines, neurotransmitters, and neuropeptides (*5*). These functions often require high energy expenditure and mitochondrial metabolism, for which microglia display metabolic flexibility in acute hypoglycemic states (*9*). An intriguing question is whether chronic mitochondrial dysfunction observed in AD and PD (*13*, *14*) impairs microglia's ability to be metabolically flexible and thus properly monitor and govern neuronal function and health. Notably, in models of neuronal mitochondrial defect and neurodegeneration (*15*), glia accumulate lipid droplets, which can modulate macrophage function. Alternatively, lipid droplets accumulate with aging in microglia (*16*), which raises the question of whether improper lipid metabolism in aged microglia underlies susceptibility to neurodegeneration in AD and PD. In support of this hypothesis, various AD and PD risk factors, including triggering receptor expressed on myeloid cells 2 (TREM2), apolipoprotein E (ApoE), GBA1, and stearoyl-CoA-desaturase (SCD), have

been found to modulate lipid metabolism, lysosomal pathways, and microglial metabolic fitness (*3*, *17*–*23*).

How microglia may mediate synaptic loss in AD

In AD, synaptic loss and dysfunction are region specific, early, and strongly correlated with cognitive impairment (*24*). Prefibrillar oligomeric β-amyloid (Aβ) and/or tau accumulate on synapses and induce pathological synaptic dysfunction and loss (*25*–*30*). More than half of the identified genetic risk factors in AD are expressed by myeloid cells (*31*). Together, these data highlight the need to understand how mutations in risk genes and alleles impair the cross-talk between microglia, the major myeloid cell population in the brain, and neurons at the synapse.

Multiple studies in animal models of AD have suggested that there is dysregulation of neuroimmune signaling pathways on synapses involving classical complement cascade, TREM2, phosphatidylserine (PtdSer), and ApoE (Fig. 1). These studies raise the intriguing question of whether the accumulation of local pathological proteins on synapses dysregulates neuron-glia interactions that are critical for synaptic health. For example, pathological Aβ or tau accumulated on synapses up-regulate C1q in surrounding microglia and promote complement activation on synapses and subsequent microglial engulfment (*26*, *29*, *32*, *33*). Blocking the activation of the classical complement cascade in AD mouse models with genetic or antibody-based means has been shown to protect synapses from loss and dysfunction and memory loss (*26*, *29*, *32*–*34*), which suggests that the microglia-synapse pruning pathway may be a potential therapeutic target. What remains unclear is whether this pruning mechanism—at least in the beginning—is a beneficial process that then becomes dysregulated in a chronic manner and impairs the very neurons it was trying to save. Microglial engulfment of synapses likely involves a fine balance of "eat me" and "don't-eat me" signals (*35*). Because many of the microglial functions including synaptic pruning appear to be activity-dependent (*7*, *8*, *12*, *36*), it will be important to determine whether neuronal hyperactivity observed in early-AD mouse models (*30*, *37*, *38*) instructs microglia to aberrantly engulf synapses (*36*). Insight into the pathways that regulate pruning, as well as the specific signals that guide microglia to engulf synapses, will be crucial for identifying potential therapeutic targets against cognitive decline and for developing biomarker candidates to quantify microglial dysfunction in relation to synaptic loss.

Another biologically and therapeutically important question is whether particular synapses are targeted for elimination by microglia. Proteomic studies in synaptosomes from human and mouse AD brains have highlighted synaptic mitochondrial dysfunction (*14*, *39*). However, whether complement factors, including C1q and C3b, target specific—i.e., dysfunctional and/or damaged—synapses is not known. Lipid signaling in neuron-glia interplay may be a pivotal determinant. For example, TREM2, a key damage sensor on microglia (*40*), has been shown to mediate synaptic refinement in the developing mouse hippocampus (*41*). A proposed ligand for TREM2 is exposed PtdSer on the outer leaflet of neuronal membranes (*19*). Thus, exposed PtdSer on synapses may be an eat-me signal for microglial TREM2 (*42*, *43*). Furthermore, recent work in a model of tauopathy has suggested a potential link between TREM2 and microglia-mediated synaptic elimination in AD (*44*). The AD risk variant of TREM2 was associated with less synaptic localization of C1q and fewer engulfed synaptic elements by microglia compared with the common variant of TREM2. Together, these data suggest a potential role for microglial TREM2 in sensing damaged synaptic membranes in AD, perhaps through PtdSer signaling. Gangliosides, a family of sialic acid–containing lipids enriched in the brain, have also been postulated to be crucial for Aβ-induced synaptic dysfunction in mice (*45*). GM1 ganglioside–bound Aβ is enriched on membranes in early-AD brains (*46*). Moreover, anti–GM1 ganglioside antibodies have been shown to fix complement on neuronal membranes, and the same antibody targeting C1q that was used in AD models (*26*, *29*) has been shown to ameliorate antiganglioside antibody-mediated neuronal injury in a mouse model of acute motor axonal neuropathy (*47*). These studies raise the question of whether brain gangliosides contribute to synaptic loss in AD and complement-mediated synaptic engulfment by microglia.

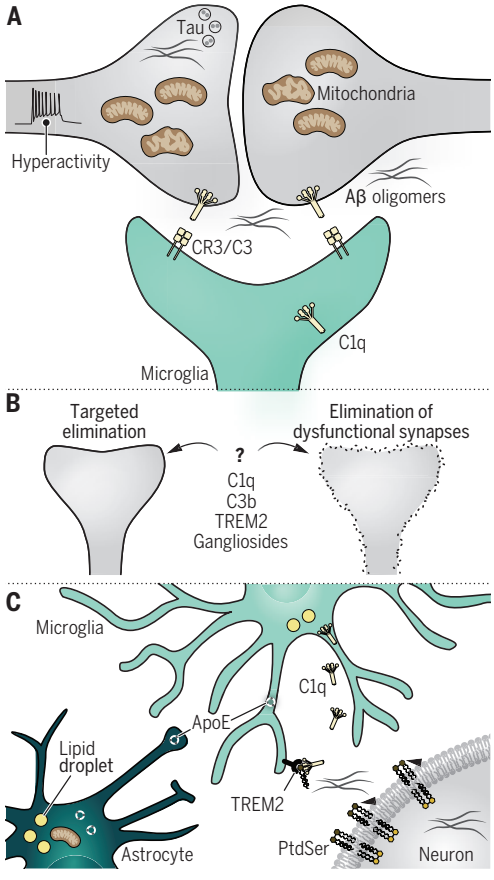


Fig. 1. Complement-mediated synapse loss by microglia in AD. Potential mechanisms leading to complement-mediated synapse elimination by microglia. (**A**) Whether microglia target specific synapses is not known. Neuronal hyperactivity and/or mitochondrial dysfunction observed in AD mouse models and patients may lead to up-regulation of complement factors (C1q and CR3) in microglia to target the dysfunctional synapses. (**B**) In AD mouse models, pathological Aβ or tau accumulated on gangliosides may up-regulate complement signaling pathways through membrane damage sensors like TREM2, which results in synaptic elimination by microglia. (**C**) Recognition of exposed PtdSer on synapses by myeloid TREM2 may lead to synaptic engulfment by microglia. Alternatively, lipid transporter ApoE potentially ameliorates hyperactivity- or membrane damage–induced lipid toxicity by transporting excess lipids to lipid droplets in astrocytes and microglia.

An additional neuroimmune and lipid-related pathway to consider is ApoE. Previous research has suggested a possible link between astrocytic ApoE and microglial synaptic pruning: The ApoE allele-dependent rate of synaptic engulfment by astrocytes appears to be important for normal synapse plasticity (*48*). This rate appears to slow down during aging, thus potentially increasing vulnerability of synapses to complement-mediated pruning by microglia. Notably, ApoE ε4 has been associated with enhanced synaptic localization of pathological Aβ in human AD brains (*25*). Furthermore, ApoE has recently been shown to bind C1q and regulate the activation of the classical complement cascade (*49*). Together, these data suggest a role for ApoE at the synapse in astrocyte-neuron-microglia cross-talk, which is of great interest, especially in light of cell type–specific dysregulation of ApoE in AD and critically linked cholesterol and other lipid metabolic pathways.

Studies involving TREM2 and ApoE, two of the major risk factors in late-onset AD, suggest that lipid metabolism in microglia may be a determinant of how well the brain's immune system can respond to the chronic buildup of amyloids. For example, TREM2-deficient microglia fail to properly metabolize lipids in a chronic demyelination paradigm (*21*). Furthermore, TREM2 appears to be a key regulator of ApoE, a major lipid transporter (*18*). ApoE has been shown to transport excess lipids from hyperactive neurons to lipid droplets in astrocytes where they are metabolized, which suggests a role for ApoE

ILLUSTRATION: MELISSA THOMAS BAUM/*SCIENCE*

in ameliorating neuronal hyperactivity–induced lipid toxicity (*20*). It will be interesting to elucidate the relationship between amyloid-related neuronal hyperactivity and lipid metabolism in astrocytes and microglia and to explore how this relationship falters in the aged brain or in brains with mutations in AD risk genes such as TREM2.

Glial cells as modulators of αSyn toxicity in PD

PD pathology is often accompanied by the distinct accumulation of the neuronal protein alpha-synuclein (αSyn) in astrocytes and microglia (*50*), which has also been recently described as a prominent feature in PD mouse models (*51*). Furthermore, manipulating microglia-astrocyte cross-talk alleviates PD-like pathology in αSyn-aggregation models (*52*). These studies suggest a direct role for glia in uptake and mediating the neurotoxicity of αSyn. Notably, neuronophagia—microglial phagocytosis of neurons—is evidenced in PD by the accumulation of neuromelanin within microglia (*50*). This could point to a synaptic engulfment mechanism analogous to that observed in AD given the fact that synapses in PD tissue are enriched for pathological αSyn aggregates (*53*). However, whether complement and microglia mediate synaptic loss in PD is not known.

From a genetics perspective, the link between PD and microglia is, at first glance, not as apparent as that between AD and microglia. Familial synucleinopathies can be tied to the expression levels of total neuronal αSyn (*54*). However, in sporadic PD, neurodegeneration strongly correlates with certain bioactive forms of αSyn rather than with total levels of αSyn (*55*). Furthermore, three synucleinopathies—PD, dementia with Lewy bodies (DLB), and multiple system atrophy (MSA)—are all characterized by amyloid αSyn burden, but they notably show distinct brain region–specific patterns of amyloid accumulation and neuronal dysfunction (*56*). This remarkable region-specific pattern of αSyn spreading is thought to be induced by a prion-like spread of specific extracellular αSyn aggregates or disease strains, analogous to prion disease (*57*). These findings collectively raise the need to understand what governs the brain region–specific distribution and local abundance of these disease strains. Recent genetic studies in PD have suggested the enrichment of genetic risk factors in sphingolipid metabolism (*58*). The risk genes *GBA1*, *SMPD1*, *GALC*, *ASAH1*, *CTSD*, *SPTLC1*, and *SLC17A5* point to the dysfunctional lysosomal degradation of aggregates as key determinants in disease manifestation. *GBA1*, in particular, has received attention as one of the biggest risk factors for PD (*17*), notably for its potential role in creating toxic variants of αSyn aggregates through defective lysosomal function (*59*). It is important to note that past studies have investigated these genes in a neuronal context, but recent mouse brain single-cell atlases indicate that most of these genes, including *Gba1*, are expressed by microglia rather than neurons (*60*).

Taken together with the prion-like spread of αSyn aggregates, an interesting question is whether—and how—glia are involved in blocking or promoting the transmission of extracellular αSyn aggregates throughout the different brain regions, thereby contributing to region-specific vulnerability in synucleinopathies. In support of this, the only known uptake receptor for extracellular αSyn aggregates, LAG3 (*61*), is mainly expressed by microglia (*60*). Furthermore, in a recent synucleinopathy model, disruption in microglial clearance of extracellular αSyn through autophagy led to dopaminergic neuron degeneration (*62*), whereas in another study, oligodendrocytes were shown to selectively enhance the toxicity of exogenous αSyn aggregates after uptake (*57*). These studies demonstrate that glial uptake and processing is critical in modulating the activity of αSyn. Thus, although glia can act in a physiological context as the waste disposal system of expelled misfolded aggregates by neurons, this is potentially a double-edged sword in disease: The uptake and processing of nontoxic αSyn by glia could actually be the process that generates the disease-specific toxic strains through autophagy and defective lysosomal degradation (Fig. 2). Pathological modification of extracellular αSyn by microglia mediated by imbalances in sphingolipid metabolism could be a key determinant for chronic αSyn dysfunction leading to PD, DLB, or MSA.

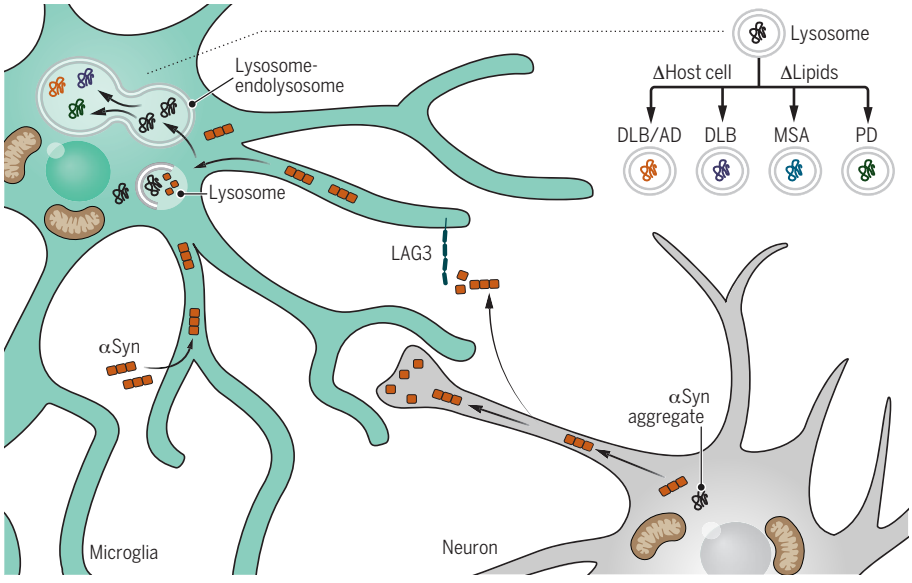


Fig. 2. Glial cells as modulators of αSyn toxicity in PD. Schematic representation of glia in the central nervous system contributing to the spreading of toxic αSyn. Extracellular αSyn is internalized by microglia, potentially through LAG3 receptor–mediated uptake, and processed via endolysosomal machinery. Defective autophagy and impairment in lysosomal degradation could potentially modulate the internalized αSyn aggregates and expel them after failed degradation. These modified disease strains then may contribute to differential region-specific pathology observed in DLB, MSA, and PD.

Looking beyond the brain in PD: Macrophage-neuron signaling in the gut

Emerging preclinical and genetic data suggest that the enteric nervous system (ENS)—the so-called little brain of the gut—may be implicated in PD pathology. αSyn aggregation has been observed in the ENS, and it is believed to spread from here to the brain in a cell-to-cell transsynaptic manner (*63*). In support of this theory, truncal vagotomy in mice has been shown to prevent transmission of pathological αSyn into the brain and related motor deficits, which suggests that the vagus nerve is a potential conduit of αSyn (*64*). Notably, gut-injected αSyn not only induced phosphorylation of αSyn in enteric neurons, but it also stimulated the production of CX3CL1 and CSF1, ligands that bind to CX3CR1 and CSF1R on gut macrophages (*65*). A recent study highlighted a specific type of tissue-resident macrophage in the ENS that is—analogous to microglia in the brain—long-lived and important for neuronal survival and function of the gut (*66*) (Fig. 3). These ENS-resident gut macrophages express high levels of transcripts involving vesicular trafficking and endolysosomal pathways including *Gba1* and *Lrrk2*. Mutations in *LRRK2* are a common cause of autosomal dominant PD; however, exactly how LRRK2 contributes to αSyn pathology and PD-like symptoms is unclear (*67*). Notably, macrophages deficient for *Lrrk2* show higher proteolytic activity and contain higher levels of lysozyme (*68*), which suggests that LRRK2 regulates lysosomal function and phagosome maturation. Furthermore, LRRK2 interacts with the

ILLUSTRATION: MELISSA THOMAS BAUM/*SCIENCE*

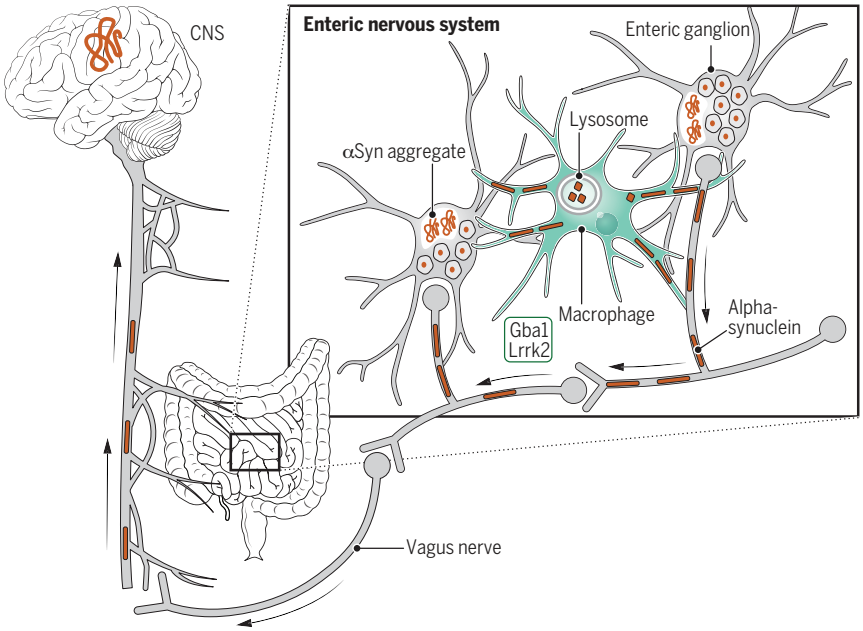


Fig. 3. Macrophage-neuron cross-talk along the gut-brain axis in PD. Schematic of the ENS harboring a specific type of tissue-resident macrophage that is long-lived and provides critical neurotrophic support. These macrophages express transcripts involving vesicular trafficking and endolysosomal pathways, including *Lrrk2* and *Gba1*, which suggests a potential role for uptake, processing, and clearance of αSyn aggregates in PD. Pathological modification of extracellular αSyn by gut macrophages could be a potential modifier of αSyn spreading between enteric ganglia in the ENS or from the ENS to the central nervous system (CNS) via the vagus nerve.

actin-remodeling factor WAVE2 to regulate phagocytosis specifically in macrophages (*69*). Thus, it will be interesting to assess how gut macrophages become affected in PD and regulate the clearance of pathological αSyn along the gut-brain axis.

Outlook

In AD and PD, failure of the tissue-resident macrophages in our brain and gut to sense dysfunctional neurons may lead to the pathological dismantling of neuronal homeostasis and function. Studies from both patient tissues and animal models pinpoint glia as more than cleaners and phagocytes of amyloids: They govern and modulate neuronal health. Therefore, it will be critical to develop methods to monitor glia-neuron cross-talk in living brains and assess which ones are relevant to disease.

Using amyloids, which are pathological hallmarks of neurodegeneration, as biomarkers and diagnostics is undoubtedly important. However, adding specific neuroimmune modulators has the potential to change how we diagnose and treat neurodegeneration, particularly during the early stages of disease when relatively few neurons may be affected but certain neuroimmune markers or relevant microglial cell states may be detectable—for example, in cerebrospinal fluid (*70*) or by brain imaging. Furthermore, macrophage-neuron cross-talk in the ENS may present an early opportunity to intervene in PD. In therapy, early screening for risk factors and the preventive application of drugs that stabilize amyloidogenic proteins—e.g., gamma-secretase modulators, aggregation inhibitors, antisense oligonucleotides, or antibodies—with modulators of neuroglial pathways (potentially through targeted modulation of lipid metabolism and enhancing the autophagosome-lysosome system) could be used.

Finally, whether neuroinflammation is beneficial or detrimental depends on the context. For example, classical complement cascade helps to reduce amyloids but also mediates synapse loss (*26*, *34*), and enhancing TREM2 activity may be beneficial in amyloid-burdened brains but not in tangle-bearing ones (*44*). Therefore, in chronic, multifaceted, and multifactorial diseases such as AD and PD, we need to carefully consider the local milieu when assessing function and impact. It will be important to use distinct strategies at various disease stages to target the appropriate biological processes for effective treatment.

ILLUSTRATION: MELISSA THOMAS BAUM/*SCIENCE*

REFERENCES AND NOTES

1. A. B. Ganz *et al.*, *Acta Neuropathol. Commun.* **6**, 64 (2018).
2. Y. Stern, *Lancet Neurol.* **11**, 1006–1012 (2012).
3. B. W. Kunkle *et al.*, *Nat. Genet.* **51**, 414–430 (2019).
4. A. Nott *et al.*, *Science* **366**, 1134–1139 (2019).
5. S. De Schepper, G. Crowley, S. Hong, *Dev. Neurobiol.* 10.1002/dneu.22777 (2020).
6. C. N. Parkhurst *et al.*, *Cell* **155**, 1596–1609 (2013).
7. R. D. Stowell *et al.*, *Nat. Neurosci.* **22**, 1782–1792 (2019).
8. Y. U. Liu *et al.*, *Nat. Neurosci.* **22**, 1771–1781 (2019).
9. L.-P. Bernier *et al.*, *Nat. Commun.* **11**, 1559 (2020).
10. C. Wang *et al.*, *Science* **367**, 688–694 (2020).
11. P. T. Nguyen *et al.*, *Cell* **182**, 388–403.e15 (2020).
12. C. Cserép *et al.*, *Science* **367**, 528–537 (2020).
13. J. N. Guzman *et al.*, *Nature* **468**, 696–700 (2010).
14. R. Hesse *et al.*, *Acta Neuropathol. Commun.* **7**, 214–218 (2019).
15. L. Liu *et al.*, *Cell* **160**, 177–190 (2015).
16. J. Marschallinger *et al.*, *Nat. Neurosci.* **23**, 194–208 (2020).
17. E. Sidransky *et al.*, *N. Engl. J. Med.* **361**, 1651–1661 (2009).
18. H. Keren-Shaul *et al.*, *Cell* **169**, 1276–1290.e17 (2017).
19. T. K. Ulland *et al.*, *Cell* **170**, 649–663.e13 (2017).
20. M. S. Ioannou *et al.*, *Cell* **177**, 1522–1535.e14 (2019).
21. A. A. Nugent *et al.*, *Neuron* **105**, 837–854.e9 (2020).
22. S. Fanning, D. Selkoe, U. Dettmer, *NPJ Parkinsons Dis.* **6**, 3–9 (2020).
23. D. Agarwal *et al.*, *Nat. Commun.* **11**, 4183 (2020).
24. S. W. Scheff, D. A. Price, F. A. Schmitt, E. J. Mufson, *Neurobiol. Aging* **27**, 1372–1384 (2006).
25. R. M. Koffie *et al.*, *Brain* **135**, 2155–2168 (2012).
26. S. Hong *et al.*, *Science* **352**, 712–716 (2016).
27. L. Zhou *et al.*, *Nat. Commun.* **8**, 15295 (2017).
28. J. McInnes *et al.*, *Neuron* **97**, 823–835.e8 (2018).
29. B. Dejanovic *et al.*, *Neuron* **100**, 1322–1336.e7 (2018).
30. B. Zott *et al.*, *Science* **365**, 559–565 (2019).
31. A. G. Efthymiou, A. M. Goate, *Mol. Neurodegen.* **12**, 43 (2017).
32. A. Litvinchuk *et al.*, *Neuron* **100**, 1337–1353.e5 (2018).
33. T. Wu *et al.*, *Cell Rep.* **28**, 2111–2123.e6 (2019).
34. Q. Shi *et al.*, *Sci. Transl. Med.* **9**, eaaf6295 (2017).
35. G. Lemke, *Nat. Rev. Immunol.* **19**, 539–549 (2019).
36. D. P. Schafer *et al.*, *Neuron* **74**, 691–705 (2012).
37. L. Verret *et al.*, *Cell* **149**, 708–721 (2012).
38. M. A. Busche *et al.*, *Nat. Neurosci.* **22**, 57–64 (2019).
39. B. A. Györffy *et al.*, *Cell. Mol. Life Sci.* **298**, 789 (2020).
40. Y. Wang *et al.*, *Cell* **160**, 1061–1071 (2015).
41. F. Filipello *et al.*, *Immunity* **48**, 979–991.e8 (2018).
42. B. A. Györffy *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **115**, 6303–6308 (2018).
43. N. Scott-Hewitt *et al.*, *EMBO J.* **39**, e105380 (2020).
44. M. Gratuze *et al.*, *J. Clin. Invest.* **130**, 4954–4968 (2020).
45. S. Hong *et al.*, *Neuron* **82**, 308–319 (2014).
46. K. Yanagisawa, A. Odaka, N. Suzuki, Y. Ihara, *Nat. Med.* **1**, 1062–1066 (1995).
47. R. McGonigal *et al.*, *Acta Neuropathol. Commun.* **4**, 23 (2016).
48. W.-S. Chung *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **113**, 10186–10191 (2016).
49. C. Yin *et al.*, *Nat. Med.* **25**, 496–506 (2019).
50. D. W. Dickson, *Cold Spring Harb. Perspect. Med.* **2**, a009258 (2012).
51. G. Tanriöver *et al.*, *Acta Neuropathol. Commun.* **8**, 133 (2020).
52. S. P. Yun *et al.*, *Nat. Med.* **24**, 931–938 (2018).
53. M. Colom-Cadena *et al.*, *Brain* **140**, 3204–3214 (2017).
54. A. Singleton, J. Hardy, *Hum. Mol. Genet.* **28**, R215–R218 (2019).
55. J. B. Sanderson *et al.*, *Brain Commun.* **2**, fcaa010 (2020).
56. J. Alegre-Abarrategui *et al.*, *Acta Neuropathol.* **138**, 681–704 (2019).
57. C. Peng *et al.*, *Nature* **557**, 558–563 (2018).
58. L. A. Robak *et al.*, *Brain* **140**, 3191–3203 (2017).
59. J. R. Mazzulli *et al.*, *Cell* **146**, 37–52 (2011).
60. A. Saunders *et al.*, *Cell* **174**, 1015–1030.e16 (2018).
61. X. Mao *et al.*, *Science* **353**, aah3374 (2016).
62. I. Choi *et al.*, *Nat. Commun.* **11**, 1386 (2020).
63. L. Klingelhoefer, H. Reichmann, *Nat. Rev. Neurol.* **11**, 625–636 (2015).
64. S. Kim *et al.*, *Neuron* **103**, 627–641.e7 (2019).
65. C. Challis *et al.*, *Nat. Neurosci.* **23**, 327–336 (2020).
66. S. De Schepper *et al.*, *Cell* **175**, 400–415.e13 (2018).
67. D. R. Alessi, E. Sammler, *Science* **360**, 36–37 (2018).
68. A. Härtlova *et al.*, *EMBO J.* **37**, e98694 (2018).
69. K. S. Kim *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **115**, E5164–E5173 (2018).
70. M. Ewers *et al.*, *Sci. Transl. Med.* **11**, eaav6221 (2019).

ACKNOWLEDGMENTS

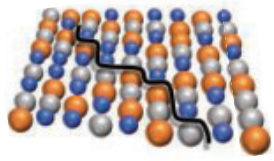
We thank T. Childs for his critical comments. Figures were originally made with Biorender. **Funding:** This work was supported by the following: grants from the UK Dementia Research Institute (DRI), which receives its funding from DRI Ltd., the UK Medical Research Council and Alzheimer's Society, and Alzheimer's Research UK (to S.H. and T.B.); the National Institute of Neurological Disorders and Stroke NIH grants (U54-NS110435, R01-NS109209, and R21-NS107950 to T.B.); the Michael J. Fox Foundation (Ken Griffin Imaging Award to T.B.); a Parkinson's Disease Foundation Stanley Fahn Award (PF-JFA-1884 to T.B.); the Eisai Pharmaceutical postdoctoral programme to T.B.; and the Chan Zuckerberg Collaborative Pairs Initiative (to T.B. and S.H.). **Competing interests:** The following patents have been granted or applied for: PCT/2015/010288, US14/988387, and EP14822330 (S.H.). All authors declare no other competing interests related to this project.

10.1126/science.abb8587

RESEARCH

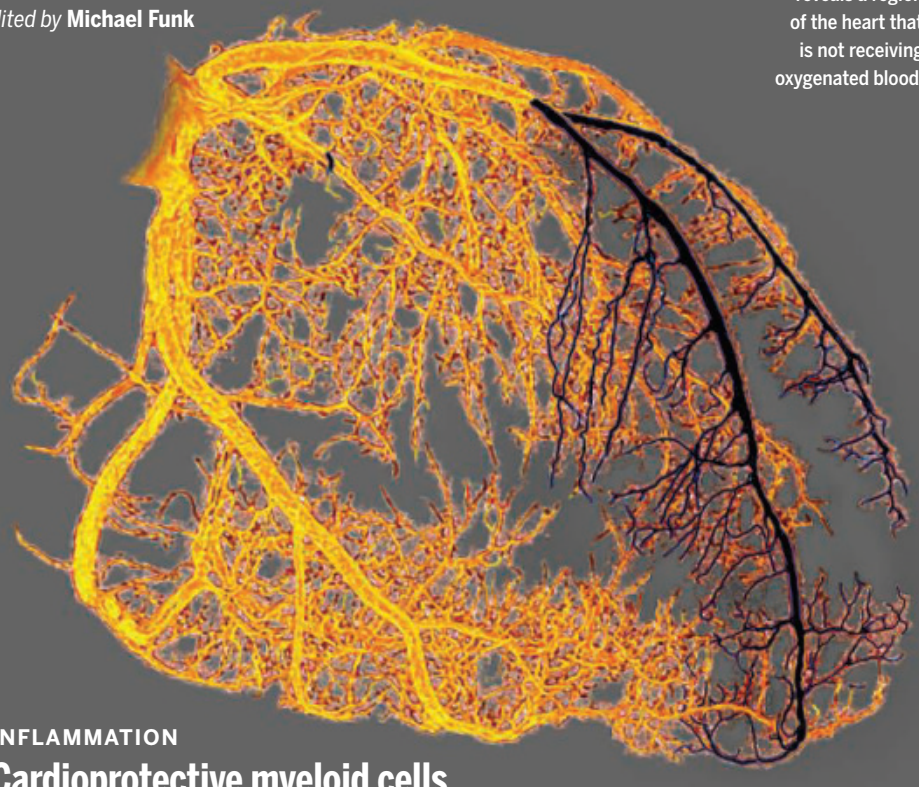
Enhanced ductility in a multicomponent alloy

F. Wang et al., *p. 95*



IN *SCIENCE* JOURNALS

Edited by **Michael Funk**



A coronary angiogram reveals a region of the heart that is not receiving oxygenated blood.

INFLAMMATION

Cardioprotective myeloid cells

Acute ischemic injury to the heart precipitates a strong inflammatory response, including influx of myeloid cells expressing type I interferon–stimulated genes (ISGs). Calcagno *et al.* used single-cell RNA sequencing to probe the origin, evolution, and heterogeneity of this response in the first 4 days after myocardial infarction using human and mouse myeloid cells. Induction of ISGs in myeloid cells was initially observed in bone marrow and blood. Postinfarct cardiac tissue in mice contained myeloid subsets with and without ISG expression and a steady-state macrophage population with anti-inflammatory activity dependent on Nrf2, a cytoprotective transcriptional regulator. On the basis of their findings, the authors developed an ISG score as a potential biomarker to assess how the vigor of type I interferon signaling influences clinical outcomes after a heart attack. —IRW

Sci. Immunol. **5**, eaaz1974 (2020).

CORONAVIRUS

Preexisting immune response to SARS-CoV-2

Robust T cell responses to the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) virus occur in most individuals with coronavirus disease 2019 (COVID-19). Several studies have reported that some people who have not been exposed to SARS-CoV-2 have preexisting reactivity to SARS-CoV-2 sequences. The immunological mechanisms underlying this preexisting reactivity are not clear, but previous exposure to widely circulating common cold coronaviruses might be involved. Mateus *et al*. found that the preexisting reactivity against SARS-CoV-2 comes from memory T cells and that cross-reactive T cells can specifically recognize a SARS-CoV-2 epitope as well as the homologous epitope from a common cold coronavirus. These findings underline the importance of determining the impacts of preexisting immune memory in COVID-19 disease severity. —PNK

Science, this issue p. 89

ORGANIC CHEMISTRY

Paper in, product out

A typical chemist running a known reaction will start by finding the method described in a published paper. Mehr *et al*. report a software platform that uses natural language processing to translate the organic chemistry literature directly into editable code, which in turn can be compiled to drive automated synthesis of the compound in the laboratory. The synthesis procedure is intended to be universally applicable to robotic systems operating in a batch reaction architecture. The full process is demonstrated for synthesis of an analgesic as well as common oxidizing and fluorinating agents. —JSY

Science, this issue p. 101

ELECTROCALORICS

Competitive cooling with capacitors

Current large-scale cooling devices use vapor compression refrigeration. The efficiency of air conditioners has been optimized, but they can be noisy and rely on problematic greenhouse gases. Two groups now present designs for electrocaloric cooling using lead scandium tantalate capacitors that change temperature under an electric field. Y. Wang *et al*.

CREDITS (FROM TOP): WANG *ET AL.*; JAMES CAVALLINI/SCIENCE SOURCE

obtained a very large heat flux using only solid materials and a cooling fan to remove heat from their device. Torello *et al*. used fluids for heat transfer, leading to a very large temperature difference between the hot side and the cold side. The new designs demonstrate the potential for devices that might be competitive with vapor compression–based appliances with further optimization. —BG

Science, this issue p. 129, p. 125

SOLAR CELLS

Relieving unwanted strain

Although the α-phase of formamidinium lead iodide ($FAPbI_3$) has a suitable bandgap for use in solar cells, it must be stabilized with additional cations. These compositions can adversely affect the bandgap and produce lattice strain that creates trap sites for charge carriers. Kim *et al*. found that substituting small, equimolar amounts of cesium and methylenediammonium cations for formamidinium reduced the lattice strain and trap densities. The enhancement in open-circuit voltage led to a certified power conversion efficiency of 24.4%, and encapsulated devices retained 90% of their initial efficiency after 400 hours of maximal power point operating conditions. —PDS

Science, this issue p. 108

SEX DETERMINATION

Two rather than one

For several decades, it has been believed that the mammalian sex-determining gene *Sry* contains a single exon. Miyawaki *et al*. have now identified a cryptic second exon of mouse *Sry*. Loss- and gain-of-function analyses revealed that the two-exon SRY (SRY-T), not the canonical single exon–encoded SRY (SRY-S), is the bona fide testis-determining factor. *Sry* exon2 is composed of retrotransposon-derived sequences. The SRY-S carboxyl terminus contains a degradation sequence (degron), whereas the SRY-T carboxyl terminus encoded in the *Sry* exon2 is degron free, thereby conferring protein stability on SRY-T. —BAP

Science, this issue p. 121

DEVELOPMENTAL BIOLOGY

Convergence of paradigms yields patterns

In embryo development, spatial patterns of distinct cell types arise reproducibly. In the zebrafish spinal cord, neural progenitors form stereotypic stripe patterns despite the noisy instructive signals and large-scale cellular rearrangement required during morphogenesis. Tsai *et al*. show that a cell type–specific adhesion code, regulated by a Shh morphogen gradient composed of three adhesion molecules, provides adhesion specificity for three neural progenitor types and mediates patterning robustness in the zebrafish spinal cord. Although insufficient on their own, the integration of the morphogen gradient and differential adhesion mechanisms enables robust pattern formation during tissue morphogenesis. —BAP

Science, this issue p. 113

CORONAVIRUS

How SARS-CoV-2 hides its mRNA

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) uses a methyltransferase to cap its messenger RNAs to prevent them from being recognized by the host immune system and ensure their translation in host cells. Rosas-Lemus *et al*. solved crystal structures for the SARS-CoV-2 methyltransferase in complex with its methyl donor and cap structure substrates, a reaction product, and an inhibitor. These structures offer potential strategies for developing therapeutics that disrupt the formation of the active methyltransferase complex or block its catalytic activity. —AMV

Sci. Signal. **13**, eabe1202 (2020).

IN OTHER JOURNALS

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

IMMUNOGENOMICS

Leprosy-associated genetics

Leprosy, caused by *Mycobacterium leprae*, occurs worldwide, but some individuals are more susceptible to the disease than others. Genes associated with leprosy have been mapped to the human leukocyte antigen (HLA) proteins, which are responsible for the regulation of the immune system. However, the HLA region is quite variable, and mapping specific disease-associated mutations is difficult. Dallmann-Sauer *et al.* have succeeded in sequencing, mapping, and genotyping 11 HLA genes in a case-control design of 1155 Vietnamese individuals. Four leprosy-associated amino acid variants were identified within these genes, which comprised two pairs of linked genes, with one set conferring susceptibility and one being protective. This study exemplifies how examination of underrepresented populations can help to identify genetic factors that may reveal vulnerabilities to disease around the world. —LMZ

PLOS Pathog. **16**, e1008818 (2020).

Leprosy is a disfiguring mycobacterial infection with a distinctive genetic signature in some human populations for either susceptibility or protection.

CATALYSIS

A sintering-avoiding support

Heterogeneous catalysts containing dispersed metal nanoparticles (NPs) on supports such as metal oxides or carbon can undergo sintering at high temperatures, a process in which the diffusion of metal atoms results in larger NPs with smaller surface areas. Chen *et al*. reduced a melt of $NaBH_4$, $NaNH_2$, and noble metal salts (Au, Pd, and Pt) to form metal

PHOTO: MANAN VATSYAYANA/GETTY IMAGES

ENVENOMATION

Recovery from paralysis

Envenoming by snake bite is recorded unsystematically. The World Health Organization reports that more than 5 million people are bitten by snakes annually. They estimate that up to 138,000 die of complications, and many victims are left disabled. Stazi *et al*. developed a mouse model of envenoming by the Papuan taipan (*Oxyuranus scutellatus*). This snake's poison, taipoxin, is a phospholipase A_2 neurotoxin akin to tetanus and botulinum toxins. Taipoxin causes the degeneration of motor axon synaptic transport but does not kill the neurons, and a slow recovery occurs if the victim is given respiratory support. Signaling by the G protein–coupled receptor CXCR4 is implicated in this synaptic pathophysiology. The authors studied the effects of a previously identified agonist of CXCR4, NUCC-390, a nontoxic compound in mice that accelerates regeneration of their synapses, and found that it significantly accelerated recovery from paralysis. NUCC-390 thus has the potential to enhance the recovery of patients hospitalized with neuroparalytic snake envenomation. —CA

PLOS Negl. Trop. Dis. **14**, e0008547 (2020).

The Paupan taipan poisons with a neurotoxin that causes synapse breakdown, resulting in paralysis.

NPs encapsulated by hexagonal boron nitride nanosheets. The supported NPs were released by washing away excess salt and could be loaded onto conventional supports such as alumina. The strong interaction of Pd NPs with these nanosheet supports allowed them to resist sintering up to 950°C and enabled their use for CO oxidation in the presence of water and hydrocarbons. —PDS

ACS Cent. Sci. **6**, 1617 (2020).

MATERIALS SCIENCE

Corrosion detection and protection in one

The corrosion or rusting of metal structures is an important problem that shortens the life span of buildings and vehicles. Coatings can be used to slow the corrosion down, but they can be damaged or worn away and do not help reverse the process. Thongchaivetcharat *et al*. created Pickering emulsions in linseed oil containing a corrosion sensor, a corrosion inhibitor, and a healing agent stabilized by a coating of solid particles. The Pickering emulsion droplets were embedded in a polymer matrix that could be applied to the surface of a metal. Corrosion of the metal caused a color change of the trapped dye molecules and the release of the corrosion inhibitors. The coating also showed self-healing after mechanical damage. —MSL

ACS Appl. Mater. Interfaces **12**, 42129 (2020).

PHYSIOLOGY

Characterizing the menstrual cycle

The endometrium, which lines the uterus, is a distinct tissue that is capable of monthly remodeling, shedding, and regeneration. Other than humans, very few species exhibit a similar menstrual cycle. Given the importance of the endometrium to fertility and women's health, Wang *et al*. undertook single-cell transcriptomic characterization of the endometrium across the menstrual cycle from 19 healthy donors. The authors defined four phases across the cycle and characterized the window of implantation and accompanying changes in gene expression in the epithelia. They also discovered the presence of ciliated epithelium and characterized how these cells and six other cell types change over the course of the cycle. These high-resolution data provide important insights into female physiology and have numerous applications in fertility and endometrial biology. —GKA

Nat. Med. 10.1038/s41591-020-1040-z (2020).

PHYSIOLOGY

Exercise circuit for training

Exercise has many clear beneficial effects. Aerobic exercise elicits a coordinated metabolic response across multiple tissues. Brandão *et al*. unraveled the cross-talk that occurs between muscle and adipose tissue and regulation by microRNAs. In mice induced to exercise, the authors found that the energy-sensing protein kinase AMPK became activated in muscle and fat cells. A signal released into the circulation increased the expression of the type III endoribonuclease DICER in fat cells, which promoted the synthesis of microRNAs that control translation of specific messenger RNAs. Enhanced microRNA production inhibited glucose use in fat cells, which then provided substrate to support muscle in exercise training. —LBR

Proc. Natl. Acad. Sci. U.S.A. **117**, 23932 (2020).

WORKFORCE DIVERSITY

Three elements to increase diversity

Women remain underrepresented in STEM career pathways. Mentoring is a promising intervention to increase the representation of women, but mentorship programs vary in complexity, cost, and scalability. Hernandez *et al*. performed a randomized experiment with female undergraduate geoscience majors to identify the critical elements of a successful mentorship program. A combination of inspiration through exposure to female role models in the geosciences, inoculation through training on how to grow a mentor network, and an introduction to a local female geoscientist mentor were all necessary for success. These three elements can be easily adapted and widely implemented across scientific disciplines to support undergraduate women at the start of, and throughout, their STEM careers. —MMc

Commun. Earth Environ. **1**, 7 (2020).

PHOTO: MATTHIJS KUIJPERS/ALAMY STOCK PHOTO

RESEARCH ARTICLE SUMMARY

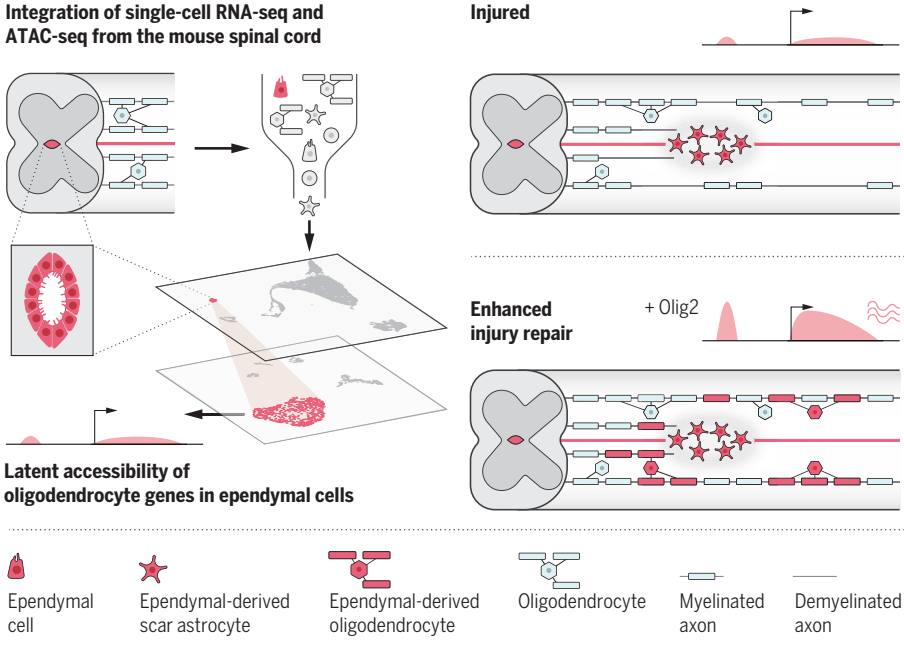
NEURODEVELOPMENT

A latent lineage potential in resident neural stem cells enables spinal cord repair

Enric Llorens-Bobadilla, James M. Chell, Pierre Le Merre, Yicheng Wu, Margherita Zamboni, Joseph Bergenstråhle, Moa Stenudd, Elena Sopova, Joakim Lundeberg, Oleg Shupliakov, Marie Carlén, Jonas Frisén*

INTRODUCTION: The capacity of a tissue to regenerate itself rests on the potential of its resident cells to replace cells lost to injury. Some tissues, such as skin or intestine, do this remarkably well through the activation of tissue-specific stem cells. Injuries to the central nervous system (CNS), in contrast, often lead to permanent functional impairment; some cells lost to injury are never replaced. Neural stem cells have been identified in the adult brain and spinal cord and are activated by injury. However, injury-activated neural stem cells predominantly produce scar-forming astrocytes, and the contribution of neural stem cells to cell replacement is insufficient for regeneration. To design regenerative strategies aimed at recruiting resident neural stem cells for repair, it is essential to know whether greater regenerative potential exists and how to elicit such potential.

RATIONALE: The spinal cord is a great system to study neural stem cell recruitment for repair. The neural stem cell potential of the spinal cord resides in a well-characterized population of ependymal cells. Ependymal cells, normally quiescent, are activated by injury to generate almost exclusively scar-forming astrocytes. Ependymal-derived astrocytes help to preserve tissue integrity, but other cell types, such as myelin-forming oligodendrocytes, are insufficiently replaced. In parallel, neural stem cell transplantation has proven to be beneficial to recovery after spinal cord injury—a benefit that is associated with the increased supply of oligodendrocytes able to remyelinate demyelinated axons. Ependymal cells share a developmental origin with spinal oligodendrocytes, which led us to explore whether a latent potential for expanded oligodendrocyte generation might exist.



Latent potential in neural stem cells. Through the integration of different layers of genomic information in single cells, we found that the genetic program for oligodendrocyte generation is latently accessible in ependymal neural stem cells of the adult spinal cord. After injury, activating the latent potential by forced OLIG2 expression unfolds efficient oligodendrocyte generation, leading to enhanced repair.

RESULTS: We integrated single-cell RNA sequencing (scRNA-seq) and single-cell assay for transposase-accessible chromatin using sequencing (scATAC-seq) to study lineage potential in adult ependymal cells of the mouse spinal cord. We found that the genetic program for oligodendrocyte generation is accessible in ependymal cells. However, this program is latent, as oligodendrocyte genes are not expressed. In particular, we found that a large fraction of binding sites for OLIG2, the transcription factor that initiates developmental oligodendrogenesis, had basal accessibility, despite OLIG2 and its key target genes not being expressed in adult ependymal cells. To study whether this latent accessibility was associated with a greater capacity to produce oligodendrocytes, we genetically engineered a mouse model to express OLIG2 in adult ependymal cells. We found that OLIG2 expression was compatible with ependymal identity during homeostasis. However, after injury, OLIG2 expression led to the increased accessibility of the latent program and subsequent expression of genes specifying oligodendrocyte identity. Unfolding of the latent program was followed by efficient oligodendrocyte production from ependymal cells, but not from astrocytes, after injury. Using scRNA-seq of ependymal-derived cells, we found that new oligodendrocytes followed the developmental program of oligodendrocyte maturation, including a self-amplifying oligodendrocyte progenitor cell–like state. These cells later matured to acquire the identity of resident mature myelinating oligodendrocytes. Further, ependymal oligodendrocyte generation occurred in parallel and not at the expense of astrocyte scarring. Newly recruited ependymal-derived oligodendrocytes migrated to sites of demyelination, where they remyelinated axons over the long term. Finally, using optogenetics, we found that ependymal-derived oligodendrocytes contributed to normalizing axon conduction after injury.

CONCLUSION: Adult neural stem cells have a greater potential for regeneration than is normally manifested. Targeted activation of such potential leads to the recruitment of neural stem cells for the generation of remyelinating oligodendrocytes in numbers comparable to those obtained via cell transplantation. Resident stem cells can thus serve as a reservoir for cellular replacement and may offer an alternative to cell transplantation after CNS injury. ■

The list of author affiliations is available in the full article online.
*Corresponding author. Email: jonas.frisen@ki.se
Cite this article as E. Llorens-Bobadilla *et al.*, *Science* **370**, eabb8795 (2020). DOI: 10.1126/science.abb8795

READ THE FULL ARTICLE AT
https://doi.org/10.1126/science.abb8795

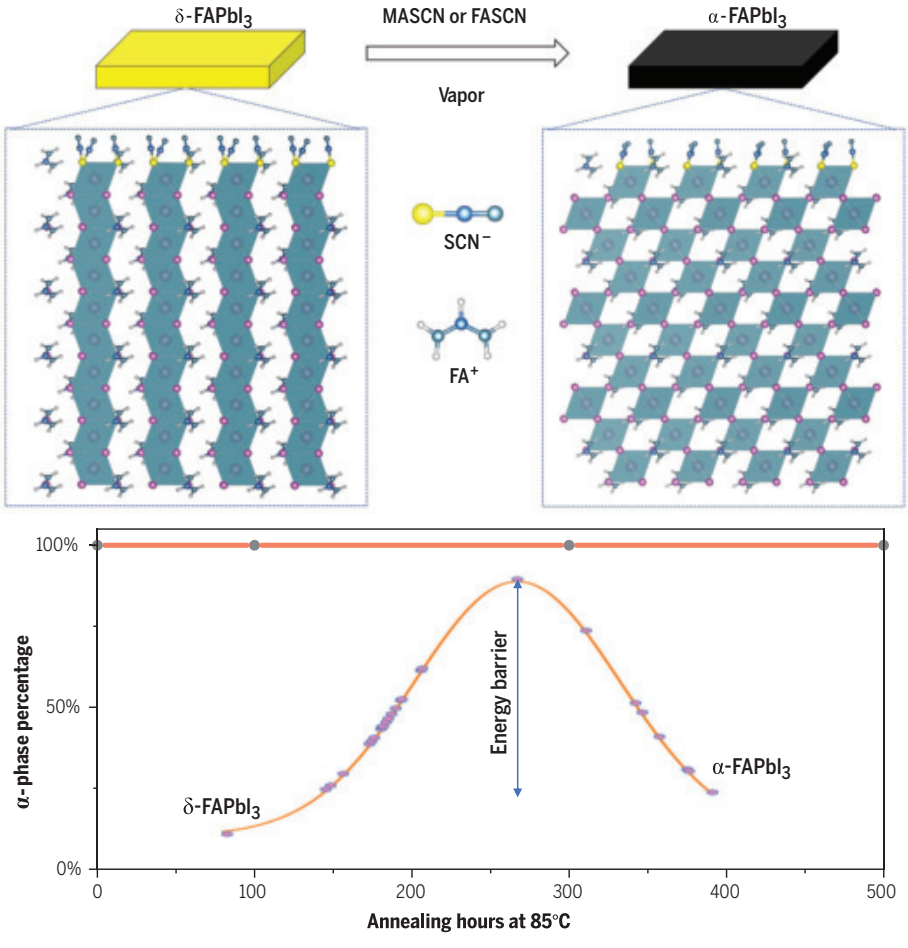
RESEARCH ARTICLE SUMMARY

SOLAR CELLS

Vapor-assisted deposition of highly efficient, stable black-phase $FAPbI_3$ perovskite solar cells

Haizhou Lu, Yuhang Liu, Paramvir Ahlawat, Aditya Mishra, Wolfgang R. Tress, Felix T. Eickemeyer, Yingguo Yang, Fan Fu, Zaiwei Wang, Claudia E. Avalos, Brian I. Carlsen, Anand Agarwalla, Xin Zhang, Xiaoguo Li, Yiqiang Zhan*, Shaik M. Zakeeruddin, Lyndon Emsley, Ursula Rothlisberger, Lirong Zheng*, Anders Hagfeldt*, Michael Grätzel*

INTRODUCTION: Metal halide perovskite solar cells (PSCs) have reached a power-conversion efficiency (PCE) of 25.2%, thus exceeding other thin-film solar cells. $FAPbI_3$ (where FA is formamidinium) has been shown to be an ideal candidate for efficient, stable PSCs. Obtaining highly crystalline, stable, and pure α-phase $FAPbI_3$ films has been of vital importance. However, $FAPbI_3$ undergoes a phase transition from the black α-phase to the photoinactive δ-phase below 150°C. Previous approaches to overcoming this problem include mixing it with MA, Cs or Br ions. Here, we report a deposition method using methylammonium thiocyanate (MASCN) vapor treatment to convert δ-$FAPbI_3$ to the desired pure α-phase below the thermodynamic phase-transition temperature. Molecular dynamics (MD) simulations show that the SCN^- anions promote the formation and stabilization of α-$FAPbI_3$. These vapor-treated $FAPbI_3$ PSCs exhibit outstanding photovoltaic and electroluminescent performance.



Stable and phase pure MASCN vapor-treated $FAPbI_3$ films. Vapor-treated $FAPbI_3$ films were annealed at 85°C for 500 hours in an N_2 environment.

RATIONALE: Although the phase transition from δ- to α-phase $FAPbI_3$ requires a high temperature, the treatment of δ-phase $FAPbI_3$ films with MASCN vapor allows the conversion to occur at temperatures below 150°C. MD simulations show that SCN^- ions preferentially adsorb on the surface of δ-$FAPbI_3$ to replace iodide ions that are bound to Pb^{2+}. This process disintegrates the top layer of face-sharing octahedra and induces the transition to the corner-sharing architecture of α-$FAPbI_3$. Once the corner-sharing α-form is formed on the top surface, this layer templates the progression of the phase transition from δ- to α-$FAPbI_3$ toward the bulk. Once the pure α-$FAPbI_3$ is formed, its back conversion to the δ-phase is prevented by a high energy barrier.

RESULTS: We show a complete conversion from δ- to α-$FAPbI_3$ at 100°C using the MASCN vapor treatment method. This phase transition can also be achieved using FASCN vapor. The vapor-treated $FAPbI_3$ film remained in its pure black phase even after 500 hours of annealing at 85°C, whereas the reference $FAPbI_3$ film formed mainly PbI_2 during the heat exposure. X-ray diffraction data showed an improved crystallinity and preferred orientation of the $FAPbI_3$ films after vapor treatment. One- and two-dimensional NMR experiments were used to probe changes in symmetry and quantify the incorporation of MA into the perovskite framework. Time-of-flight secondary ion mass spectrometry measurements confirmed that the MASCN content was mostly located near the surface region of the $FAPbI_3$ films. We used these low-defect-density α-$FAPbI_3$ films to make PSCs with >23% PCE, long-term operational stability, low (330 mV) open-circuit voltage (V_{oc}) loss, and low (0.75 V) turn-on voltage of electroluminescence.

CONCLUSION: SCN^- anions play a key role in promoting the formation and stabilization of α-$FAPbI_3$. Vapor-treated $FAPbI_3$ films showed long-term thermal stability. MD simulations showed that the pure α-$FAPbI_3$ remained kinetically stable. These findings are important for developing stable and pure black-phase $FAPbI_3$-based PSCs. Our vapor-treated $FAPbI_3$ PSCs showed high efficiency and good long-term stability under maximum power point tracking conditions. Because of its high V_{oc} and high external quantum efficiency electroluminescence yield, pure α-$FAPbI_3$ will be useful for other applications such as light-emitting diodes and photodetectors. ■

The list of author affiliations is available in the full article online.
*Corresponding author. Email: yqzhan@fudan.edu.cn (Y.Z.); lrzheng@fudan.edu.cn (L.Z.); anders.hagfeldt@epfl.ch (A.H.); michael.graetzel@epfl.ch (M.G.)
Cite this article as H. Lu *et al.*, *Science* **370**, eabb8985 (2020). DOI: 10.1126/science.abb8985

READ THE FULL ARTICLE AT
https://doi.org/10.1126/science.abb8985

RESEARCH ARTICLES

MUTATION

Extensive heterogeneity in somatic mutation and selection in the human bladder

Andrew R. J. Lawson[1], Federico Abascal[1], Tim H. H. Coorens[1], Yvette Hooks[1], Laura O'Neill[1], Calli Latimer[1], Keiran Raine[1], Mathijs A. Sanders[1,2], Anne Y. Warren[3], Krishnaa T. A. Mahbubani[4,5], Bethany Bareham[4,5], Timothy M. Butler[1], Luke M. R. Harvey[1], Alex Cagan[1], Andrew Menzies[1], Luiza Moore[1,3], Alexandra J. Colquhoun[6], William Turner[6], Benjamin Thomas[7,8], Vincent Gnanapragasam[9,10], Nicholas Williams[1], Doris M. Rassl[11], Harald Vöhringer[12], Sonia Zumalave[13], Jyoti Nangalia[1], José M. C. Tubío[13,14,15], Moritz Gerstung[12], Kourosh Saeb-Parsy[4,5], Michael R. Stratton[1], Peter J. Campbell[1,16], Thomas J. Mitchell[1,6], Iñigo Martincorena[1]*

The extent of somatic mutation and clonal selection in the human bladder remains unknown. We sequenced 2097 bladder microbiopsies from 20 individuals using targeted (n = 1914 microbiopsies), whole-exome (n = 655), and whole-genome (n = 88) sequencing. We found widespread positive selection in 17 genes. Chromatin remodeling genes were frequently mutated, whereas mutations were absent in several major bladder cancer genes. There was extensive interindividual variation in selection, with different driver genes dominating the clonal landscape across individuals. Mutational signatures were heterogeneous across clones and individuals, which suggests differential exposure to mutagens in the urine. Evidence of APOBEC mutagenesis was found in 22% of the microbiopsies. Sequencing multiple microbiopsies from five patients with bladder cancer enabled comparisons with cancer-free individuals and across histological features. This study reveals a rich landscape of mutational processes and selection in normal urothelium with large heterogeneity across clones and individuals.

Recent technological developments have started to enable the detection of somatic mutations in normal tissues (*1*–*15*). One observation derived from these studies is that as we age, some tissues are colonized by mutant clones carrying driver mutations in cancer genes (*2*, *3*, *6*–*8*, *11*, *15*). These mutations confer a growth advantage driving clonal expansions, some of which are thought to represent the earliest steps toward cancer. However, the extent of this phenomenon remains unclear as driver mutations appear to be rare in other tissues (*4*, *9*, *10*, *12*).

Bladder urothelium is an interesting tissue in this context. It is one of the slowest-dividing epithelia in the human body, being largely quiescent in homeostasis but able to regenerate quickly upon injury (*16*). However, bladder cancers arising from the urothelium have some of the highest mutation burdens of all major cancer types (*17*) and a rich landscape of driver mutations (*18*, *19*). Bladder urothelium is also constantly bathed in urine, which can contain mutagenic and carcinogenic molecules known to increase the risk of bladder cancer, such as aromatic amines from tobacco smoking; aristolochic acid from certain herbal medicines; and compounds present in dyes, solvents, and fumes from occupational and environmental exposures (*20*, *21*).

Somatic mutations in the normal bladder

To characterize the mutational landscape of normal bladder urothelium both within and across individuals, we performed laser microdissection of small strips of urothelium. Microbiopsies had a median length of 855 μm and typically contained a few hundred cells (Fig. 1A). In total, we studied 1647 microbiopsies from 15 deceased transplant organ donors (ranging from 25 to 78 years of age) and 450 microbiopsies from five patients with bladder cancer (49 to 75 years of age; table S1) (*22*). Formalin-free fixation and paraffin embedding were used to ensure high-quality morphology and genome sequencing (*22*).

To search for mutant clones, we performed targeted sequencing of 321 cancer-associated genes for 1914 microbiopsies (median coverage of 89×) (*22*). To study mutation burden and signatures, copy number changes, and selection outside of cancer genes, we performed whole-exome sequencing of 655 microbiopsies (median coverage of 72×) (*22*) and whole-genome resequencing of 88 microbiopsies dominated by large clones (median coverage of 33×; Fig. 1A) (*22*). By sequencing many biopsies per individual, we were able to study the heterogeneity in drivers, burden, and signatures across clones and individuals.

In histologically normal urothelium, we detected a median number of 40 mutations per exome and 1879 mutations per genome, although the numbers varied considerably across microbiopsies (Fig. 1B and fig. S1) (*22*). Variant allele fractions (VAFs) were moderately low (median exome VAF = 0.13), and most mutations were detected in a single microbiopsy with few shared by adjacent microbiopsies (fig. S2), which indicates that mutant clones are typically smaller than the microbiopsy sizes used in this study. Considering the allele fractions and the length of each microbiopsy, we estimate that most mutant clones are smaller than a few hundred micrometers in one-dimensional sections of urothelium (Fig. 1C) (*22*), consistent with estimates derived from mitochondrial markers (*23*). This shows that histologically normal bladder urothelium is a patchwork of small—typically microscopic—mutant clones. Below, we first describe the mutational landscape of the healthy bladder by focusing on data from the 15 transplant organ donors (Figs. 2 and 3), followed by an analysis of microbiopsies from the five patients with bladder cancer (Fig. 4).

Widespread positive selection in normal urothelium

To determine whether positive selection on certain genes drives these clonal expansions, we used the ratio of nonsynonymous to synonymous mutation rates (dN/dS). Mutations driving clonal expansions become overrepresented among mutant clones reaching detectable sizes, which manifests as an excess of nonsynonymous mutations in driver genes (*22*). We used the dNdScv algorithm, an implementation of dN/dS that corrects for trinucleotide mutation rates, sequence composition, and variable rates across genes (*19*). Applying it to the 321 cancer genes sequenced in 1500 microbiopsies of normal urothelium from the transplant organ donors revealed significant positive selection on 12 genes (*22*): *KMT2D* (also known as *MLL2*), *KDM6A* (also known as *UTX*), *ARID1A*, *RBM10*, *EP300*, *STAG2*, *NOTCH2*,

[1]Cancer, Ageing and Somatic Mutation Programme, Wellcome Sanger Institute, Hinxton CB10 1SA, UK. [2]Department of Hematology, Erasmus University Medical Center, Rotterdam 3015 GD, Netherlands. [3]Department of Histopathology, Cambridge University Hospitals NHS Foundation Trust, Cambridge Biomedical Campus, Cambridge CB2 0QQ, UK. [4]Department of Surgery, University of Cambridge, Cambridge CB2 0QQ, UK. [5]NIHR Cambridge Biomedical Research Centre, Cambridge Biomedical Campus, Cambridge CB2 0QQ, UK. [6]Department of Urology, Cambridge University Hospitals NHS Foundation Trust, Cambridge CB2 0QQ, UK. [7]The Royal Melbourne Hospital, Parkville, Victoria 3010, Australia. [8]Department of Surgery, The University of Melbourne, Parkville, Victoria 3010, Australia. [9]Academic Urology Group, Department of Surgery and Oncology, University of Cambridge, Cambridge CB2 0QQ, UK. [10]Cambridge Urology Translational Research and Clinical Trials Office, University of Cambridge CB2 0QQ, UK. [11]Department of Pathology, Royal Papworth Hospital NHS Foundation Trust, Cambridge Biomedical Campus, Cambridge CB2 0AY, UK. [12]European Molecular Biology Laboratory, European Bioinformatics Institute (EMBL-EBI), Hinxton CB10 1SD, UK. [13]Mobile Genomes and Disease, Center for Research in Molecular Medicine and Chronic Diseases (CiMUS), Universidade de Santiago de Compostela, Santiago de Compostela 15706, Spain. [14]Department of Zoology, Genetics and Physical Anthropology, Universidade de Santiago de Compostela, Santiago de Compostela 15706, Spain. [15]The Biomedical Research Centre (CINBIO), University of Vigo, Vigo 36310, Spain. [16]Department of Haematology, University of Cambridge, Cambridge CB2 2XY, UK.
*Corresponding author. Email: im3@sanger.ac.uk

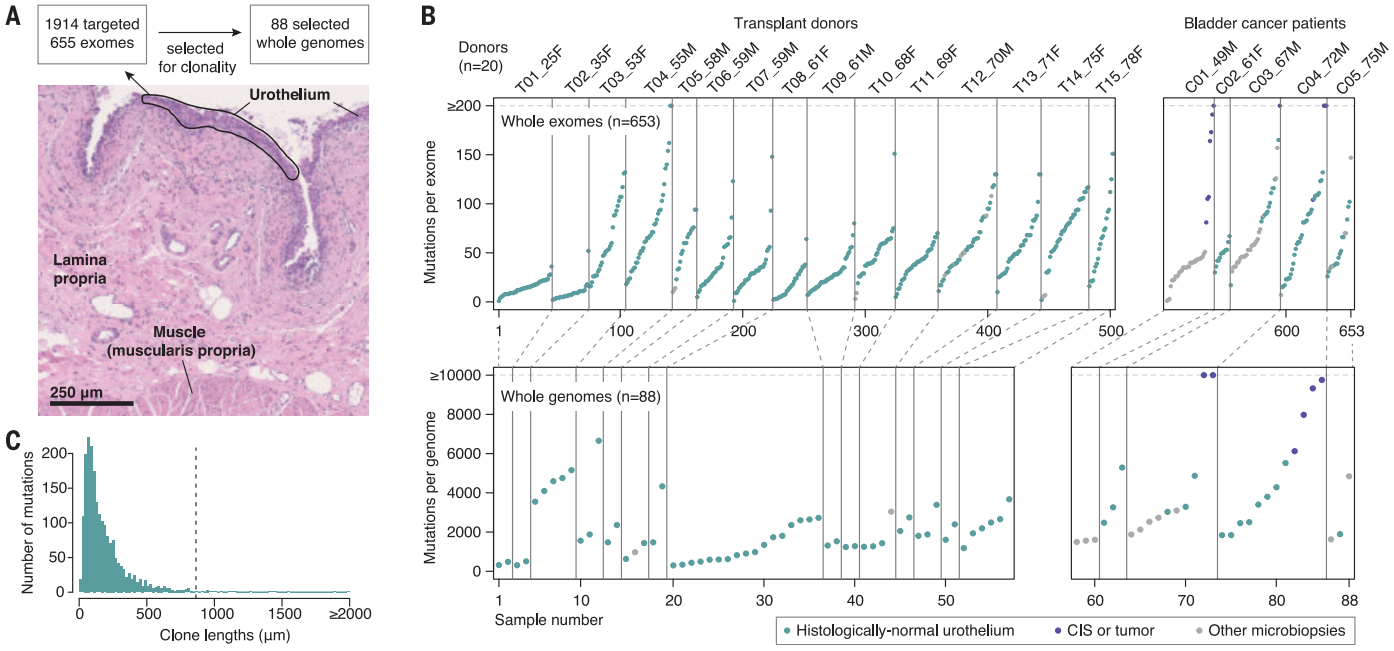


Fig. 1. Detection of somatic mutations in bladders by laser microdissection and low-input sequencing. (**A**) Sequencing strategy and histology image of bladder mucosa (hematoxylin and eosin staining). (**B**) Combined number of substitutions and indels detected per exome (top) and whole genome (bottom) across 15 transplant organ donors and five patients with bladder cancer. Donor identifiers contain age and gender information in suffix. Two exomes without mutations are not shown. (**C**) Distribution of estimated clone lengths for histologically normal urothelium (median indicated by a dashed line) (*22*).

CDKN1A, *CREBBP*, *FOXQ1*, *RHOA*, and *ERCC2* (Fig. 2A). Using restricted hypothesis testing on known bladder cancer genes and a dN/dS model at the level of single hotspots, we identified an additional five genes under selection: *KLF5*, *ZFP36L1*, *ELF3*, *GNA13*, and *PTEN* (*22*). Overall, 17 genes were found to be under clear, positive selection, which implies that their mutation confers a competitive advantage on the mutant cells over neighboring cells in normal urothelium.

The enrichment of nonsynonymous mutations in positively selected genes was large, with dN/dS ratios higher than 10 or even 100 (Fig. 2B). In most genes, selection on protein-truncating mutations (indels, nonsense, and essential splice site substitutions) was stronger than that on missense mutations—a pattern characteristic of tumor suppressor genes (*19*). In fact, although indels contributed just under 8% of all detected mutations across exomes and genomes, they accounted for 39% of all driver mutations. Clear exceptions were *RHOA*, *ERCC2*, and *GNA13*, which displayed higher frequencies of missense mutations, typically at known oncogenic hotspots (Fig. 2B and fig. S3). Overall, on the basis of the excess of nonsynonymous mutations measured by dN/dS, we detected a total of 385 (95% confidence interval: 357 to 401) driver mutations across all microbiopsies (*22*).

We can integrate allele fractions to estimate the proportion of cells in bladder urothelium that carry a driver mutation, while also accounting for the possibility of undetected copy number losses and mutations occurring in one or two alleles per cell (Fig. 2C) (*22*). This conservatively estimates that between 8 and 19% of cells carry a driver mutation in normal bladder urothelium in middle-aged and elderly individuals.

Chromatin remodeling genes dominate the driver landscape

Of the 17 positively selected genes, all but *NOTCH2* have been identified as bladder cancer genes from TCGA (The Cancer Genome Atlas) data (*18*, *19*) (Fig. 2D). In contrast to the case of *NOTCH1* in the normal esophagus (*7*, *8*), the mutation frequency of these 17 genes is higher in bladder cancers than in normal urothelium from middle-aged and elderly individuals in our cohort. This suggests that these mutations confer an increased tumorigenic potential on the mutant cells, even if the risk of progression of individual clones is extremely small. Most common bladder cancer genes can be classified into three functional groups: the RTK-Ras-PI3K pathway (such as *PIK3CA*, *FGFR3*, *ERBB2*, and *ERBB3*), the p53-Rb pathway (such as *TP53*, *RB1*, and *ATM*), and genes involved in chromatin remodeling (*18*, *24*). Five of the top six most-mutated driver genes in the normal bladder are involved in chromatin remodeling, whereas mutations in RTK-Ras-PI3K or p53-Rb genes that are very common in bladder cancer are much rarer in normal urothelium (Fig. 2E).

The absence of mutations in some of the main bladder cancer genes was noteworthy. Across 1500 microbiopsies, we found only three independent mutations in *TP53*, which is mutated in nearly 50% of muscle-invasive bladder cancers, and no mutations in *FGFR3*, which is mutated in 60 to 80% of non–muscle-invasive bladder cancers (*25*). We also did not detect any *TERT* promoter mutations across 55 whole genomes of normal urothelium, despite it being mutated in ~70 to 80% of bladder cancers, including early-stage bladder cancers (*26*). This suggests that these driver mutations may not confer large clonal advantages in normal urothelium but are key drivers of bladder cancer development. Detection of mutations in these genes in liquid biopsies may prove informative for the early detection of bladder cancer (*26*).

The analyses above were restricted to the targeted panel of 321 known cancer genes. The extent of selection in normal tissues outside of known cancer genes is less understood. It is conceivable that the mutation of certain genes could drive benign clonal expansions in healthy tissues without contributing to tumorigenesis or could even push cells down evolutionary paths away from cancer. Running dNdScv on all genes using 483 whole exomes from normal urothelium yielded seven genes under clear, positive selection, all within the list of 17 genes above (*22*). This confirms that the most common drivers of clonal expansions

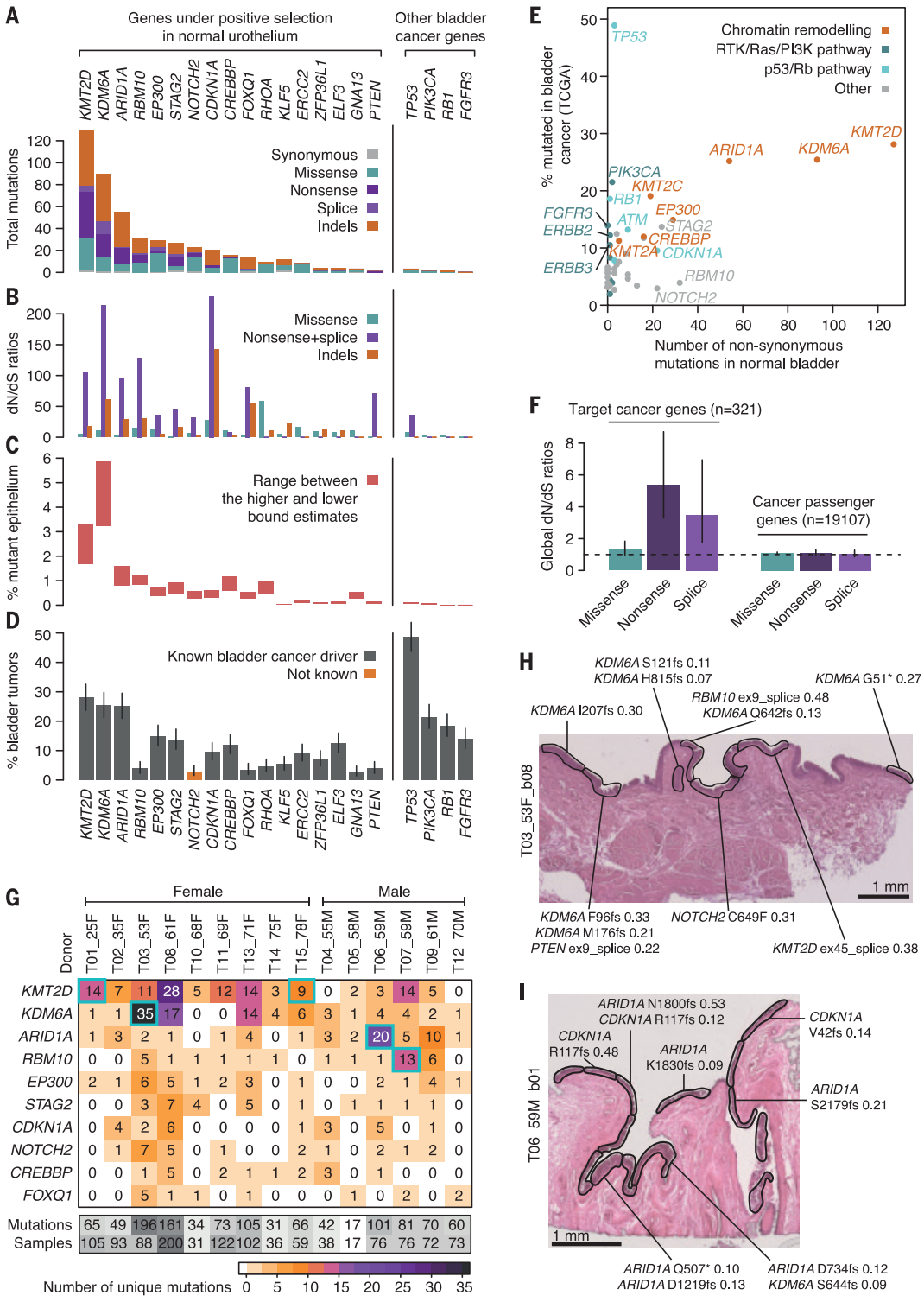


Fig. 2. Positive selection of bladder cancer genes in normal urothelium from organ donors. In (A) to (D), analyses are shown for 17 genes under positive selection in normal urothelium and for four other genes frequently mutated in bladder cancer. (**A**) Number and consequence of mutations detected in histologically normal urothelium. (**B**) Observed-to-expected ratios for missense substitutions, truncating (nonsense and essential splice site) substitutions, and indels. (**C**) Estimated percentage of urothelial cells bearing a mutation for donors aged ≥50 years from samples with median on-target coverage ≥50×. (**D**) Percentage of urothelial carcinomas in TCGA with a nonsynonymous substitution or indel. Error bars depict 95% binomial confidence intervals. (**E**) Scatterplot comparing mutation frequency in bladder cancer (D) and the number of nonsynonymous mutations in normal urothelium (A) for driver genes (colored by biological function) identified in this study and in previous studies (*18*, *19*). (**F**) Comparison of dN/dS values for the 321 cancer genes in the targeted panel to 19,107 cancer passenger genes [defined in (*19*)]. Dashed line indicates a dN/dS value of 1, which indicates neutral expectation. (**G**) Heatmap showing the number of unique nonsynonymous mutations in abundant (≥10 mutations) driver genes across transplant organ donors. Sample numbers refer to samples with at least one mutation. Blue boxes indicate statistically significant combinations of gene and donor (*22*). (**H** and **I**) Histology images annotated with driver mutations and their cellular fractions in sequenced microbiopsies from two transplant organ donors exhibiting enrichment of drivers in *KDM6A* and *ARID1A*, respectively.

in normal urothelium are all known cancer genes. Somatic mutations in certain genes could also lead to cellular death or differentiation, which would lead to a depletion of protein-altering mutations in surviving clones. Although this dataset is not powered to detect negative selection at the level of individual genes, exome-wide dN/dS ratios excluding known cancer genes were close to, and not significantly lower than, 1 (Fig. 2F). This is consistent with the vast majority of somatic coding point mutations being tolerated by normal cells and accumulating passively, in line with observations in cancer genomes (*19*). Similar nonsignificant results were obtained when focusing on putative antigenic regions of the exome, thereby providing no clear evidence of immune editing against these mutant clones (fig. S4) (*22*).

Extreme variation in driver preference across individuals

Having identified many independent mutant clones per donor, we were able to study

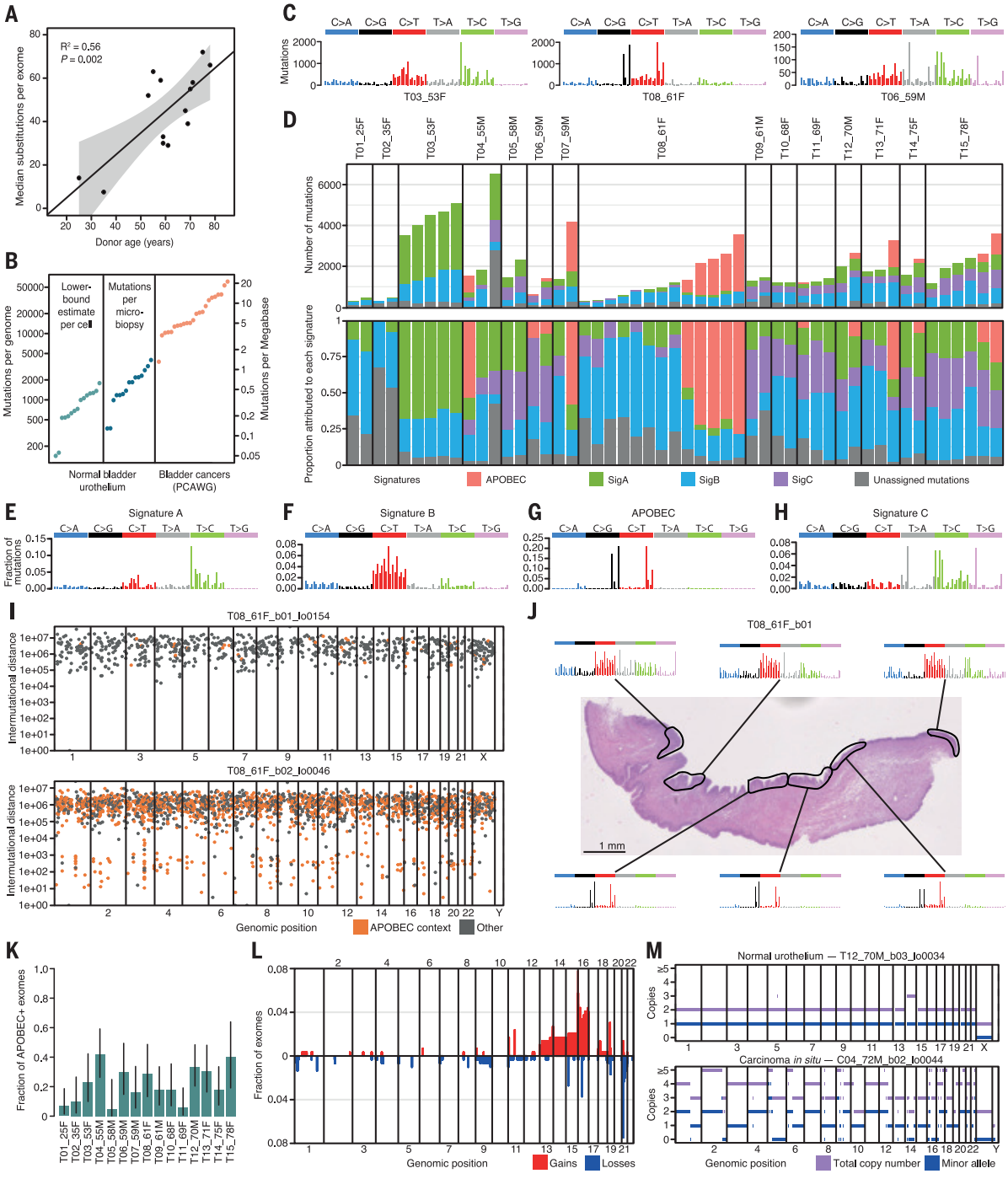


Fig. 3. Mutation burden and signatures in normal urothelium. (**A**) Scatterplot of donor age versus the median number of substitutions in high-coverage exomes (≥40× for ≥80% of the exome). The fitted line, R^2 (coefficient of determination) value, and *P* value were obtained by linear regression. (**B**) Comparison of mutation burden between normal bladder urothelium and bladder cancers. To account for subclonality, both a mean lower-bound estimate per cell (*22*) and the mean number of mutations per microbiopsy are shown for whole genomes from the 15 transplant organ donors. Bladder cancer data reflect total mutations per genome from PCAWG (*48*). (**C**) Raw mutational spectra for all urothelial genomes combined for three donors. (**D**) Number (top) and proportion (bottom) of mutations assigned to the four most abundant signatures extracted using a Bayesian hierarchical Dirichlet process (*22*) for urothelial genomes from transplant organ donors. The weak attribution of signature C (SigC) to genomes from T08 may reflect overfitting to residual ATT > AAT alignment errors. (**E** to **H**) Bar plots depicting mutational spectra, split by type and trinucleotide context, of extracted signatures, as in (*17*). (**I**) Intermutational distance plots for urothelial clones free from and affected by APOBEC activity, respectively, as in (*17*). (**J**) Histology image depicting variability in mutational processes between nearby urothelial microbiopsies. Mutational spectra are from independent clones. (**K**) Fraction of exomes with evidence of APOBEC mutagenesis (*22*). Error bars depict 95% binomial confidence intervals. (**L**) Proportion of exomes from normal urothelium with large-scale copy number alterations in autosomes (*22*). Gains (red) and losses (blue) are shown above and below the *x* axis respectively. (**M**) Copy number plots for representative whole genomes of normal urothelium (top) and CIS (bottom).

differences in selection across individuals. We used a dN/dS-based likelihood ratio test that compares the relative enrichment of nonsynonymous mutations in particular genes, while also correcting for differences in mutation rates, mutation signatures, coverage, and selection at other genes (*22*). This analysis revealed notable differences in the landscape of clonal selection across donors (Fig. 2G and fig. S4). For example, one individual (T03_53F) had 35 different *KDM6A* mutations and two *ARID1A* mutations, whereas another (T06_59M) had four *KDM6A* mutations and 20 *ARID1A* mutations (Fig. 2, G to I). The four most-frequently mutated genes in our dataset, *KMT2D*, *KDM6A*, *ARID1A*, and *RBM10*, all showed highly significant differences in selection across donors (Fig. 2G; $Q < 0.05$ from dN/dS likelihood ratio tests) (*22*).

It is unclear whether these differences are driven by variability in environmental exposures or by the genetic background of each individual. No clear evidence of pathogenic germline mutations was found in these genes (*22*). *KDM6A* and *RBM10* are both located on the X chromosome, and *KDM6A* is known to escape X-chromosome inactivation, with some evidence suggesting that both *KDM6A* and *RBM10* are more-frequently mutated in males across cancer types (*27*). However, in our limited cohort, *KDM6A* appears to be more-frequently mutated in women than men, which is in line with previous observations in non–muscle-invasive bladder cancers (*28*). Larger cohorts would be required to establish robust associations between epidemiological factors and differences in somatic mutation rates and selection.

Large heterogeneity in burden and signatures across clones and donors

The whole-exome data showed an increase in the number of mutations detected with age, which is consistent with continual, irreversible accumulation of mutations during life (Fig. 3A). To estimate the mutation burden per cell despite the presence of multiple clones per microbiopsy, we used two alternative approaches to obtain lower bounds from the whole-genome data: integration of allele frequencies and deconvolution of the major subclone (*22*). We estimate that, by middle age (50 to 65 years), cells in normal urothelium carry more than 500 to 2000 mutations per genome. This burden is within the range observed for other normal tissues (*1*, *4*, *7*), but it is an order of magnitude lower than the typical burden of bladder cancers (Fig. 3B).

Analysis of the mutational spectra revealed notable differences across donors (Fig. 3C). To better understand this variation, we performed de novo mutational signature decomposition in 80 genomes of normal urothelium from all 20 individuals using a Bayesian hierarchical Dirichlet process, and we matched these signatures to known signatures from cancer genomes (figs. S5 and S6) (*22*). This identified four main signatures that contribute >89% of all mutations in the dataset (Fig. 3, D to H). The same four signatures were found using nonnegative matrix factorization (SigProfiler) (fig. S7A) (*22*).

One signature, the third most abundant, was clearly attributable to APOBEC mutagenesis (cosine similarity with SBS2 + SBS13 = 0.995) (*29*). The high mutation burden in bladder cancers is largely driven by activation of APOBEC3 cytidine deaminases, which preferentially generate C > G and C > T changes in a TCN context (Fig. 3G) (*17*). APOBEC mutagenesis has been reported only rarely in normal tissues sequenced to date (*8*, *9*, *15*), but it occurs frequently in normal urothelium and contributes hundreds to thousands of mutations in the clones in which it is active (Fig. 3D).

The other three signatures did not match known signatures (fig. S6). Signatures A and B may contain a fraction of SBS5 mutations, which are common in bladder cancers (*17*), but they were stably extracted as separate from small amounts of SBS5 when using known signatures as priors or when adding cancer genomes to the signature extraction (figs. S7 and S8) (*22*). Signature A is dominated by T > C changes, with a clear transcriptional strand bias suggestive of transcription-coupled damage or repair (Fig. 3E and fig. S9). Reanalysis of whole-genome data from the PCAWG (Pan-Cancer Analysis of Whole Genomes) consortium suggests a high contribution of signature A to some bladder cancer genomes (fig. S6, C to E) (*22*). Signature B is dominated by C > T changes (Fig. 3F) and shares some resemblance with SBS5 in combination with a signature rich in C > T with a modest transcriptional strand bias (figs. S6 and S9). Signature C has distinct peaks at T > A and T > G in an ATT context (Fig. 3H) and does not resemble any known signature or combination of signatures (fig. S6). It has a strong transcriptional strand asymmetry with lower mutation rates in transcribed regions (fig. S9)—a pattern indicative of this signature being generated by DNA damage to thymines by adducts and subject to transcription-coupled repair (*9*). Signature C also has an extended sequence context dominated by adenines and thymines (fig. S10).

The relative contribution of different signatures within each individual was particularly noteworthy. APOBEC mutations are responsible for large differences in mutation burden and spectra between clones (Fig. 3D). This contrasts with signatures A to C, which show little variation across clones from the same individual but large differences between individuals (Fig. 3D). For example, signature A contributes ~70% of mutations in all clones from a 53-year-old woman (T03_53F), but it is scarcely present (~5% of all mutations) in all clones from a 61-year-old woman (T08_61F). Similarly, signature C contributes >25% of all mutations in 6 of the 15 donors, but it is undetectable in others (Fig. 3D). The interindividual differences in mutational signatures, together with the diverse etiology of bladder cancers, is suggestive of variable mutagenic exposures through the urine. This is exemplified by the presence of aristolochic acid mutagenesis in normal urothelium from Chinese patients (*30*). Smoking is a major risk factor of bladder cancer, increasing risk three- to fourfold (*20*). No evidence of the smoking-associated signature (SBS4) was found in any of the individuals, including the heavy smokers (table S1), which is a pattern consistent with the lack of SBS4 in bladder cancers from smokers (*31*). We used a linear mixed-effect regression model to test whether any of the four signatures found was statistically associated with smoking or alcohol consumption. Despite the small cohort size, signature A was significantly associated with smoking history (linear mixed-effect regression, $P = 9.4 \times 10^{-5}$; fig. S11) (*22*), which raises the possibility that signature A may result from tobacco smoke mutagens excreted in the urine. Signature A may thus offer a mechanistic link between smoking and bladder cancer risk.

One additional source of heterogeneity across clones was exemplified by the microbiopsy with the highest mutation burden of the cohort, which contained ~6500 mutations (Fig. 3D and fig. S12). This genome carried a hotspot mutation (N238T) in *ERCC2*, which is known to cause hypermutation in some bladder cancers through aberrant nucleotide excision repair (*32*). A total of eight different *ERCC2* mutations were identified in the targeted and exome data, with clear, positive selection acting on *ERCC2* (Fig. 2), which suggests that this mechanism is relatively common in normal urothelium.

Frequency and spatial distribution of APOBEC clones

APOBEC-induced mutations in normal urothelium displayed the characteristic replicational strand bias observed in human cancers and an extended sequence context that suggested that APOBEC3A might be the main contributing enzyme (fig. S10) (*22*, *33*). Analysis of APOBEC-positive genomes revealed extensive evidence of mutational clusters, known as kataegis (Fig. 3I) (*17*). These clusters were modest in size and displayed the typical strandedness observed in cancer genomes. Although kataegis in cancers is often reported to occur near rearrangement breakpoints (*17*), this was not the case in normal urothelium. Overall, the patterns observed were consistent with replication-associated APOBEC mutagenesis (*34*).

Analysis of the distribution of APOBEC-positive genomes in their tissue context revealed

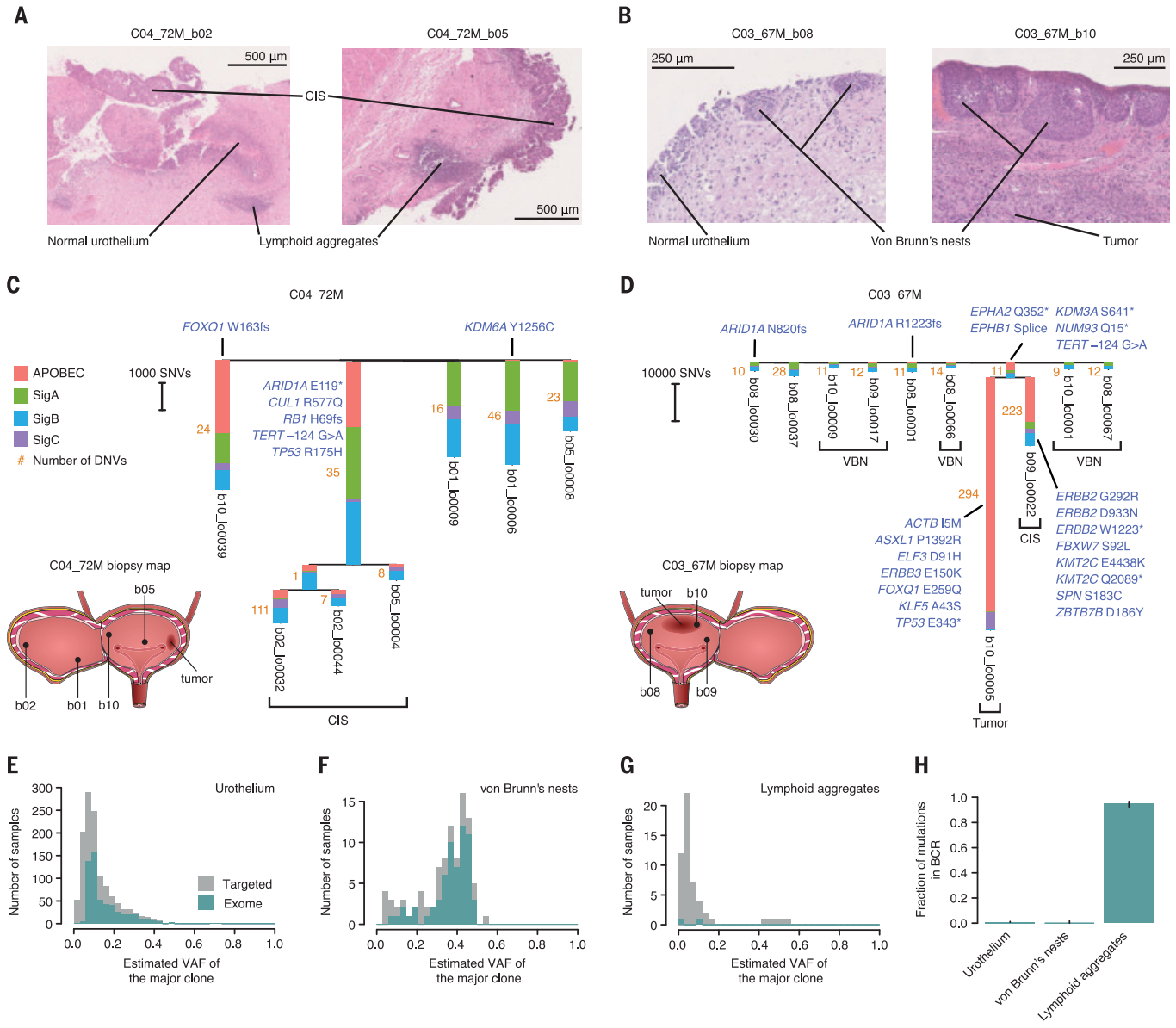


Fig. 4. The mutational landscape across histological features in patients with bladder cancer. (**A** and **B**) Histology images depicting features microdissected from cystectomy material for two patients with bladder cancer (C04_72M and C03_67M). (**C** and **D**) Phylogenetic reconstruction of the evolution of cancer and CIS clones (*22*). Only microbiopsies with a high degree of clonality (mean VAF ≥ 0.25) were included. Biopsy maps show the relative positions of macroscopic biopsies (b01 to b10) within the bladder. Branches without a feature indicated are histologically normal urothelium. Branch lengths depict single nucleotide variant (SNV) counts, and the number next to each branch denotes assigned dinucleotide variants (DNVs). Driver genes identified in this study and in previous studies (*18*, *19*) are annotated. Truncating mutations in *EPHB1* and *KDM3A* are annotated in the branch shared by the CIS and tumor for C03_67M. VBN, von Brunn's nest. (**E** to **G**) Histograms showing the estimated VAF of the major clone in targeted and exome sequencing data for three different histological features: urothelium, von Brunn's nests, and lymphoid aggregates. (**H**) Proportion of mutations located within the B cell receptor (BCR) across histological features.

a suggestive example of spatial clustering of three APOBEC-positive clones (Fig. 3J). To study the frequency and spatial distribution of APOBEC-positive clones, we used signature fitting and a likelihood ratio test to annotate all exomes according to their evidence of APOBEC mutagenesis (*22*). Across donors, 22% of all microbiopsies of normal urothelium showed evidence of APOBEC mutagenesis (likelihood ratio test, $Q < 0.05$; Fig. 3K). To determine whether APOBEC-positive clones tend to cluster in space, for each positive clone we calculated the fraction of positive clones surrounding it (Euclidean distance <1 mm) both in the real data and in random permutations of the data (fig. S13) (*22*). This analysis suggests that APOBEC clones appear to be scattered uniformly in the tissue (permutation test, $P = 0.92$), without evidence of spatial clustering of unrelated clones, which suggests that APOBEC mutagenesis is typically triggered independently in individual cells across the urothelium.

Copy number and rearrangement analyses of normal urothelium revealed that the majority of clones carry no structural variants (*22*). Copy number alterations were detected in only 28% of urothelial exomes, with the most common changes involving whole or arm-level gains of chromosomes 13, 14, 15, and 16

and losses of chromosomes 9 and 21 (Fig. 3L). Across 55 genomes of normal urothelium, only 30 rearrangements and three retrotransposition events were detected (tables S7 and S11) (*22*). This is in stark contrast with bladder cancers, which display extensive aneuploidy with an average of ~200 segmental alterations per exome and 1.7 retrotransposition events per genome (*35*, *36*). This pattern is similar to that observed in other normal tissues (*3*, *4*, *7*, *9*, *37*), and it suggests that extensive structural changes are characteristic of later stages of carcinogenesis across a wide range of cancer types.

The mutational landscape in bladder cancer patients

Bladder cancer often presents with multiple synchronous tumors in different parts of the bladder. It remains unclear to what extent this is due to large premalignant clones colonizing distant parts of the bladder or to widespread changes in multiple independent clones across the bladder (*38*). To explore the mutational landscape of histologically normal urothelium in bladder cancer patients, and to study the genomic changes underlying histologically abnormal areas, we performed laser microdissection of 450 microbiopsies from 19 distant biopsies from five bladder cancer patients.

Analysis of histologically normal urothelium from bladder cancer patients revealed patterns similar to those observed in healthy bladders. As in transplant organ donors, mutant clones were small—typically constrained to single microbiopsies (fig. S2). There seems to be a modest increase in the number of mutations detected per exome (linear mixed-effect regression, P = 0.0068) and in their allele frequencies (P = 0.00048) in some cystectomy samples (fig. S14) (*22*). However, differences should be interpreted with caution given the limited cohort size and the considerable interindividual variation. The fraction of APOBEC-positive microbiopsies was similar in cystectomies and in age-matched transplant organ donors (25 versus 24%, Fisher's exact test, P = 0.91). Driver discovery in 223 microbiopsies of normal urothelium from bladder cancer patients yielded a very similar driver landscape to that observed in the 15 transplant organ donors, and the density of driver mutations detected per microbiopsy appeared comparable (*22*). Although a much larger number of patients would be required to accurately quantify differences between cohorts, these results suggest that the mutational landscape of histologically normal urothelium from bladder cancer patients broadly resembles the patchwork of microscopic clones observed in healthy donors. The results also suggest that widespread mutational changes in independent clones are unlikely to explain the emergence of multiple tumors in bladder cancer, which is consistent with the observation that synchronous tumors tend to be clonally related (*38*–*40*).

Areas of carcinoma in situ (CIS) were observed in three of the five cystectomies studied. CIS of the bladder is a flat, high-grade urothelial carcinoma restricted to the epithelial layer, which often appears concomitantly with more-advanced tumors. A total of 44 CIS microbiopsies were sequenced, including 11 whole exomes and 5 whole genomes. Phylogenetic analysis revealed that all CIS areas sequenced within a patient were clonally related (Fig. 4, A to D, and figs. S15 and S16). In a 72-year-old patient (C04_72M), the same CIS clone was detected in two biopsies several centimeters away from the tumor and from one another, with most mutations shared across distant biopsies (Fig. 4C). The phylogenetic tree provides a snapshot of the genome of the most recent common ancestor cell that gave rise to this clone. This cell had an only modestly increased burden—largely due to APOBEC—compared with other clones in normal urothelium, but it had already acquired driver mutations in *ARID1A*, *RB1*, and *TP53* as well as a hotspot promoter mutation in *TERT* (Fig. 4C). In contrast to histologically normal clones, the CIS showed extensive aneuploidy, including evidence of whole-genome duplication (Fig. 3M). Notably, one of the terminal branches of the CIS clone showed an unusually high number of CC > AA dinucleotide changes of uncertain origin (Fig. 4C and fig. S17). In a 67-year-old patient (C03_67M), we sequenced an area of CIS and an area of tumor from two separate biopsies. This revealed that the tumor and the CIS had originated from a common ancestor cell that had already acquired putative driver mutations in *NUP93*, *EPHA2*, and *TERT*. The CIS and the tumor diverged early, and each subsequently acquired an entirely different complement of driver mutations (Fig. 4D), which provides a window into the early evolution of this tumor. This analysis corroborates that CIS clones are genetically highly aberrant and can colonize distant areas of the bladder, forming a hotbed from which invasive tumors can evolve (*40*, *41*). A systematic analysis of tumor and noninvasive areas combining laser microdissection and genome sequencing could help to shed light on the order of events in early bladder cancer evolution.

Laser microdissection also enabled us to study other histological changes observed in bladder cancer patients. Von Brunn's nests are groups of urothelial cells in the lamina propria, which are believed to arise from invagination of the surface urothelium (*42*). Although they are common in histological sections from bladder cancer patients (Fig. 4B), they can also be seen in small numbers in healthy individuals. Sequencing of 98 microbiopsies revealed that most von Brunn's nests are single clones, with all cells within a nest derived from a single cell (Fig. 4, E and F). Phylogenetic reconstruction reveals that adjacent nests are clonally unrelated (Fig. 4D). The vast majority of von Brunn's nests sequenced did not carry a driver mutation; their driver landscape, mutation burden, and largely diploid genomes resembled those of the adjacent histologically normal urothelium. Overall, this is consistent with von Brunn's nests being benign ectopic growths that are not actively driven by specific mutations (*22*). Lymphoid aggregates are also common in cystectomy biopsies (Fig. 4A), which reflects adaptive immunity in the tumor microenvironment, and they can also occur in healthy samples with evidence of inflammation (*43*). We microdissected 82 lymphoid aggregates for deep targeted sequencing, as the targeted gene panel contained probes for the B cell and T cell receptor loci (*22*). Unlike von Brunn's nests, lymphoid aggregates were highly polyclonal, with nearly all of the mutations detected at low allele fractions (Fig. 4G). The only exception was one clonal lymphoid aggregate, which also carried a lymphoid driver IgH-BCL2 translocation (fig. S18). This biopsy was from a donor who had previously been investigated for a possible lymphoma, although the relationship between the clonal lymphoid aggregate and the donor's clinical history is unclear. Across all lymphoid aggregates, 95% of mutations detected with the panel clustered in the *IGH* locus and had the characteristic signature of somatic hypermutation (SBS9) (Fig. 4H), which confirmed the presence of multiple clones of mature B lymphocytes in each aggregate sequenced. These examples showcase the power of laser microdissection and low-input sequencing to inform on the clonal composition and genetic changes underlying different histological structures.

Discussion

These data reveal a rich mutational landscape in healthy and diseased bladder urothelium, with widespread positive selection; extensive APOBEC mutagenesis; and large differences in mutation burden, signatures, and selection across clones and individuals. The heterogeneity in mutational signatures and driver mutations across donors is particularly notable and appears larger than that reported in other tissues. Epidemiological studies have linked bladder cancer risk to a diversity of carcinogens, such as smoking, occupational or environmental exposures, and recurrent infections (*20*, *44*). Whether carcinogens are genotoxic (inducing mutations) or nongenotoxic (affecting cellular growth or the microenvironment), they are expected to leave distinct marks on the mutational landscape of normal tissues—altering mutation rates, mutation signatures, driver frequencies, or clone sizes. Thus, the differences in the mutational landscapes across individuals observed here may be expected to reflect the interplay between

genetics and a lifetime of different exposures. The differences across donors might also raise the possibility of developing personalized risk models (*45*). However, our results suggest that differences in normal urothelium between healthy individuals and cancer patients may be subtle, consistent with theories predicting that modest differences in mutation and selection could have considerable effect on risk (*46*, *47*). Systematic analyses of large cohorts of individuals will be needed to quantify the relationship between epidemiological factors, germline variants, changes in the mutational and selective landscape, and risk. Such analyses might enable the development of mechanistic risk models of cancer development.

Although somatic mutations have traditionally been studied in the context of cancer, the growing realization that some human tissues become colonized by mutant clones throughout life raises questions about their potential impact in aging and other diseases. Laser microdissection and low-input sequencing enable the study of somatic mutations associated with histological changes and could shed light on somatic evolution in cancer, aging, and nonmalignant disease.

REFERENCES AND NOTES

1. J. S. Welch *et al.*, *Cell* **150**, 264–278 (2012).
2. S. Jaiswal *et al.*, *N. Engl. J. Med.* **371**, 2488–2498 (2014).
3. I. Martincorena *et al.*, *Science* **348**, 880–886 (2015).
4. F. Blokzijl *et al.*, *Nature* **538**, 260–264 (2016).
5. M. A. Lodato *et al.*, *Science* **359**, 555–559 (2018).
6. K. Suda *et al.*, *Cell Rep.* **24**, 1777–1789 (2018).
7. I. Martincorena *et al.*, *Science* **362**, 911–917 (2018).
8. A. Yokoyama *et al.*, *Nature* **565**, 312–317 (2019).
9. H. Lee-Six *et al.*, *Nature* **574**, 532–537 (2019).
10. S. F. Brunner *et al.*, *Nature* **574**, 538–542 (2019).
11. L. Moore *et al.*, *Nature* **580**, 640–646 (2020).
12. K. Yizhak *et al.*, *Science* **364**, eaaw0726 (2019).
13. I. Franco *et al.*, *Nat. Commun.* **9**, 800 (2018).
14. J. D. Krimmel *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **113**, 6005–6010 (2016).
15. K. Yoshida *et al.*, *Nature* **578**, 266–272 (2020).
16. C. Wang, W. T. Ross, I. U. Mysorekar, *Dev. Dyn.* **246**, 336–343 (2017).
17. L. B. Alexandrov *et al.*, *Nature* **500**, 415–421 (2013).
18. A. G. Robertson *et al.*, *Cell* **171**, 540–556.e25 (2017).
19. I. Martincorena *et al.*, *Cell* **171**, 1029–1041.e21 (2017).
20. S. Letašiová *et al.*, *Environ. Health* **11**, S11 (2012).
21. S. L. Poon *et al.*, *Genome Med.* **7**, 38 (2015).
22. See materials and methods in the supplementary materials.
23. N. T. Gaisa *et al.*, *J. Pathol.* **225**, 163–171 (2011).
24. Y. Gui *et al.*, *Nat. Genet.* **43**, 875–878 (2011).
25. S. Hernández *et al.*, *J. Clin. Oncol.* **24**, 3664–3671 (2006).
26. Y. Allory *et al.*, *Eur. Urol.* **65**, 360–366 (2014).
27. A. Dunford *et al.*, *Nat. Genet.* **49**, 10–16 (2017).
28. C. D. Hurst *et al.*, *Cancer Cell* **32**, 701–715.e7 (2017).
29. L. B. Alexandrov *et al.*, *Nature* **578**, 94–101 (2020).
30. R. Li *et al.*, *Science* **370**, 82–89 (2020).
31. L. B. Alexandrov *et al.*, *Science* **354**, 618–622 (2016).
32. J. Kim *et al.*, *Nat. Genet.* **48**, 600–606 (2016).
33. K. Chan *et al.*, *Nat. Genet.* **47**, 1067–1072 (2015).
34. H. Vöhringer, A. van Hoeck, E. Cuppen, M. Gerstung, Learning mutational signatures and their multidimensional genomic properties with TensorSignatures. bioRxiv 850453 [Preprint]. 19 June 2020. https://doi.org/10.1101/850453.
35. Cancer Genome Atlas Research Network, *Nature* **507**, 315–322 (2014).
36. B. Rodriguez-Martin *et al.*, *Nat. Genet.* **52**, 306–319 (2020).
37. L. Moore *et al.*, *Nature* **580**, 640–646 (2020).
38. Ö. Acar *et al.*, *BMC Cancer* **15**, 871 (2015).
39. M. B. H. Thomsen *et al.*, *Sci. Rep.* **7**, 11702 (2017).
40. T. Majewski *et al.*, *Cell Rep.* **26**, 2241–2256.e4 (2019).
41. K. Shin *et al.*, *Nat. Cell Biol.* **16**, 469–478 (2014).
42. K. E. Volmar, T. Y. Chan, A. M. De Marzo, J. I. Epstein, *Am. J. Surg. Pathol.* **27**, 1243–1252 (2003).
43. M. Koti *et al.*, *Bladder Cancer* **3**, 259–267 (2017).
44. S. H. Vermeulen *et al.*, *Br. J. Cancer* **112**, 594–600 (2015).
45. S. Abelson *et al.*, *Nature* **559**, 400–404 (2018).
46. P. Armitage, R. Doll, *Br. J. Cancer* **8**, 1–12 (1954).
47. C. Tomasetti, L. Marchionni, M. A. Nowak, G. Parmigiani, B. Vogelstein, *Proc. Natl. Acad. Sci. U.S.A.* **112**, 118–123 (2015).
48. ICGC/TCGA Pan-Cancer Analysis of Whole Genomes Consortium, *Nature* **578**, 82–93 (2020).
49. I. Martincorena, im3sanger/Lawson_etal_2020_somatic_mutations_bladder: Manuscript code and auxiliary files, version v1.0.0, Zenodo (2020); http://doi.org/10.5281/zenodo.3966023.

ACKNOWLEDGMENTS

We are grateful to the families of the deceased transplant organ donors and the patients with bladder cancer for their consent and to the Cambridge Biorepository for Translational Medicine for access to human tissue. We thank P. H. Jones and J. C. Fowler for their early help with wholemounts; L. Alexandrov for advice on mutational signatures; K. Haase and P. van Loo for their advice on calling copy number changes in exome data using ASCAT (allele-specific copy number analysis of tumors); J. M. A. Lawson for artistic contribution to figures; D. Phillips for advice on carcinogen exposure in urine; P. Ellis, P. Nicola, M. Maddison, E. Anderson, S. Gamble, K. Roberts, and A. Dooner for technical assistance; J. Hewinson and C. Hardy for their assistance with project management; J. Field-Rayner for consenting patients; and E. Cromwell for tissue processing. **Funding:** I.M. is funded by Cancer Research UK (C57387/A21777) and the Wellcome Trust. P.J.C. is a Wellcome Trust Senior Clinical Fellow. T.J.M. is funded by Cancer Research UK, Royal College of Surgeons Clinician Scientist Fellowship (C63474/A27176). L.M. is a recipient of a CRUK Clinical Ph.D. fellowship (C20/A20917). Fresh cystectomy samples were acquired as part of the DIAMOND study "Evaluation of biomarkers in urological disease - NHS National Research Ethics Service reference 03/018," whose infrastructure is partially funded by the Cambridge NIHR BRC and CRUK Cambridge Cancer Centre Urological Malignancies program. **Author contributions:** A.R.J.L. and I.M. conceptualized the project with support from P.J.C., M.R.S., and T.J.M. A.R.J.L. and I.M. led the data analysis with support from F.A., T.H.H.C., H.V., and S.Z. A.R.J.L. led the experimental work with support from Y.H., L.M.R.H., and A.C. T.M.B. and L.M. contributed to method development. K.R., M.A.S., A.M., N.W., H.V., J.N., M.G., and I.M. developed algorithms and software. L.O., C.L., and K.T.A.M. helped with samples and project administration. A.Y.W., K.T.A.M., B.B., A.J.C., W.T., B.T., V.G., and K.S.-P. collected samples. J.N., J.M.C.T., M.G., K.S.-P., M.R.S., P.J.C., T.J.M., and I.M. provided supervision. D.M.R. provided histology support. I.M. and A.R.J.L. wrote the manuscript, and all authors contributed to reviewing and editing it. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** Sequencing data are available in the European Genome-phenome Archive (EGA) under accession numbers EGAD00001006113, EGAD00001006114, EGAD00001006115, EGAD00001006116, and EGAD00001006117. Reproducible code is available in the supplementary materials and on Zenodo (*49*).

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/370/6512/75/suppl/DC1
Materials and Methods
Figs. S1 to S18
Tables S1 to S11
References (*50*–*76*)
MDAR Reproducibility Checklist

9 January 2020; accepted 5 August 2020
10.1126/science.aba8347

MUTATION

Macroscopic somatic clonal expansion in morphologically normal human urothelium

Ruoyan Li[1]*, Yiqing Du[2]*, Zhanghua Chen[1]*, Deshu Xu[1]*, Tianxin Lin[3]*, Shanzhao Jin[1], Gongwei Wang[4], Ziyang Liu[1], Min Lu[5], Xu Chen[3], Tao Xu[2]†, Fan Bai[1]†

Knowledge of somatic mutation accumulation in normal cells, which is essential for understanding cancer development and evolution, remains largely lacking. In this study, we investigated somatic clonal events in morphologically normal human urothelium (MNU; epithelium lining the bladder and ureter) and identified macroscopic clonal expansions. Aristolochic acid (AA), a natural herb-derived compound, was a major mutagenic driving factor in MNU. AA drastically accelerates mutation accumulation and enhances clonal expansion. Mutations in MNU were widely observed in chromatin remodeling genes such as *KMT2D* and *KDM6A* but rarely in *TP53*, *PIK3CA*, and *FGFR3*. *KMT2D* mutations were found to be common in urothelial cells, regardless of whether the cells experience exogenous mutagen exposure. Copy number alterations were rare and largely confined to small-scale regions, along with copy-neutral loss of heterozygosity. Single AA-associated clones in MNU expanded to a scale of several square centimeters in size.

Over the course of their life span, cells inevitably acquire somatic mutations that mainly result from unrepaired or incorrectly repaired DNA replication errors that occur during cell division (*1*, *2*). Although most somatic mutations in normal cells do not have any phenotypic consequence, mutations that affect essential genes, especially those related to cell proliferation and death, may trigger mutant clonal expansions (*3*). A well-recognized example is human cancer, in which progressive accumulation of somatic mutations drives clonal expansions and the eventual malignant transformation of cells (*4*). Although genomic sequencing of various human malignancies has revolutionized our understanding of the molecular and genetic bases of cancer development and evolution (*5*–*7*), little is known about the patterns and driving factors of somatic mutations in normal cells before malignant transformation. Recent studies have shed light on the mutational landscapes of different normal tissues, including skin epidermis (*8*), esophageal tissue (*3*, *9*),

colorectal tissue (*10*), liver (*11*), endometrial epithelium (*12*), bronchial epithelium (*13*), brain (*14*, *15*), embryonic tissue (*16*), and blood cells (*17*), thus contributing to our understanding of mutation rates, driver genes, and mutagenic driving forces in normal cells (*18*, *19*). In particular, previous studies have highlighted the critical roles of the aging-related endogenous mutational process in normal cells, evidenced by the positive correlation between mutation load and age (*3*, *9*, *15*). Also, ultraviolet light, as an exogenous mutagenic factor, has been reported to trigger mutagenesis in normal skin cells and induce skin cancer onset (*8*, *20*, *21*). Whether other underlying mutational processes, both endogenous and exogenous, operate early in normal cells warrants further investigation.

The urothelium is the epithelium that lines the urinary bladder and ureters. It is classified as a transitional epithelium because its properties lie between stratified squamous and simple nonstratified epithelia (*22*). It is highly regenerative in response to damage, thus guaranteeing its barrier function (*23*). Given its direct contact with urine, the urothelium is continually exposed to an array of potentially carcinogenic metabolic products and environmental factors that can cause tissue damage and pose genotoxic stress to urothelial cells. Under these conditions, the urothelium may accumulate somatic mutations through recurrent cell turnover. In this study, using a combination of laser-capture microdissection and exome sequencing, we systematically investigated somatic mutant clonal events in morphologically normal urothelium (MNU), including both bladder and ureter urothelium, from 120 patients with urothelial cell carcinoma (UCC).

Somatic mutations in MNU tissues

In total, we sequenced 161 MNU samples from 120 UCC patients with radical cystectomy or nephroureterectomy (table S1). Urothelium layers of each sample were dissected from consecutive tissue sections using laser-capture microdissection to provide a urothelial surface area of ~2 mm^2 (fig. S1A). Independent pathological examinations confirmed that MNU samples (125 from bladder and 36 from ureter), which were extracted far from tumors, were free of histological changes (fig. S1B). DNA from white blood cells of each patient was sequenced as the germline comparator. We also sequenced 126 tumors (93 bladder, 17 ureter, and 16 renal pelvis tumors) from the 120 patients. On average, we obtained 138-fold, 129-fold, and 138-fold coverage depth of target regions in UCC, MNU, and blood samples, respectively (table S2). Overall, the median mutational burden of UCC was higher than those of prostate, breast, and kidney clear cell carcinomas and comparable to The Cancer Genome Atlas (TCGA) bladder cancer data (fig. S2A and tables S3 and S4). Unexpectedly, while the median mutational burden was low, the overall mutational burden of MNU displayed a wide range (fig. S2A and tables S4 and S5). Several urothelium samples were even hypermutated (for example, sample P65U had >6000 mutations). This finding illustrates that detectable somatic mutations have accumulated in some MNU samples.

Next, we combined our cohort (including both UCC and MNU samples) with a bladder cancer cohort (Chinese population) (n = 99 individuals) (*24*) to catalog significantly mutated genes (SMGs). We identified 19 SMGs with significant recurrent mutation rates, including canonical cancer genes such as *TP53*, *ARID1A*, and *PIK3CA* (table S6). Mutations in these genes have high clonalities in tumors (fig. S2B). All 19 SMGs identified here have been reported by TCGA as potential driver genes in bladder cancer (*25*, *26*). To further investigate the occurrence of mutations in putative driver genes in MNU, we focused on both the 19 SMGs and nine additional genes that were reported by TCGA as potential driver genes and were frequently, but not significantly, mutated in our cohort (e.g., *ATM*, *KMT2C*, and *FAT1*) (Fig. 1A and tables S7 and S8). These 28 genes recapitulated key pathways (e.g., cell cycle and p53 pathways) that have been implicated in urothelial tumorigenesis (fig. S2, C and D). Overall, we found that ~37% of MNU samples had a somatic mutation in at least one of the 28 putative driver genes (Fig. 1A and tables S7 and S8). Meanwhile, we found that 28 MNU samples shared origins with their paired tumors from the same patients (Fig. 1A and table S9). When we excluded these samples, MNU with mutations in *KMT2D* (16/133, 12.0%), *KDM6A* (15/133, 11.3%), *ATM* (11/133, 8.3%), *CREBBP* (11/133, 8.3%), *FAT1* (12/133, 9.0%), and *KMT2C* (10/133, 7.5%) remained widely observed. Although *TP53* was the second most frequently mutated gene in UCC, *TP53* mutations were relatively rare in MNU (5/133, 3.8%) (Fig. 1A). Notably, the mutation rates of *FGFR3* (0/133, 0.0%) and *PIK3CA* (1/133, 0.8%) were lower than those of *CREBBP*, *ATM*, and *KMT2C* in MNU (Fig. 1A). This observation was substantially different from that in UCC, suggesting that different molecular mechanisms underlie early clonal expansion and final cancer development. Putative driver genes in MNU, such as *KMT2D* and *FAT1*, are also frequently mutated in normal skin and esophageal tissues (*3*, *8*, *9*). However, we did not observe enrichment of *NOTCH1* mutations in MNU, although it has been reported as the most frequently mutated gene in skin and esophageal tissues. This observation may reflect intrinsic biological differences among various cell types.

Widespread mutagenesis related to aristolochic acid in MNU

To explore the underlying mutagenic driving factors, we used a nonnegative matrix factorization algorithm on the MNU and UCC samples to extract potential mutational signatures (table S10). We identified three mutational signatures through the de novo extraction (Fig. 1B, fig. S3A, and table S11). Signature B closely resembled the Catalogue of Somatic Mutations in Cancer (COSMIC) signature SBS1 and SBS5 (Fig. 1B and fig. S3B). Signature C exhibited dominant C>G and C>T substitutions in the 5′-TpCpA-3′ and 5′-TpCpT-3′ context and largely conformed to COSMIC SBS2 and SBS13 (Fig. 1B and fig. S3B), which are associated with the activity of APOBEC cytidine deaminases.

Signature A displayed predominant T>A transversions with conspicuous biases in the local sequence context and a markedly high proportion in the 5′-CpTpG-3′ context (Fig. 1B and fig. S3C). This signature matched COSMIC SBS22 with the underlying etiological factor being aristolochic acid (AA), a natural herb-derived compound that is known as a notorious mutagen (*27*–*33*) (fig. S3B). Our finding demonstrates that AA mutagenesis is prevalent in normal human tissues (fig. S3, D and E, and table S11), although it was reported in noncancerous tissues in a patient with alcohol-related liver disease (*11*). Widespread AA mutagenesis in MNU was further confirmed using another mutational signature analysis approach (figs. S4 and S5A). AA-associated samples, both tumors and MNU, exhibited significantly higher mutation numbers ($P < 0.001$, Wilcoxon rank-sum test), demonstrating the strong mutagenic effect of AA (fig. S5B). We also found that AA mutagenesis was more prevalent in females than in males (fig. S5C, $P < 0.001$, Fisher's exact test). This gender bias has been reported in upper tract urothelial carcinoma, but the underlying mechanism is unclear (*34*). Our findings demonstrate that AA exposure poses strong genotoxic stress to urothelial cells and widely triggers mutagenesis in normal urothelium.

Copy number alterations in MNU tissues

We assessed copy number alterations (CNAs) in MNU and UCC samples using exome sequencing data (fig. S6). Overall, we observed marked differences in CNAs between tumors and MNU. As expected, tumors harbored extensive CNAs across the whole genome, with

[1]Biomedical Pioneering Innovation Center (BIOPIC), School of Life Sciences, Peking University, Beijing, China. [2]Department of Urology, Peking University People's Hospital, Beijing, China. [3]Department of Urology, Sun Yat-sen Memorial Hospital, Sun Yat-sen University, Guangzhou, China. [4]Department of Pathology, Peking University People's Hospital, Beijing, China. [5]Department of Pathology, School of Basic Medical Sciences, Peking University Third Hospital, Peking University Health Science Center, Beijing, China.
*These authors contributed equally to this work.
†Corresponding author. Email: xutao@pkuph.edu.cn (T.X.); fbai@pku.edu.cn (F.B.)

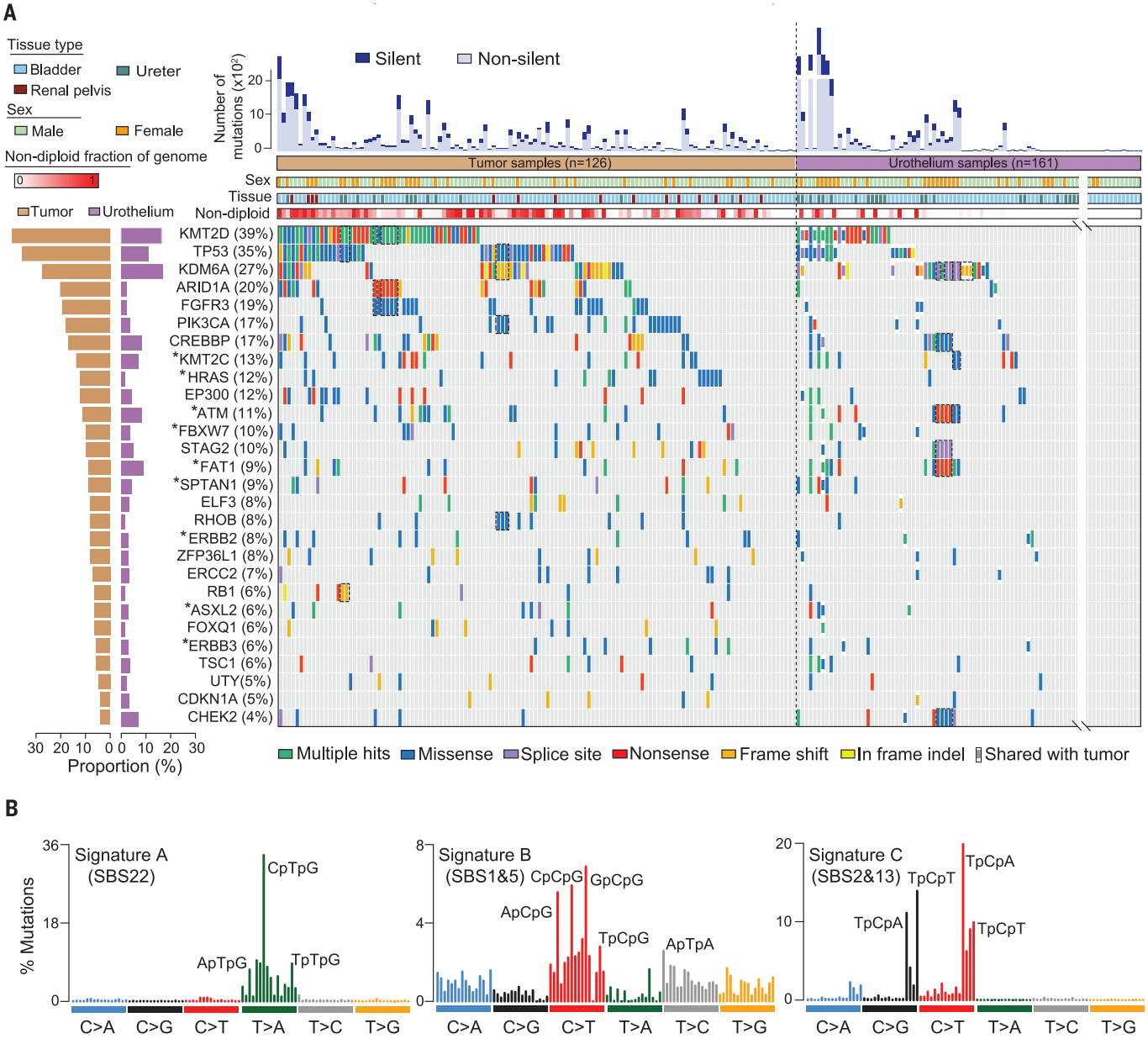


Fig. 1. Somatic mutations and mutational signatures. (**A**) Mutational landscapes of UCC and MNU samples showing mutations in putative driver genes, including 19 SMGs and nine frequently, but not significantly, mutated genes (indicated with asterisks), ordered by their mutation frequency in UCC (percent of tumors with mutations in each gene is in parentheses). Black dashed boxes show that mutations are shared in samples from the same patient. indel, insertion or deletion. (**B**) Mutational spectrum of the three de novo mutational signatures extracted by the SigProfilerExtractor analysis. Representative 3–base pair (bp) mutational contexts are labeled. Corresponding COSMIC signatures are labeled in parentheses.

recurrent CNA regions consistent with TCGA data (e.g., chromosome 5p, 8q, and 3p amplifications; and chromosome 8p and 9p deletions) (Fig. 2A). However, CNAs were rare across the genomes of MNU samples, even those with mutational burdens comparable to tumors and with mutations in *TP53*, *KMT2D*, and *KDM6A* (Fig. 2, A and B). For example, sample P6U5 had 1978 mutations and harbored driver mutations in *KMT2D* and *TP53* but displayed no obvious CNAs across its genome (fig. S6; fraction of nondiploid genome: 0.1%). CNAs in MNU were sporadic and largely confined to small-scale genomic regions, along with copy-neutral loss of heterozygosity (LOH) (Fig. 2C and fig. S6). We further explored how accumulation of somatic mutations and CNAs coordinate in UCC and MNU tissues. Notably, we found that some MNU tissues, especially those exposed to AA, had mutational burdens similar to or even higher than those of tumors, but the vast majority of their genomes remained diploid (Fig. 2D). This finding implies that acquisition of CNAs occurs late in clonal expansion in the urothelium and that genomic stability is a choke point for the final malignant transformation.

Mutational burden and mutant clone expansion in MNU

To gain deeper insights into the mutational burden and mutant clonal expansion in MNU, we used the variant allele fractions of somatic mutations to estimate the mutant cell fraction (MCF) and clone size in MNU samples. Overall, the mutational burden was markedly

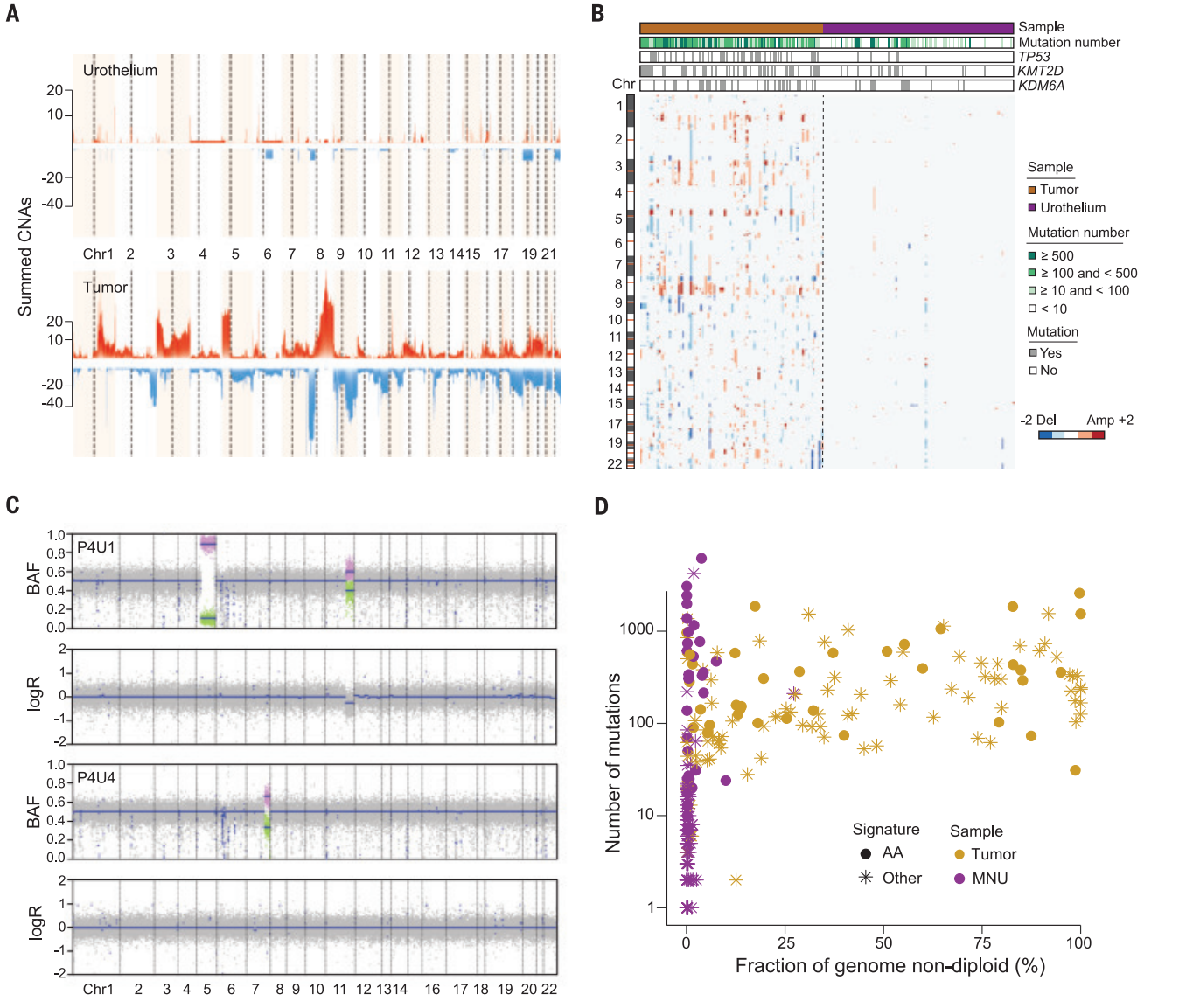


Fig. 2. CNAs in UCC and MNU samples. (**A**) Stacked mountain plots comparing summed CNAs in UCC and MNU samples. Red stacks represent amplifications, and blue stacks represent deletions. Chr, chromosome. (**B**) Comparison of CNAs in UCC and MNU samples. The mutational burdens and mutation statuses of driver genes are indicated. Del, deletion; Amp, amplification. (**C**) CNA log ratios (logR) and B allele frequency (BAF) of two representative MNU samples showing typical copy number statuses and copy-neutral LOH. Blue lines represent the fitted values of logR and BAFs calculated by the circular binary segmentation algorithm. (**D**) The relationship between somatic mutation loads and CNAs. UCC ("Tumor") and MNU samples were assigned to one of the two groups according to their mutational signatures: AA-associated and non-AA-associated ("Other").

different between AA-associated and non-AA-associated MNU (Fig. 3A). Mutational burdens in the AA-associated MNU ranged broadly, with the median (2.2 mutations per Mb) being higher than that in breast cancer (0.9 mutations per Mb) and kidney clear cell carcinoma (1.5 mutations per Mb). The mutational burden of non-AA-associated MNU was more than one order of magnitude lower than that of AA-associated MNU (Fig. 3A). P65U stood out among non-AA-associated MNU with an extraordinarily high mutational burden (~62 mutations per Mb) (Fig. 3A). We further explored the mutational signature of this hypermutator and found that it conformed to the COSMIC SBS10a and SBS10b (cosine similarity = 0.86) (Fig. 3B and fig. S7, A and B). The mutational process underlying this signature has been implicated in altered activity of the central DNA polymerase POLE (*35*, *36*). We detected a canonical mutation in the *POLE* proofreading domain (Pro^{286}→Arg) that could lead to this mutational signature in the sample (fig. S7C). To our knowledge, there have been no previous reports of a mutational process associated with mutated *POLE* occurring in normal human tissues.

Next, we characterized mutant clonal expansion in MNU tissues. Overall, the distribution of mutant clone sizes exhibited a long tail (Fig. 3C and table S12). Upon further dividing samples into AA-associated and non-AA-associated MNU, we observed a bimodal distribution of mutant clone sizes in both groups (Fig. 3C). The peak corresponding to larger clone sizes in AA-associated MNU was

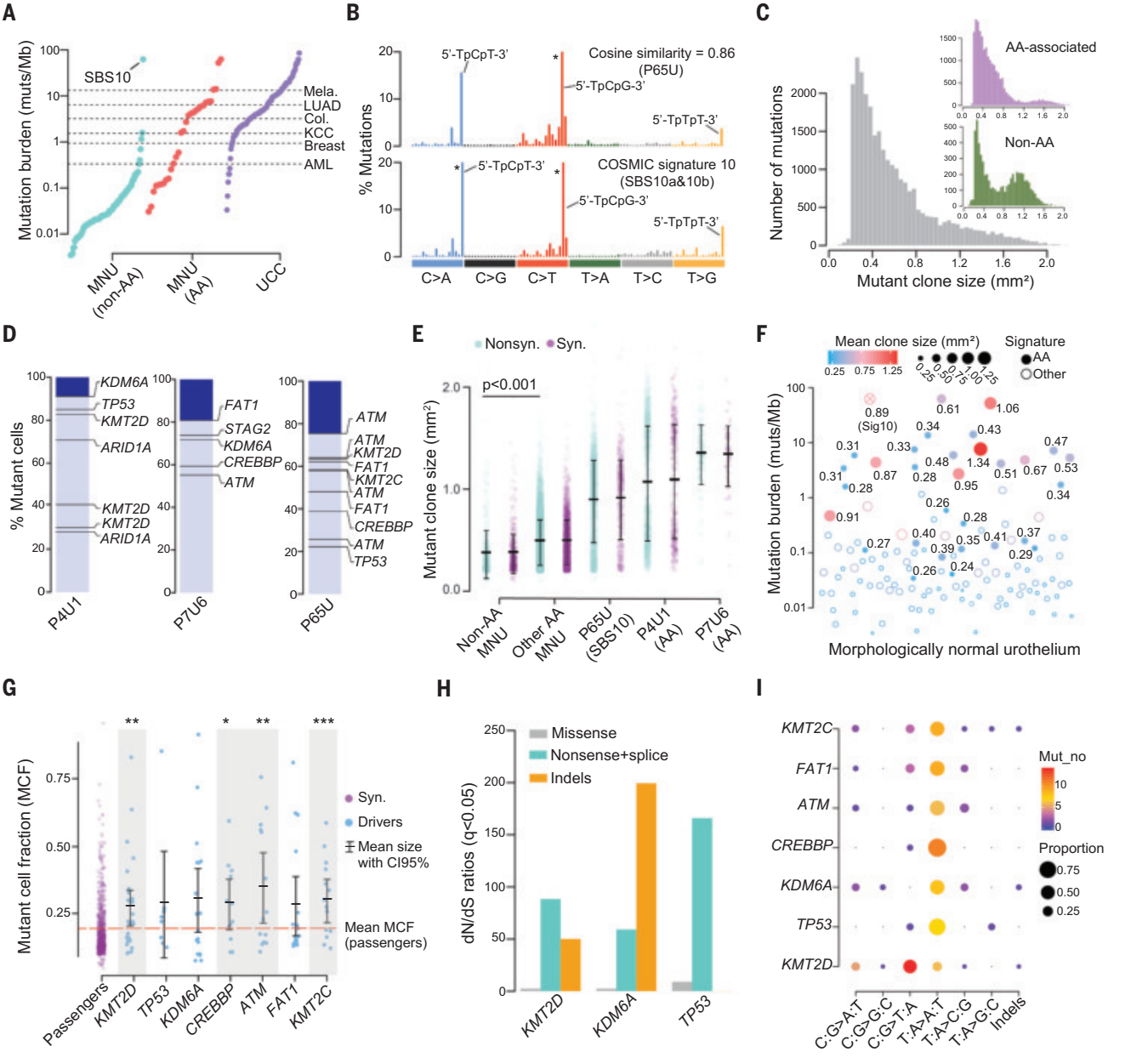


Fig. 3. Mutational burden and mutant clone expansion in MNU.
(**A**) Comparison of mutational burdens (muts/Mb, mutations per megabase) in MNU (both AA-associated and non-AA-associated) and UCC samples. Median mutational burdens of six other cancer types are indicated by dashed lines. AML, acute myeloid leukemia; KCC, kidney clear cell carcinoma; Col., colorectal carcinoma; LUAD, lung adenocarcinoma; Mela., melanoma. SBS10 resembles the COSMIC mutational signature 10 (SBS10a and 10b). (**B**) Comparison of mutational contexts of sample P65U and COSMIC SBS10a and 10b. Representative 3-bp mutational contexts are labeled. (**C**) Distributions of mutant clone sizes of mutations in AA-associated and non-AA-associated MNU samples. The *y* axis represents the number of mutations. (**D**) Bar plots displaying the co-occurrence of driver mutations in the same clones deduced on the basis of the pigeonhole principle. (**E**) Comparison of clone sizes among different MNU samples. Nonsyn., nonsynonymous; Syn., synonymous. (**F**) Mutational burdens and average mutant clone sizes in MNU samples. The average mutant clone sizes of AA-associated samples are labeled. (**G**) Comparison of clone sizes between putative driver and passenger mutations. Wilcoxon rank-sum test was used. $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$. CI95%, 95% confidence interval. (**H**) dN/dS ratios for the three genes under significant positive selection in MNU. (**I**) Mutational spectra of putative driver mutations. Mut_no, mutation number.

due mainly to two samples, P4U1 and P7U6, while in the non-AA-associated distribution, the peak corresponding to larger clone sizes was almost exclusively due to sample P65U. This observation demonstrates that drastic clonal expansions have occurred in some individual MNU tissues. One possible explanation for this observation is that urothelial cells from the same clone have acquired more than one driver mutation, which confers proliferative and competing advantages. To test this hypothesis, we deduced the co-occurrences of driver mutations in those clones using the pigeonhole principle (*37*). Indeed, more than one driver mutation was nested in single clones in each of the above three samples (P4U1, P7U6, and P65U) (Fig. 3D). For example, mutations in *KDM6A*, *TP53*, *KMT2D*, and *ARID1A* were simultaneously acquired by a single clone in P4U1 (Fig. 3D). Compared with non-AA-associated samples, AA-associated MNU had a significantly

larger median mutant clone size ($P < 0.001$, Wilcoxon rank-sum test), even when we excluded the three samples with drastic clonal expansions (Fig. 3E). Additionally, we found that AA-associated MNU displayed greater mutational burdens (Fig. 3, A and F). Altogether, these data indicate that AA exposure considerably accelerates somatic mutation accumulation and enhances clonal expansions in normal urothelium.

Positive selection of somatic mutations provides the necessary fuel for clonal expansion. We compared the clone size of driver mutations with those of synonymous mutations in non-AA-associated MNU, which were considered as passengers under neutral selection (table S13). As expected, clones with driver mutations were larger than those with passengers (Fig. 3G). However, statistically significant differences were observed only in *KMT2D*, *CREBBP*, *ATM*, and *KMT2C*, not in canonical driver genes such as *TP53* (Fig. 3G, $P < 0.05$, Wilcoxon rank-sum test). This observation is similar to previous findings and is likely attributable to putative passengers co-occurring with driver mutations in individual clones being hijacked by positive clonal selection (*8*). We next estimated genes under positive selection using a context-dependent dN/dS model (dN/dS is the ratio of the rate of substitution at nonsilent sites versus silent sites) (*38*). Genes with a significant global dN/dS ratio included *KMT2D*, *KDM6A*, and *TP53*, which are the top three recurrently mutated genes in UCC (Fig. 3H and table S14).

A predominance of T:A>A:T transversions was observed in most mutations in driver genes, suggesting that AA mutagenesis in MNU can explain the occurrence of most driver mutations observed in the current study (Fig. 3I). This was further confirmed by analyzing the probability of each mutational signature underlying the driver mutations (table S15). This finding rationalizes our observation that AA exposure largely boosts mutant clone sizes in MNU (Fig. 3, E and F). Unexpectedly, mutations in *KMT2D* were dominant for C:G>T:A transitions rather than T:A>A:T transversions (Fig. 3I). Nearly half of the *KMT2D* mutations (13/28) occurred in non-AA-associated MNU. Even in AA-associated MNU, ~60% of the *KMT2D* mutations were not T:A>A:T transversions, which contrasted greatly with other driver genes (fig. S7, D to F), although these *KMT2D* mutations were still most likely caused by AA mutagenesis (table S15). This finding implies that *KMT2D* mutations may be widely carried by urothelial cells through both intrinsic (e.g., SBS2, SBS13, and SBS5) and exogenously triggered mutational processes (e.g., SBS22). Mutations in this gene may be essential for clonal expansion in urothelial cells, regardless of whether they experience exogenous mutagen exposure.

Competitive mutant clones evolve independently under AA exposure

We sequenced two tumors and three MNU samples from the ureter tract of patient P4. Somatic mutations harbored by the five samples were different from each other, indicating that they evolved independently, as reflected by the phylogenetic tree (Fig. 4A). All five samples displayed clear AA-associated mutational signatures (fig. S8A). Given the different sampling sites, we concluded that AA-triggered mutational processes can spread throughout the entire ureter tract (Fig. 4A). Similar results were also observed in other patients (fig. S8, B and C). Forming competing clones in the ureter tract, each sample independently accumulated driver mutations that were most likely triggered by AA mutagenesis (table S15). These driver mutations may confer competitive advantages on these clones. Among the three MNU samples, we observed putative driver mutations convergent in *KDM6A*, suggesting that mutations in this gene were widespread in MNU and important for early mutant clonal evolution (Fig. 4A). Interestingly, in P4U1, we identified an obvious bimodal distribution of MCFs of somatic mutations (Fig. 4B). The smaller peak, with driver mutations in *KMT2D* and *ARID1A*, was estimated to be a subclone originating from the major clone (the larger peak) on the basis of the pigeonhole principle. Given that most driver mutations in P4U1 were caused by AA mutagenesis (table S15), this observation suggests that mutant subclones originate and evolve in MNU under AA mutagenic stress.

A single AA-associated clone in MNU can expand to a scale of several square centimeters in size

We have demonstrated that AA mutagenesis drives mutant clonal expansion in MNU. However, to what scale an AA-associated mutant clone can expand remains to be elucidated. We sequenced six MNU samples from patient P7 that were extracted from different sites in the bladder (fig. S9A). The mutational burdens of these samples ranged from 1.7 to 7.6 mutations per Mb (fig. S9B). Somatic mutations detected in these samples were entirely different from those in the patient's tumor, indicating their independent clonal origins (fig. S9C). Additionally, we found that the mutational spectra largely matched the AA-associated signature, which was ubiquitous among the six samples (Fig. 4C).

Next, we compared somatic mutations and their MCFs among the six MNU samples from patient P7. Except for P7U3, the other five MNU samples shared 161 somatic mutations with variable MCFs (Fig. 4D), demonstrating that those five MNU samples may originate from the large expansion of a single mutant clone. In this case, we observed five independent *KDM6A* mutations, further demonstrating that mutations in this gene are widely carried by urothelial cells (Fig. 4D). We next clustered somatic mutations into mutant clones using a Dirichlet process (fig. S10). A single mutant clone with putative driver mutations in *FAT1* and *ATM* was shared by the five MNU samples. This shared clone seemingly derived from a small clone in P7U1 which progressively evolved and acquired additional driver mutations in *CREBBP*, *KDM6A*, and *STAG2* (Fig. 4D and fig. S10A). Another possibility is that an independent mutant clone in P7U1 intermingled with the large clone, which could explain the low MCFs of the shared mutations in P7U1 (Fig. 4D and fig. S10A). A competing mutant clone in P7U3 originated independently and evolved in parallel, acquiring two mutations in *KDM6A* and one in *EGFR* (Fig. 4D and fig. S10A). Taken together and combined with the sampling distances, our results reveal that a single AA-associated mutant clone can expand massively to a scale of several square centimeters in size (Fig. 4E). We observed similar results in another patient sample (fig. S11).

Discussion

Accumulation of mutations in somatic cells has long been implicated in various pathological processes, including human cancer (*2*). However, how and in what patterns somatic mutations occur and drive clonal expansion in normal cells remain largely uncharacterized. Recent genome sequencing studies have revealed landscapes of somatic mutations in various normal tissues, thus broadening our knowledge of mutagenesis in somatic cells (*3*, *8*–*12*, *14*–*17*). Although previous studies identified some mutations with low allele frequencies in normal-appearing urothelium in a limited number of bladder cancer patients (*39*, *40*), our study depicts a comprehensive mutational landscape of human normal urothelium from UCC patients, especially under exogeneous mutagen exposure, and a study by Lawson *et al.* published in this issue investigated somatic mutations in normal bladder urothelium mostly from cancer-free individuals (*41*).

Overall, we observed variable numbers of somatic mutations in MNU using a relatively large sampling size and moderate sequencing depth. We found that macroscopic mutant clones originated in at least some MNU tissues. Acquisition of putative driver mutations in MNU may explain why some mutant clones can expand to a detectable size. Our mutational signature analysis revealed an underlying prevalence of AA mutagenesis in MNU, demonstrating that the mutational process triggered by AA can occur widely in normal human somatic cells in vivo. As AA is prevalent in traditional herbal medicine in Asia (*42*–*45*), our results may reflect the specificity of Chinese and Asian populations. AA exposure boosts somatic mutation accumulation and clonal

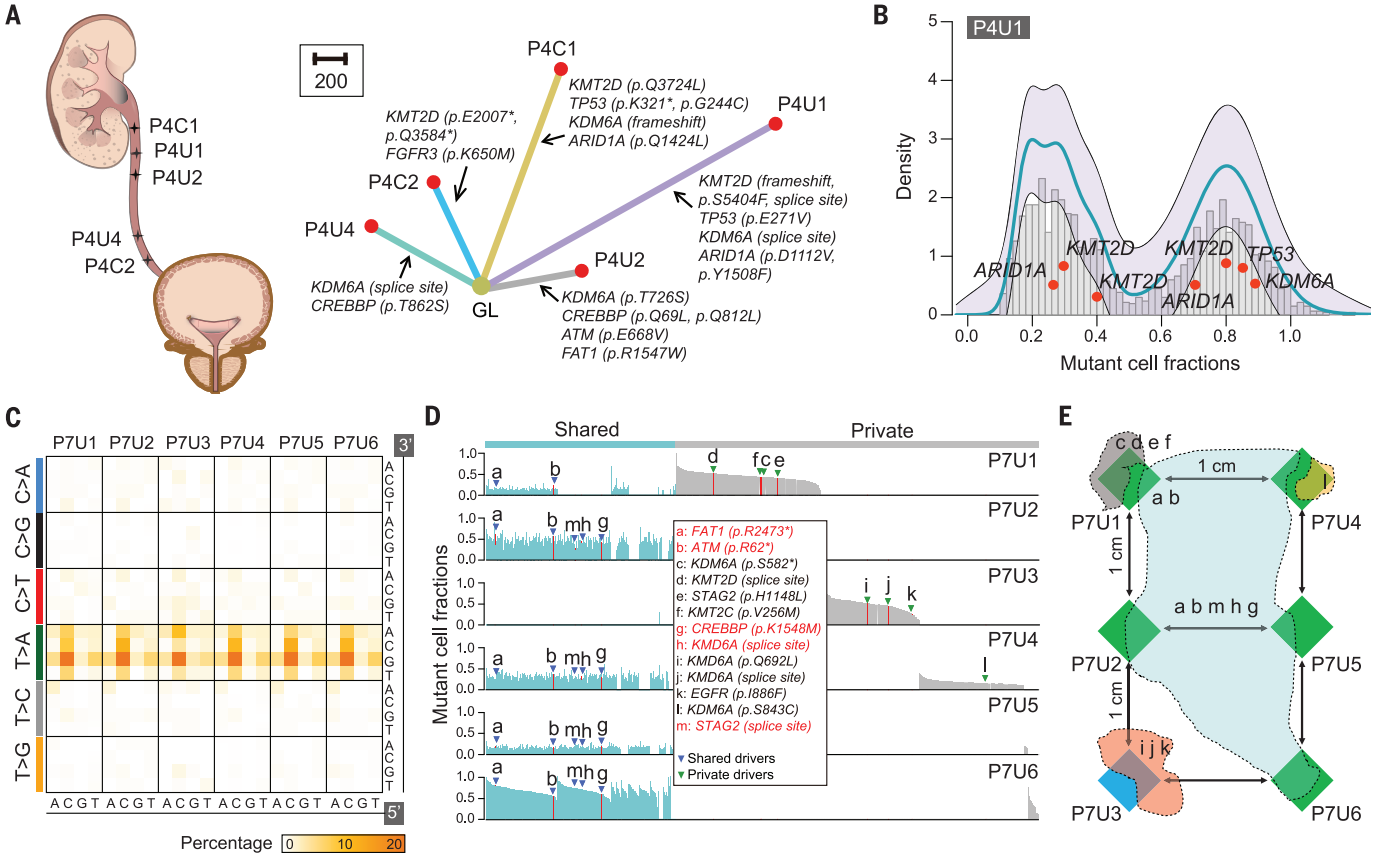


Fig. 4. Mutant clonal evolution in MNU exposed to AA. (**A**) (Left) Sampling sites in patient P4. (Right) Phylogenetic tree depicting the clonal relationships of five samples from P4. Putative driver mutations are labeled on the branches. Asterisks indicate stop codons. Scale bar, 200 mutations. (**B**) Clustering of clone sizes in P4U1 using a Dirichlet process. The blue line indicates the fitted distribution, and the purple region represents the 95% posterior confidence interval. Red dots indicate mutations in putative driver genes. (**C**) Heatmaps showing the proportion of each mutation type within 96 mutational contexts. (**D**) MCFs of somatic mutations in the six MNU samples. Putative driver mutations are highlighted by red vertical lines. (**E**) Schematic displaying the area of different mutant clones, as inferred by mutation overlapping and clustering results. Lowercase letters indicating driver mutations correspond to the table in (D). The differently colored regions represent different mutant clones.

expansions in MNU; the latter most likely because the mutagenic effects of AA widely cause driver mutations in MNU, facilitating clonal expansion through positive clonal selection. Unexpectedly, we identified drastic expansion of single AA-associated clones to a scale of several square centimeters in size. However, nailing down a precise size of AA-implicated mutagenesis and AA-triggered clonal expansion in urothelium requires further investigation using denser sampling. A previous study used *N*-butyl-*N*-(4-hydroxybutyl)nitrosamine (BBN) to trigger bladder carcinoma in mice and found that a single basal stem cell could proliferate and colonize the entire urothelium (*46*). Here, we reported multiple independent clones originating in human urothelium under AA exposure. These differences in findings could be explained by the fact that tumorigenesis in humans is considerably more complicated than in mice or by the different degrees of mutagenic stress caused by BBN versus AA.

Chromatin remodeling–related genes, such as *KMT2D* and *KDM6A*, were frequently mutated in MNU, whereas canonical driver genes in UCC, such as *PIK3CA* and *FGFR3*, were rarely mutated. This suggests that epigenetic dysfunction may be critical for early clonal expansion in human urothelium. *TP53* had a clearly low mutation rate. A total of five MNU samples (independent clonal origin) harbored *TP53* mutations, four of which were AA-associated and one of which was the hypermutator P65U (COSMIC SBS10). This finding suggests that either (i) *TP53* is rarely mutated in early clonal evolution in MNU unless there are certain strong mutagenic driving forces or (ii) mutant clones with *TP53* mutations are too small to be detected. The latter possibility can be examined using ultra-sensitive and deep sequencing strategies (*47*). Although the overwhelming mutagenicity of AA caused most of the driver mutations in our study, *KMT2D* mutations were widely observed in other MNU lacking the AA signature, which suggests that mutations in this gene widely occur in urothelial cells, regardless of whether they experience exogenous mutagen exposure.

Unlike somatic mutations, CNAs are relatively rare and less extensive in MNU than in tumors. Even in those MNU with obvious clonal expansion (such as P65U), copy number remained diploid across the entire genome. Similar results were observed in previous studies of normal skin and esophageal tissues (*8*, *9*). These findings may reveal a universal principle in somatic clone expansion, namely that genomic instability is fundamental for malignant transformation.

One possible clinical implication of our study is that more radical treatment strategies may be appropriate in AA-associated UCC patients because their normal-appearing urothelium may harbor high mutational burdens and have undergone mutant clonal expansions. Such conditions may substantially contribute to tumor relapse.

REFERENCES AND NOTES

1. I. Martincorena, P. J. Campbell, *Science* **349**, 1483–1489 (2015).
2. R. A. Risques, S. R. Kennedy, *PLOS Genet.* **14**, e1007108 (2018).

3. I. Martincorena *et al.*, *Science* **362**, 911–917 (2018).
4. C. Tomasetti, B. Vogelstein, G. Parmigiani, *Proc. Natl. Acad. Sci. U.S.A.* **110**, 1999–2004 (2013).
5. M. R. Stratton, P. J. Campbell, P. A. Futreal, *Nature* **458**, 719–724 (2009).
6. M. Greaves, C. C. Maley, *Nature* **481**, 306–313 (2012).
7. K. A. Hoadley *et al.*, *Cell* **173**, 291–304.e6 (2018).
8. I. Martincorena *et al.*, *Science* **348**, 880–886 (2015).
9. A. Yokoyama *et al.*, *Nature* **565**, 312–317 (2019).
10. H. Lee-Six *et al.*, *Nature* **574**, 532–537 (2019).
11. S. F. Brunner *et al.*, *Nature* **574**, 538–542 (2019).
12. L. Moore *et al.*, *Nature* **580**, 640–646 (2020).
13. K. Yoshida *et al.*, *Nature* **578**, 266–272 (2020).
14. T. Bae *et al.*, *Science* **359**, 550–555 (2018).
15. M. A. Lodato *et al.*, *Science* **359**, 555–559 (2018).
16. Y. S. Ju *et al.*, *Nature* **543**, 714–718 (2017).
17. H. Lee-Six *et al.*, *Nature* **561**, 473–478 (2018).
18. K. Yizhak *et al.*, *Science* **364**, eaaw0726 (2019).
19. P. E. García-Nieto, A. J. Morrison, H. B. Fraser, *Genome Biol.* **20**, 298 (2019).
20. D. E. Brash *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **88**, 10124–10128 (1991).
21. A. Ziegler *et al.*, *Nature* **372**, 773–776 (1994).
22. S. C. Baker, J. Southgate, in *Electrospinning for Tissue Regeneration*, L. A. Bosworth, S. Downes, Eds. (Woodhead Publishing Series in Biomaterials, Woodhead Publishing, 2011), pp. 225–241.
23. C. Varley *et al.*, *Exp. Cell Res.* **306**, 216–229 (2005).
24. G. Guo *et al.*, *Nat. Genet.* **45**, 1459–1463 (2013).
25. A. G. Robertson *et al.*, *Cell* **171**, 540–556.e25 (2017).
26. Cancer Genome Atlas Research Network, *Nature* **507**, 315–322 (2014).
27. J. Michl, M. J. Ingrouille, M. S. Simmonds, M. Heinrich, *Nat. Prod. Rep.* **31**, 676–693 (2014).
28. U. Mengs, *Arch. Toxicol.* **61**, 504–505 (1988).
29. A. W. T. Ng *et al.*, *Sci. Transl. Med.* **9**, eaan6446 (2017).
30. Y. Du *et al.*, *Eur. Urol.* **71**, 841–843 (2017).
31. S. L. Poon *et al.*, *Genome Med.* **7**, 38 (2015).
32. L. B. Alexandrov *et al.*, *Nature* **500**, 415–421 (2013).
33. L. B. Alexandrov *et al.*, *Nature* **578**, 94–101 (2020).
34. C. H. Chen *et al.*, *Int. J. Cancer* **133**, 14–20 (2013).
35. Cancer Genome Atlas Network, *Nature* **487**, 330–337 (2012).
36. C. Kandoth *et al.*, *Nature* **497**, 67–73 (2013).
37. S. Nik-Zainal *et al.*, *Cell* **149**, 994–1007 (2012).
38. I. Martincorena *et al.*, *Cell* **171**, 1029–1041.e21 (2017).
39. M. B. H. Thomsen *et al.*, *Sci. Rep.* **7**, 11702 (2017).
40. T. Majewski *et al.*, *Cell Rep.* **26**, 2241–2256.e4 (2019).
41. A. R. J. Lawson *et al.*, *Science* **370**, 75–82 (2020).
42. L. Vaclavik, A. J. Krynitsky, J. I. Rader, *Food Addit. Contam. Part A Chem. Anal. Control Expo. Risk Assess.* **31**, 784–791 (2014).
43. H. Y. Yang, P. C. Chen, J. D. Wang, *BioMed Res. Int.* **2014**, 569325 (2014).
44. T. P. Cheung, C. Xue, K. Leung, K. Chan, C. G. Li, *Clin. Toxicol.* **44**, 371–378 (2006).
45. M. N. Lai, S. M. Wang, P. C. Chen, Y. Y. Chen, J. D. Wang, *J. Natl. Cancer Inst.* **102**, 179–186 (2010).
46. K. Shin *et al.*, *Nat. Cell Biol.* **16**, 469–478 (2014).
47. J. J. Salk *et al.*, *Cell Rep.* **28**, 132–144.e3 (2019).
48. R. Li *et al.*, supplementary of UTUC-Normal, Version V1, Zenodo (2020); https://doi.org/10.5281/zenodo.3966801.
49. R. Li *et al.*, supplementary of UTUC-Normal 2, Version V1, Zenodo (2020); https://doi.org/10.5281/zenodo.3726413.

ACKNOWLEDGMENTS

We thank all the patients for their consent and participation in this study. We thank X. M. Wang from Peking University for the helpful discussion on mutational signature analysis. We thank T. S. Sun from Peking University for helping create the bladder cartoon. **Funding:** This work was financially supported by the National Key Research and Development Program (2018YFA0902802), the National Science and Technology Major Project (2018ZX10302205 and 2019YFC1315702), the National Natural Science Foundation of China (31722003, 31770925, and 81802533), the Guangdong Province Key Research and Development Program (2019B020226002), and the Beijing Municipal Science and Technology Commission (Z191100006619010). R.L. was funded by the BoYa Postdoctoral Fellowship of Peking University. **Author contributions:** F.B., T.X., R.L., Y.D., Z.C., and D.X. designed the experiments. R.L. led the data analysis with help from Z.C. and Z.L. Y.D., T.L., and X.C. collected the samples with help from G.W. Y.D., M.L., and G.W. performed pathological examinations. Z.C. and D.X. performed experiments with help from S.J. and Y.D. R.L. and F.B. wrote the manuscript with help from Y.D. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** The sequencing raw data generated in this study have been deposited in the Genome Sequence Archive (GSA) in BIG Data Center (https://bigd.big.ac.cn), Beijing Institute of Genomics (BIG), Chinese Academy of Sciences, under accession number HRA000137 (https://bigd.big.ac.cn/gsa-human/s/A88e89NG). Data from the Chinese bladder cancer study that were used in the MutSigCV analysis are deposited in the Sequence Read Archive (SRA) (www.ncbi.nlm.nih.gov/sra) under accession number SRA063495 (*24*). The input data for Sequenza, ReCapSeq, ABSOLUTE, mutational signature analysis, Dirichlet process analysis, and VCF files of all somatic mutations; the raw outputs from Sequenza, ABSOLUTE, and Dirichlet process analysis; and the code for Dirichlet process analysis are all available at Zenodo (*48*, *49*).

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/370/6512/82/suppl/DC1
Materials and Methods
Figs. S1 to S11
Tables S1 to S15
References (*50*–*65*)
MDAR Reproducibility Checklist

30 December 2019; accepted 4 August 2020
10.1126/science.aba7300

CORONAVIRUS

Selective and cross-reactive SARS-CoV-2 T cell epitopes in unexposed humans

Jose Mateus[1], Alba Grifoni[1], Alison Tarke[1], John Sidney[1], Sydney I. Ramirez[1,3], Jennifer M. Dan[1,3], Zoe C. Burger[3], Stephen A. Rawlings[3], Davey M. Smith[3], Elizabeth Phillips[2], Simon Mallal[2], Marshall Lammers[1], Paul Rubiro[1], Lorenzo Quiambao[1], Aaron Sutherland[1], Esther Dawen Yu[1], Ricardo da Silva Antunes[1], Jason Greenbaum[1], April Frazier[1], Alena J. Markmann[4], Lakshmanane Premkumar[5], Aravinda de Silva[5], Bjoern Peters[1,3], Shane Crotty[1,3], Alessandro Sette[1,3]*†, Daniela Weiskopf[1]*†

Many unknowns exist about human immune responses to the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) virus. SARS-CoV-2–reactive CD4$^+$ T cells have been reported in unexposed individuals, suggesting preexisting cross-reactive T cell memory in 20 to 50% of people. However, the source of those T cells has been speculative. Using human blood samples derived before the SARS-CoV-2 virus was discovered in 2019, we mapped 142 T cell epitopes across the SARS-CoV-2 genome to facilitate precise interrogation of the SARS-CoV-2–specific CD4$^+$ T cell repertoire. We demonstrate a range of preexisting memory CD4$^+$ T cells that are cross-reactive with comparable affinity to SARS-CoV-2 and the common cold coronaviruses human coronavirus (HCoV)-OC43, HCoV-229E, HCoV-NL63, and HCoV-HKU1. Thus, variegated T cell memory to coronaviruses that cause the common cold may underlie at least some of the extensive heterogeneity observed in coronavirus disease 2019 (COVID-19) disease.

The emergence of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) in late 2019 and its subsequent global spread has led to millions of infections and substantial morbidity and mortality (*1*). Coronavirus disease 2019 (COVID-19), the clinical disease caused by SARS-CoV-2 infection, can range from mild, self-limiting disease to acute respiratory distress syndrome and death (*2*). The mechanisms underlying the spectrum of COVID-19 disease severity states and the nature of protective immunity against COVID-19 remain unclear.

[1]Center for Infectious Disease and Vaccine Research, La Jolla Institute for Immunology, La Jolla, CA 92037, USA. [2]Institute for Immunology and Infectious Diseases, Murdoch University, Perth, WA 6150, Australia. [3]Department of Medicine, Division of Infectious Diseases and Global Public Health, University of California, San Diego, La Jolla, CA 92037, USA. [4]Department of Medicine, Division of Infectious Diseases, University of North Carolina School of Medicine, Chapel Hill, NC 27599, USA. [5]Department of Microbiology and Immunology, University of North Carolina School of Medicine, Chapel Hill, NC 27599, USA.
***Corresponding author. Email: alex@lji.org (A.S.); daniela@lji.org (D.W.)** †These authors contributed equally to this work.

Studies investigating the human immune response against SARS-CoV-2 have begun to characterize SARS-CoV-2 antigen-specific T cell responses (*3*–*8*), and multiple studies have described marked activation of T cell subsets in acute COVID-19 patients (*9*–*13*). Unexpectedly, antigen-specific T cell studies performed with five different cohorts reported that 20 to 50% of people who had not been exposed to SARS-CoV-2 had significant T cell reactivity directed against peptides corresponding to SARS-CoV-2 sequences (*3*–*7*). The studies were from geographically diverse cohorts (the United States, the Netherlands, Germany, Singapore, and the United Kingdom), and the general pattern observed was that the T cell reactivity found in unexposed individuals was predominantly mediated by CD4$^+$ T cells. It was speculated that this phenomenon might be due to preexisting memory responses against human "common cold" coronaviruses (HCoVs) such as HCoV-OC43, HCoV-HKU1, HCoV-NL63, and HCoV-229E. These HCoVs share partial sequence homology with SARS-CoV-2, are

widely circulating in the general population, and are typically responsible for mild respiratory symptoms (*14*–*16*). However, the hypothesis of cross-reactive immunity between SARS-CoV-2 and common cold HCoVs still awaits experimental trials. This potential preexisting cross-reactive T cell immunity to SARS-CoV-2 has broad implications because it could explain aspects of differential COVID-19 clinical outcomes, influence epidemiological models of herd immunity (*17*, *18*), or affect the performance of COVID-19 candidate vaccines.

Epitope repertoire in SARS-CoV-2–unexposed individuals

To define the repertoire of CD4$^+$ T cells recognizing SARS-CoV-2 epitopes in previously unexposed individuals, we used in vitro stimulation of peripheral blood mononuclear cells (PBMCs) for 2 weeks with pools of 15-mer peptides. This method is known to be robust for detecting low-frequency T cell responses to allergens and bacterial or viral antigens (*19*, *20*), including naive T cells (*21*). For screening SARS-CoV-2 epitopes, we used PBMC samples from unexposed subjects collected between March 2015 and March 2018, well before the global circulation of SARS-CoV-2 occurred. The unexposed subjects were confirmed to be seronegative for SARS-CoV-2 (fig. S1A).

SARS-CoV-2–reactive T cells were expanded, with one pool of peptides spanning the entire sequence of the spike protein (CD4-S) and the other a nonspike "megapool" (CD4-R) of predicted epitopes from the nonspike regions (i.e., "remainder") of the viral genome (*4*). In total, 474 15-mer SARS-CoV-2 peptides were screened. After 14 days of stimulation, T cell reactivity against intermediate "mesopools," each encompassing ~10 peptides, was assayed using a FluoroSPOT assay (e.g., 22 CD4-R mesopools; fig. S2A). Positive mesopools were further deconvoluted to identify specific individual SARS-CoV-2 epitopes. Representative results from one donor show the deconvolution of mesopools P6 and P18 to identify seven different SARS-CoV-2 epitopes (fig. S2B). Intracellular cytokine-staining assays specific for interferon γ (IFN-γ) determined whether antigen-specific T cells responding to the SARS-CoV-2 mesopools were CD4$^+$ or CD8$^+$ T cells (fig. S2C). Results from the 44 donors/CD4-R mesopool and 40 donors/CD4-S mesopool combinations yielding a positive response are shown in fig. S2, D and E, respectively. In 82/88 cases (93.2%), the cells responding to SARS-CoV-2 mesopool stimulation were clearly CD4$^+$ T cells, as judged by the ratio of CD4/CD8–responding cells; in four cases (4.5%), the responding cells were CD8$^+$ T cells; and in two cases (2.3%), the responses were mediated by both CD4$^+$ and CD8$^+$ T cells. The fact that CD8$^+$ T cells were rarely detected was not surprising because the peptides used in CD4-R encompassed predicted class II epitopes and the CD4-S is composed of 15-mer peptides (9- to 10-mer peptides are optimal for CD8$^+$ T cells). Furthermore, the 2-week restimulation protocol was originally designed to expand CD4$^+$ T cells (*20*). Overall, these results indicated that the peptide-screening strategy used mapped SARS-CoV-2 epitopes recognized by CD4$^+$ T cells in unexposed individuals.

A total of 142 SARS-CoV-2 epitopes were identified, 66 from the spike protein (CD4-S) and 76 from the remainder of the genome (CD4-R) (table S1). For each combination of epitope and responding donor, potential human leukocyte antigen (HLA) restrictions were inferred on the basis of the predicted HLA-binding capacity of the particular epitope for the specific HLA alleles present in the responding donor (*22*). Each donor recognized an average of 11.4 epitopes (range 1 to 33, median 6.5; fig. S3A). Forty of the 142 epitopes were recognized by two or more donors (fig. S3B), accounting for 55% of the total response (fig. S3C). These 142 mapped SARS-CoV-2 epitopes may prove

A

SARS-CoV-2 proteins	Protein length	Peptides tested	Epitopes identified		Positive responses		Response magnitude	
	aa	*n*	*n*	%	*n*	%	SFCs 10^6/ PBMCs	%
M	222	8	0	0	0	0	0	0
N	419	7	4	3	4	2	6573	2
S (Non-RBD)	1050	207	54	38	90	44	149,300	39
S (RBD)	223	46	12	8	23	11	40,873	11
E	75	8	2	1	2	1	2240	1
ORF3a	275	10	6	4	7	3	17,587	5
ORF7a	121	4	3	2	3	1	4080	1
ORF6	61	6	6	4	7	3	16,453	4
ORF8	121	3	3	2	4	2	13,553	4
nsp1	180	2	1	1	1	0	387	0
nsp2	638	14	6	4	7	3	32,673	9
nsp3	1945	34	10	7	11	5	7307	2
nsp4	500	22	8	6	10	5	5320	1
nsp5	306	6	2	1	4	2	52,907	14
nsp6	290	18	6	4	9	4	5613	1
nsp7	83	4	0	0	0	0	0	0
nsp8	198	5	2	1	2	1	6473	2
nsp9	113	2	0	0	0	0	0	0
nsp10	139	2	1	1	2	1	440	0
nsp12	932	19	7	5	9	4	6850	2
nsp13	601	14	3	2	4	2	5360	1
nsp14	527	19	3	2	4	2	2193	1
nsp15	346	4	1	1	1	0	533	0
nsp16	298	9	2	1	2	1	1853	0
ORF10	38	1	0	0	0	0	0	0
	Total	474	142	100	206	100	378,570	100

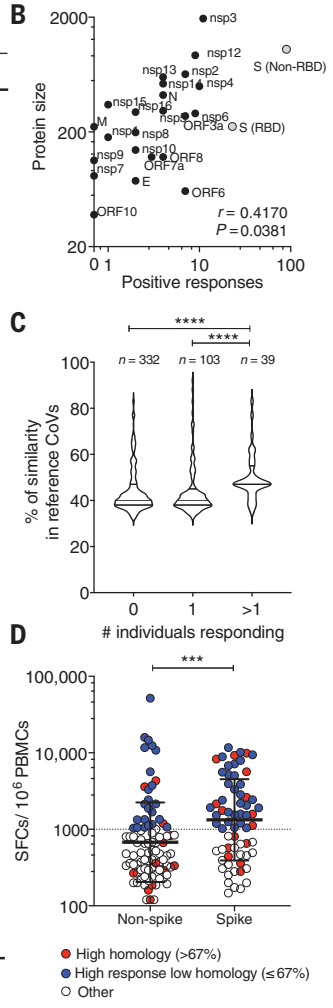


Fig. 1. Characteristics of SARS-CoV-2 epitopes identified in unexposed donors. Reactivity was determined by FluoroSPOT assay after 17 days of in vitro stimulation of unexposed donor PBMCs (*n* = 18) with one pool of peptides spanning the entire sequence of the spike protein (CD4-S) or a nonspike "megapool" (CD4-R) of predicted epitopes from the nonspike (i.e., "remainder") regions of the viral genome. (**A**) Summary of the responses as a function of the protein of origin. (**B**) Spearman correlation of positive responses per SARS-CoV-2 protein size. (**C**) Percent similarity of the identified epitopes with common cold coronavirus peptides as a function of the number of responding donors. (**D**) Each dot shows the reactivity of a donor-epitope combination derived from either nonspike (CD4-R) or spike (CD4-S) protein. Black bars indicate the geometric mean and geometric SD. Red indicates donor-epitope combinations with sequence identity >67% with common cold coronaviruses, and blue indicates highly reactive donor-epitope combinations (>1000 SFCs*10^6) with sequence identity ≤67%. In (C) and (D), statistical comparisons were performed with a two-tailed Mann–Whitney test. ***P < 0.001, ****P < 0.0001.

useful in future studies as reagents for tracking CD4+ T cells in SARS-CoV-2–infected individuals and in COVID-19 vaccine trials.

Epitope distribution by ORF of origin

Although a broad range of different SARS-CoV-2 antigens were recognized, several of the epitopes yielding the most frequent (i.e., recognized in multiple donors) or most vigorous [i.e., the most spot-forming cells (SFCs)/10^6 cells] responses were derived from the SARS-CoV-2 spike antigen (table S1). We therefore assessed the overall distribution of the 142 T cell epitopes mapped among all SARS-CoV-2 proteins compared with the relative size of each SARS-CoV-2 antigen (Fig. 1, A and B). Fifty-four percent of the total positive response was associated with spike-derived epitopes [Fig. 1A; 11% for receptor-binding domain (RBD), and 44% for the non-RBD portion of spike]. Of relevance for COVID-19 vaccine development, only 20% of the spike responses were derived from the RBD region (Fig. 1A; comparing 11 versus 44%, as described above), and the RBD region accounted for only 11% of the overall CD4+ T cell reactivity (Fig. 1A). Mapped epitopes were fairly evenly distributed across the SARS-CoV-2 genome in proportion to the size of each protein (Fig. 1B; P = 0.038, r = 0.42). In addition to the strong responses directed to spike, responses were also seen for open reading frame 6 (ORF6), ORF3a, N, ORF8, and within Orf1a/b, where nsp3, nsp12, nsp4, nsp6, nsp2, and nsp14 were more prominently recognized. These mapped epitope results at the ORFeome level partially overlap with the ORFs targeted by CD4+ T cells in COVID-19 cases (*4*). No epitopes derived from the membrane protein (M) were identified in unexposed individuals (Fig. 1B), but M is robustly recognized by SARS-CoV-2–specific CD4+ T cell responses in COVID-19 cases (*4*). The lack of quality class II epitopes in M was unsurprising based on M molecular biology: M is a small protein with three transmembrane domains. Combined, the data indicate that class II epitopes are relatively broadly available across the SARS-CoV-2 genome but that SARS-CoV-2 memory CD4+ T cells preferentially target proteins highly expressed during infection, as exemplified by M and S (spike) epitope-mapping results.

Sequence homology of the identified SARS-CoV-2 epitopes to other common HCoVs

When this epitope-mapping study was initiated, an assumption was that the in vitro T cell culture epitope mapping would reveal an epitope repertoire associated with de novo generation of responses from naïve T cells. However, while these epitope-mapping studies were in progress, we and others detected significant ex vivo reactivity against bulk pools of SARS-CoV-2 peptides (*3*–*7*) and speculated that this might reflect the presence of memory T cells cross-reactive between HCoVs and SARS-CoV-2. These other HCoVs circulate widely in human populations and are typically responsible for mild, usually undiagnosed, respiratory illnesses

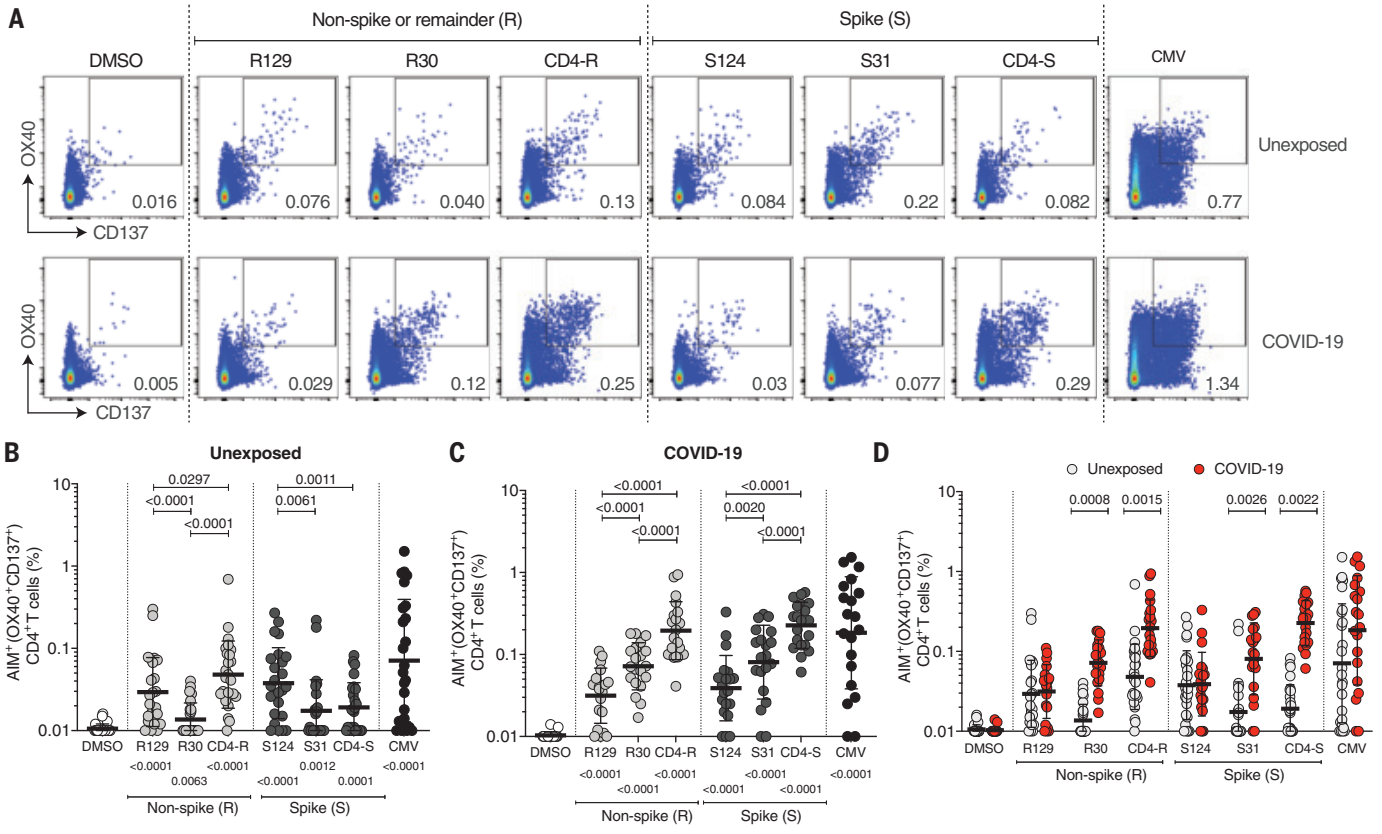


Fig. 2. CD4+ T cells in SARS-CoV-2–unexposed and recovered COVID-19 patients against HCoV epitopes homologous to SARS-CoV-2 epitopes. (**A**) Example of flow cytometry gating strategy for antigen-specific CD4+ T cells based on activation-induced marker assays (OX40+ and CD137+ double expression) after stimulation of PBMCs with HCoV or SARS-CoV-2 peptides. (**B** to **D**) Antigen-specific CD4+ T cells measured as the percentage of activation-induced marker assay–positive (OX40+CD137+) CD4+ T cells after stimulation of PBMCs with HCoV epitopes homologous to SARS-CoV-2 epitopes. Samples were derived from SARS-CoV-2–unexposed donors (n = 25) and recovered COVID-19 patients (n = 20). Black bars indicate the geometric mean and geometric SD. Each dot is representative of an individual subject. Statistical pairwise comparisons [(B) and (C)] were performed with the Wilcoxon test. P values related to comparisons with the DMSO controls are listed at the bottom of the graphs, and any significant P values related to intergroup comparisons are listed on top of the graphs. Statistical comparisons across cohorts were performed with the Mann–Whitney test (D). See also figs. S5 and S6.

such as the common cold (*14*–*16*). However, there is currently a lack of experimental data addressing whether memory CD4$^+$ T cells that are cross-reactive between SARS-CoV-2 and other HCoVs do indeed exist.

We therefore next determined the degree of homology for all four widely circulating HCoVs for all 142 SARS-CoV-2 epitopes identified herein. For the analysis, we split the peptides into three groups based on immunogenicity as follows: (i) never immunogenic, (ii) immunogenic in one individual, or (iii) immunogenic in two or more individuals (Fig. 1C). There was significantly higher sequence similarity in peptides recognized by more than one individual compared with peptides recognized by a single individual or not recognized at all ($P < 0.0001$, two-tailed Mann–Whitney test). Additionally, almost all donors from the unexposed cohort used for the epitope screen were seropositive for three widely circulating HCoVs (HCoV-NL*63*, HCoV-OC42, and HCoV-HKU1) (fig. S1B). Thus, epitope homology and seropositivity data suggest that T cell cross-reactivity is plausible between SARS-CoV-2 and HCoVs already established in the human population.

To select the epitope subsets to be analyzed in more detail, we plotted the T cell response magnitude of each positive epitope per donor (Fig. 1D). This analysis confirmed the dominance of the spike antigen over the epitopes derived from the remainder of the genome ($P < 0.001$, two-tailed Mann–Whitney test).

Next, we selected two categories of SARS-CoV-2 epitopes of interest. The first category was epitopes with potential cross-reactivity from HCoVs. We initially selected the 67% arbitrary cutoff because we reasoned that a 9-mer is the epitope region involved in binding to class II (*23*) and that one or two residues in addition to the 9-mer core region are often required for optimal recognition (*24*) (Fig. 1D, red). Second, we independently filtered for any epitopes associated with high responses (top ~30%; Fig. 1D, blue). This resulted in the selection of 31 epitopes from spike (six with high homology and 25 for dominant responses) organized in a new CD4-[S31] pool. Similarly, we generated a new CD4-[R30] pool composed of 30 epitopes from the remainder of the genome (nine with high homology and 21 associated with strong responses; Fig. 1D). These epitope pools were then used for further CD4$^+$ T cell studies.

Direct evidence of reactivity to HCoV epitopes homologous to SARS-CoV-2 epitopes

To directly address whether reactivity against SARS-CoV-2 in unexposed donors could be ascribed to cross-reactivity against other HCoVs, we designed a peptide pool encompassing peptides homologous to CD4-R30 epitopes derived from HCoV-229E, HCoV-NL*63*, HCoV-OC*43*, HCoV-HKU1, and several other HCoVs (see the materials and methods), for a total of 129 HCoV homologs (HCoV-R129; table S2). Similarly, we synthesized a pool that encompassed peptides homologous to the SARS-CoV-2 CD4-S31 epitope pool consisting of potential

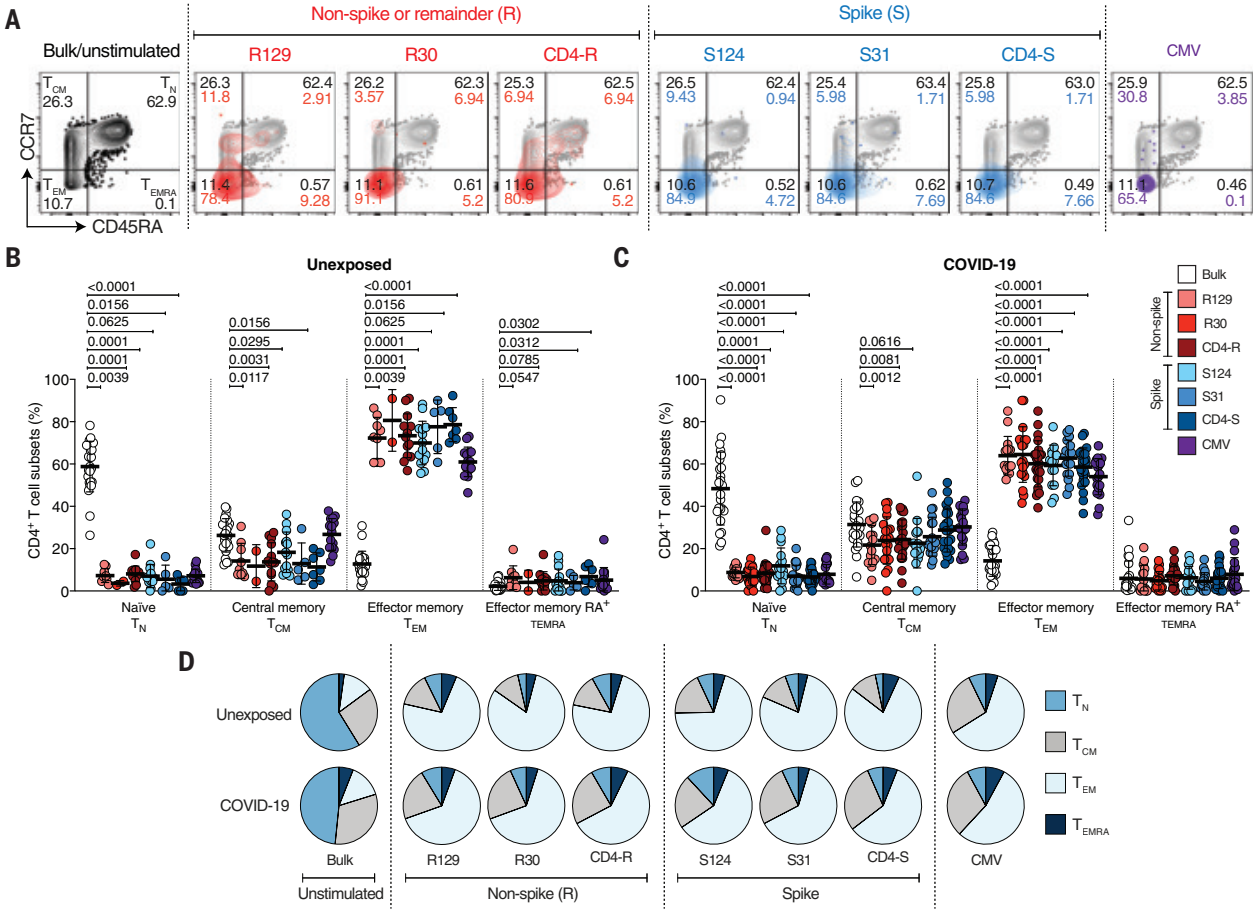


Fig. 3. Phenotypes of antigen-specific CD4$^+$ T cells from SARS-CoV-2–unexposed and recovered COVID-19 patients responding to HCoV epitopes homologous to SARS-CoV-2 epitopes. (**A**) Example of flow cytometry gating strategy for antigen-specific CD4$^+$ T cell subsets after overnight stimulation of PBMCs with HCoV or SARS-CoV-2 peptides ex vivo. (**B** and **C**) Phenotype of antigen-specific CD4$^+$ T cells (OX40$^+$CD137$^+$) responding to the indicated pools of SARS-CoV-2 and HCoV epitopes in unexposed subjects and recovered COVID-19 patients. Data are shown as mean ± SD. Each dot represents an individual subject. Statistical pairwise comparisons in (B) and (C) were performed with the Wilcoxon test. (**D**) Overall averages of antigen-specific CD4$^+$ T cell subsets detected in unexposed subjects and recovered COVID-19 patients. See also fig. S5.

epitopes derived from other HCoVs, for a total of 124 HCoV homologs (HCoV-S124; table S3).

Next, we used an activation-induced marker assay (*25*–*27*) to detect virus-specific T cells in a new set of unexposed donors not used for the epitope identification studies (Fig. 2A and table S4) and a set of convalescent COVID-19 patients (table S5). We detected significant ex vivo $CD4^+$ T cell responses against the SARS-CoV-2 nonspike (CD4-R) and spike (CD4-S) peptides compared with the negative control [dimethyl sulfoxide (DMSO)] (Fig. 2, B and C; $P < 0.0001$ and $P < 0.0001$, respectively, two-tailed Mann–Whitney test). These responses were increased in COVID-19 cases compared with unexposed subjects (Fig. 2D; $P = 0.0015$ and $P = 0.0022$, respectively, two-tailed Mann–Whitney test), as previously reported (*4*). In the unexposed subjects, significant frequencies of $CD4^+$ T cells were detected against the CD4-R30 and CD4-S31 SARS-CoV-2 epitope pools compared with the negative control (Fig. 2B; $P = 0.0063$ and $P = 0.0012$, respectively, two-tailed Mann–Whitney test). Significant $CD4^+$ T cell reactivity was also seen against the corresponding HCoV-R129 and HCoV-S124 pools of matching homologous peptides from other HCoVs (Fig. 2D; $P < 0.0001$ and $P < 0.0001$, two-tailed Mann–Whitney test). Detection of $CD4^+$ T cells with peptide pools selected on the basis of homology was consistent with the hypothesis that cross-reactive $CD4^+$ T cells between SARS-CoV-2 and other HCoVs exist in many individuals.

Reactivity against CD4-R30 and CD4-S31 (Fig. 2D; $P = 0.0008$ and $P = 0.0026$, respectively), but not against HCoV-R129 and HCoV-S124, was increased in COVID-19 cases compared with unexposed individuals (Fig. 2C). Thus, preexisting $CD4^+$ T cell reactivity to HCoV epitopes is modulated by COVID-19 and exposure to cross-reactive SARS-CoV-2 epitopes in COVID-19. These data from COVID-19 cases do not support the hypothesis that the HCoV exposure might induce an original antigenic sin phenomenon, impairing subsequent T cell responses to SARS-CoV-2 epitopes (*28*, *29*), at least for COVID-19 cases of average disease severity.

Next, we examined the ex vivo memory phenotype of the T cells responding to the various

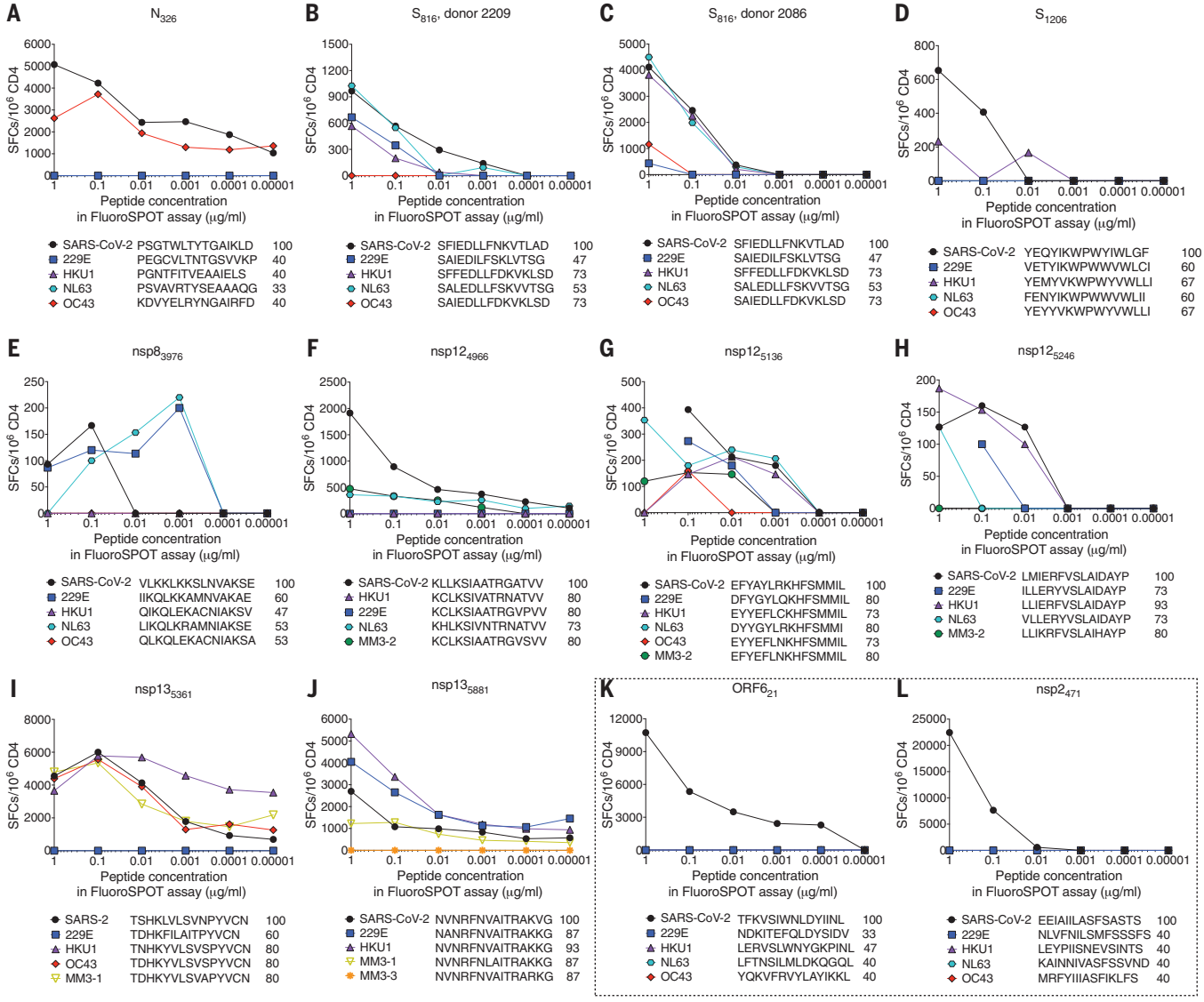


Fig. 4. Cross-reactivity of SARS-CoV-2 and homologous HCoV peptides. Twelve short-term cell lines were generated using specific SARS-CoV-2 donor-epitope combinations selected on the basis of the primary screen. After 14 days of in vitro expansion, each T cell line was tested with the SARS-CoV-2 epitope used for stimulation and peptides corresponding to analogous sequences from other HCoVs at six different concentrations (1, 0.1, 0.01, 0.001, 0.0001, and 0.00001 μg/ml). SFCs/10^6 PBMCs are plotted for T cell lines stimulated with each peptide. See also fig. S7.

epitope megapools. Results from one representative unexposed donor are shown in Fig. 3A. Responding cells in unexposed donors were predominantly found in the effector memory $CD4^+$ T cell population ($CD45RA^{neg}CCR7^{neg}$), followed by the central memory T cells ($CD45RA^{neg}CCR7^{pos}$) (*30*) (Fig. 3, A, B, and D). Comparable patterns of effector and central memory cells were observed among the antigen-specific $CD4^+$ T cells detected in the COVID-19 cases (Fig. 3, C and D). The $CD4^+$ T cells in unexposed donors that recognize SARS-CoV-2 epitopes and epitopes from other HCoVs have a memory phenotype. Overall, these data are consistent with the SARS-CoV-2–reactive $CD4^+$ T cells in unexposed subjects being HCoV-specific memory $CD4^+$ T cells with cross-reactivity to SARS-CoV-2.

Identification of SARS-CoV-2 epitopes cross-reactive with other common HCoVs

The epitopes derived from the CD4-R30 and CD4-S31 pools were used to generate short-term T cell lines derived by stimulation of PBMCs from unexposed subjects. PBMCs were stimulated with an individual SARS-CoV-2 cognate epitope demonstrated to be recognized by T cells from that subject (Fig. 1 and table S1). Overall, T cell lines could be derived that were specific for a total of 42 SARS-CoV-2 epitopes.

These T cell lines were next tested for cross-reactivity against various coronavirus homologs, analogous to an approach previously successful in flavivirus studies (*31*). Cross-reactivity between SARS-CoV-2 epitope recognition and other HCoV epitope recognition was detected for 10/42 (24%) of the T cell lines (Fig. 4, A to J). Cross-reactivity was associated with epitopes derived from SARS-CoV-2 spike, N, nsp8, nsp12, and nsp13. In three cases, HCoV analogs were better antigens than the SARS-CoV-2 peptide, suggesting that they may be the cognate immunogen (Fig. 4, E, I, and J). One SARS-CoV-2 spike epitope was tested in two different donors with similar findings, suggesting that HCoV cross-reactivity patterns are recurrent across individuals. Non–cross-reactive SARS-CoV-2 T cell lines are also shown (Fig. 4, K to L, and fig. S4). It is possible that cross-reactivity to these epitopes might be detected if T cell lines from additional individuals were to be tested. In addition, these epitopes might be homologous to some other, as yet unidentified viral sequence or be recognized by cognate naive T cells expanding in the in vitro culture (*32*). In addition, only 3/18 cases of strong response epitopes (defined in Fig. 1D) were cross-reactive compared with 4/5 of weaker epitopes ($P = 0.02$, Fisher's exact test). To further demonstrate that the cross-reactive responses in unexposed donors are indeed derived from memory T cells, we stimulated purified memory and naïve $CD4^+$ T cells with the CD4-[S31] epitope pool. After 14 days, we detected responses to the CD4-[S31] peptide pool from cultures of memory $CD4^+$ T cells but not naïve $CD4^+$ T cells (fig. S8). These data demonstrate that memory $CD4^+$ T cells recognizing common cold coronaviruses including HCoV-OC43, HCoV-HKU1, HCoV-NL63, and HCoV-229E can exhibit substantial cross-reactivity to the homologous epitope in SARS-CoV-2.

Next we examined, for each SARS-CoV2:HCoV epitope pair, the degree of amino acid sequence homology and any relationship between homology and T cell cross-reactivity, considering different ranges of potentially relevant homology. Only 1% (1/99) of peptide pairs with 33 to 40% homology were cross-reactive. In the 47 to 60% epitope homology range, we observed cross-reactivity in 21% of cases (7/33). Epitope homology ≥67% was associated with cross-reactivity in 57% of cases (21/37; $P = 0.0001$ or $P = 0.0033$ by two-tailed Fisher's exact test compared with the 33 to 40% range epitopes or the 47 to 60% range, respectively). A relationship was observed between epitope homology and $CD4^+$ T cell cross-reactivity. The data demonstrated that the arbitrary selection used as described in Fig. 1D was indeed supported by the experimental data. Thus, ~67% amino acid homology appears to be a useful benchmark for consideration of potential cross-reactivity between class II epitopes. In summary, we have identified more than 140 human T cell epitopes derived from across the genome of SARS-CoV-2. We provide direct evidence that numerous $CD4^+$ T cells that react to SARS-CoV-2 epitopes actually cross-react with corresponding homologous sequences from any of the many different commonly circulating HCoVs, and that these reactive cells are largely canonical memory $CD4^+$ T cells. These findings of cross-reactive HCoV T cell specificities are in stark contrast to HCoV-neutralizing antibodies, which are HCoV species specific and did not show cross-reactivity against SARS-CoV-2 RBD (*33*–*35*). On the basis of these data, it is plausible to hypothesize that preexisting cross-reactive HCoV $CD4^+$ T cell memory in some donors could be a contributing factor to variations in COVID-19 patient disease outcomes, but at present this is highly speculative (*36*).

REFERENCES AND NOTES

1. E. Dong, H. Du, L. Gardner, *Lancet Infect. Dis.* **20**, 533–534 (2020).
2. C. Huang *et al.*, *Lancet* **395**, 497–506 (2020).
3. N. Le Bert *et al.*, *Nature* (2020).
4. A. Grifoni *et al.*, *Cell* **181**, 1489–1501.e15 (2020).
5. B. J. Meckiff *et al.*, *bioRxiv* 2020.06.12.148916 (2020).
6. D. Weiskopf *et al.*, *Sci. Immunol.* **5**, eabd2071 (2020).
7. J. Braun *et al.*, Presence of SARS-CoV-2 reactive T cells in COVID-19 patients and healthy donors. medRxiv 2020.2004.2017.20061440 [Preprint]. 22 April 2020; https://doi.org/10.1101/2020.04.17.20061440.
8. Y. Peng *et al.*, *bioRxiv* 2020.06.05.134551 (2020).
9. L. Kuri-Cervantes *et al.*, *bioRxiv* 2020.05.18.101717 (2020).
10. I. Thevarajan *et al.*, *Nat. Med.* **26**, 453–455 (2020).
11. L. Rodriguez *et al.*, Systems-level immunomonitoring from acute to recovery phase of severe COVID-19. medRxiv 2020.2006.2003.20121582 [Preprint]. 7 June 2020; https://doi.org/10.1101/2020.06.03.20121582.
12. J. Liu *et al.*, *EBioMedicine* **55**, 102763 (2020).
13. D. Mathew *et al.*, *Science* eabc8511 (2020).
14. M. E. Killerby *et al.*, *J. Clin. Virol.* **101**, 52–56 (2018).
15. G. J. Gorse, G. B. Patel, J. N. Vitale, T. Z. O'Connor, *Clin. Vaccine Immunol.* **17**, 1875–1880 (2010).
16. E. E. Walsh, J. H. Shin, A. R. Falsey, *J. Infect. Dis.* **208**, 1634–1642 (2013).
17. S. Nickbakhsh *et al.*, *J. Infect. Dis.* **222**, 17–25 (2020).
18. S. M. Kissler, C. Tedijanto, E. Goldstein, Y. H. Grad, M. Lipsitch, *Science* **368**, 860–868 (2020).
19. D. Weiskopf *et al.*, *J. Infect. Dis.* **214**, 1117–1124 (2016).
20. C. Oseroff *et al.*, *J. Immunol.* **185**, 943–955 (2010).
21. D. Weiskopf *et al.*, *J. Infect. Dis.* **212**, 1743–1751 (2015).
22. H. Voic *et al.*, Identification and characterization of CD4+ T cell epitopes after Shingrix vaccination. bioRxiv 2020/227082 [Preprint]. 29 July 2020; https://doi.org/10.1101/2020.07.29.227082.
23. D. R. Madden, *Annu. Rev. Immunol.* **13**, 587–622 (1995).
24. R. T. Carson, K. M. Vignali, D. L. Woodland, D. A. Vignali, *Immunity* **7**, 387–399 (1997).
25. J. M. Dan *et al.*, *J. Immunol.* **197**, 983–993 (2016).
26. C. Havenar-Daughton *et al.*, *J. Immunol.* **197**, 994–1002 (2016).
27. S. Reiss *et al.*, *PLOS ONE* **12**, e0186998 (2017).
28. A. Vatti *et al.*, *J. Autoimmun.* **83**, 12–21 (2017).
29. K. Kadkhoda, *MSphere* **5**, e00344-20 (2020).
30. F. Sallusto, A. Langenkamp, J. Geginat, A. Lanzavecchia, *Curr. Top. Microbiol. Immunol.* **251**, 167–171 (2000).
31. A. Grifoni *et al.*, *J. Virol.* **94**, e00089-20 (2020).
32. E. J. Hensen, B. G. Elferink, *Nature* **277**, 223–225 (1979).
33. L. Premkumar *et al.*, *Sci. Immunol.* **5**, eabc8413 (2020).
34. M. Yuan *et al.*, *Science* **368**, 630–633 (2020).
35. A. Z. Wec *et al.*, *Science* eabc7424 (2020).
36. A. Sette, S. Crotty, *Nat. Rev. Immunol.* **20**, 457–458 (2020).

ACKNOWLEDGMENTS

We thank the Flow Cytometry Core Facility at the La Jolla Institute for Immunology for technical assistance provided during FACS experiments. **Funding:** This work was funded by the National Institutes of Health (NIH) (NIAID award no. AI42742 from the Cooperative Centers for Human Immunology to S.C. and A.S.; NIH contract no. 75N9301900065 to A.S. and D.W.; grant no. U19 AI118626 to A.S. and B.P.; NIAID K08 award no. AI135078 to J.M.D.); by the UCSD (T32 grant nos. AI007036 and AI007384 from the Infectious Diseases Division to S.I.R. and S.A.R.); and by the John and Mary Tu Foundation (D.M.S). **Author contributions:** Conceptualization: D.W., S.C., and A.S.; Data curation and bioinformatic analysis, J.G. and B.P.; Formal analysis: J.M, A.G., D.W., J.M.D., A.J.M., L.P., and S.C.; Funding acquisition: S.C., A.S., D.W., S.I.R., S.A.R., and J.M.D.; Investigation: J.M., A.G., A.T., J.S., E.P., S.M., M.L., P.R., L.Q., A.S., E.D.Y., S.A.R., A.J.M., L.P., and D.W.; Project administration, A.F.; Resources: S.I.R., Z.C.B., S.A.R., D.M.S., S.C., and A.S.; Supervision: B.P., A.d.S., S.C., A.S., and D.W.; Writing: S.C., A.S., and D.W. **Competing interests:** A.S. and S.C. are inventors on patent application no. 63/012,902, submitted by the La Jolla Institute for Immunology, that covers the use of the megapools and peptides thereof for therapeutic and diagnostic purposes. A.S. is a consultant for Gritstone and Flow Pharma. A.S. and S.C. are consultants for Avalia. The remaining authors declare no competing interests. **Data and materials availability:** All datasets generated for this study are included in the supplementary materials. All epitopes identified in this study have been submitted to the Immune Epitope Database and Analysis Resource (http://www.iedb.org/submission/1000855). Epitope pools used in this study will be made available to the scientific community upon request and execution of a material transfer agreement directed to D.W.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/370/6512/89/suppl/DC1
Materials and Methods
Figs. S1 to S8
Tables S1 to S6
References (*37*–*47*)
MDAR Reproducibility Checklist

25 June 2020; accepted 30 July 2020
Published online 4 August 2020
10.1126/science.abd3871

METALLURGY

Multiplicity of dislocation pathways in a refractory multiprincipal element alloy

Fulin Wang[1], Glenn H. Balbus[1], Shuozhi Xu[2], Yanqing Su[3], Jungho Shin[1], Paul F. Rottmann[4], Keith E. Knipling[5], Jean-Charles Stinville[1], Leah H. Mills[1], Oleg N. Senkov[6], Irene J. Beyerlein[1,2,3], Tresa M. Pollock[1], Daniel S. Gianola[1]*

Refractory multiprincipal element alloys (MPEAs) are promising materials to meet the demands of aggressive structural applications, yet require fundamentally different avenues for accommodating plastic deformation in the body-centered cubic (bcc) variants of these alloys. We show a desirable combination of homogeneous plastic deformability and strength in the bcc MPEA MoNbTi, enabled by the rugged atomic environment through which dislocations must navigate. Our observations of dislocation motion and atomistic calculations unveil the unexpected dominance of nonscrew character dislocations and numerous slip planes for dislocation glide. This behavior lends credence to theories that explain the exceptional high temperature strength of similar alloys. Our results advance a defect-aware perspective to alloy design strategies for materials capable of performance across the temperature spectrum.

The history of materials advancement over centuries has been anchored by the tenet of utilizing one principal element and adding dilute concentrations of alloying elements to achieve the properties of interest. The past decade has witnessed a shift toward an alloy design strategy focusing on the center of multicomponent compositional space. Termed multiprincipal element alloys (MPEAs), complex concentrated alloys, or most commonly as a subclass of these materials known as high-entropy alloys, some of these materials exhibit exceptional combinations of strength, ductility, and damage tolerance (*1*–*5*). Refractory alloys are attractive candidates for use at extremely high temperatures as is demanded by many technology applications, particularly in the aerospace and power-generation sectors (*6*, *7*). Progress in alloy development in this material class, however, has been slow and stands to benefit from the MPEA design paradigm.

The desire for materials with thermally stable microstructures and temperature-insensitive properties has inspired the development of a family of body-centered cubic (bcc) refractory MPEAs (*7*, *8*). These alloys made up a near equiatomic mixture of refractory metal elements. The high strengths displayed, at high temperatures in particular, are very attractive because they surpass state-of-the-art Ni-based superalloys in some cases, which often lose strength at temperatures above 1200°C (*3*, *7*). Alloys that can remain strong at high temperatures enable increased operating temperatures with improved efficiency in a variety of energy, aerospace, and nuclear applications. Alloys with the bcc crystal structure, including steels, are ubiquitous. These alloys are the foundation of a vast array of structures and technologies because of their economies of scale. However, conventional bcc alloys are plagued by a pronounced dependence of the mechanical properties on temperature, which often manifests as a ductile-to-brittle transition with decreasing temperature. The origin of this behavior is linked to the sluggish motion of screw dislocations, which are linear crystalline defects that allow plastic strain accommodation, owing to the nature of the atomic bonding at the dislocation core, and/or embrittlement by interstitial elements such as C, O and N. Increasing temperature mobilizes screw dislocations (in a catch-up game with their edge dislocation counterparts) to enable noncatastrophic shape changes, albeit at the expense of strength. In contrast to many bcc pure metals and dilute alloys, some bcc refractory MPEAs such as MoNbTaW and MoNbTaVW exhibit a gradual decrease of strength with increasing temperature and even a strength plateau in the intermediate temperature range of 600° to 1000°C (*9*) (Fig. 1).

The high strength of single-phase bcc MPEAs is fundamentally related to (i) solute strengthening by the concentrated compositions and (ii) the variation of core structure along a screw dislocation due to local chemical fluctuations (*10*, *11*). Both phenomena suggest that the thermally activated kink nucleation on a screw dislocation is not necessarily the rate-limiting step for dislocation motion, as it is for simple bcc metals (*12*), leading to the prediction of a weak temperature dependence of strength. However, the diversity of macroscopic behavior of different MPEAs (Fig. 1A) has eluded theoretical interpretations and, thereby, a clear mechanistic picture. Furthermore, unlike their face-centered cubic MPEA counterparts, there is limited experimental evidence of the underlying deformation mechanisms in bcc MPEAs, which thus far does not sufficiently support some analytical models (*11*). For instance, dislocations in the deformed bcc MPEA HfNbTaTiZr were observed to exhibit a strong screw character (*13*), a feature indicating the distinctly easy glide of edge compared with screw dislocations, in accordance with the classical kink mechanism of screw dislocations in bcc metals. The dominance of screw dislocations appears to be incommensurate with theoretical predictions of dislocations in MPEAs, such as preexisting kinks on a tortuous screw dislocation and the retarded motion of nonscrew dislocations. Both predictions would render dislocation lines with appreciable deviations from a pure screw orientation. Taken as a whole, the current picture depicted by experiments suggests that classical bcc dislocation mechanisms (Fig. 1C), with only subtle augmentations, are operative in MPEAs. This would be an unexpected finding, given the ruggedness of the atomic landscape that dislocations must navigate in these alloys (Fig. 1D).

A robust understanding of the mechanistic origin of the distinctive properties in the bcc MPEAs that exhibit weak temperature-dependent behavior would greatly inform alloy design principles. We experimentally demonstrated a striking departure from conventional bcc dislocation behavior in the MPEA MoNbTi, an alloy that shows decent strength at low temperature and an intermediate temperature-strength plateau yet a considerably lower density (ρ = 7.67 g/cm^3) (Fig. 1, A and B) (*14*). Moreover, the combination of these three elements reflects one of the most frequently used base blocks of the reported refractory MPEAs (*7*), among which are the ductile and strong examples of HfMoNbTiZr and MoNbTiV. We focused on elucidating the intrinsic capacity for plastic deformation by dynamically probing dislocations in the single bcc phase and single-crystal environment, at equiatomic composition with global randomness (Fig. 1, E and F). Experiments were performed at room temperature [0.12 melting temperature (T_m)], which is below the classical transition temperature of ~0.2 T_m (*12*, *15*), at which thermally activated kink-pair nucleation ceases to be the rate-limiting step. This temperature enabled us to probe differences in the dislocation slip behavior in MPEAs and conventional bcc metals. Our results highlight multiplanar, multicharacter dislocation slip in MoNbTi, encouraged by the broad dispersion in the glide resistance for dislocations, due to the atomic-scale chemical fluctuations. The ability of

[1]Materials Department, University of California, Santa Barbara, CA, USA. [2]California NanoSystems Institute, University of California, Santa Barbara, CA, USA. [3]Department of Mechanical Engineering, University of California, Santa Barbara, CA, USA. [4]Department of Chemical and Materials Engineering, University of Kentucky, KY, USA. [5]Materials Science and Technology Division, U. S. Naval Research Laboratory, Washington, DC, USA. [6]Air Force Research Laboratory, Wright-Patterson AFB, OH, USA.
*Corresponding author. Email: gianola@ucsb.edu

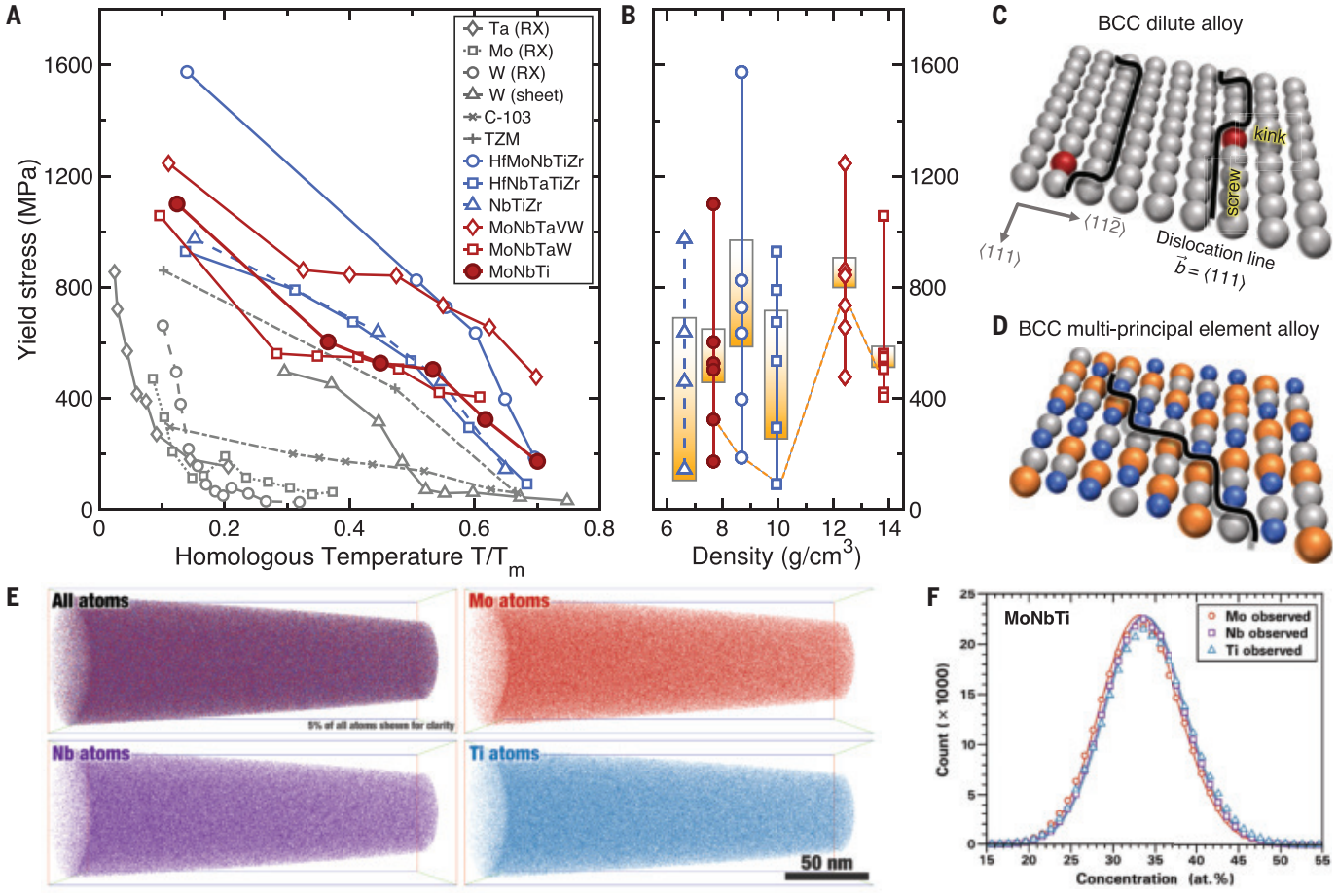


Fig. 1. The temperature dependence of the yield stress of the equiatomic MoNbTi alloy. (**A**) Representative refractory MPEAs are bcc phase polycrystals tested in compression (*22*, *25*). For comparison, the tensile yield stresses of pure bcc metals in either recrystallized (RX) or rolled condition (sheet) are included (*26*, *27*), as are those of the commercial dilute alloys C-103 (a Nb-based alloy) and TZM (a Mo-based alloy, Mo-Ti-Zr) (*28*, *29*). (**B**) Densities are from (*22*). The topmost data are at room temperature. The boxes highlight the yield strengths in the temperature range of 600° to 1000°C. Those at 1200°C are connected by the dashed orange line. (**C** and **D**) Schematic depictions of the dislocation morphologies on the $\{1\bar{1}0\}$ slip plane in bcc dilute alloy and in bcc MPEA, respectively. (**E**) Atom probe tomography reconstructions (87 nm by 87 nm by 246 nm) containing 29.5×10^6 identified ions (*14*), showing the spatial distribution of all the atoms and of Mo, Nb, or Ti atoms individually. (**F**) Concentration of alloying elements in this analyzed volume is 33.181 Mo–33.272 Nb–33.188 Ti [atomic % (at.%)], with trace amounts of interstitial N, O, and C. The comparison with a theoretical binomial distribution (solid line) confirms that Mo, Nb, and Ti are homogeneously distributed.

dislocations to choose the easy gliding direction and plane in the random field of multiple atomic species enables an excellent combination of strength and homogeneous plasticity in this alloy, traits that are not simultaneously observed in conventional metallic alloys.

Characterization of dislocations after nanoindentation

We first select the multiaxial stress state imposed during indentation, as it provides an avenue to investigate all potential configurations of dislocations. The microstructure of the as-processed MoNbTi is initially free from dislocations, as demonstrated by the uniform contrast outside the plastic zone created by the indentation (Fig. 2); thus all the dislocations we observed underwent glide to their rest positions. Detailed analyses of the dislocation Burgers vector and line direction were performed in an area containing discernable individual dislocations (tables S2 and S3). We identified two groups of dislocations with the Burgers vectors of $1/2[\bar{1}11]$ and $1/2[11\bar{1}]$, respectively. The segments of the $1/2[\bar{1}11]$ dislocations that lie roughly vertical in the images are close to screw character, whereas the segments that are approximately horizontal are close to edge character. The morphological deviation away from a straight pure screw orientation indicates a substantial tortuosity with segments of nonscrew character along the otherwise rectilinear screw dislocation (fig. S2). Dislocations #2 to #6 are close to pure edge ones and exist on distinct slip planes, including (211), (321), and (110). Dislocations #7 and #8 have one segment of edge or mixed character (roughly horizontal in the images) and the other segment close to screw character (roughly vertical in the images), with both habiting the (110) plane. The dislocations in the second group have a Burgers vector of $1/2[11\bar{1}]$ and are parallel to each other, as represented by dislocation #9 (Fig. 2F). These dislocations appear at an angle to the projection of their Burgers vector $1/2[11\bar{1}]$ at the diffraction condition of $\mathbf{g}_3$ (Fig. 2D), indicating that they are all mixed dislocations. Both the presence of substantial nonscrew segments and high-order slip planes are unexpected for bcc metals at such low homologous temperature ($0.12\ T_m$). Furthermore, dislocations #1 to #8 have the same Burgers vector and are subjected to a similar stress state on the basis of their close proximity to each other, suggesting that the propensity of

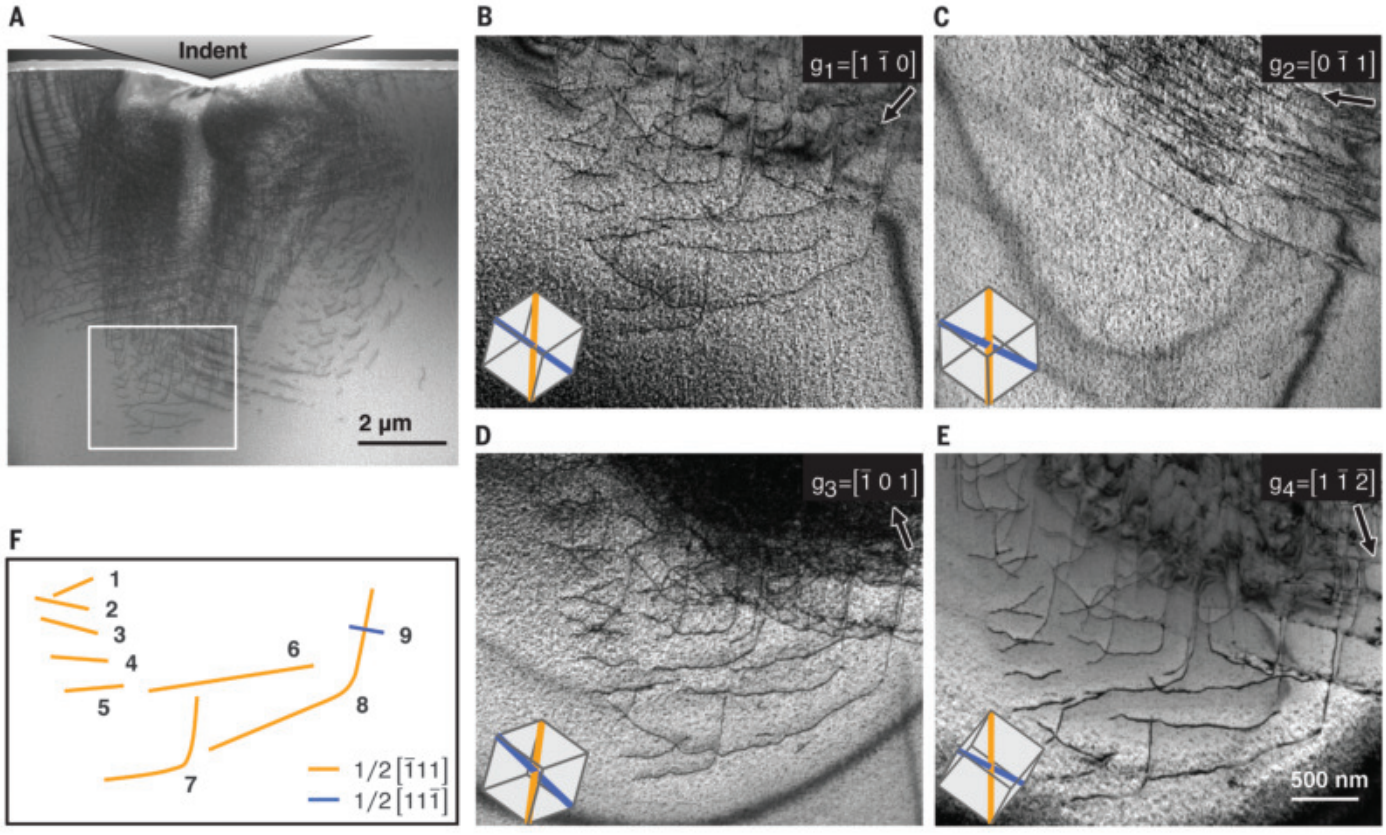


Fig. 2. Dislocations induced by nanoindentation. The scanning transmission electron microscopy (TEM) image in (**A**) shows an overview of the dislocations under the indent. (**B** to **E**) Two-beam bright-field TEM images of the boxed area in (A). The diffraction vector, **g**, is annotated in each image with the direction shown by the arrow. The crystal orientation is indicated with a cubic lattice in the respective image. The Burgers vectors are drawn in the lattices, with the orange line denoting $1/2[\bar{1}11]$ and the blue line denoting $1/2[11\bar{1}]$. (**F**) shows schematically the dislocations numbered 1 to 9 for dislocation line direction analysis. They are colored according to the respective Burgers vector.

dislocations to glide on various planes may be similar.

Dislocations during single-crystal tensile deformation

We assessed the dynamic behavior of dislocations during plastic deformation using quantitative in situ tensile tests, which revealed cross-slip and bowing of dislocations, providing key insights on the mechanisms mediating dislocation glide. We conducted tensile loading of the single-crystal sample close to the [001] crystal direction, resulting in equal resolved shear stresses (RSSs) along all four possible ⟨111⟩ Burgers vectors. We quantitatively determined the full crystallographic characteristics of the gliding dislocations, including Burgers vector, slip plane, and line direction, with high fidelity (14) and uncovered the substantial nonscrew character of gliding dislocations in MoNbTi (Fig. 3). All of the dislocations shown in Fig. 3 glide along the traces at the same angle of −61° to the tensile direction, corresponding to a distinctive $1/2[1\bar{1}1](\bar{2}13)$ slip system (table S4). We reconstructed the line morphology of dislocation (iii) at three sequential positions of the line during glide in Fig. 3G, demonstrating the evolution from almost pure screw dislocation to a serrated line and eventually to a more smoothly curved line. These observations suggest that the critical stress to move edge or mixed dislocations is not distinctly lower than that to move screw dislocations, which is not the case for conventional bcc metals at low homologous temperature. Rather, in conventional bcc metals, dislocations remain in a pure screw orientation as they migrate, because the nonscrew segments can glide away easily (*12*).

With increased stress from 812 to 938 MPa (Fig. 3, A and B), we observed highly variable motion among seemingly identical dislocations. Many dislocations [such as dislocation (i)] glide to the upper edge of the specimen and create long traces, whereas other dislocations [such as dislocations (ii) and (iii)] glide for only a short distance. Usually heterogeneous stress fields created by nearby defects would explain such variability, yet the defect environment in this region is not appreciably different. The different gliding behaviors of the dislocations in the same slip system suggest variations in the local lattice resistance for dislocation slip, originating at the atomic scale and not detectable at the current magnification.

Distribution of slip activities on high-order planes

The experimental observations of dislocations possessing largely nonscrew character on high-order slip planes are statistically robust over the course of loading and are homogeneously distributed across the entire single crystal (region that measured ~5 by 5 μm gauge) (Fig. 4), suggesting that successive source operation on a single plane leading to heterogeneous avalanche behavior is not favored. The single crystal yields at 812 MPa under tensile loading along the [001] crystal direction, which, upon considering the largest Schmid factor of m = 0.49 in the single crystal, agrees well with the measured bulk polycrystal compressive yield strength of 1100 MPa for MoNbTi (Fig. 1A). This comparison results in a Taylor factor of 2.76, which is in the range of 2.75 to 3.06 for

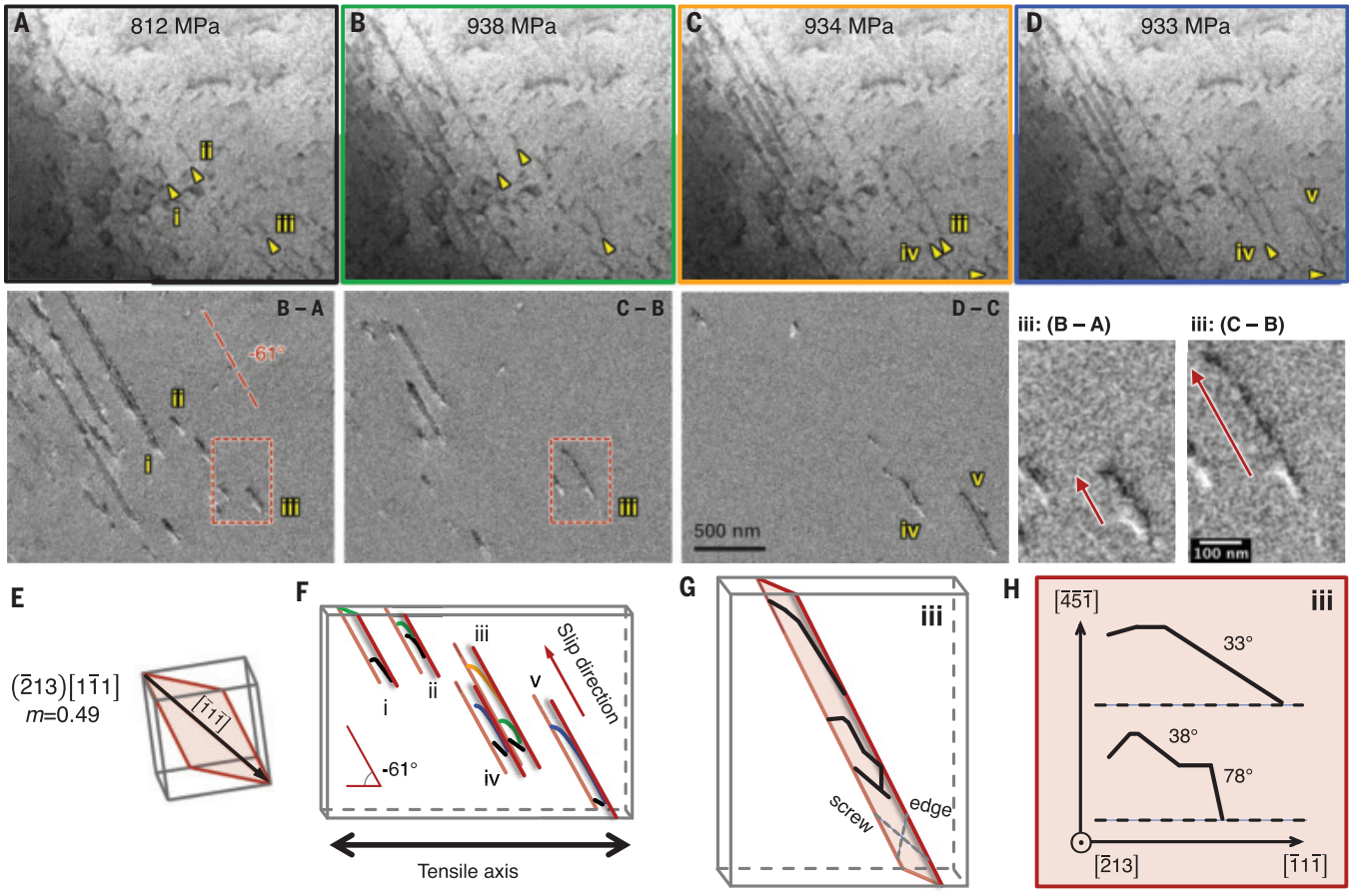


Fig. 3. Dynamic observation of dislocation bowing in the scanning electron microscope operating in transmission mode. (**A** and **B**) During loading. (**C** and **D**) At load-hold. The difference images between each of the two raw images are shown below. The parallel slip traces are measured to be at −61° from the horizontal (tension) direction. (**E**) Corresponding slip plane and Burgers vector shown from the same perspective as the images, including the Schmid factor *m*. The schematic isometric view in (**F**) shows the positions of selected dislocations colored by the respective images. The long and parallel lines in red are the slip traces. The evolution of dislocation (iii) is presented by the magnified difference images (B − A) and (C − B). The arrows indicate the slip direction from the original position (bright line) to the current position (dark line). These dislocation lines are reconstructed as thick black lines in (**G**), with pure edge and screw orientations indicated with dashed lines. (**H**) Line morphology of dislocation (iii) from the perspective of slip plane face-on. The horizontal and vertical axes are along the screw and edge orientations, respectively. The dislocation character angles for each segment are indicated.

typical bcc metals and close to 2.73 calculated for polycrystalline bcc metals deforming by pencil glide (*16*). Therefore, the activation of slip systems in the microtensile specimen is representative of that in a bulk polycrystalline material. The individually determined slip systems are active across the stress range without exhaustion (Fig. 4F). All of the measured slip traces are linked to high-order slip planes of {112}, {123}, and even {134} types.

To determine the origin of the glide-plane selection, we carried out a Schmid factor analysis for all recorded slip traces. Most of the active slip systems have high Schmid factors (0.42 to 0.49), with only a few systems having low Schmid factors (0.26 or 0.29) (table S5). We did not observe traces corresponding to the commonly considered {110} slip planes, even though some of these systems exhibit relatively high Schmid factors of 0.36, 0.40, or 0.45. The unexpected result of the selection of high-order slip planes with high Schmid factors could be rationalized by composite (effective) slip on various {110} planes through facile cross-slip. However, the absence of the traces of {110} slip systems with equally high Schmid factor (0.45) indicates that the slip resistances on the various high-order planes are comparable if not even lower, thereby favoring high-order planes. Additionally, our observations of mixed-character gliding dislocations would preclude the copious cross-slip of screw dislocations required to give rise to slip on effective (noncrystallographic) high-order planes according to the maximum RSS plane criterion (*17*). Considering the complex atomic distribution in MPEAs, our results motivate the hypothesis that the probability of moving dislocations on the observed high-order crystallographic slip planes is similar to or even higher than that on {110} planes, promoting slip selection on a widespread basis and uniformly over the crystal.

Slip resistance for dislocations in MoNbTi

In light of our experimental evidence for the nonscrew character of dislocations and the multiplicity of operative slip systems, we focused on the origin of the slip resistance for the edge and screw dislocations on different slip planes. Atomistic simulations reported the decreased ratio of the critical stress required to move screw and edge dislocations on {110} planes in NbTiZr, relative to conventional bcc metals and alloys, owing to the fluctuations of solute concentration (*10*) and the varying energy barrier for kink migration in a random solute environment of MPEAs (*18*). Our results suggest that examining only the common {110} slip planes in bcc MPEAs does not suffice in

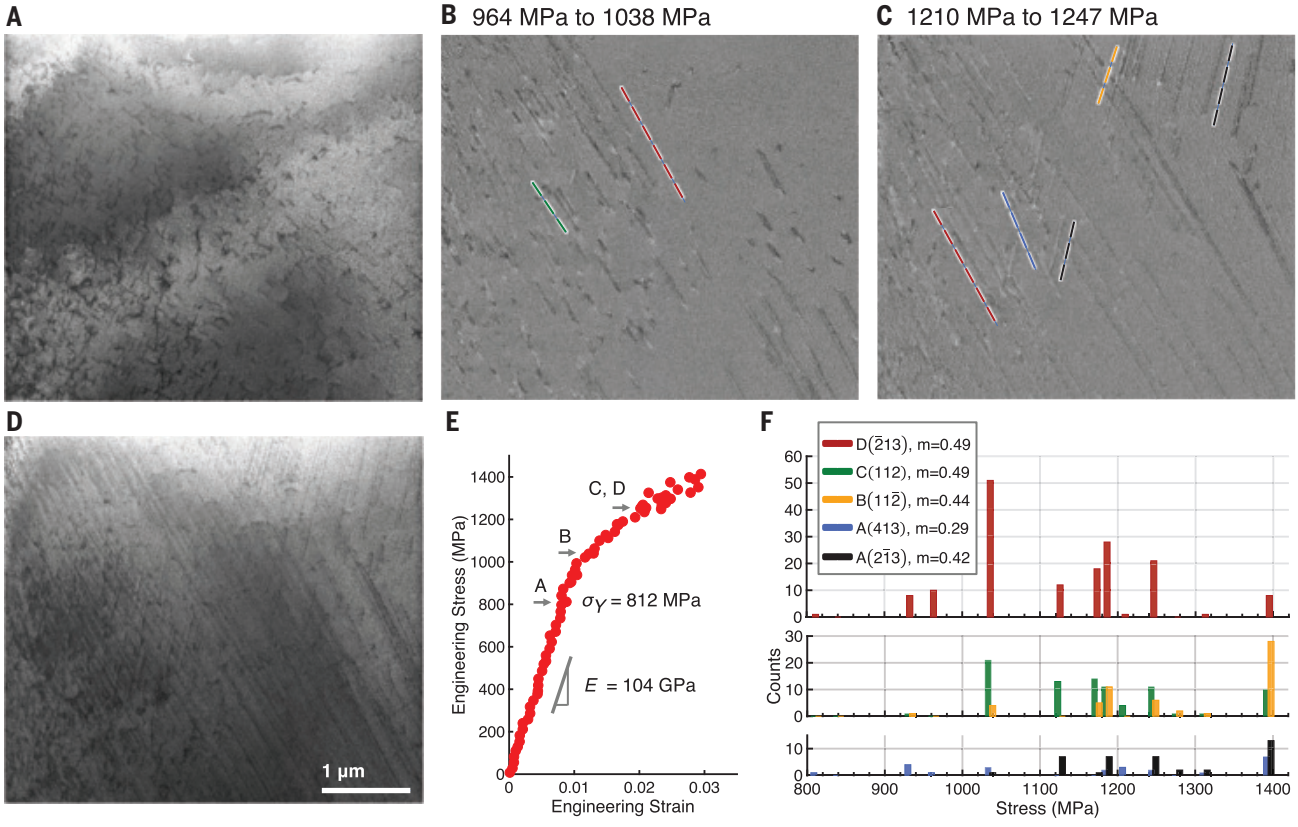


Fig. 4. The distribution of slip activities in the gauge region and across the tested stress values. (**A**) Initial microstructure of the gauge region. (**B** and **C**) Difference images showing the change in microstructure for the respective stress increments. (**D**) Deformed microstructure at 1247 MPa. (**E**) Engineering stress-strain data. (**F**) Distribution of the occurrence of slip activities at different stress values. The Burgers vectors are expressed according to the Schmid-Boas notation: $A = [\bar{1}11]$, $B = [111]$, $C = [11\bar{1}]$, and $D = [1\bar{1}1]$. The Schmid factors m are indicated.

accurately representing the three-dimensional fluctuation of chemistry. Reminiscent effects have been reported in NiAl hosting the B2 crystal structure (*19*), in which the Peierls stresses for screw and edge ⟨111⟩ dislocations were shown to be very similar on {112} planes, whereas they have a ratio of ~29 on {110} planes, and the Peierls stress for screw dislocations on {112} was slightly smaller than that on {110} planes.

We performed atomistic simulations to calculate the slip resistance in MoNbTi for screw and edge dislocations on {110}, {112}, and {123} planes. To reflect the influence of the local atomic environment and account for statistical fluctuations, calculations were performed by enforcing a short dislocation segment of length 3*b* to 4*b* (*b* is the magnitude of Burgers vector) to glide on the respective slip plane by one lattice periodicity, with multiple enumerations using different simulation cells. Each cell contains atoms that were randomly assigned to a bcc lattice on the basis of the equiatomic composition. The calculated slip resistance is akin to a Peierls stress in elemental and dilute alloys but, in MPEAs, is more accurately termed a local slip resistance (LSR), as it accounts for the periodic energy landscape of the crystal lattice as well as any local solute environment sampled by a dislocation as it moves in the globally random structure.

The computed LSR values have a broad distribution (Fig. 5) in contrast to the deterministic value in pure metals (the Peierls stress in a homogeneous lattice). The substantial variations result in dislocation glide of a probabilistic nature. Among the simulated dislocations, the lowest stress to move a dislocation (of edge character) can occur on either {110}, {112}, or {123} planes depending on the local atomic configuration. We observe that on the {112} planes, certain local atomic configurations can bear an LSR that is lower for screw dislocations than for edge dislocations. The dislocation slip asymmetry, owing either to twinning–anti-twinning asymmetry known for {112} planes (*17*) or to a shear in a positive or negative sense on other planes (Fig. 5B), also takes on a probabilistic nature. Slip in the anti-twinning direction is not always harder than in the twinning direction on {112} planes, as is the case for conventional bcc metals. Taken as a whole, the diminished distinctions between slip planes, dislocation characters, and even slip directions imply multiple pathways for dislocation slip, which is desirable for plastic formability and damage tolerance.

Discussion

Despite the decreased difference between the LSR to move screw and edge dislocations in MoNbTi, slip still occurs primarily by edge dislocations. What is distinctive to bcc MPEAs is that the straight screw dislocation could have a varying core structure along its line due to the local chemical fluctuations. As a result, the barrier for kink nucleation may be lowered, and kinks could even be preexisting (*10*, *20*, *21*). Once kinks are present on a screw dislocation on any of the {110}, {112}, and {123} planes, the edge character dislocation segments would glide on the slip plane that locally have the lowest resistance. On the basis of the series of simulations performed in MoNbTi, it appears that gliding on {123} planes is statistically the easiest (Fig. 5C), which agrees with the experimental observation that the active slip system at yield is $1/2[1\bar{1}1](\bar{2}13)$ (Fig. 3). Using the 812-MPa tensile yield stress of the experimental specimen and the Schmid factor of 0.49, we determined the RSS on the slip

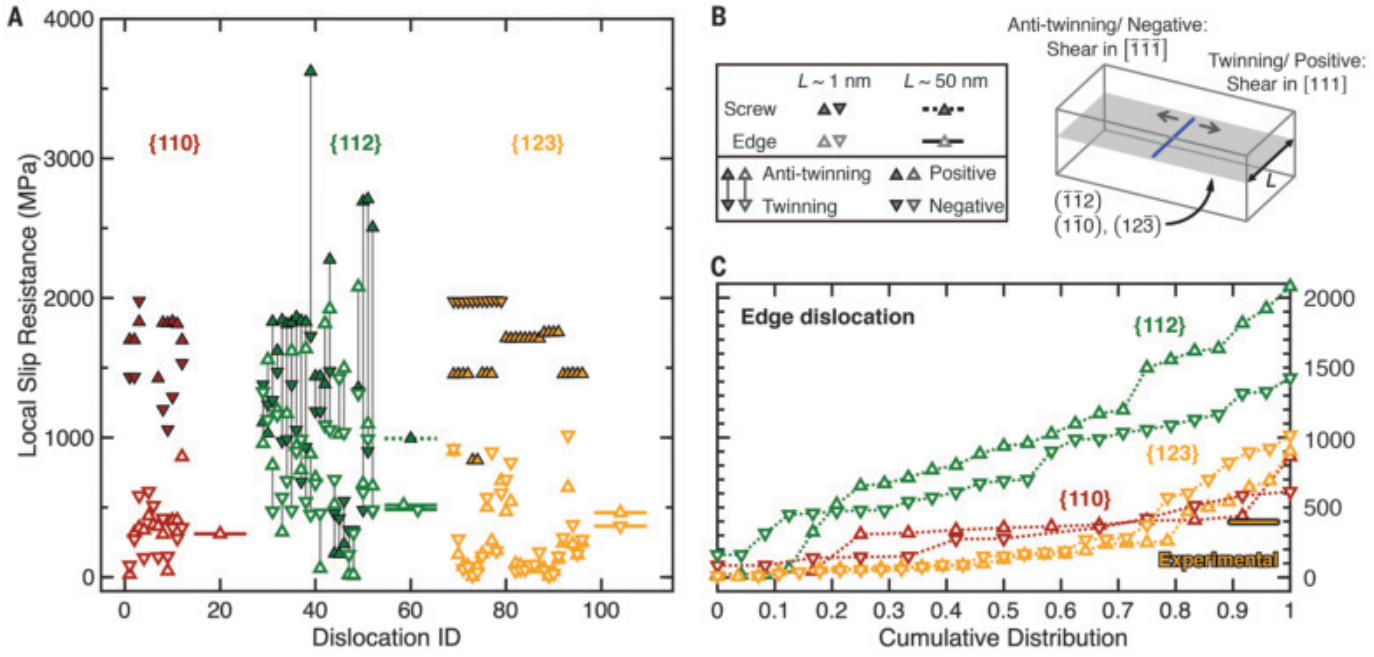


Fig. 5. Simulated LSR for screw and edge dislocations on {110}, {112}, and {123} planes in MoNbTi. (**A**) Simulations were performed respectively on screw (filled markers) and edge (open markers) character dislocations and on different types of planes (marker colors). Both short [length (L) ~ 1 nm] and long dislocations (L ~ 50 nm) were used to assess the effect of local chemistry. Each simulated dislocation is identified by an ID number. The same initial dislocation [shown as a thick blue line in the schematic in (**B**)] was moved to create shear in opposite directions to assess the slip sense asymmetry. If the shear created by dislocation glide is in the [111] direction, it is the twinning sense on ($\bar{1}\bar{1}$2) plane and is denoted as the positive sense on (1$\bar{1}$0) or (12$\bar{3}$) plane. If the shear is in the [$\bar{1}\bar{1}\bar{1}$] direction, it is the anti-twinning sense and negative sense on the respective planes. Only the results of stably gliding dislocations are shown in the plot. (**C**) Cumulative distribution of the LSR values for edge dislocations is plotted. The experimentally determined resolved shear stress on the active slip system 1/2[$\bar{1}\bar{1}$1]($\bar{2}$13) (in Fig. 3) is shown as a thick horizontal line.

system to be 398 MPa. Compared with the computed LSRs, it is highly likely (at a probability of ~0.8) that the RSS is sufficient to activate edge dislocation motion on {123} planes. Together with the dominance of nonscrew segments on the observed dislocation lines, this quantitative comparison demonstrates that the high slip resistance of edge dislocations controls the yield stress in this alloy. On the other hand, rather than material softening by opting for the weakest link of gliding dislocations, the variation of the LSR in the same glide plane serves to suppress dislocation runaway and localized plastic flow, which is exhibited both as the increased tortuosity of the dislocation line and as multiple slip events in the crystal. Consistent with the statistical analysis of our experiments (Fig. 4), the traces of various slip systems are homogeneously distributed in the region on multiple high-order planes, and no localization is observed up to 600 MPa beyond yielding.

The probabilistic glide on {110}, {112}, {123}, and even {134} planes can contribute to retained strength at elevated temperatures, before the onset of diffusion-controlled dislocation glide/climb around 0.5 to 0.6 T_m (*22*). The efficacy of the solute-strengthening mechanism decreases with increasing temperature, because it is a thermally activated process by nature. The alternative strengthening mechanism of dipole dragging on screw dislocations depends weakly on a temperature below ~0.6 T_m, because it is based on the elastic interaction of dislocation segments. Frequent kink formation on different planes and the subsequent glide of the edge or mixed kink dislocation on different planes (*10*, *23*) are prerequisites for edge dipole formation, so the chances are greatly enhanced in MoNbTi by the activities on a large number of slip planes. The operation of dipole dragging strengthening, paired with the low extent of thermal elastic softening in the refractory bcc alloys (*24*), is likely responsible for the strength plateau in the intermediate temperature regime of MoNbTi (Fig. 1). However, the multiple operative slip planes may not be common in all bcc MPEAs, as evidenced by the diverse strength-temperature dependence (Fig. 1). This promotes the need to understand and even tailor the extent of dipole dragging and hence its contribution to hardening through a defect-aware alloy design strategy.

Our study reports clear, albeit unexpected, evidence for (i) a large fraction of nonscrew segments on the gliding dislocations in a bcc refractory MPEA at a low homologous temperature of 0.12 T_m, (ii) the importance of high-order slip planes in bcc MPEAs, and (iii) a probabilistic description of the slip resistance in random alloys. These results constitute the mechanistic basis that can explain the high strength and homogeneous plasticity of this alloy at a low homologous temperature, and they contribute to understanding its weak temperature dependence of strength. All three features are desirable for applications exposing materials to extreme temperatures yet are not simultaneously attainable in traditional refractory metals or alloys. These insights pave the way for theory-guided exploration of new alloys in the vast compositional space of MPEAs, as the multicharacter and multiplanar nature of dislocation slip could be employed as a mechanistic fingerprint for material screening. At a fundamental level, a probabilistic description of the traditional Peierls stress of dislocations owing to the fluctuating chemical landscape at the atomic scale should be brought to the forefront when designing new MPEAs.

REFERENCES AND NOTES

1. J. Yeh *et al.*, *Adv. Eng. Mater.* **6**, 299–303 (2004).
2. B. Cantor, I. Chang, P. Knight, A. Vincent, *Mater. Sci. Eng. A* **375-377**, 213–218 (2004).
3. D. Miracle, O. Senkov, *Acta Mater.* **122**, 448–511 (2017).
4. B. Gludovatz *et al.*, *Science* **345**, 1153–1158 (2014).
5. B. Gludovatz *et al.*, *Nat. Commun.* **7**, 10602 (2016).
6. B. Butler *et al.*, *Int. J. Refract. Met. Hard Mater.* **75**, 248–261 (2018).
7. O. Senkov, D. Miracle, K. Chaput, J.-P. Couzinie,*J. Mater. Res.* **33**, 3092–3128 (2018).

8. O. Senkov, G. Wilks, D. Miracle, C. Chuang, P. Liaw, *Intermetallics* **18**, 1758–1765 (2010).
9. O. Senkov, G. Wilks, J. Scott, D. Miracle, *Intermetallics* **19**, 698–706 (2011).
10. S. Rao *et al*., *Acta Mater.* **168**, 222–236 (2019).
11. F. Maresca, W. Curtin, *Acta Mater.* **182**, 235–249 (2020).
12. B. Sestak, A. Seeger, *Z. Metallk.* **69**, 195–202 (1978).
13. L. Lilensten *et al*., *Acta Mater.* **142**, 131–141 (2018).
14. Materials and methods are available as supplementary materials.
15. M. Butt, *Philos. Mag.* **87**, 3595–3614 (2007).
16. J. Rosenberg, H. Piehler, *Metall. Trans.* **2**, 257–259 (1971).
17. C. Weinberger, B. Boyce, C. Battaile, *Int. Mater. Rev.* **58**, 296–314 (2013).
18. F. Maresca, W. Curtin, *Acta Mater.* **182**, 144–162 (2020).
19. P. Gumbsch, R. Schroll, *Intermetallics* **7**, 447–454 (1999).
20. H. Suzuki, "Solid solution hardening in body-centered cubic alloys" in vol. 4 of *Dislocations in Solids* (1980), 191–217.
21. S. Rao *et al*., *Acta Mater.* **125**, 311–320 (2017).
22. O. Senkov, S. Gorsse, D. Miracle, *Acta Mater.* **175**, 394–405 (2019).
23. D. Caillard, *Acta Mater.* **61**, 2808–2827 (2013).
24. G. Laplanche, P. Gadaud, L. Perriere, I. Guillot, J. Couzinie, *J. Alloys Compd.* **799**, 538–545 (2019).
25. J. P. Couzinié, O. N. Senkov, D. B. Miracle, G. Dirras, *Data Brief* **21**, 1622–1641 (2018).
26. J. Bechtold, *Acta Metall.* **3**, 249–254 (1955).
27. F. Schmidt, H. Ogden, "The engineering properties of tungsten and tungsten alloys" (DMIC Report 191, Battelle Memorial Institute, 1963).
28. ATI Wah Chang Nb/Nb Alloy C-103. MatWeb (2020); http://www.matweb.com/search/DataSheet.aspx?MatGUID=ed50a3a07706450590669cedc7784150.
29. P/M Molybdenum TZM (Mo-0.5Ti-0.1Zr; Mo Alloy 364), Stress Relieved. MatWeb (2020); http://www.matweb.com/search/DataSheet.aspx?MatGUID=1868b5a70b8240999ccc55468cc90c9e.

ACKNOWLEDGMENTS

We thank C. Woodward and S. Rao at the Air Force Research Laboratory for fruitful discussions and for providing detailed comments on the manuscript, and P. Callahan at the U. S. Naval Research Laboratory for assistance with atom-probe specimen preparation. We also thank K. Dang for helpful discussions on the periodic array of dislocations model. **Funding:** This work was funded by the Office of Naval Research, Basic Research Challenge Program (grant no. N00014-18-1-2392). The work of S.X. was supported in part by the Elings Prize Fellowship in Science offered by the California NanoSystems Institute (CNSI) on the UC Santa Barbara campus. The work of K.E.K. was supported in part by a Department of Defense Basic Research Office Laboratory University Collaboration Initiative (LUCI) fellowship. The work of O.N.S. was supported through the United States Air Force on-site contract (FA8650-15-D-5230). The experimental work made use of the MRL Shared Experimental Facilities at UC Santa Barbara supported by the Materials Research Science and Engineering Center (MRSEC) Program of the National Science Foundation (NSF) (award no. DMR 1720256), a member of the NSF-funded Materials Research Facilities Network (www.mrfn.org). Use was made of computational facilities purchased with funds from the NSF (CNS-1725797) and administered by the Center for Scientific Computing (CSC). The CSC is supported by the CNSI and the MRSEC. This work used the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by NSF (grant no. ACI-1053575). **Author contributions:** O.N.S. synthesized the material and conducted bulk mechanical tests. L.H.M. performed chemical analyses. K.E.K. performed atom probe tomography experiment and analysis. G.H.B. performed nanoindentations. F.W. and P.F.R. performed ex situ TEM analyses of the dislocations after nanoindentation. F.W., G.H.B., and J.-C.S prepared samples and performed in situ tests in scanning electron microscopy. F.W. and J.S. analyzed the in situ test data. S.X. and Y.S. performed the atomistic simulations. F.W., O.N.S., I.J.B., T.M.P., and D.S.G. jointly designed and interpreted the study and prepared the manuscript. All authors contributed to the discussion and revision of the manuscript. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** All data are available in the manuscript or the supplementary materials.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/370/6512/95/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S6
Tables S1 to S8
References (*30*–*53*)

27 November 2019; accepted 2 July 2020
10.1126/science.aba3722

ORGANIC CHEMISTRY

A universal system for digitization and automatic execution of the chemical synthesis literature

S. Hessam M. Mehr*, Matthew Craven*, Artem I. Leonov*, Graham Keenan, Leroy Cronin†

Robotic systems for chemical synthesis are growing in popularity but can be difficult to run and maintain because of the lack of a standard operating system or capacity for direct access to the literature through natural language processing. Here we show an extendable chemical execution architecture that can be populated by automatically reading the literature, leading to a universal autonomous workflow. The robotic synthesis code can be corrected in natural language without any programming knowledge and, because of the standard, is hardware independent. This chemical code can then be combined with a graph describing the hardware modules and compiled into platform-specific, low-level robotic instructions for execution. We showcase automated syntheses of 12 compounds from the literature, including the analgesic lidocaine, the Dess-Martin periodinane oxidation reagent, and the fluorinating agent AlkylFluor.

Synthetic chemistry remains labor intensive, and some protocols suffer from errors or ambiguity (*1*, *2*). Recently, there has been rapid growth in the development of robotic synthesis of molecules (*3*–*5*), but new developments are limited to specific reaction types, and a universal approach for the automatic encoding and validation of the chemical synthesis literature is lacking, which means that automation currently just displaces effort from manual labor to programming (*6*). The burden is further increased by the plethora of robotic solutions, which lack a common standard architecture. What is needed is an abstraction that can not only implement the literature (*7*) but also adapt to new synthetic methods (*8*–*11*), in accordance with a standard that ensures interoperability between hardware systems. Currently, the standard of the recording and subsequent reporting of the synthesis of new chemical compounds varies greatly, and procedures are often available only as incomplete and ambiguous prose, relying on the expert to fill in any gaps (*12*). This means that the quality of the data stored in many reaction databases is highly variable (*13*), posing many problems for reproducibility (*14*), as well as preventing the development of reliable digital methods for prediction of reactivity (*15*), new structures (*16*), and functions (*17*). These limitations have also prevented the practical digitization of chemistry (*6*)—i.e., the development of automated systems that could run reactions and make molecules—because of the lack of standards linking the reaction dependencies to a standard hardware control and specification, as well as a machine-readable standard for recording synthetic procedures.

School of Chemistry, University of Glasgow, Glasgow G12 8QQ, UK.
*These authors contributed equally to this work.
†**Corresponding author. Email: lee.cronin@glasgow.ac.uk**

A key factor hindering the digitization of chemistry is the lack of a universal chemical programming language despite the recent proliferation of chemical automation platforms. For example, we have recently developed the Chemputer (*18*), a programmable modular system with hardware capable of performing the fundamental processes of chemical synthesis. The Chemputer was able to automate batch synthetic procedures but was limited to executing a set of specialized low-level hardware instructions, and no uniform development environment or universal hardware interface or specification was provided to allow the code to run on other systems. Consequently, previously automated syntheses involved laborious and error-prone manual translation of the synthesis procedures to these low-level instructions, which precludes portability to other platforms and requires programming expertise as well as detailed knowledge of the system's robotic operations. Such implementations demonstrate the capabilities of the hardware but are not a suitable or sustainable way of automating chemical synthesis. Far from being restricted to the Chemputer, the absence of a universal chemical programming language threatens to undermine the feasibility of the nascent automated synthesis ecosystem. Automated platforms from different companies or research groups all have bespoke instruction sets with no obvious semantic link among them or to the literature. This broken link has prevented the digitization of chemistry: Our vast repertoire of synthetic knowledge cannot be directly executed by robots today. To address this, we envisaged that a new architecture must rely on hardware-independent instructions represented in a standard chemical language that can express the synthesis of virtually any molecule (Fig. 1A).

Whether carried out manually by a chemist or automatically by a robot, the execution of batch synthesis procedures follows a fixed

set of steps, with parallels between the manual actions in the traditional workflow and their automated counterparts in the digital workflow (Fig. 1B).

Comparing the two workflows reveals the steps necessary for realizing the digital workflow. These steps all operate on or refer to an intermediate representation of the actions in the procedure. To use a sufficiently capable batch synthesis robot for the universal digitization of chemistry, we thus hypothesized that the robot's hardware abstraction must be connected to practical synthesis by an executable hardware-independent chemical programming language. This universal chemistry-oriented domain-specific markup language allows us to map the commonly reported methodological instructions to hardware-agnostic chemical instructions. Additionally, the semantic connection between this language and laboratory batch synthesis enables robotic execution

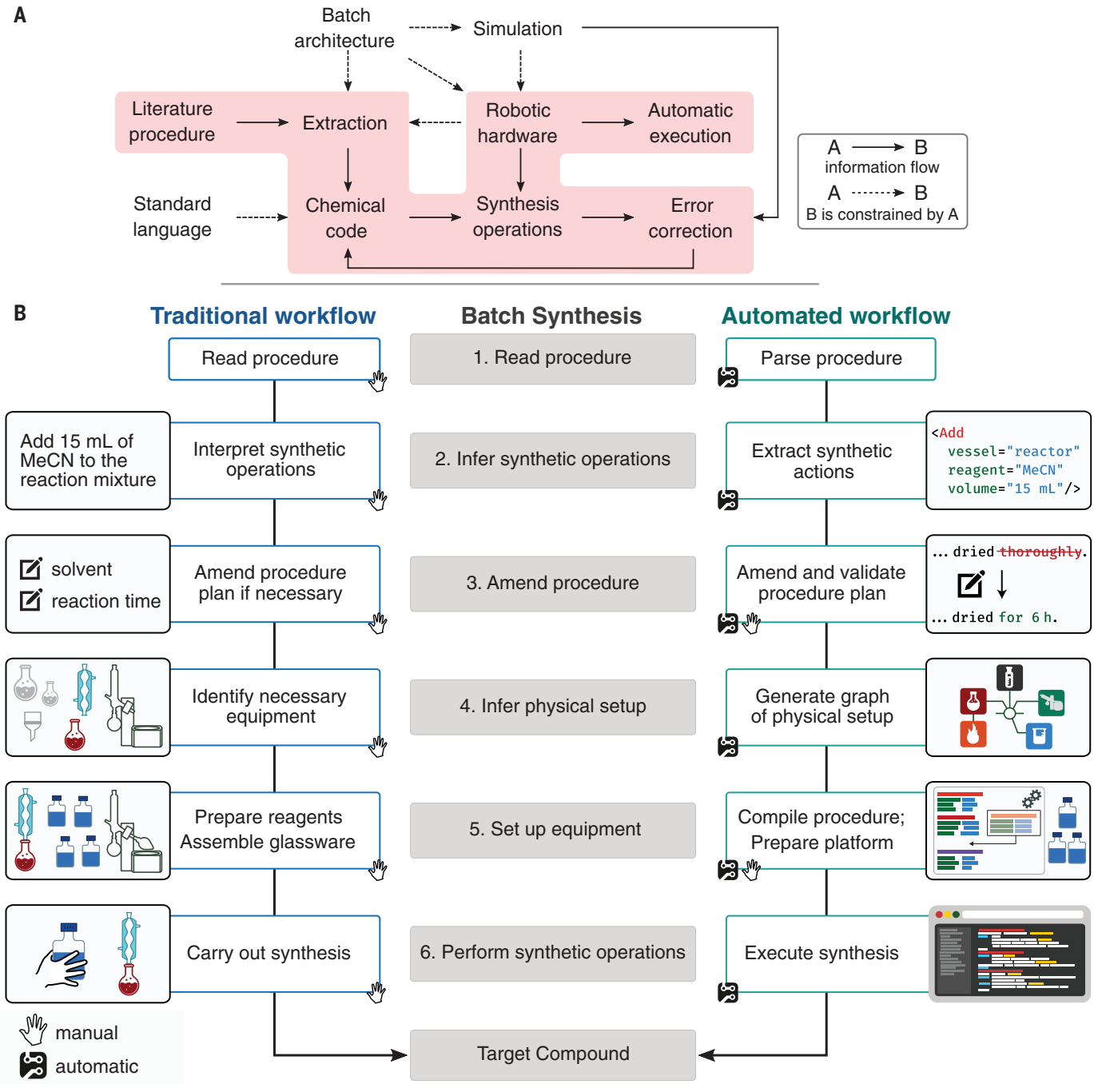


Fig. 1. Conceptual representation of chemical synthesis execution. (**A**) Outline of the universal system for the automatic execution of chemical synthesis from the literature. A method of extracting the procedure is needed, followed by an algorithmic process for producing the code that conforms to a standard hardware and software architecture. Manual error correction by the expert chemist and simulated execution ensure reliability and safety. The execution route is highlighted. (**B**) Abstraction of batch synthesis and two possible workflows for its execution. In the traditional workflow (left), all operations specified in the middle column are carried out manually by the chemist. Most operations are automated in the digital workflow (right), although the chemist can inspect and override the system's output without having to modify the software or learn a programming language.

of laboratory synthetic procedures without manual adaptation or modification. Instructions can thus be translated between the two without loss of information, as we no longer need to cross the large semantic gap between chemistry and robotic operations.

Herein, we present a system for the digitization of chemistry by making the chemical literature directly executable on automated synthesis platforms such as the Chemputer. To implement the system, we required a language with which synthetic procedures can be expressed succinctly by using vocabulary similar to that in the literature and executed on any compatible robotic platform. To this end, we have devised a chemical programming language and associated visual development environment that allows users to encode procedures without ambiguity and exchange them using a standard format. Our Chemical Description Language, XDL (the X pronounced "kai" in reference to "χημία," the Greek word for chemistry), achieves this goal by representing syntheses as sequences of processes taking place in abstract vessels with abstract hardware and is based on the ubiquitous XML format (*19*). Our chemical development environment then inspects this representation and determines which hardware components are necessary to execute it on a virtual machine, producing specifications for a platform capable of running the procedure. Given a robotic platform with the required hardware modules, which conform to the standard architecture, the generic XDL description of the procedure is compiled into an executable specific to the platform

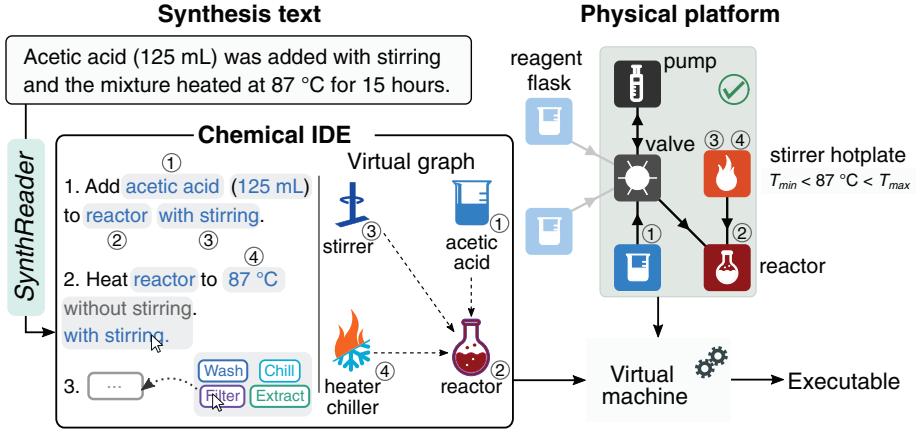


Fig. 2. System workflow. SynthReader converts synthetic procedures described using natural language into hardware-independent chemical code (XDL) within ChemIDE. The digitized procedure is represented and can be visually edited as natural language but internally stored as XML. Using the software, the user can amend the procedure or fix any potential translation errors. Once the correct chemical code has been produced, the virtual hardware used in the procedure is mapped to the user's physical platform. Compilation then combines the code with the mapped graph in a virtual machine to produce a hardware-specific executable suitable for immediate execution on the target platform. T_{max}, maximum temperature; T_{min}, minimum temperature.

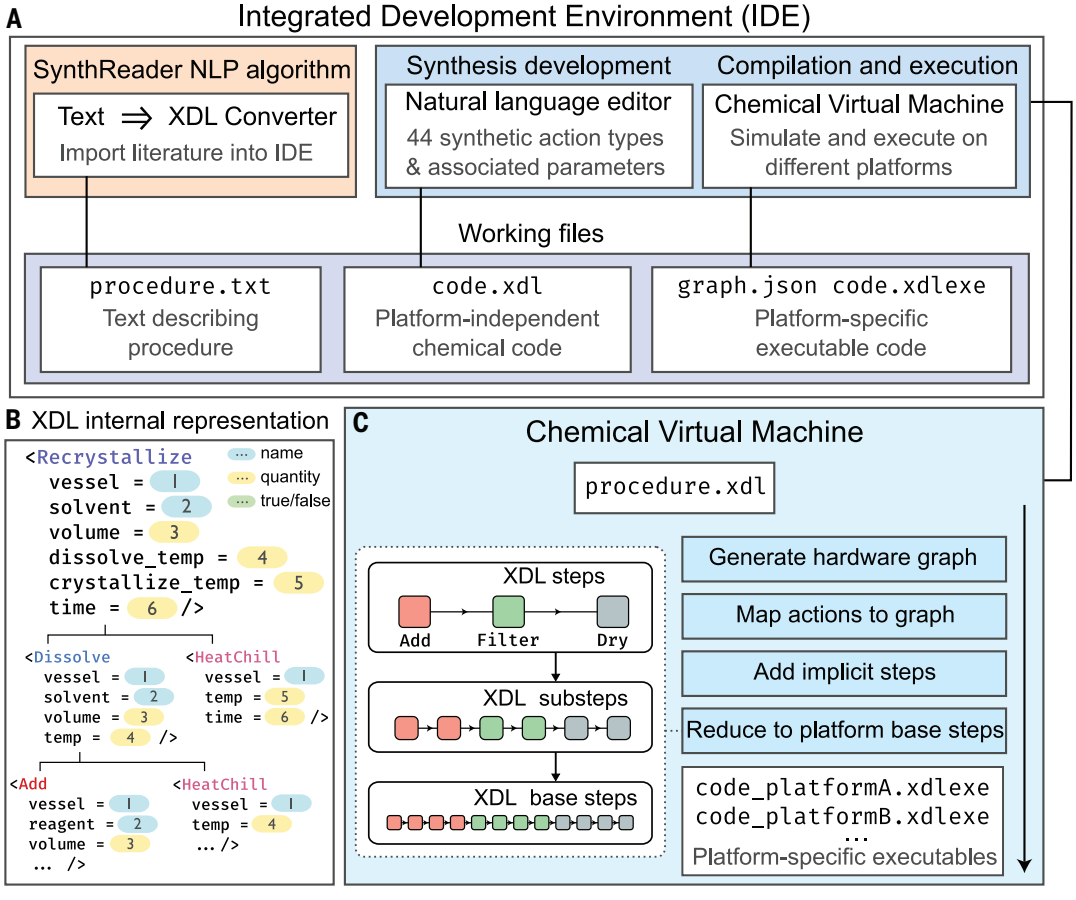


Fig. 3. An overview of our system's architecture and operation. (**A**) ChemIDE provides a visual user interface in which chemical code can be edited by using natural language. Existing literature procedures can be imported into this environment by using the SynthReader NLP algorithm, then inspected and enhanced by the user before being compiled to execute on a specific hardware target by the virtual machine. (**B**) XDL's internal XML-based representation propagates process information from steps to substeps. (**C**) A multi-step process within the virtual machine maps chemical synthesis steps to the relevant hardware modules within the target platform and recursively expands each step until it is reduced to basic hardware operations understood by the hardware.

and run on the target hardware to carry out the automated synthesis (Fig. 2).

The synthesis procedure, once written using our standard chemical programming language, provides a universal and hardware-independent way of digitizing chemical synthesis. However, there must also be a way of easily transferring syntheses written in natural language into code without programming knowledge or duplication of effort, while leveraging the expertise of the synthetic chemist. To do this, our system includes a chemical integrated development environment (ChemIDE) that facilitates importing literature procedures using a natural language processing (NLP) algorithm called SynthReader. In this context, other groups have recently applied NLP-based text mining to unstructured data from chemical synthesis texts to extract synthesis actions for both organic and inorganic reactions. This has been demonstrated by using pattern-matching techniques as well as machine learning (*20*–*22*).

Although these approaches are useful for mining vast literature datasets, we needed a system that could output a machine-readable representation of a procedure with sufficient process details to unambiguously execute the procedure on an automated platform. This goes beyond simply tagging chemical entities found in literature procedures, as it also requires a structured declaration of the location of different reagents throughout the procedure, inference of implicit process details such as separation of phases and reflux temperatures, and a development environment in which the expert chemist can unambiguously edit the output. SynthReader achieves this goal by tagging text with relevant entities, converting the tagged text to a list of actions, and then adding implicit process information and concrete reagent locations and outputting the procedure in the XDL format, which contains all of the necessary information to unambiguously execute the procedure on an automated platform. We have demonstrated the efficacy of this approach experimentally by converting literature syntheses to XDL using SynthReader and synthesizing the target molecules by executing the produced XDL.

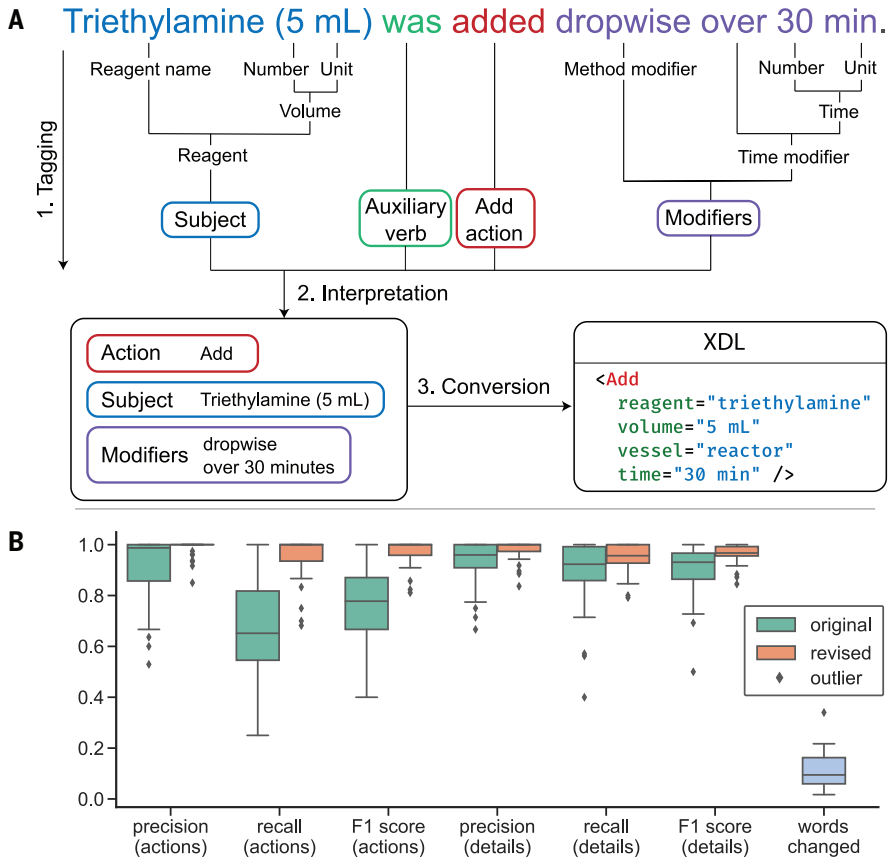


Fig. 4. Overview of SynthReader operation and performance. (**A**) Overview of the process by which text is converted to XDL. First, the text is hierarchically tagged by pattern matching. Pattern matching is then used again to extract all actions from the labeled text with their accompanying subjects and modifiers. Finally, the extracted actions are converted to XDL. The example text here contains only one action, but the system can handle multiple actions in one sentence. (**B**) Accuracy statistics for the latest version of SynthReader measured against 42 literature procedures before and after making revisions to the text. The final column shows the fraction of words in the procedure text that were modified in the process of editing. The lower and upper edge of each box represent the first and third quartile of values. The lower and upper horizontal lines (whiskers) represent the lowest and highest data points within 1.5 times the interquartile range of the box edges. Outliers (data points outside this range) are represented by diamonds. For benchmarking details, see supplementary materials section 2.5.

Design and implementation of the system

The key observation underlying our system is that any synthetic step is composed of a connected series of processes (add, heat, filter, etc.). Building on this observation, our system integrates the following components to realize automated synthesis from the literature: (i) a markup language capable of representing these extracted chemical processes and combining them in a context in which they can be executed as a chemical synthesis; (ii) an IDE allowing nonprogrammers to easily edit this representation of a chemical synthesis; (iii) a tool to automatically import existing procedures into the IDE, directly from the literature; and (iv) a virtual machine capable of transforming these chemical processes into basic operations that can be directly executed on a given automated platform. The integration of these parts is shown in Fig. 3. Below, we describe each component.

The XDL markup language was created to describe chemical synthesis in a robust, machine-readable way. The representation of a chemical synthesis as a sequence of discrete operations in XDL is the bridge between SynthReader, ChemIDE, and the virtual machine and the physical hardware operations necessary to perform that synthesis. The core of the XDL language is the XDL step. Each high-level step has a name and associated properties, and these steps and properties define the standard by which chemical syntheses can be described. Examples of top-level steps implemented in XDL are "separate," "evaporate," and "add," and a total of 44 such top-level steps have been implemented so far. These steps can be combined in a linear sequence, or branched sequences can be created and executed concurrently by using the asynchronous capabilities of XDL.

A disparate list of actions alone would hardly be sufficient for automated execution. XDL is notable for also providing the necessary experimental context, a stateful model of hardware and associated chemicals at every point in time. We designed XDL to capture information about synthetic procedures at multiple levels of abstraction, thereby allowing processes to be specified at a high level directly comparable to published methods sections but

executed through low-level hardware instructions. From this representation, a graph specification of the physical platform required can be inferred. This is achieved by producing a set of hardware requirements based on the procedure, such as required hardware modules and their connectivity as well as any reagents and temperatures involved. In this respect, XDL is distinct from common chemical data interchange formats such as CML (Chemical Markup Language) in that it provides a complete executable abstraction of chemical synthesis.

As most synthetic chemists are not experienced in programming, we created ChemIDE to facilitate user-friendly editing of XDL procedures. Similar to IDEs used by software developers, this environment helps the chemist adjust the procedure, add new operations to the procedure, and use the full capabilities of the software presented here within a graphical user interface. ChemIDE allows the user to program chemical synthesis without any coding knowledge by representing each step in natural language. This means that each step is shown to the user as a simple English sentence, highlighting words and phrases that they can edit, with input options and validation built in. For example, a dry step might show "Dry contents of **reactor** for **1 h**," in which "**reactor**" and "**1 h**" are editable, allowing the user to edit the vessel being dried and drying time. Changes made to these editable sentences are concretely linked to the underlying XDL representation, which can also be viewed in ChemIDE. Thus, users with no programming experience can interactively resolve ambiguities in the original text or amend any missing or implicit steps or process variables (movie S1).

The SynthReader NLP algorithm can extract sequences of processes from synthesis descriptions and represent them in XDL. Many recent advances in NLP have relied on the use of machine learning with large datasets of labeled text (*23*). However, to the best of our knowledge, there is no large dataset of labeled synthetic protocols in which the labels correspond to the final list of actions and details required for this application; therefore, we decided to build SynthReader as a domain-specific algorithm with predefined rules. This is a viable design thanks to the regular, almost mechanical language in which synthetic protocols tend to be written (*24*).

The text-to-XDL algorithm used in SynthReader was structured around three phases: (i) tagging [using a similar process to that used in previously published work, such as the ChemicalTagger (*22*)], (ii) interpretation, and (iii) conversion (Fig. 4A). In the tagging phase, different parts of the text are assigned labels, such as reagent names, volumes, or temperatures. This is achieved by using pattern-matching techniques. For example, one of the patterns used for finding solutions is "a solution of **Reagent** in **Reagent**" in which "**Reagent**" represents a phrase previously tagged as a reagent. An example phrase that would be matched by this pattern is "a solution of Oxone (181.0 g, 0.29 mol, 1.3 equiv) in deionized water (650 mL)." SynthReader contains a total of 16,582 patterns, some hard-coded, most programmatically generated from smaller hard-coded fragments.

Users are free to add, edit, and remove steps in the procedure as they wish; however, if trying to run a procedure from the literature, it is cumbersome to add every step individually. To conveniently import a literature procedure for further editing, we designed an NLP algorithm that automatically converts synthesis text to XDL. One component of the tagging process in which simple pattern matching is insufficient is the tagging of reagent names, as compiling an exhaustive list of these is not feasible. To tag reagent names, we use pattern matching with a database of common reagent names and then use a naïve Bayes classifier to determine whether candidate phrases are reagent names, excluding some specific phrases. The classifier uses probabilities that certain two-, three-, and four-letter word fragments are contained in a reagent name, calculated by using reagent names from the Reaxys database and non–reagent-name text from the Brown English language corpus. In the interpretation phase, pattern matching is again used to find common sentence structures in the tagged text and extract actions described by these sentences, along with subjects of the actions and any action-modifying phrases, resulting in a list of actions with associated information. The final conversion stage takes the list of actions produced by the interpretation stage, standardizes the details associated with each action, and converts every action to one or more XDL steps, producing a final XDL file. XDL can track the movement of reagents throughout the procedure and SynthReader can combine this information with a built-in table of physical properties to deduce process variables such as reflux temperatures and separation phases automatically. The entire text-to-XDL process is visualized for an example sentence in Fig. 4A; however, the same process is extendable to multiple paragraphs of text. SynthReader was designed with the goal of providing an accurate translation of any given synthetic protocol. However, we acknowledge that because of the flexibility of natural language, there will always be cases where the algorithm fails and information is lost, and

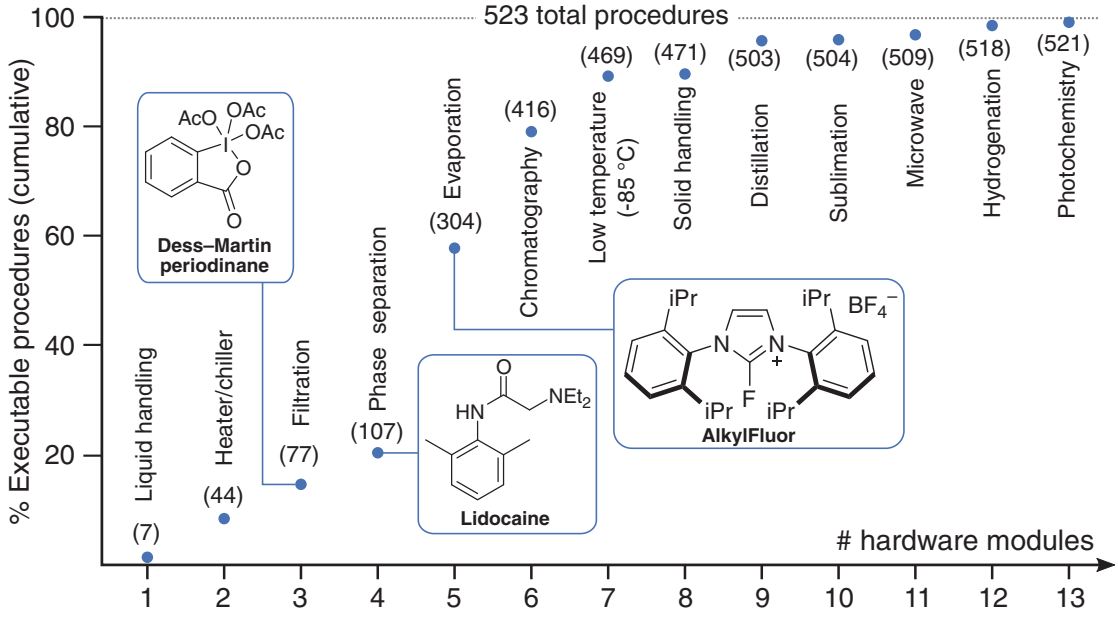


Fig. 5. Automatic analysis of the hardware requirements of literature procedures using SynthReader. The dataset used consisted of every procedure from *Organic Syntheses*, volume 77 onward (559 procedures), that was analyzed by SynthReader without a fatal error (523 procedures) (*9*, *38*). The cumulative number of executable procedures with the successive addition of each hardware module is shown in parentheses. OAc, acetate; Et, ethyl; iPr, isopropyl.

thus, the output of any such algorithm cannot be blindly trusted. ChemIDE makes it trivial to spot and amend any such omissions or errors. Figure 4B shows the results of benchmarks measuring SynthReader's ability to recover synthetic actions and associated details from literature procedures.

Once users have imported a procedure into ChemIDE using SynthReader and edited it to their satisfaction, the resulting chemical code can be compiled and executed on various hardware platforms. This functionality, enabled by a chemical virtual machine, combines the hardware-independent XDL representation with the graph description of the target hardware and automatically associates each step with the hardware modules required for execution.

This means that one XDL file, without alteration, can be compiled and executed on multiple different platforms as long as the hardware can execute the necessary base steps. Otherwise, any missing hardware modules or modules that do not meet the required specifications are reported.

The hardware graph can be automatically generated from a template graph of a specific platform or, alternatively, from the built-in default template. The hardware requirements are checked for compatibility with the template. If they are compatible, the template graph is altered to produce a procedure graph capable of executing the procedure. If they are incompatible, the user has the option to update either the template graph or the procedure. To produce platform-specific code, the virtual machine expands each XDL step in the procedure into a tree data structure, as shown in Fig. 3C. The leaves of this tree are hardware-specific unit operations, which are directly executable on the robot. Every step effectively has a decision tree that decides, on the basis of the parameters, which lower-level steps should be executed to achieve the specified behavior. Details (e.g., vessel names, process variables) assigned to a top-level step such as "Recrystallize" are propagated to those of the appropriate substeps (heat, stir, cool). For example, recrystallization time (variable 6, Fig. 3B) is correctly attributed to the cool-down period after the dissolution of a solid.

In addition to reducing generic synthetic steps to hardware-specific unit operations, the virtual machine is responsible for supplying the necessary implicit steps. These are operations that are specific to the targeted execution platform rather than intrinsic to the chemical process. In the case of the Chemputer, for instance, residual chemicals must be removed from the liquid backbone by using an appropriate cleaning solvent regularly to prevent cross-contamination. The placement of these cleaning steps and the type of solvent used for cleaning is algorithmically determined from the sequence of operations in the synthetic protocol and a series of chemical compatibility rules. The automatic addition of these implicit steps makes digitizing chemical synthesis much simpler as the user only needs to think about the synthesis as they would on the bench, not the operation of the platform. At the moment, the high-level steps included with XDL describe common synthetic operations such as "Evaporate" and "Separate." In the future, XDL steps can be devised to express

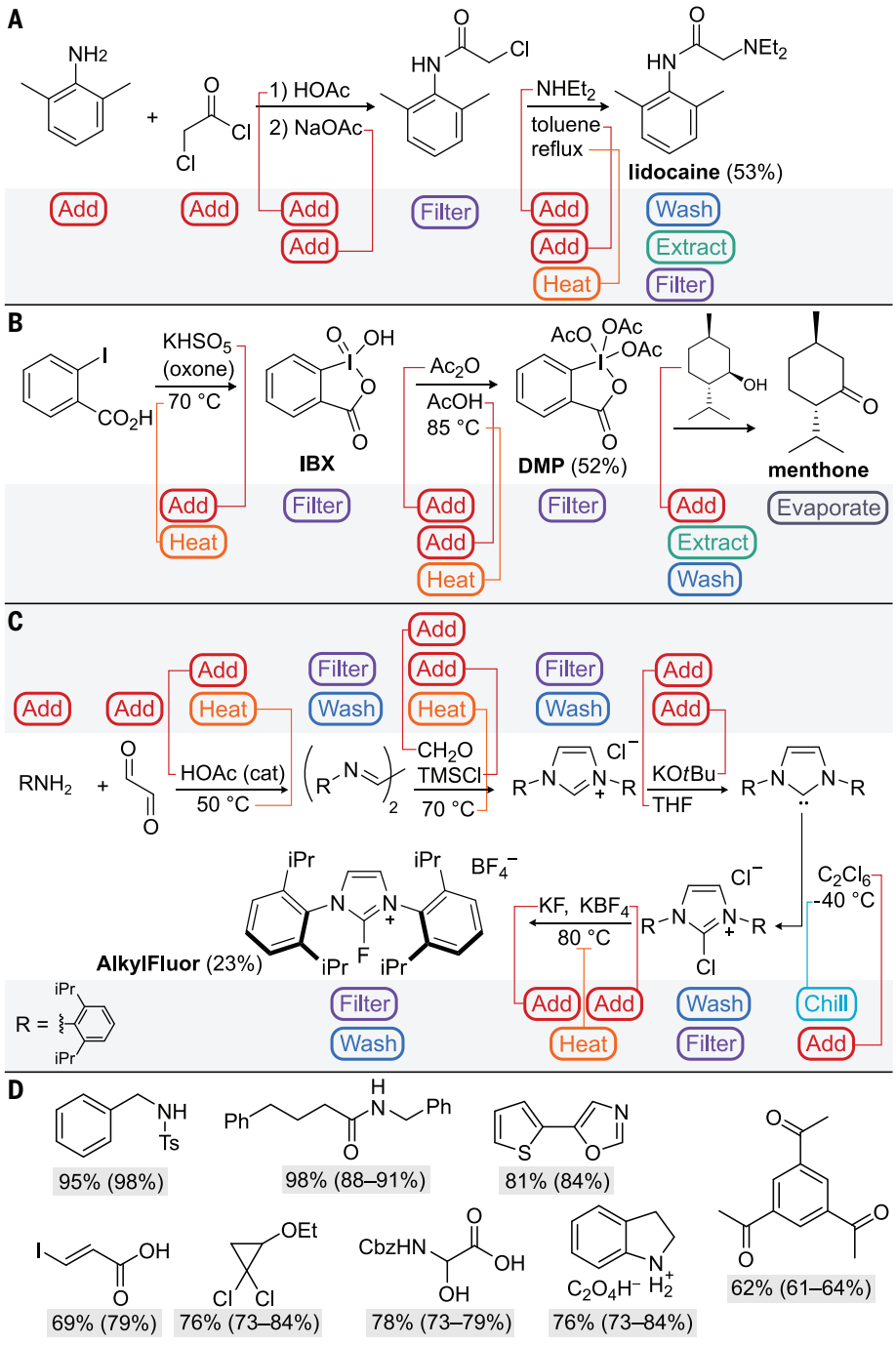


Fig. 6. Chemical schemes and the corresponding abstract chemical processes for Chemputer-implemented syntheses. (**A**) Lidocaine. (**B**) Menthone. (**C**) AlkylFluor. Published synthetic procedures in the literature make these abstract processes concrete by providing a textual representation and supplying experimental details. (**D**) Other molecules automatically synthesized from literature procedures by using our system (literature yields are shown in parentheses for reference). See supplementary materials section 6 for details. Cbz, benzyloxycarbonyl; tBu, tert-butyl; THF, tetrahydrofuran; TMS, trimethylsilyl; Ts, toluenesulfonyl; Ph, phenyl.

higher-level concepts. For instance, we envision a library of XDL steps to represent name reactions, such as "SuzukiCoupling" and "DessMartinOxidation." XDL also provides support for asynchronously executed steps and steps that execute dynamically on the basis of live feedback from analytical devices. These capabilities mean that branched syntheses and more advanced laboratory techniques such as adding until a color change occurs or maintaining a certain temperature during an addition will be possible by using the XDL framework.

As the most mature hardware target available to us, the Chemputer was used to execute the syntheses in the experimental section below. However, because the chemical code is written against a chemical virtual machine rather than any specific hardware, our system can use any hardware platform with a batch synthesis architecture and an open application programming interface as a first-class target, as demonstrated by the port to a second platform (supplementary materials). The ability to target diverse robotic systems is a boon to the viability of these proliferating platforms, as it ensures that digitized synthetic knowledge is not tied to specific hardware. Furthermore, as these platforms mature and add support for more hardware, they become capable of running a larger subset of published chemical syntheses. We simulate this scenario in Fig. 5 by using SynthReader to parse and analyze the hardware requirements of 523 literature procedures. For instance, ~60% of the procedures surveyed can be executed by using the most basic six modules: the addition of the low-temperature module raises this figure by 30%. The universality of our paradigm thus extends beyond currently available hardware.

Experimental validation of the system

We have used our approach to automatically execute 12 literature procedures on the Chemputer without any additional programming or hardware changes (*25*). To exemplify the process, we will detail the synthesis of three compounds here: lidocaine, the Dess-Martin periodinane (DMP), and AlkylFluor. Lidocaine is used as a local anesthetic and to treat arrhythmia and epilepsy (*26*). The literature procedure that we consulted for the synthesis of lidocaine (*27*) describes a simple two-step process involving the formation of an α-chloroamide intermediate and its subsequent nucleophilic substitution reaction with diethylamine. These steps map in a straightforward fashion to the process diagram illustrated in Fig. 6A. We fed the unmodified procedure for the synthesis of lidocaine to our system to run on the Chemputer.

On the basis of the procedure described by the XDL file, the Chemputer operated the backbone pumps and valves to automatically transfer acetic acid solvent to the jacketed filter module—which the system had identified as a suitable reactor—followed by 2,6-dimethylaniline, chloroacetyl chloride, and saturated aqueous sodium acetate. During the process, our system correctly found points at which two chemicals are mixed and controlled stirring appropriately to ensure proper mixing. On the basis of the XDL instructions, the Chemputer then performed a filtration and routed the filtrate into a waste container. The next step was executed similarly by adding diethylamine and toluene solvent, heating the jacketed filter up to reflux by using a circulation chiller to effect the substitution reaction, and using the liquid-liquid separation module to perform an acidic extraction with an aqueous hydrochloric acid solution. The detection of the liquid-liquid phase boundary is facilitated by a conductivity sensor exploiting the high conductivity of the aqueous phase compared with the organic phase (*18*). Finally, lidocaine was precipitated from the aqueous solution in the jacketed filter by addition of sodium hydroxide solution, filtered, and dried in the jacketed filter under vacuum. Automated execution of the literature procedure in this manner produced lidocaine in 53% yield. The chemist is responsible for ensuring that the experimental setup matches the hardware graph. For the Chemputer, this preparation step involves connecting reagent bottles for each reagent, solvent, and piece of glassware to the correct position on the liquid backbone.

The second example, the DMP, is a versatile oxidation reagent that is prized for its specificity and functional group tolerance, despite its relatively high price and moisture sensitivity. Both the preparation and use of this reagent, as well as its precursor, 2-iodoxybenzoic acid (IBX), have been the subject of recent reproducibility debates (*28*–*32*). The synthesis of DMP is a prime candidate for automation, as it is often prepared fresh, a process that is time-consuming and bears a non-negligible risk of explosion due to impurities. We applied SynthReader to three separate literature procedures: (i) a modern synthesis of IBX by oxidation of 2-iodobenzoic acid with aqueous oxone (potassium peroxymonosulfate) (*33*); (ii) acetylation of IBX by using acetic anhydride to form DMP (*29*); (iii) oxidation of menthol to menthone by using DMP to determine its activity (*34*). The resulting XDL files, representing the entire process, were joined together and executed on a Chemputer, giving DMP product (in 52% overall yield), which subsequently showed quantitative oxidizing activity when reacted with excess menthol (Fig. 6B) (*35*).

To demonstrate that our text-to-molecule machinery is not limited to short syntheses, we also converted text describing the five-step synthesis of the fluorinating agent AlkylFluor to XDL and executed it on a Chemputer platform, obtaining AlkylFluor in 23% overall yield (75% average stepwise yield) (Fig. 6C) (*36*, *37*). Though the proliferation of automated chemical synthesis systems holds much promise, differences in the instruction set provided by the various platforms make it impractical to write portable chemical code. The approach described here is hardware-universal, meaning the software can execute a given synthetic procedure on any hardware platform as long as the platform provides the hardware instructions necessary to express the processes described in the procedure. To demonstrate this, we successfully executed the literature synthesis of the polyoxometalate $(C_2H_8N)_8Na_3[W_{19}Mn_2O_{61}Cl(SeO_3)_2(H_2O)_2]Cl_2 \cdot 6H_2O$ on a bespoke high-throughput chemistry robot used in our group that relies on a completely different instruction set to the Chemputer. Because this robot lacks hardware modules for heating and filtering, only a subset of the procedures executable on the Chemputer will run on it. Any unsupported actions are automatically flagged by our system when encountered.

In summary, we present an important step toward the goal of automating all aspects of synthetic chemistry—from text to molecule—with the introduction of an abstraction that allows the digitization of chemical synthesis. Although it is not yet possible to convert all the literature with our system without some manual intervention, ChemIDE allows the user to correct errors by easily inspecting the original text and confirming translation into the process steps. In the future, we will automate this verification step using a chemical autocorrect function. The NLP capabilities of SynthReader are comparable to the current state of the art (supplementary materials section 2.6) and can be easily augmented with new rules, as the design is deterministic. Real-time feedback from analytical instruments can be used to confirm that processes proceed as described by the XDL, making the system adaptive and fault tolerant. In addition, it is possible to interface the IDE with other NLP engines or hardware compilation targets, further increasing the possibilities to interface our system with any chemical robot and the broader literature.

REFERENCES AND NOTES

1. K. C. Nicolaou, *Proc. R. Soc. A.* **470**, 20130690 (2014).
2. J. J. Irwin, B. K. Shoichet, *J. Chem. Inf. Model.* **45**, 177–182 (2005).
3. O. J. Plante, E. R. Palmacci, P. H. Seeberger, *Science* **291**, 1523–1527 (2001).
4. J. Li *et al.*, *Science* **347**, 1221–1226 (2015).
5. H. Saneii, M. R. Boroomand, M. R. Ferriell, U.S. Patent 5,746,982 (1998).
6. I. W. Davies, *Nature* **570**, 175–181 (2019).
7. Reaxys (Elsevier, 2019); www.reaxys.com
8. S. Barata-Vallejo, S. Bonesi, A. Postigo, *Org. Biomol. Chem.* **14**, 7150–7182 (2016).
9. M. M. Morgan, W. E. Piers, *Dalton Trans.* **45**, 5920–5924 (2016).

10. D. Xu *et al.*, *J. Am. Chem. Soc.* **142**, 5785–5792 (2020).
11. B. Mahjour, Y. Shen, W. Liu, T. Cernak, *Nature* **580**, 71–75 (2020).
12. V. Tshitoyan *et al.*, *Nature* **571**, 95–98 (2019).
13. P. Murray-Rust, *Nature* **451**, 648–651 (2008).
14. R. G. Bergman, R. L. Danheiser, *Angew. Chem. Int. Ed.* **55**, 12548–12549 (2016).
15. C. W. Coley, R. Barzilay, T. S. Jaakkola, W. H. Green, K. F. Jensen, *ACS Cent. Sci.* **3**, 434–443 (2017).
16. C. Mellot-Draznieks, J. Dutour, G. Férey, *Angew. Chem. Int. Ed.* **43**, 6290–6296 (2004).
17. J. Ghasemi, S. Saaidpour, *Chem. Pharm. Bull.* **55**, 669–674 (2007).
18. S. Steiner *et al.*, *Science* **363**, eaav2211 (2019).
19. T. Bray *et al.*, "Extensible Markup Language (XML)" (W3C, 2006).
20. A. C. Vaucher *et al.*, ChemRxiv [Preprint] (2020); https://doi.org/10.26434/chemrxiv.11448177.v2.
21. E. Kim *et al.*, *J. Chem. Inf. Model.* **60**, 1194–1201 (2020).
22. L. Hawizy, D. M. Jessop, N. Adams, P. Murray-Rust, *J. Cheminform.* **3**, 17 (2011).
23. T. Young, D. Hazarika, S. Poria, E. Cambria, *IEEE Comput. Intell. Mag.* **13**, 55–75 (2018).
24. P. Corbett, P. Murray-Rust, in *International Symposium on Computational Life Science*, M. R. Berthold, R. C. Glen, I. Fischer, Eds. (Springer, 2006), pp. 107–118.
25. Materials and methods are available as supplementary materials.
26. L. A. Slaughter, A. D. Patel, J. L. Slaughter, *J. Child Neurol.* **28**, 351–364 (2013).
27. T. J. Reilly, *J. Chem. Educ.* **76**, 1557 (1999).
28. R. E. Ireland, L. Liu, *J. Org. Chem.* **58**, 2899 (1993).
29. R. K. Boeckman Jr., P. Shao, J. J. Mullins, *Org. Synth.* **77**, 141 (2000).
30. L. Kürti, B. Czakó, *Strategic Applications of Named Reactions in Organic Synthesis: Background and Detailed Mechanisms* (Elsevier Academic, 2005).
31. P. J. Stevenson, A. B. Treacy, M. Nieuwenhuyzen, *J. Chem. Soc. Perkin Trans. 2* **1997**, 589–592 (1997).
32. G. Tojo, M. Fernández, *Oxidation of Alcohols to Aldehydes and Ketones: A Guide to Current Common Practice* (Springer, 2010).
33. M. Frigerio, M. Santagostino, S. Sputore, *J. Org. Chem.* **64**, 4537–4538 (1999).
34. N. A. Reed, R. D. Rapp, C. S. Hamann, P. G. Artz, *J. Chem. Educ.* **82**, 1053 (2005).
35. The original literature procedure (*29*) uses a similar validation method, except with excess benzyl alcohol. Because detailed steps were not provided, we used a similar literature procedure (*34*) based on the oxidation of menthol instead. Our method of validation was not aimed at approximating how DMP is used in practice.
36. N. W. Goldberg, X. Shen, J. Li, T. Ritter, *Org. Lett.* **18**, 6102–6104 (2016).
37. P. Tang, W. Wang, T. Ritter, *J. Am. Chem. Soc.* **133**, 11482–11484 (2011).
38. Royal Society of Chemistry, *ChemSpider SyntheticPages* (2019); http://cssp.chemspider.com/.
39. ChemIDE; https://croningroup.gitlab.io/chemputer/xdlapp/
40. S. H. M. Mehr, M. Craven, A. I. Leonov, G. Keenan, L. Cronin, Code for XDL, Synthreader and ChemputerXDL, Zenodo (2020); doi:10.5281/zenodo.3955103.
41. S. H. M. Mehr, M. Craven, A. I. Leonov, G. Keenan, L. Cronin, Benchmarking results and the XDL XML schema, Zenodo (2020); doi:10.5281/zenodo.3955107.

ACKNOWLEDGMENTS

We thank D. Angelone and S. Zalesskiy for contributing to Chemputer hardware development, A. Hammer for fruitful synthetic discussions, D. Salley for help with implementing XDL on Chemobot, and P. J. Kitson for contributing ideas and helping with manuscript preparation. We also thank M. Symes, N. Bell, S. Rohrbach, R. Hartley, and L. Wilbraham for comments on the manuscript. **Funding:** We thank the following funders: EPSRC (grant nos. EP/H024107/1, EP/J00135X/1, EP/J015156/1, EP/K021966/1, EP/K023004/1, EP/L023652/1) and ERC (project 670467 SMART-POM). This research was developed with funding from the Defense Advanced Research Projects Agency (DARPA). The views, opinions, and/or findings expressed are those of the author and should not be interpreted as representing the official views or policies of the Department of Defense or the U.S. Government. **Author contributions:** L.C. conceived the initial idea and designed the architecture, the standard, and the ontology. M.C. developed the XDL/ChemIDE and SynthReader software. S.H.M.M. helped develop the software to run the system and demonstrated the SynthReader with lidocaine, DMP, and two additional syntheses. A.I.L. demonstrated the SynthReader with AlkylFluor and six additional syntheses. G.K. helped develop and maintain the Chempiler software. The manuscript was written by L.C. together with M.C. and S.H.M.M. with input from all the authors. **Competing interests:** L.C. is the founder of DeepMatter Group PLC and Chemify Ltd., which aims to commercialize various aspects of the digitization of chemistry and synthesis using robotic platforms. L.C., M.C., S.H.M.M., and A.I.L. are listed as inventors on the UK patent 2006243.6, which describes this system. **Data and materials availability:** The ChemIDE app is available online (*39*). We have also made available the source code (*40*), as well as benchmarking results, literature survey data, and the XDL XML schema (*41*).

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/370/6512/101/suppl/DC1
Materials and Methods
Figs. S1 to S65
Tables S1 to S8
References (*42–49*)
Movie S1

16 April 2020; resubmitted 16 June 2020
Accepted 27 July 2020
10.1126/science.abc2986

REPORTS

SOLAR CELLS

Impact of strain relaxation on performance of α-formamidinium lead iodide perovskite solar cells

Gwisu Kim*, Hanul Min*, Kyoung Su Lee, Do Yoon Lee, So Me Yoon, Sang Il Seok†

High-efficiency lead halide perovskite solar cells (PSCs) have been fabricated with α-phase formamidinium lead iodide ($FAPbI_3$) stabilized with multiple cations. The alloyed cations greatly affect the bandgap, carrier dynamics, and stability, as well as lattice strain that creates unwanted carrier trap sites. We substituted cesium (Cs) and methylenediammonium (MDA) cations in FA sites of $FAPbI_3$ and found that 0.03 mol fraction of both MDA and Cs cations lowered lattice strain, which increased carrier lifetime and reduced Urbach energy and defect concentration. The best-performing PSC exhibited power conversion efficiency >25% under 100 milliwatt per square centimeter AM 1.5G illumination (24.4% certified efficiency). Unencapsulated devices maintained >80% of their initial efficiency after 1300 hours in the dark at 85°C.

The lead halide perovskite structure is represented by the general chemical formula $APbX_3$, where A denotes an organic ammonium or inorganic cation such as methylammonium (MA^+), formamidinium (FA^+), or Cs^+; and X denotes a halide (I^-, Br^-, or Cl^-). Corner-shared PbX_6 octahedra form a cuboctahedral cage to accommodate the A cation that satisfies the steric requirements and stabilize three-dimensional perovskite structure. The optoelectrical properties of $APbX_3$ perovskites are dominated by the inorganic lead halide lattice, while the cation in the A site contributes to the stabilization of the structure (*1*). Therefore, the bandgap of the perovskite depends primarily on halides and is relatively less affected by the A cation. Thus, the factors affecting device performance, such as dielectric properties (*2*) and distortion (*3*), can be easily controlled by the proper selection or combination of A cations without compromising the bandgap.

Although MA^+, FA^+, and Cs^+ are the most well studied A cations for the $APbI_3$ perovskite with the narrowest bandgap capable of absorbing a wide range of sunlight (*4–7*), it is difficult to use $APbI_3$ with FA^+ and Cs^+ cations in a solar cell because it is thermodynamically unstable at room temperature and crystallizes into the δ-phase (a very wide bandgap nonperovskite) from the α-phase (*6*, *8*, *9*). Although the structural stabilization of lead halide perovskite with various cations was explained by the Goldschmidt tolerance (*t*) factor [$t = r_A + r_I = \sqrt{2}(r_{Pb} + r_I)$, where r_A, r_{Pb}, and r_I are the radii of the A cation, Pb cation, and I anion, respectively] (*1*, *10*), it can be achieved by distorting PbI_6 octahedron or by N–H···I hydrogen bonding with organic cations (*1*). Currently, high-efficiency perovskite solar cells (PSCs) are predominately fabricated with $FAPbI_3$, for which phase stabilization is achieved by controlling *t* through the synergistic entropic effect of mixing MA^+, Cs^+, and Br^- (*5*, *11–14*). However, as the composition of $FAPbI_3$ is changed to stabilize the perovskite structure of the α-phase, the bandgap may widen. Moreover, the local strain can increase because of the distortion of the ideal structure, including tilting, deformation, expansion, and shrinkage of the octahedral network (*15–17*).

The residual strain in halide perovskites substantially affects the performance of PSCs by reducing structural stability (*18*, *19*), decreasing

Department of Energy Engineering, School of Energy and Chemical Engineering, Ulsan National Institute of Science and Technology, 50 UNIST-gil, Eonyang-eup, Ulju-gun, Ulsan 44919, Korea.
*These authors contributed equally to this work.
†Corresponding author. Email: seoksi@unist.ac.kr

the carrier mobility (*20*), and suppressing atomic vacancies (*21*). Lattice strain was directly related to increased defect concentration and nonradiative recombination, which is associated with the efficiency (*22*). Huang and colleagues reported the reduction of residual strain by thermal annealing of perovskite films produced by a solution process enhanced the intrinsic stability of the films under illumination by decreasing ion migration (*23*). Chen and colleagues (*24*) demonstrated that α-$FAPbI_3$ can be stabilized by growing a single crystal on the substrate and that the bandgap and hole mobility changed when a compressive strain was applied.

Alloyed mixed-halide perovskites are nonuniform because of the phase separation through segregation of ions, which causes local lattice mismatch and leads to residual deformation. The efficiency was increased by improving carrier transport and extraction at the interface of the perovskite absorber and hole transport material (HTM) by controlling the vertically strained gradient with flipped annealing method (*24*). Xue *et al*. (*25*) observed that the stability considerably improved with added compressive strain by using an HTM with a high coefficient of thermal expansion for the perovskite film. Tsai *et al*. (*26*) reported improved PSC device performance under continuous light illumination owing to the uniform lattice expansion in the perovskite film.

The strain in $APbI_3$ perovskite can also be reduced by substituting some of the Pb^{2+} ions with isovalent Cd^{2+} ions of a small ionic radius (*21*, *27*). As a result, both the efficiency and stability improved with the relaxation of the local lattice strain. Similarly, it was shown for mixed $(FASnI_3)_{0.5}(MAPbI_3)_{0.5}$ PSCs that the addition of 2.5 mol % of Cs^+ ions led to the relaxation of the lattice strain, resulting in a lower concentration of defects, which in turn improved efficiency (*28*). Therefore, the strain engineering of lead halide perovskites has attracted attention as a method to further improve both the efficiency and stability of PSCs.

Recently, we reported the stabilization of α-$FAPbI_3$ by substituting FA^+ with the slightly larger methylenediammonium (MDA^{2+}) (*29*). Compared with stabilizing α-phase by adding MA^+ or Cs^+ (both of which have a smaller ionic radius than FA^+), the change in the bandgap resulting from MDA^{2+} substitution is very small (~0.01 eV), and a high short-circuit current (J_{sc}) was obtained with relatively better stability (retaining >90% of initial performance over 600 hours of irradiation). Nevertheless, substituting only MDA^{2+} cations with a larger ionic radius, or Cs^+ with a smaller ionic radius, than FA^+ can distort Pb–I–Pb bonds by tilting the PbI_6 octahedron. One of the most common strain-compensation strategies is to introduce larger and smaller ions together to reduce the local tensile and compressive strain in the perovskite lattice.

In this study, we used the dual substitution of FA^+ sites with MDA^{2+} and Cs^+ in the same molar ratio to relax the lattice strain of MDA-stabilized α-$FAPbI_3$. The alloyed $FAPbI_3$ with 0.03 mol fraction of both MDA and Cs cations effectively reduced the lattice strain and the trap density in the PSCs, resulting in the fabrication of PSCs with 24.4 and 21.6% certified

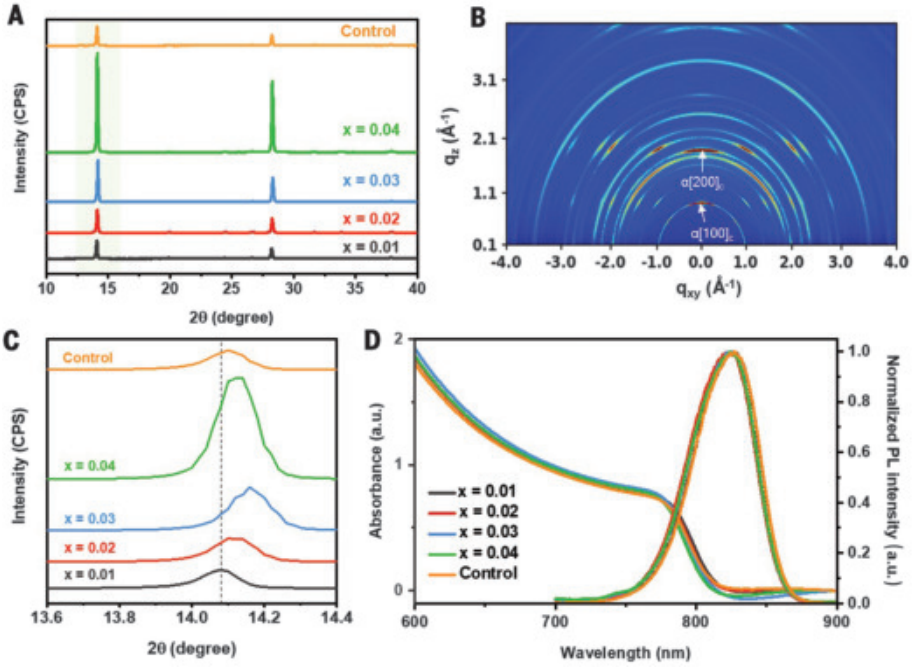


Fig. 1. Characterization of perovskite thin films deposited on mp-TiO_2 for the $(FAPbI_3)_{1-x}(MC)_x$ and control perovskite films. (**A**) XRD patterns of perovskite film. CPS, counts per second. (**B**) GIWAXS pattern for x = 0.04. (**C**) Magnified (100) plane diffraction peaks in the region shaded in green in (A). (**D**) UV-vis absorption and normalized PL spectra. a.u., arbitrary units.

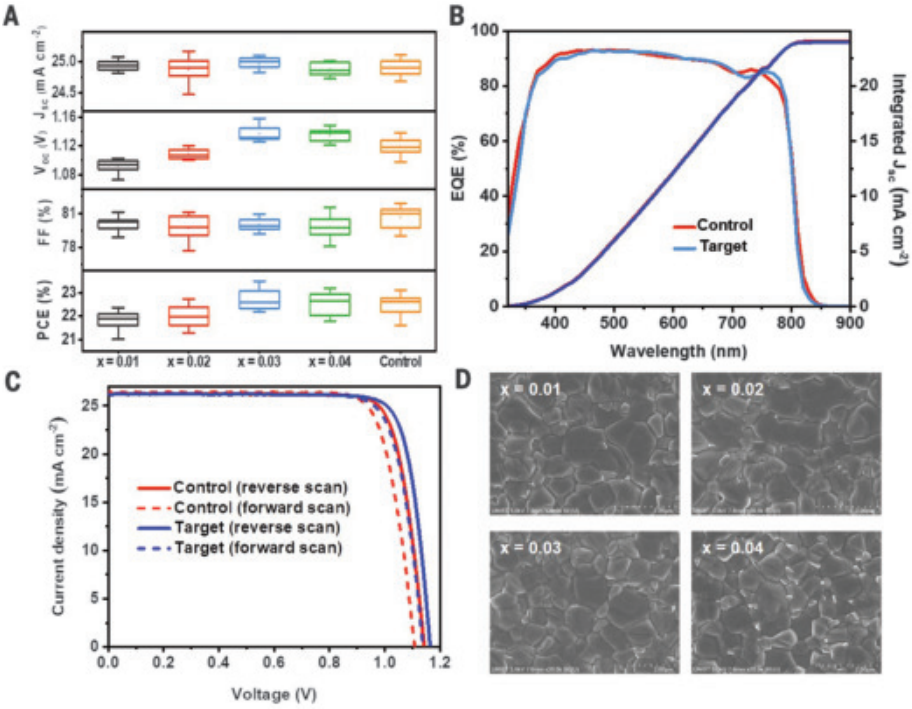


Fig. 2. Performance and surface morphologies of PSCs fabricated with x in $(FAPbI_3)_{1-x}(MC)_x$ and control perovskite films. (**A**) J_{sc}, V_{oc}, FF, and PCE statistics of 24 PSCs. (**B**) EQE curves of target and control PSCs. (**C**) *J-V* curves of target and control PSCs. (**D**) Surface SEM images of perovskite thin layers.

efficiency for small and large (1 cm by 1 cm) cells, respectively, through the suppression of nonradiative carrier recombination. Moreover, the corresponding devices exhibited superior thermal stability and maintained >80% of their initial efficiency after 1300 hours in storage at 85°C (at 15 to 25% relative humidity).

The $(FAPbI_3)_{1-x}(MC)_x$ perovskite thin films (with x = 0.01, 0.02, 0.03, or 0.04 mol fraction) were deposited by our solvent engineering process (*29*, *30*) using precursor solutions that dissolved the FAI and PbI_2 with the desired number of MDA^{2+} and Cs^+ cations. Here, MC is an abbreviation that means that MDA^{2+} and Cs^+ were mixed in equimolar amounts (i.e., MDA:Cs = 1:1). Figure 1A shows the x-ray diffraction (XRD) patterns of the samples with different x values for the perovskite thin layers prepared directly on a mesoporous-TiO_2 (mp-TiO_2) electron-transporting layer. In the XRD patterns, two dominant peaks could be seen at about 14° and 28°, which we assigned to the characteristic (001) and (002) crystal planes of the α-$FAPbI_3$ phase, and no δ-phase appeared at 11.6°.

As x increased in $(FAPbI_3)_{1-x}(MC)_x$, the diffraction intensity of the two peaks also increased, and no new peaks appeared, which was likely caused by preferred orientation during crystallization on the substrate from the comparison with XRD patterns of powders scraped from the films (fig. S2A). The average crystalline sizes and full width at half maximum of the diffraction peak at the (100) plane did not show a large variation with x and became smaller and broader with the shift of the diffraction angle to the low angle in powder case, owing to residual stress present in the films (fig. S2, B and C). Because crystallographic data measured from one-dimensional XRD typically only provide limited structural information, particularly for highly oriented perovskite films, we obtained two-dimensional grazing-incidence wide-angle x-ray scattering (GIWAXS) patterns to probe changes in the crystallographic orientations of the perovskite film as x was varied.

Distinct and relatively strong spots were observed in the ring patterns (fig. S1A) as x increased (i.e., increased substitution) in $(FAPbI_3)_{1-x}(MC)_x$. The GIWAXS pattern for x = 0.04 (Fig. 1B) exhibited a strong diffraction intensity and was similar to that of x = 0.03 (fig. S1C) but had slightly greater intensity and appeared more clearly as a mixture of scattered secondary spots and rings. The preferential crystal orientations for α-$FAPbI_3$ were observed out of the plane in the direction $[100]_c$ and $[200]_c$ (Fig. 1B, white arrows). We concluded that the diffraction intensity increased with increasing x because of the highly oriented crystal domains, not the improvements in crystallinity.

In addition, compared with $(FAPbI_3)_{0.972}(MDACl_2)_{0.038}$ (denoted as control), the peak position at ~14° gradually shifted to higher angles from 14.07° to 14.16° as x increased to 0.03, then it slightly decreased to 14.12° at x = 0.04 (Fig. 1C). In the same crystal, because the diffraction angle (2θ) reflected the expansion and contraction of the lattice, the diffraction angle could shift depending on the proportion of the relatively smaller Cs^+ and larger MDA^{2+} cations to the FA^+ cations. This result suggested that the incorporation of Cs^+ and MDA^{2+} into the lattice of $FAPbI_3$ formed a solid-state alloy. As shown in fig. S3, this shift in diffraction angle was most consistent with the change in the radius of cations calculated using the ionic radii of individual FA^+, Cs^+, and MDA^{2+}, except at x = 0.04 and when considering the FA^+ vacancies for the charge balance without also considering the insertion of Cl^- ions that may result from the bivalent MDA (*29*).

Inferring the deviation from this trend at x = 0.04, we expected the composition of the actual perovskite thin film to differ slightly from the composition of the precursor solution. Nevertheless, the change in optical properties with x was negligible. In Fig. 1D, the ultraviolet-visible (UV-vis) absorption spectra and normalized photoluminescence (PL) data with different x mole fraction in $(FAPbI_3)_{1-x}(MC)_x$ and control are compared. A slight blue-shift was observed in the absorption onset as x increased, however, the shift was very small compared with the composition in which a single Cs^+ cation was added to pure $FAPbI_3$ (*11*). The corresponding shifts are consistent with the PL emission peaks at 826, 825, 825, and 824 nm for x = 0.01, 0.02, 0.03, and 0.04, respectively, and the peak at 827 mm for the control.

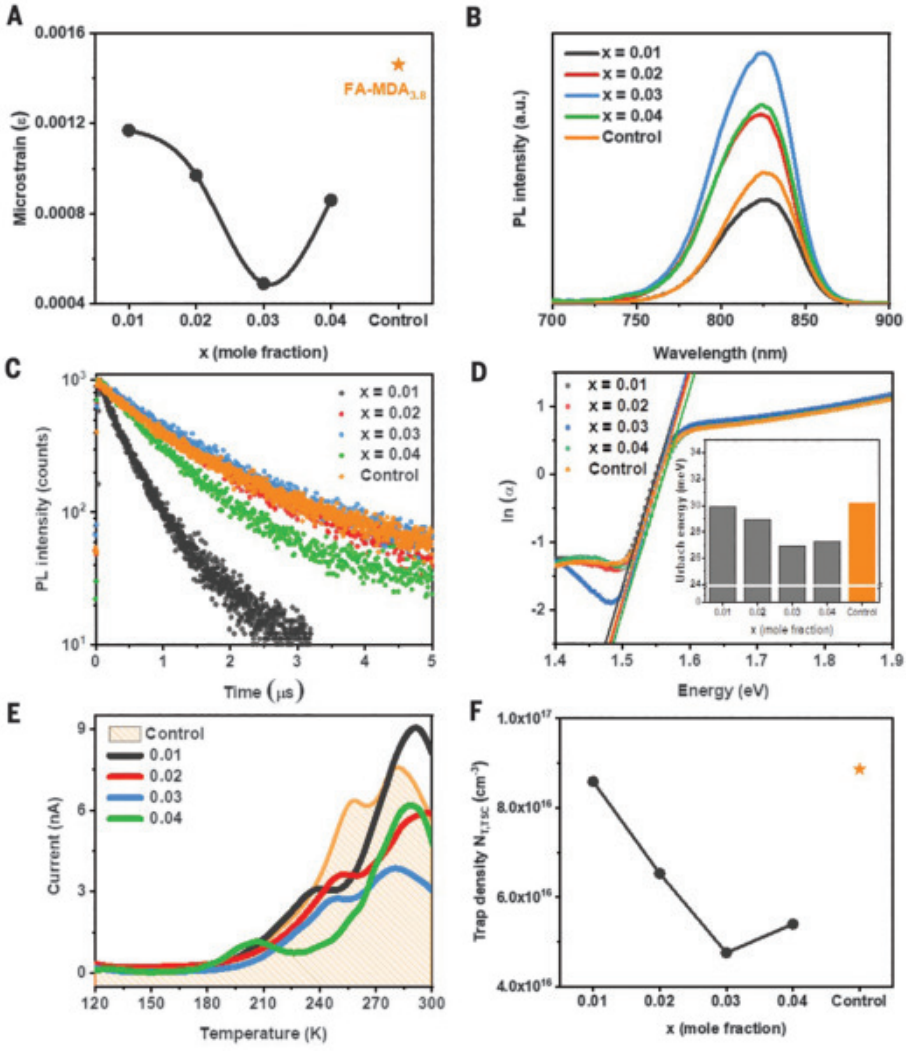


Fig. 3. Defect analysis of perovskite films deposited with x in $(FAPbI_3)_{1-x}(MC)_x$ and control. (**A**) Residual strain calculated in perovskites consisting of FTO/mp-TiO_2/perovskite. (**B**) Steady-state photoluminescence and (**C**) time-resolved photoluminescence spectra of films deposited on glass. (**D**) Logarithm of absorption coefficient (α) versus photon energy and (inset) Urbach energy calculated in perovskites consisting of FTO/mp-TiO_2/perovskite. (**E**) Thermally stimulated current spectra. (**F**) Trap density. The orange star in (A) and (F) indicates control perovskite.

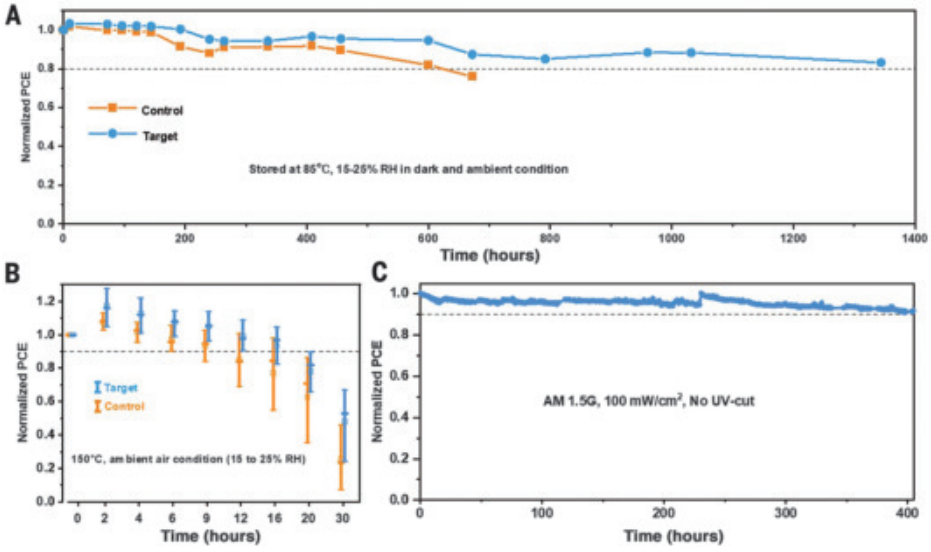


Fig. 4. Long-term stability tests. Comparison of the thermal stability at (**A**) 85°C and (**B**) 150°C of unencapsulated control and target PCSs with minimum, maximum, and average values for eight devices. (**C**) Maximum power point tracking measured with the encapsulated target device under full solar illumination (AM 1.5G, 100 mW/cm^2 in ambient conditions at 25° to 35°C) without a UV filter.

Figure 2A shows the variation in solar cell characteristics with different x in $(FAPbI_3)_{1-x}(MC)_x$ and control. The devices comprised multiple layers: glass/fluorine-doped tin oxide (FTO)/compact-TiO_2/thin mp-TiO_2/perovskite/2,2',7,7'-tetrakis[*N*,*N*-di(4-methoxyphenyl)amino]-9,9'-spirobifluorene (Spiro-OMeTAD)/Au. To confirm the effect of cosubstitution of MDA^{2+} and Cs^+ on solar cell performance, at least 24 devices in each category were fabricated and evaluated in one experiment (Fig. 2A). The statistical distribution of factors such as J_{sc}, open-circuit voltage (V_{oc}), and fill factor (FF) determining power conversion efficiency (PCE) showed that a relatively better performance was noticed for $x = 0.03$ (denoted as target) with a narrow distribution of J_{sc}, V_{oc}, FF, and PCE and an average PCE of 22.72 ± 0.45%.

In general, when the composition or coating process of the perovskite thin film is changed, the surface morphology may change and in turn change the efficiency. Thus, we imaged the surface morphology in $(FAPbI_3)_{1-x}(MC)_x$ and the control films using top-view scanning electron microscopy (SEM) (Fig. 2D). Regardless of the amounts of additives, all films showed a similar grain size without apparent pinholes on the surface. These results indicate that the incorporation with small amounts of Cs^+ and MDA^{2+} cations did not affect the morphological features of the perovskite layers, such as grain sizes and surface roughness.

For simplicity, the control and target were compared to find out why the efficiency improved by substituting the same amount of Cs^+ and MDA^{2+} in $FAPbI_3$. Figure 2B shows the external quantum efficiency (EQE) spectra for one of the representative control and target devices. The EQE onset of the target is slightly blue-shifted, which was consistent with the tuned bandgap (Fig. 1D). Nevertheless, as can be seen in the integrated J_{sc} (Fig. 2B), this small blue shift in the bandgap was not appreciable. The cross-sectional thickness of the control and target devices were compared by using SEM images (fig. S4). There was no noticeable difference in the thicknesses between two representative layers (415 nm for the control and 425 nm for the target). The similarities in J_{sc} implied that there was no substantial difference in the charge collection in the two comparison groups.

The current density-voltage (*J-V*) characteristics of the best-performing control and target devices in a reverse and forward bias sweep with antireflective coatings (fig. S5) on the device surface are compared in Fig. 2C. The J_{sc}, V_{oc}, and FF values calculated from the *J-V* curve of the target were 26.23 mA cm^{-2}, 1.168 V, and 82.15%, respectively; these factors correspond to a PCE of 25.17% under standard air mass (AM) 1.5 conditions, whereas the control exhibited a PCE of 24.48% with J_{sc} = 26.25 mA cm^{-2}, V_{oc} = 1.138 V, and FF = 81.95%. The PCE of the target devices shown in Fig. 2C was certified by an accredited laboratory (Newport, USA) using a quasi-steady-state (QSS) method. The stabilized PCE measured by QSS was 24.37% with J_{sc} = 26.17 mA cm^{-2}, V_{oc} = 1.162 V, and FF = 80.13% for the small cell (fig. S6), and 21.63% for the large cell (1 cm by 1 cm) (fig. S7). Within the scope of this study, the effect of compositional changes on J_{sc} was very limited, and most efficiency improvements were attributable to an increase in V_{oc}. As noted earlier, because control and target devices have almost the same bandgap, surface morphology, and thickness, the large increase in V_{oc} (low V_{oc} loss) arose from the changes inside the crystalline perovskites. The V_{oc} loss for x and control occurred at $x = 0.03$ (fig. S8).

Generally, V_{oc} loss is directly related to reduction in defect concentration and nonradiative losses (*31*–*34*). In this regard, there have been many reports that lattice strain in perovskites increases defect concentrations and nonradiative recombination (*22*, *24*, *28*, *35*). We estimated the variation in lattice strain of perovskite films with x in $(FAPbI_3)_{1-x}(MC)_x$ using the Williamson–Hall (WH) plot (individual plots with x are displayed in fig. S9) from the XRD patterns of Fig. 1A. As can be seen in Fig. 3A, the strain decreased as x increased from 0.01 to 0.03 and then increased again at 0.04. Here, the WH method considers the broadening of the peak as a function of the diffraction angle, which is assumed to be the combined effect of broadening induced by both the crystalline size and strain (*36*, *37*). Furthermore, the strain of perovskite films can also be attributed from the preferred crystalline phase and oriented domain boundaries, etc. However, the change in crystallite size was not very large (fig. S2), and the variation of strain obtained from the XRD patterns of powders scraped from the thin films was almost similar to that of the films (fig. S10). In particular, when the mole fraction of MDA and Cs incorporated in $FAPbI_3$ was 3:3, the lowest strain appeared. This result proposes that 3:3 substitution of Cs^+ and MDA^{2+} in $FAPbI_3$ effectively mitigated the lattice strain in the perovskite structure (fig. S11). Thus, reducing the strain of the lattice minimized defect centers or traps that can capture charge carriers and negatively affect solar cell performance.

Steady-state PL and time-resolved photoluminescence (TRPL) measurements were carried out to investigate the nonradiative carrier recombination of the perovskite thin films with different x [$(FAPbI_3)_{1-x}(MC)_x$] and control. Figure 3B shows PL spectra for the thin perovskite layers deposited on glass substrate. The addition of equal amounts of Cs^+ and MDA^{2+} increased the PL intensity; the intensity maximized at $x = 0.03$ and then decreased again at $x = 0.04$. Under the same conditions, an increase in PL intensity implied a decrease

in nonradioactive recombination. For a more quantitative comparison of charge recombination, TRPL was measured using time-correlated single photon counting (TCSPC) under low-intensity pulsed excitation, for which decay dynamics mostly reflect nonradiative trap-assisted recombination. From the PL decay curves in Fig. 3C, the carrier lifetime values were obtained by using the biexponential equation $Y = A_1\exp(-t/\tau_1) + A_2\exp(-t/\tau_2)$, where τ_1 and τ_2 denote the fast and slow decay time and are related to the trap-assisted nonradiative and radiative recombination processes, respectively. As seen in the average carrier lifetime (table S1) measured using six different samples, both τ_1 and τ_2 at $x = 0.03$ are the longest and longer than control. This increased lifetime was consistent with the reduction of lattice strain, indicating that the relaxation of the strain suppressed nonradiative recombination by reducing the number of trap states.

Changes in the lattice strain can also affect the Urbach energy (E_u). A lower E_u of the perovskite film indicates a higher structural quality of the film, as well as a lower voltage loss between V_{oc} and the bandgap voltage. Figure 3D shows E_u for the perovskite films calculated from UV-vis absorption spectra using the equation $\alpha = \alpha_0 \exp(h\nu/E_u)$, where α is absorption coefficient and $h\nu$ is photon energy. The E_u of the control and target films were 30.18 and 26.88 meV, respectively (Fig. 3D and fig. S12). Interestingly, the change in E_u of $(FAPbI_3)_{1-x}(MC)_x$ perovskite films shows a similar trend with the change in the microstrain of the films. This result implies that the lattice strain is closely related to the trap states in the perovskite films.

To further understand the total quantity and energetic levels of the trap states, thermally stimulated current (TSC) analysis was performed in complete devices. Figure 3E presents the TSC spectra for temperatures from 120 to 300 K for $(FAPbI_3)_{1-x}(MC)_x$–based PSCs compared with the control device. The TSC signal could be integrated to estimate the lower limit of trap densities over the elapsed time according to the equation

$$\int_{\text{signal}} I_{TSC}\,dt \le eN_{T,TSC}Vol$$

where $N_{T,TSC}$ is the trap density, e is the elementary charge, and Vol is the volume of the perovskite layer (*38*, *39*). The trap density of the control device (8.86×10^{16} cm^{-3}) was the highest, which indicates that the strain in the perovskite structure induced an increase in defects. The trap density decreased in the $(FAPbI_3)_{1-x}(MC)_x$ devices from $x = 0.01$ to $x = 0.03$, where it was smallest (4.76×10^{16} cm^{-3}), and then increased at $x = 0.04$ (Fig. 3F).

To extract the activation energy of the trap states, the slope of the initial rise of the TSC current, which was attributed to the start of trap release, in the Arrhenius plot for each condition was fitted to the following equation

$$I_{TSC} \propto \exp\left(-\frac{E_A}{k_B T}\right)$$

where E_A, k_B, and T are the activation energy, Boltzmann constant, and temperature, respectively (*40*, *41*). For the control device, trap states with an activation energy $E_{A1} = 150$ meV were estimated (fig. S13). For $x = 0.01$, a higher activation energy (224 meV) was observed, but it was drastically reduced at $x = 0.02$ (131 meV) and $x = 0.03$ (130 meV) and then rapidly increased to 272 meV at $x = 0.04$. A higher activation energy indicates traps formed deeper in the bandgap, which promotes nonradiative recombination. At very low temperatures (<250 K), the activation energy E_{A2} of $x = 0.03$ (169 meV) was slightly lower than that of the control (183 meV); however, it is speculated that the E_{A1} value is more critical to real operational conditions, which are higher than room temperature (298 K). The defect analysis, PL, TCSPC, Urbach energy, and TSC results were in good agreement with the changes in the lattice strain of the $(FAPbI_3)_{1-x}(MC)_x$ perovskites, which implies that the enhanced device performance with improved V_{oc} was closely related to defect passivation induced by the strain relaxation of the perovskite structure.

Long-term stability is also very important, even if the efficiency of PSCs is improved by reducing strain. The thermal stability of PSCs with thermally stable poly(triarylamine) as the HTM were compared to rule out thermal degradation from the HTM layer for the control and target devices. Unencapsulated devices were stored at 85°C in an oven under 15 to 25% relative humidity under dark and ambient conditions. The control device retained at least 80% of its initial efficiency up to 600 hours. In contrast, the target device maintained >80% of its initial efficiency even after 1300 hours (Fig. 4A). In addition, control and target devices using copper phthalocyanine as HTM were tested at 150°C without any encapsulation. The target device retained almost 80% of initial PCE after 20 hours, unlike the control device, for which PCE decreased to ~60% of its initial value (Fig. 4B). This improved thermal stability by Cs^+ incorporation is consistent with the reported results (*4*, *5*, *42*–*44*). The long-term operational stability of an encapsulated device (using Spiro-OMeTAD as HTM) was also tested with maximum power point tracking under ambient air and full solar illumination without an ultraviolet cut-off filter (Fig. 4C). The target device maintained >90% of its initial efficiency after 400 hours, which is comparable with other efficient mp-TiO_2–based PSCs (*29*, *45*, *46*).

REFERENCES AND NOTES

1. F. Brivio, A. B. Walker, A. Walsh, *APL Mater.* **1**, 042111 (2013).
2. E. J. Juarez-Perez *et al.*, *J. Phys. Chem. Lett.* **5**, 2390–2394 (2014).
3. D. Ghosh, A. R. Smith, A. B. Walker, M. S. Islam, *Chem. Mater.* **30**, 5194–5204 (2018).
4. D. P. McMeekin *et al.*, *Science* **351**, 151–155 (2016).
5. M. Saliba *et al.*, *Energy Environ. Sci.* **9**, 1989–1997 (2016).
6. N. Pellet *et al.*, *Angew. Chem. Int. Ed.* **53**, 3151–3157 (2014).
7. J.-W. Lee *et al.*, *Adv. Energy Mater.* **5**, 1501310 (2015).
8. Q. Han *et al.*, *Adv. Mater.* **28**, 2253–2258 (2016).
9. C. C. Stoumpos, C. D. Malliakas, M. G. Kanatzidis, *Inorg. Chem.* **52**, 9019–9038 (2013).
10. V. M. Goldschmidt, *Naturwissenschaften* **14**, 477–485 (1926).
11. C. Yi *et al.*, *Energy Environ. Sci.* **9**, 656–662 (2016).
12. M. Saliba *et al.*, *Science* **354**, 206–209 (2016).
13. Z. Wang *et al.*, *Nat. Energy* **2**, 17135 (2017).
14. D. Luo *et al.*, *Science* **360**, 1442–1446 (2018).
15. C. Grote, R. F. Berger, *J. Phys. Chem. C* **119**, 22832–22837 (2015).
16. N. Rolston *et al.*, *Adv. Energy Mater.* **8**, 1802139 (2018).
17. L. Zhang *et al.*, *Sci. Rep.* **8**, 7760 (2018).
18. X. Zheng *et al.*, *ACS Energy Lett.* **1**, 1014–1020 (2016).
19. Y. Chen *et al.*, *Nature* **577**, 209–215 (2020).
20. K. Nishimura *et al.*, *ACS Appl. Mater. Interfaces* **11**, 31105–31110 (2019).
21. M. I. Saidaminov *et al.*, *Nat. Energy* **3**, 648–654 (2018).
22. T. W. Jones *et al.*, *Energy Environ. Sci.* **12**, 596–606 (2019).
23. J. Zhao *et al.*, *Sci. Adv.* **3**, eaao5616 (2017).
24. C. Zhu *et al.*, *Nat. Commun.* **10**, 815 (2019).
25. D.-J. Xue *et al.*, *Nat. Commun.* **11**, 1514 (2020).
26. H. Tsai *et al.*, *Science* **360**, 67–70 (2018).
27. H. Wang *et al.*, *Adv. Mater.* **31**, e1904408 (2019).
28. G. Kapil *et al.*, *ACS Energy Lett.* **4**, 1991–1998 (2019).
29. H. Min *et al.*, *Science* **366**, 749–753 (2019).
30. N. J. Jeon *et al.*, *Nat. Mater.* **13**, 897–903 (2014).
31. T. Leijtens *et al.*, *Energy Environ. Sci.* **9**, 3472–3481 (2016).
32. X. Zheng *et al.*, *Nat. Energy* **2**, 17102 (2017).
33. J.-F. Liao *et al.*, *J. Mater. Chem. A Mater. Energy Sustain.* **7**, 9025–9033 (2019).
34. J. J. Yoo *et al.*, *Energy Environ. Sci.* **12**, 2192–2199 (2019).
35. J. T.-W. Wang *et al.*, *Energy Environ. Sci.* **9**, 2892–2901 (2016).
36. G. K. Williamson, W. H. Hall, *Acta Metall.* **1**, 22–31 (1953).
37. B. D. Cullity, S. R. Stock, *Elements of X-Ray Diffraction* (Prentice-Hall, ed. 3, 2001).
38. G. Gordillo, C. A. Otálora, M. A. Reinoso, *J. Appl. Phys.* **122**, 075304 (2017).
39. Y. Hu *et al.*, *Adv. Energy Mater.* **8**, 1703057 (2018).
40. G. F. J. Garlick, A. F. Gibson, *Proc. Phys. Soc.* **60**, 574–590 (1948).
41. A. Baumann *et al.*, *J. Phys. Chem. Lett.* **6**, 2350–2354 (2015).
42. X. Liu *et al.*, *J. Mater. Chem. A Mater. Energy Sustain.* **4**, 17939–17945 (2016).
43. G. Niu, W. Li, J. Li, X. Liang, L. Wang, *RSC Advances* **7**, 17473–17479 (2017).
44. C. Wang *et al.*, *Sustain. Energy Fuels* **2**, 2435–2441 (2018).
45. N. J. Jeon *et al.*, *Nat. Energy* **3**, 682–689 (2018).
46. T.-Y. Yang *et al.*, *Adv. Sci.* **6**, 1900528 (2019).

ACKNOWLEDGMENTS

Funding: This work was supported by the Basic Science Research Program (NRF-2018R1A2A3074921), the Climate Change Program (NRF2015M1A2A2056542), and the Global Frontier Program (2012M3A6A7054861) through the National Research Foundation of Korea (NRF) funded by the Ministry of Science, ICT & Future Planning (MSIP). This work was also supported by the Defense Challengeable Future Technology Program of the Agency for Defense Development, Republic of Korea, and by a brand project (1.200030.01) of UNIST. Finally we thank the beamline staff at Pohang Accelerator Laboratory for supporting GIWAXS measurement. **Author contributions:** S.I.S. designed and supervised the research. G.K. and H.M. fabricated and characterized the perovskite films and devices. K.S.L. measured TSC. D.Y.L. conducted PL and TCSPC measurement. S.M.Y. measured UV-vis absorption. G.K. performed the stability tests of perovskite devices. S.I.S., H.M., and G.K. wrote the draft of the manuscript, and all authors contributed to writing the paper. **Competing interests:** G.K., H.M., and S.I.S. are inventors on a patent application (KR 10-2020-0074349) submitted by the Ulsan National Institute of Science and Technology that covers the $MDACl_2$ and Cs costabilized a-$FAPbI_3$. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper or the supplementary materials.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/370/6512/108/suppl/DC1
Materials and Methods
Figs. S1 to S13
Table S1

24 April 2020; accepted 10 August 2020
10.1126/science.abc4417

DEVELOPMENTAL BIOLOGY

An adhesion code ensures robust pattern formation during tissue morphogenesis

Tony Y.-C. Tsai[1], Mateusz Sikora[2]*, Peng Xia[2]†, Tugba Colak-Champollion[3], Holger Knaut[3], Carl-Philipp Heisenberg[2]‡, Sean G. Megason[1]‡

Animal development entails the organization of specific cell types in space and time, and spatial patterns must form in a robust manner. In the zebrafish spinal cord, neural progenitors form stereotypic patterns despite noisy morphogen signaling and large-scale cellular rearrangements during morphogenesis and growth. By directly measuring adhesion forces and preferences for three types of endogenous neural progenitors, we provide evidence for the differential adhesion model in which differences in intercellular adhesion mediate cell sorting. Cell type–specific combinatorial expression of different classes of cadherins (N-cadherin, cadherin 11, and protocadherin 19) results in homotypic preference ex vivo and patterning robustness in vivo. Furthermore, the differential adhesion code is regulated by the sonic hedgehog morphogen gradient. We propose that robust patterning during tissue morphogenesis results from interplay between adhesion-based self-organization and morphogen-directed patterning.

Spatial patterns of distinct cell types arise reproducibly in development. The classic French flag model posits that a morphogen gradient forms across a naïve and static field of cells to provide positional information to specify patterned cell fates (*1*). The vertebrate spinal cord has been a textbook example of the French flag model: neural progenitors interpret opposing gradients of sonic hedgehog (Shh) and bone morphogenetic protein and commit to 13 distinct fates to form stereotypic stripe-like patterns that are highly conserved across vertebrate species (*2*–*4*). However, our recent work showed that the Shh signal is noisy, resulting in specification of neural progenitors in a mixed pattern at the onset of morphogenesis (*5*, *6*). Additionally, the extensive cell rearrangements during convergent extension could further disrupt pattern (*7*). Nevertheless, the stereotypic stripe patterns still form reproducibly. This makes the zebrafish spinal cord an attractive system to study how robust patterning can be achieved despite imprecision in morphogen signaling and extensive cell-cell neighbor exchange during tissue morphogenesis and growth.

We focused on patterning of three neural progenitor types (p3, pMN, and p0), which are distinguishable by individual transcription factors (*nkx2.2a*, *olig2*, and *dbx1b*, respectively) (fig. S1A) (*8*–*12*). At the neural tube stage, these three cell types form stereotypical stripe-like domains, as visualized by using transgenic zebrafish carrying fluorescent reporters of *nkx2.2a*, *olig2*, and *dbx1b* or by fluorescent in situ hybridization based on a hybridization chain reaction (in situ HCR) (Fig. 1, A and B, and fig. S1B) (*13*–*16*). Live imaging of fluorescent reporters, together with single-cell tracking, suggests that these neural progenitors form imprecise initial patterns during morphogenesis, and the mixed patterns are resolved predominantly through cell sorting (figs. S1, C to O, and S2; movies S1 to S4; and supplementary text S1).

The cell-sorting behavior suggests differences in adhesion properties among different neural progenitor types. The differential adhesion hypothesis (*17*) is a long-standing model for cell sorting, but direct biophysical evidence showing that endogenous cell types use differential adhesion to assist patterning in vivo is still lacking. Therefore, we set out to measure adhesion forces between different types of endogenous neural progenitors with the dual pipette aspiration assay (Fig. 1C, movie S5, and methods) (*18*, *19*). We measured the adhesion forces of six different doublet types (Fig. 1D), including three homotypic contacts (contacts between cells of the same type) and three heterotypic contacts (contacts between cells of different types). The average adhesion forces at the homotypic contacts between two pMN cells (7.6 ± 3.6 nN) and two p3 cells (4.0 ± 2.7 nN) were significantly greater than that of the pMN-p3 heterotypic contact (2.5 ± 2.2 nN). Similarly, the adhesion forces at the homotypic contacts between two pMN cells (7.6 ± 3.6 nN) and two p0 cells (7.2 ± 6.3 nN) were also greater than those for the pMN-p0 heterotypic contacts (4.4 ± 3.3 nN) (Fig. 1D).

To enable more direct comparison of adhesion preferences, we developed a triplet competition assay (Fig. 1E). The triplet is composed of two cells of the same type and one cell of a different type, forming one homotypic and one heterotypic contact. When the triplet is pulled apart, it mimics the challenge faced by the cells in vivo when they are pulled by neighboring cells in different directions (movie S6 and methods). All three neural progenitor cell types showed a clear preference for homotypic contacts, with pMN-pMN and p3-p3 homotypic contacts winning over pMN-p3 heterotypic contacts and pMN-pMN and p0-p0 homotypic contacts winning over pMN-p0 heterotypic contacts by a ratio of about 2:1 (Fig. 1F). Thus, each of the three neural progenitor types exhibit a homotypic preference, a term we use to describe the phenomenon of cells selectively stabilizing homotypic contacts over heterotypic contacts.

To identify the molecular mechanisms underlying this adhesion specificity, we obtained the transcriptomes of p3, pMN, and p0 cells and used CRISPR-Cas9–mediated genome editing to knock out candidate adhesion molecules to look for patterning phenotypes (*20*). N-cadherin (*cdh2*), cadherin 11 (*cdh11*), and protocadherin 19 (*pcdh19*) stood out as genes with significant loss-of-function phenotypes (fig. S3 and supplementary text S2). *cdh2* is the most abundant cadherin in p3, pMN, and p0 cells (table S1) (*21*). Using a fluorescent reporter, *TgBAC(cdh2:cdh2-mCherry)* (*22*), we found that Cdh2 exhibits a protein gradient along the ventral-dorsal (V-D) axis of the entire spinal cord that increases by twofold from the normalized V-D position of 0 to 0.8 (Fig. 2, A and B; fig. S4, A and B; and supplementary text S3). Similar Cdh2 profiles are also observed by antibody staining of sectioned neural tube and in a different fluorescent reporter fish of Cdh2, *TgBAC(cdh2:cdh2-tFT)* (fig. S4, C to F) (*23*). The expression of *cdh11* appears as one stripe along the entire length of the spinal cord (fig. S5A) (*24*), largely overlapping with the *olig2*-positive domain but with a wider distribution along the V-D axis (Fig. 2, C and D). *pcdh19* is expressed as two stripes along the entire spinal cord (fig. S5, C and D), with the ventral stripe colocalizing with the p3 domain and the medial floor plate (Fig. 2, E and F, and fig. S5E). The two *pcdh19* stripes flank the pMN domain, and expression of *pcdh19* and *olig2* appears to be mutually exclusive (Fig. 2G and fig. S5F). The correlated expressions between *cdh11* and *olig2* and between *pcdh19* and *nkx2.2a*, as well as the mutual exclusivity between *olig2* and *pcdh19*, were verified by single cell coexpression analysis (fig. S6, A to C; supplementary text S3; and methods). Together, quantitative analyses of *cdh2*, *cdh11*, and *pcdh19* expression revealed an adhesion code specific to each of the three cell types, with *cdh2*, *cdh11*, and *pcdh19* enriched in p0, pMN, and p3 cells, respectively (Fig. 2H and supplementary text S3). The differential expression patterns of *cdh2*, *cdh11*, and *pcdh19* are present at the onset of spinal cord morphogenesis, making these genes

[1]Department of Systems Biology, Harvard Medical School, 200 Longwood Avenue, Boston MA 02115, USA. [2]Institute of Science and Technology Austria, Am Campus 1, 3400 Klosterneuberg, Austria. [3]Skirball Institute of Biomolecular Medicine, New York University School of Medicine, 540 First Avenue, New York, NY 10016, USA.
*Present address: Max Planck Institute of Biophysics, Frankfurt am Main, Germany.
†Present address: Life Sciences Institute, Zhejiang University, 310058 Hangzhou, China.
‡Corresponding author. Email: megason@hms.harvard.edu (S.G.M.); heisenberg@ist.ac.at (C.-P.H.)

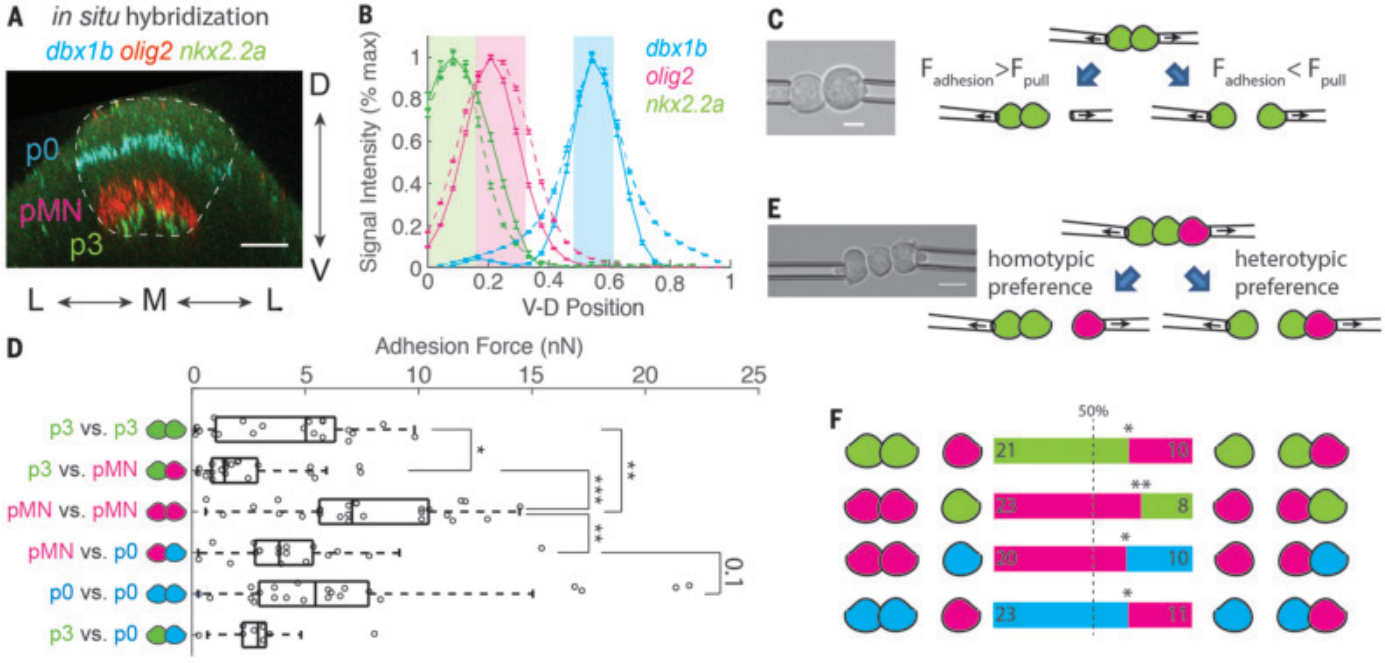


Fig. 1. Neural progenitor cells exhibit homotypic preference. Green, magenta, and blue represent p3, pMN, and p0 cells, respectively. (**A**) Cross section of an 11-somite–stage neural tube, stained with in situ HCR probes against *nkx2.2a* (green), *olig2* (red), and *dbx1b* (blue). Scale bar, 20 μm. (**B**) Average intensity profiles of *nkx2.2a*, *olig2*, and *dbx1b* along the V-D axis of the spinal cord, showing good agreement between the fluorescent reporters (dashed lines) and in situ HCR (solid lines). Error bars indicate SEM. (**C**) Illustration of the dual pipette aspiration assay. F, force (of adhesion or pull). Scale bar, 10 μm. (**D**) The adhesion force measured from six doublet types. The boxes represent 25th, 50th, and 75th percentiles of the data. *P < 0.05; **P < 0.01; ***P < 0.001 (Student's t test). (**E**) Illustration of the triplet assay. Scale bar, 10 μm. (**F**) Triplet assay results showing homotypic preference. *P < 0.05; **P < 0.01 (binomial test).

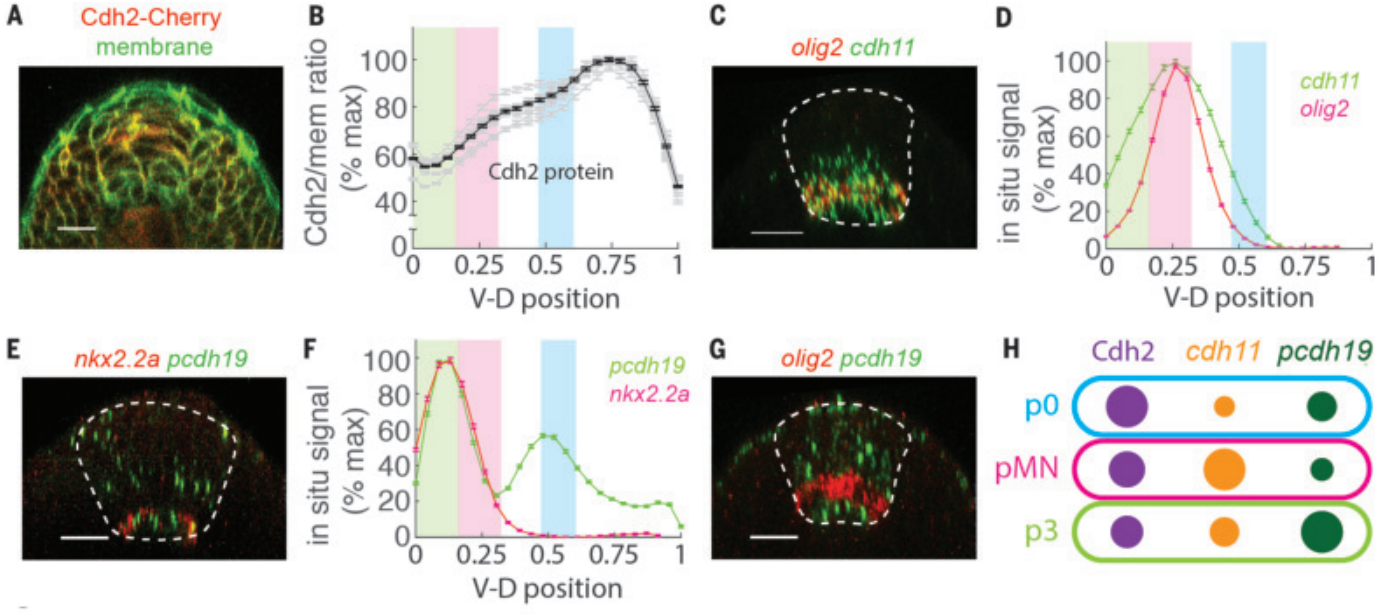


Fig. 2. A combinatorial adhesion code composed of *cdh2*, *cdh11*, and *pcdh19*. (**A**) Confocal cross section of a neural tube from a 10-somite–stage embryo with *TgBAC(cdh2:cdh2-mCherry);Tg(actb2:membrane-citrine)*. Scale bar, 20 μm. (**B**, **D**, and **F**) V-D intensity profile of *cdh2*-Cherry (B), normalized to the membrane citrine signal and V-D intensity profiles of *olig2* and *cdh11* (D) or *nkx2.2a* and *pcdh19* (F) based on in situ HCR. Shaded regions of green, magenta, and blue are identical to Fig. 1B, representing the positions of the p3, pMN, and p0 domains, respectively. Gray lines represent the average profiles from individual embryos. Error bars indicate SEM. (**C**, **E**, and **G**) Cross sections of 10-somite–stage embryos after in situ HCR staining against *olig2* and *cdh11* (C), *nkx2.2a* and *pcdh19* (E), and *olig2* and *pcdh19* (G). Scale bars, 20 μm. (**H**) Adhesion code summary. Circled areas represent the relative abundance of each adhesion molecule, quantified from the normalized *cdh2*-Cherry level (B) or the in situ HCR signals for *cdh11* (D) and *pcdh19* (F).

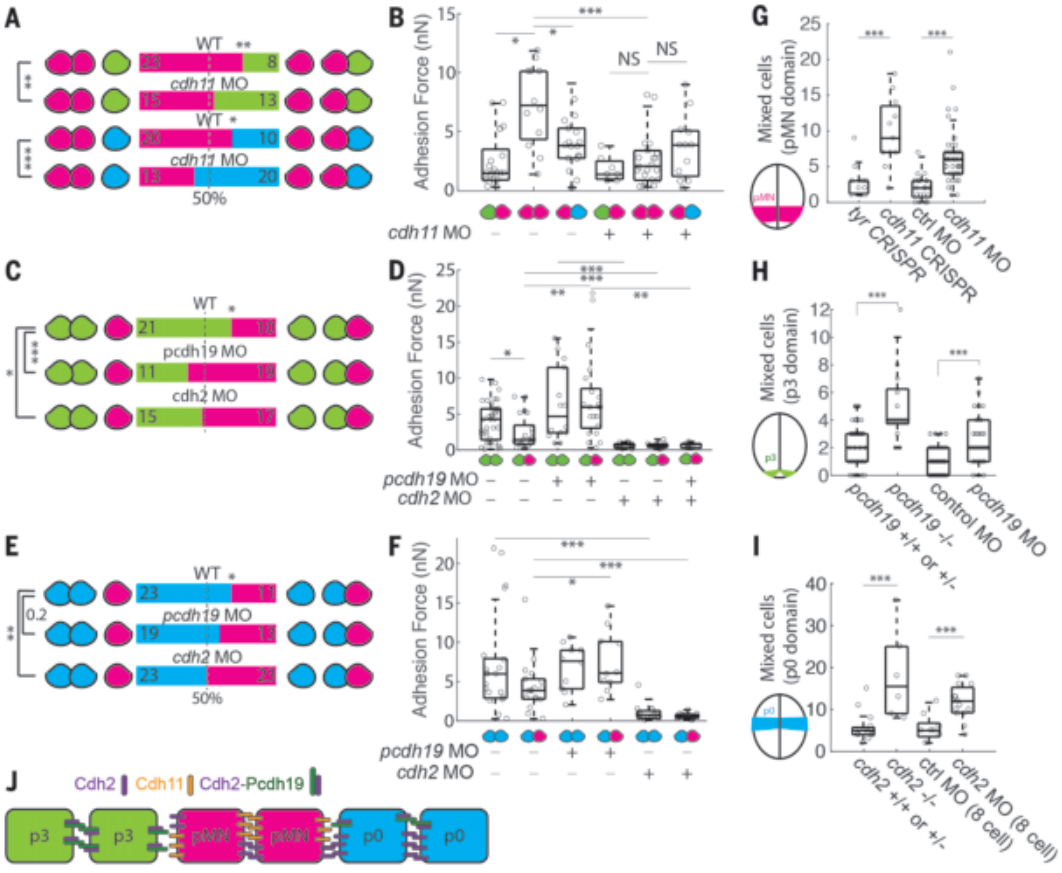


Fig. 3. The adhesion code mediates homotypic preference ex vivo and patterning robustness in vivo. Green, magenta, and blue represent p3, pMN, and p0 cells, respectively. (**A**, **C**, and **E**) Results of triplet assays. WT, wild type; MO, morpholino. *P < 0.05; **P < 0.01; ***P < 0.001 (binomial test). (**B**, **D**, and **F**) Adhesion force measurements in the doublet assays. (**G**, **H**, and **I**) The number of mixed cells in the pMN (G), p3 (H), or p0 (I) domains in control embryos or embryos lacking *cdh11* (G), *pcdh19* (H), or *cdh2* (I). For (B), (D), and (F) to (I): *P < 0.05; **P < 0.01; ***P < 0.001; NS, not significant (Student's t test). (**J**) Cartoon model for molecular mechanisms mediating homotypic preference between p3, pMN, and p0 cells. Cdh2-Pcdh19 denotes the Pcdh19-regulated form of Cdh2 (supplementary text S4).

plausible regulators of cell sorting during the formation and maintenance of neural progenitor patterns (fig. S6, D to G).

We next investigated how the adhesion code mediates the homotypic preference of the three neural progenitor types. *cdh11* is enriched in the pMN cells. Cdh11 belongs to type II cadherins, a cadherin family that is structurally incompatible to bind in trans with type I cadherins (e.g., Cdh2) (*25*). Cdh11-expressing cells and Cdh2-expressing cells segregate in culture, indicating that Cdh11 could mediate pMN-specific homotypic adhesion in a tissue dominated by Cdh2-based adhesion (*26*). Indeed, without Cdh11, pMN cells no longer exhibited homotypic preference against p3 or p0 cells (Fig. 3A). Adhesion force measurements of cell doublets further showed that loss of Cdh11 specifically lowered pMN-pMN homotypic adhesion without affecting pMN-p3 or pMN-p0 heterotypic adhesion (Fig. 3B).

Next, *pcdh19* is enriched in p3 cells. Although Pcdh19 alone exhibits very weak adhesion, it belongs to the δ2-protocadherin family, which is known to modulate Cdh2-based adhesion by multiple mechanisms (*27*, *28*). Our data suggested that the Pcdh19-Cdh2 interaction creates a specific adhesion mode that is distinct from Cdh2-based adhesion without Pcdh19 (supplementary text S4). Specifically, loss of *pcdh19* enrichment in the p3 cells, either by *pcdh19* knockdown or addition of *pcdh19* in the pMN cells, disrupted the homotypic preference of p3 cells in the p3-p3-pMN triplet assay (Fig. 3C and fig. S7). Loss of *pcdh19* does not change the p3-p3 homotypic adhesion force significantly; instead, the predominant effect of the loss of *pcdh19* is an increase of the p3-pMN heterotypic adhesion to a level that matches the p3-p3 homotypic adhesion (Fig. 3D). The increase of p3-pMN adhesion without Pcdh19 is Cdh2 dependent, consistent with the role of Pcdh19 to generate a Pcdh19-Cdh2–dependent adhesion mode that lowers adhesion to the Cdh2-exclusive mode in pMN cells.

Finally, Pcdh19 and Cdh2 are enriched in p0 cells relative to pMN cells. As in p3-pMN heterotypic adhesion, loss of Pcdh19 also led to a Cdh2-dependent increase of pMN-p0 heterotypic adhesion. However, this only slightly disrupted the homotypic preference of p0 cells (Fig. 3, E and F). Instead, loss of Cdh2 completely abolished the homotypic preference of p0 cells against pMN cells. We conclude that the homotypic preference of p0 cells against pMN cells is achieved primarily by higher levels of Cdh2-based adhesion. The weaker Pcdh19 dependence of p0 cells compared with p3 cells is consistent with a fourfold-smaller Pcdh19-to-Cdh2 ratio in p0 cells than p3 cells (Fig. 2F).

To further investigate the role of the adhesion code in patterning in vivo, we examined the cohesion of pMN, p3, or p0 domains at the neural tube stage in the mutants or morphants of *cdh11*, *pcdh19*, or *cdh2* (*21*, *29*). We found a significant increase of patterning errors in each domain when the corresponding adhesion molecule was perturbed (Fig. 3, G to I; fig. S8; movie S7; and supplementary text S5). Together, our results show that the adhesion code not only mediates the adhesion specificity underlying homotypic preference of neural progenitors ex vivo but also mediates patterning robustness in vivo (Fig. 3J).

The cell type–specific adhesion code requires coordinated regulation of cell fate specification and the expression of adhesion molecules, possibly through a common upstream regulator. The Shh morphogen gradient instructs specification of ventral neural progenitor fates (e.g., p3 and pMN cells) in a dose-dependent fashion (*30*, *31*). We therefore perturbed Shh signaling and measured changes in expression patterns of both cell fate markers for

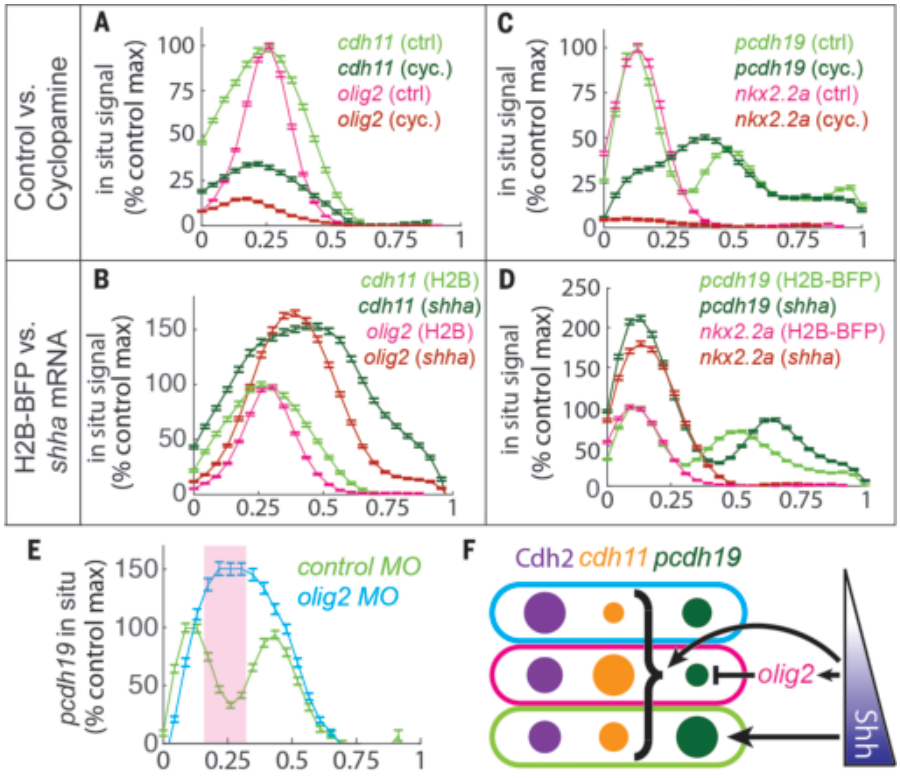


Fig. 4. The adhesion code is regulated by Shh signaling and *olig2*. (**A** to **D**) The V-D in situ HCR profile of *olig2* and *cdh11* [(A) and (B)] or *nkx2.2a* and *pcdh19* [(C) and (D)] from embryos treated with vehicle control or 100 μM cyclopamine [(A) and (C)] or injected with 90 pg of H2B-BFP or *shha* mRNA [(B) and (D)]. Error bars indicate SEM. (**E**) The V-D in situ HCR profile of *pcdh19* from embryos injected with control or *olig2* morpholino. Error bars indicate SEM. (**F**) Summary of how the Shh gradient and *olig2* regulate *pcdh19* and *cdh11*.

pMN and p3 cells (*olig2* and *nkx2.2a*) and their corresponding adhesion molecules (*cdh11* and *pcdh19*). We down-regulated Shh signaling by treatment with cyclopamine, a chemical inhibitor of the Shh activator Smoothened, and up-regulated Shh signaling by injection of *shha* mRNA into one-cell–stage embryos. Our data suggest that the cell fate and adhesion gene pairs (*olig2* versus *cdh11* and *nkx2.2a* versus ventral stripe of *pcdh19*) all respond to Shh perturbation in a dose-dependent fashion (Fig. 4, A to D; figs. S9 and S10; and supplementary text S6). The relative patterns between the cell fate and adhesion genes are preserved under up- or down-regulation of Shh, suggesting shared regulatory logic. Finally, to further connect cell fate regulators downstream of Shh signaling with the adhesion code, we knocked down *olig2*. We found that the two *pcdh19* stripes merged into one continuous stripe with higher amplitude (Fig. 4E and fig. S9, I and J), suggesting that *olig2* is responsible for *pcdh19* repression in the pMN domain. Together, our data show that the combinatorial adhesion code in the ventral spinal cord is instructed by the Shh morphogen gradient (Fig. 4F).

The origin of patterning robustness in tissues undergoing morphogenesis is an important open question. In the zebrafish spinal cord, we show that robust patterning requires a previously unappreciated interplay between two classic ideas for patterning—morphogen gradients and differential adhesion. The morphogen gradient, though not precise, allows cells with similar adhesion properties to be specified near each other; the differential adhesion mechanism then drives local self-organization of cells to correct any imperfection in the initial pattern and to remain organized in domains throughout tissue morphogenesis. With the recent discovery of a Toll receptor code in *Drosophila* convergent extension (*32*), we suggest that morphogen gradient–instructed cell type–specific codes of adhesion and tension could represent a general mechanism for robust pattern formation and morphogenesis.

REFERENCES AND NOTES

1. L. Wolpert, *J. Theor. Biol.* **25**, 1–47 (1969).
2. T. M. Jessell, *Nat. Rev. Genet.* **1**, 20–29 (2000).
3. E. Dessaud, A. P. McMahon, J. Briscoe, *Development* **135**, 2489–2503 (2008).
4. S. F. Gilbert, *Developmental Biology* (Sinauer, ed. 10, 2013).
5. F. Xiong *et al.*, *Cell* **153**, 550–561 (2013).
6. F. Xiong, A. R. Tentner, T. W. Hiscock, P. Huang, S. G. Megason, bioRxiv 412858 [Preprint]. 10 September 2018. https://doi.org/10.1101/412858.
7. C. Yin, B. Ciruna, L. Solnica-Krezel, in *Current Topics in Developmental Biology*, vol. 89, *Tissue Remodeling and Epithelial Morphogenesis*, T. Lecuit, Ed. (Elsevier, 2009), pp. 163–192.
8. J. Briscoe, A. Pierani, T. M. Jessell, J. Ericson, *Cell* **101**, 435–445 (2000).
9. S. L. Gribble, O. B. Nikolaus, R. I. Dorsky, *Dev. Dyn.* **236**, 3472–3483 (2007).
10. B. G. Novitch, A. I. Chen, T. M. Jessell, *Neuron* **31**, 773–789 (2001).
11. H.-C. Park, A. Mehta, J. S. Richardson, B. Appel, *Dev. Biol.* **248**, 356–368 (2002).
12. A. Pierani *et al.*, *Neuron* **29**, 367–384 (2001).
13. A. Kinkhabwala *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **108**, 1164–1169 (2011).
14. B. B. Kirby *et al.*, *Nat. Neurosci.* **9**, 1506–1511 (2006).
15. J. Shin, H.-C. Park, J. M. Topczewska, D. J. Mawdsley, B. Appel, *Methods Cell Sci.* **25**, 7–14 (2003).
16. H. M. T. Choi *et al.*, *Development* **145**, dev165753 (2018).
17. M. S. Steinberg, *J. Exp. Zool.* **173**, 395–433 (1970).
18. M. Biro, J.-L. Maître, in *Biophysical Methods in Cell Biology*, vol. 125, *Methods in Cell Biology*, E. K. Paluch, Ed. (Elsevier, 2015), pp. 255–267.
19. J.-L. Maître *et al.*, *Science* **338**, 253–256 (2012).
20. J. A. Gagnon *et al.*, *PLOS ONE* **9**, e98186 (2014).
21. Z. Lele *et al.*, *Development* **129**, 3281–3294 (2002).
22. T. Colak-Champollion *et al.*, *Curr. Biol.* **29**, 2570–2579.e7 (2019).
23. C. Revenu *et al.*, *Development* **141**, 1282–1291 (2014).
24. J. I. Franklin, T. D. Sargent, *Dev. Dyn.* **206**, 121–130 (1996).
25. S. D. Patel *et al.*, *Cell* **124**, 1255–1268 (2006).
26. Y. Kimura *et al.*, *Dev. Biol.* **169**, 347–358 (1995).
27. S. Hayashi, M. Takeichi, *J. Cell Sci.* **128**, 1455–1464 (2015).
28. M. R. Emond, S. Biswas, C. J. Blevins, J. D. Jontes, *J. Cell Biol.* **195**, 1115–1121 (2011).
29. S. R. Cooper *et al.*, *J. Cell Biol.* **211**, 807–814 (2015).
30. E. Dessaud *et al.*, *Nature* **450**, 717–720 (2007).
31. J. Ericson, J. Briscoe, P. Rashbass, V. van Heyningen, T. M. Jessell, *Cold Spring Harb. Symp. Quant. Biol.* **62**, 451–466 (1997).
32. A. C. Paré *et al.*, *Nature* **515**, 523–527 (2014).

ACKNOWLEDGMENTS

We thank the members of the Megason and Heisenberg labs for critical discussions of and technical assistance during the work and B. Appel, S. Holley, J. Jontes, and D. Gilmour for transgenic fish. This work is supported by the Damon Runyon Cancer Foundation, a NICHD K99 fellowship (1K99HD092623), a Travelling Fellowship of the Company of Biologists, a Collaborative Research grant from the Burroughs Wellcome Foundation (T.Y.-C.T.), NIH grant R01GM107733 (T.Y.-C.T. and S.G.M.), NIH grant R01NS102322 (T.C.-C. and H.K.), and an ERC advanced grant (MECSPEC) (C.-P.H.). **Author contributions:** T.Y.-C.T., C.-P.H., and S.G.M. conceived and designed the projects. T.Y.-C.T. performed the majority of the experiments. M.S. and P.X. assisted in the initial and final stages of the mechanical measurement of cell adhesion. T.C.-C. and H.K. generated the *TgBAC(cdh2:cdh2-mCherry2)* transgenic line. T.Y.-C.T. performed the data analysis. T.Y.-C.T., C.-P.H., and S.G.M. wrote the manuscript, with input from all coauthors. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** The RNA sequencing data are deposited at GEO (GSE154885). All other data are available in the main text or the supplementary materials.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/370/6512/113/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S13
Tables S1 and S2
References (*33*–*47*)
MDAR Reproducibility Checklist
Movies S1 to S7

20 December 2019; accepted 27 July 2020
10.1126/science.aba6637

CONSERVATION PLANNING

Integrated terrestrial-freshwater planning doubles conservation of tropical aquatic species

Cecília G. Leal[1,2]*†, Gareth D. Lennox[3]*†, Silvio F. B. Ferraz[1], Joice Ferreira[4], Toby A. Gardner[5], James R. Thomson[6], Erika Berenguer[3,7], Alexander C. Lees[8,9], Robert M. Hughes[10,11], Ralph MacNally[12], Luiz E. O. C. Aragão[13,14], Janaina G. de Brito[15], Leandro Castello[16], Rachael D. Garrett[17], Neusa Hamada[18], Leandro Juen[19], Rafael P. Leitão[20], Julio Louzada[2], Thiago F. Morello[21], Nárgila G. Moura[22], Jorge L. Nessimian[23], José Max B. Oliveira-Junior[24], Victor Hugo F. Oliveira[2], Vívian C. de Oliveira[18], Luke Parry[3], Paulo S. Pompeu[2], Ricardo R. C. Solar[20], Jansen Zuanon[18], Jos Barlow[2,3]

Conservation initiatives overwhelmingly focus on terrestrial biodiversity, and little is known about the freshwater cobenefits of terrestrial conservation actions. We sampled more than 1500 terrestrial and freshwater species in the Amazon and simulated conservation for species from both realms. Prioritizations based on terrestrial species yielded on average just 22% of the freshwater benefits achieved through freshwater-focused conservation. However, by using integrated cross-realm planning, freshwater benefits could be increased by up to 600% for a 1% reduction in terrestrial benefits. Where freshwater biodiversity data are unavailable but aquatic connectivity is accounted for, freshwater benefits could still be doubled for negligible losses of terrestrial coverage. Conservation actions are urgently needed to improve the status of freshwater species globally. Our results suggest that such gains can be achieved without compromising terrestrial conservation goals.

Freshwater ecosystems occupy less than 1% of the Earth's surface, make up only 0.01% of all water, yet host ~10% of all known species, including a third of all vertebrates (*1*). They also deliver vital ecosystem services, such as climate regulation and the provision of food, fuel, and fiber (*2*). Nevertheless, freshwater ecosystems are far more imperiled than their terrestrial or marine counterparts; since 1970, for example, populations of freshwater vertebrates have declined by 83% compared with a ~40% decline of terrestrial and marine vertebrates (*3*, *4*). A range of threats have long been linked to this collapse in freshwater biodiversity, including habitat loss and degradation, overexploitation, eutrophication, flow modification, and the introduction of nonnative species (*5*). These are now amplified by emerging stressors, including climate change and contamination from microplastics and biochemicals (*3*).

Despite the freshwater biodiversity crisis (*6*), freshwater species are rarely considered in broad-scale conservation strategies (*7*–*9*). Although distributions of terrestrial and freshwater vertebrates display a degree of spatial congruence (*10*), there are three key reasons why freshwater conservation based on terrestrial priorities cannot be taken for granted. First, studies that reveal terrestrial-freshwater congruence rely on coarse-grained data, and such congruence might not occur at local scales, where conservation decisions are implemented. Second, assessments of the distribution of freshwater biota are often restricted to small scales or specific taxonomic groups (*11*). Third, terrestrial prioritizations do not account for aquatic connectivity, which strongly affects the distribution of freshwater species, facilitates nutrient flows, and mediates the cumulative effects of stressors along watercourses (*12*–*15*). Given these limitations, there is an urgent need to understand the extent to which freshwater biodiversity can benefit from terrestrial conservation actions and whether freshwater protection can be increased through integrated planning for both realms. This is particularly critical in tropical regions, which harbor >80% of the world's freshwater fish and are undergoing the most rapid land-use changes on Earth (*16*).

We addressed these knowledge gaps using data from extensive terrestrial and freshwater biodiversity surveys in two biogeographically distinct regions of Brazilian Amazonia: Paragominas and Santarém (fig. S1) (*17*). With >40% of their forests having been converted to agricultural land uses, these regions typify the agricultural-forest frontier in the Amazon (*18*). In terrestrial sites (n = 377 sites) (fig. S2), we sampled plants (n = 812 species), birds (n = 327 species), and dung beetles (n = 141 species). In freshwater sites (n = 99 streams) (fig. S3), we sampled fish (n = 143 species); Odonata (dragonflies and damselflies; n = 134 species); and Ephemeroptera (mayflies), Plecoptera (stoneflies), and Trichoptera (caddisflies; hereafter, "EPT"), which are frequently used as a measure of freshwater ecosystem health (*19*). We could identify EPT individuals only to genus level (n = 59 genera) (*17*). All taxa are referred to as "species" hereafter.

Using these data, we first investigated the extent to which one species group (for example, fish) is protected under conservation strategies directed at another species group (for example, plants), which we refer to as "incidental conservation." To do so, we built regional species distribution maps with an array of biophysical predictors (table S1) (*17*). We then used the distribution maps and the Zonation conservation planning framework (*20*) to simulate terrestrial and freshwater conservation at the catchment scale, a natural landscape unit that integrates hydrological processes. Zonation selects catchments that maximize the weighted average proportion of species distributions under conservation while accounting for species complementarity, and we used this as our conservation benefit function (*17*). For the freshwater analyses, we used the "directed-connectivity" algorithm, which produces aquatically connected conservation networks appropriate for freshwater species (*21*). To focus on biodiversity (without socioeconomic

[1]Luiz de Queiroz College of Agriculture, University of São Paulo, CEP 13418-900, Piracicaba, SP, Brazil. [2]Departamento de Ecologia e Conservação, Universidade Federal de Lavras, CEP 37200-900, Lavras, MG, Brazil. [3]Lancaster Environment Centre, Lancaster University, Lancaster, UK. [4]EMBRAPA Amazônia Oriental, CEP 66095-100, Belém, Pará, Brazil. [5]Stockholm Environment Institute, Linegatan 87D, 11523, Stockholm Sweden. [6]Department of Environment, Land, Water, and Planning, Arthur Rylah Institute for Environmental Research, Heidelberg, Vic, Australia. [7]Environmental Change Institute, University of Oxford, Oxford, UK. [8]Department of Natural Sciences, Manchester Metropolitan University, Manchester M1 5GD, UK. [9]Cornell Lab of Ornithology, Cornell University, Ithaca, NY, USA. [10]Amnis Opes Institute, Corvallis, OR, USA. [11]Department of Fisheries and Wildlife, Oregon State University, Corvallis, OR, USA. [12]School of BioSciences, The University of Melbourne, Parkville, VIC 3052, Australia. [13]Tropical Ecosystems and Environmental Sciences Group (TREES), Remote Sensing Division, National Institute for Space Research–INPE, Avenida dos Astronautas, São José dos Campos, SP, Brazil. [14]College of Life and Environmental Sciences, University of Exeter, Exeter, UK. [15]Escola Estadual Maria Miranda Araújo, Secretaria de Educação do Estado de Mato Grosso, Av. Aeroporto, s/n, CEP 78336-000, Colniza, MT, Brazil. [16]Department of Fish and Wildlife Conservation, Virginia Polytechnic Institute and State University, Blacksburg, VA, USA. [17]Environmental Policy Lab, Departments of Environmental System Science and Humanities, Social, and Political Science, ETH Zürich, 8092 Zürich, Switzerland. [18]Coordenação de Biodiversidade, Instituto Nacional de Pesquisas da Amazônia, Avenida André Araújo, 2.936, Petrópolis, CEP 69067-375, Manaus, AM, Brazil. [19]Laboratório de Ecologia e Conservação, Instituto de Ciências Biológicas, Universidade Federal do Pará, Rua Augusto Correia, No. 1, Bairro Guamá, CEP 66075-110, Belém, PA, Brazil. [20]Departamento de Genética, Ecologia e Evolução, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, Avenida Antônio Carlos 6627, CP 486, CEP 31270-901, Belo Horizonte, MG, Brazil. [21]Universidade Federal do ABC, São Bernardo do Campo, SP, Brazil. [22]Museu Paraense Emílio Goeldi, Belém, PA, Brazil. [23]Departamento de Zoologia, Instituto de Biologia, Universidade Federal do Rio de Janeiro, Av. Carlos Chagas Filho 373, CEP 21941-590, Rio de Janeiro, RJ, Brazil. [24]Instituto de Ciências e Tecnologia das Águas, Universidade Federal do Oeste do Pará, Rua Vera Paz, s/n (Unidade Tapajós), Bairro Salé, CEP 68040-255, Santarém, PA, Brazil.

*These authors contributed equally to this work.

†Corresponding authors. Email: c.gontijoleal@gmail.com (C.G.L.); celtlen@gmail.com (G.D.L.)

considerations), we first ran the optimization analyses constrained by the proportion of the landscape that could be conserved. We then tested the robustness of these findings to budget-constrained analyses by incorporating two region-specific estimates of agricultural opportunity costs (fig. S4) (*17*). Last, we undertook sensitivity analyses by varying available conservation resources. We report results for the area-constrained analysis in which 20% of landscape could be conserved, which aligns with the Aichi target to conserve at least 17% of terrestrial and inland water areas (*4*). An overview of all analyses is available in fig. S1.

Terrestrially focused conservation planning provided limited incidental conservation benefits for freshwater species (Fig. 1). Among taxa and regions, on average just 22% (range, 14 to 29%) of the freshwater benefits achieved through freshwater conservation were secured through terrestrial conservation. By contrast, freshwater species prioritizations achieved on average 84% (range, 70 to 96%) of the terrestrial benefits achieved through terrestrial prioritizations. Within both freshwater and terrestrial realms, prioritizing for any one taxonomic group provided >92% of the maximum achievable benefits to other groups in the same realm. These results were similar whether the optimizations were constrained by area or financial budgets (Fig. 1, A to C).

Differences in the incidental conservation outcomes can be explained by (i) the correlations in catchment priority rankings among species groups (figs. S5 and S6) and (ii) the spatial distribution of conservation priorities (Fig. 2 and figs. S7 and S8). Terrestrial and freshwater groups act as good surrogates for, respectively, other terrestrial and freshwater groups because of the strong correlation in catchment priority rankings: A catchment with high marginal conservation value for one terrestrial group is likely to be of high marginal conservation value for other terrestrial groups, and the same holds for freshwater taxa. Catchment priority ranking correlations were somewhat weaker between terrestrial and freshwater groups, leading to smaller but nonetheless high incidental terrestrial benefits when focused on freshwater species. However, the failure to incorporate aquatic connectivity into terrestrial planning produced conservation network designs that were inadequate for freshwater species (Fig. 2 and figs. S7 and S8), resulting in poor freshwater outcomes from terrestrial planning.

Next, we considered the extent to which freshwater benefits could be increased through conservation planning mechanisms targeted at both terrestrial and freshwater species. To do so, we developed two integrated planning techniques (*17*). Our first approach used both terrestrial and freshwater biodiversity data to determine a prioritization optimized for species from both realms (hereafter, "joint planning"). Given the general paucity of freshwater biodiversity data, our second approach incorporated aquatic connectivity into the terrestrial optimizations to account for freshwater species habitat requirements (hereafter, "terrestrial-plus-connectivity"). Using these approaches, we undertook two trade-off analyses. We first determined the increase in freshwater benefits that could be achieved for a given reduction in terrestrial benefits from their optimum. We focus on this trade-off analysis in the main text. We also considered the increase in freshwater benefits for a given resource increase (such as an increase in landscape covered or

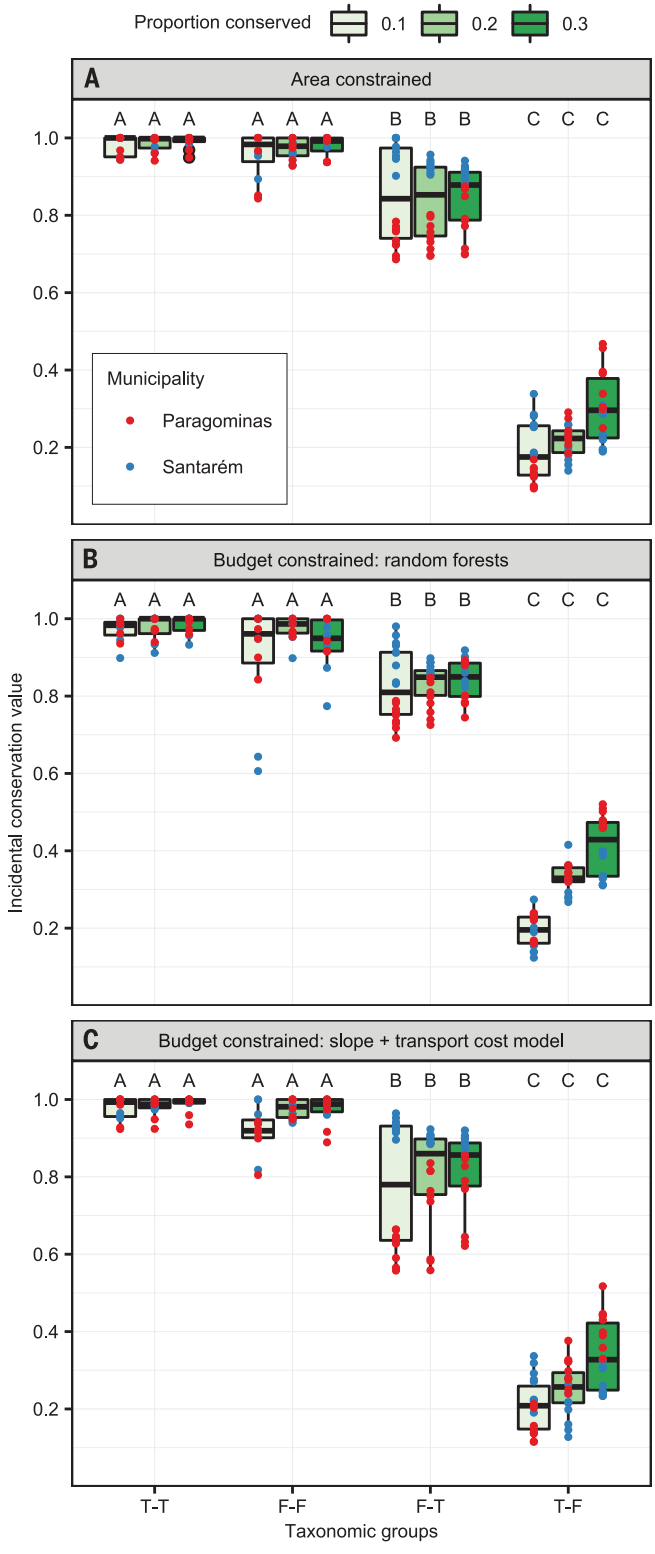


Fig. 1. Incidental conservation. (**A** to **C**) The incidental conservation benefits achieved for one species group when focused on another. The *x* axis ticks are labeled with the focal group first. For example, "T-F" shows the incidental conservation benefits achieved for a freshwater group when prioritizing for a terrestrial group. Points show results for each taxonomic pair. Boxplots show the interquartile range. The center line shows the median. Results are shown for (A) the area-constrained analysis with the constraint that 10, 20, or 30% of landscape can be conserved, and [(B) and (C)] the budget-constrained analyses with two opportunity cost estimates and with budget levels so that approximately 10, 20, and 30% of the landscape can be conserved (*17*). Letters next to the boxplots show results of pairwise comparisons of group means within resource levels (*17*). Variables not sharing a letter have statistically different means.

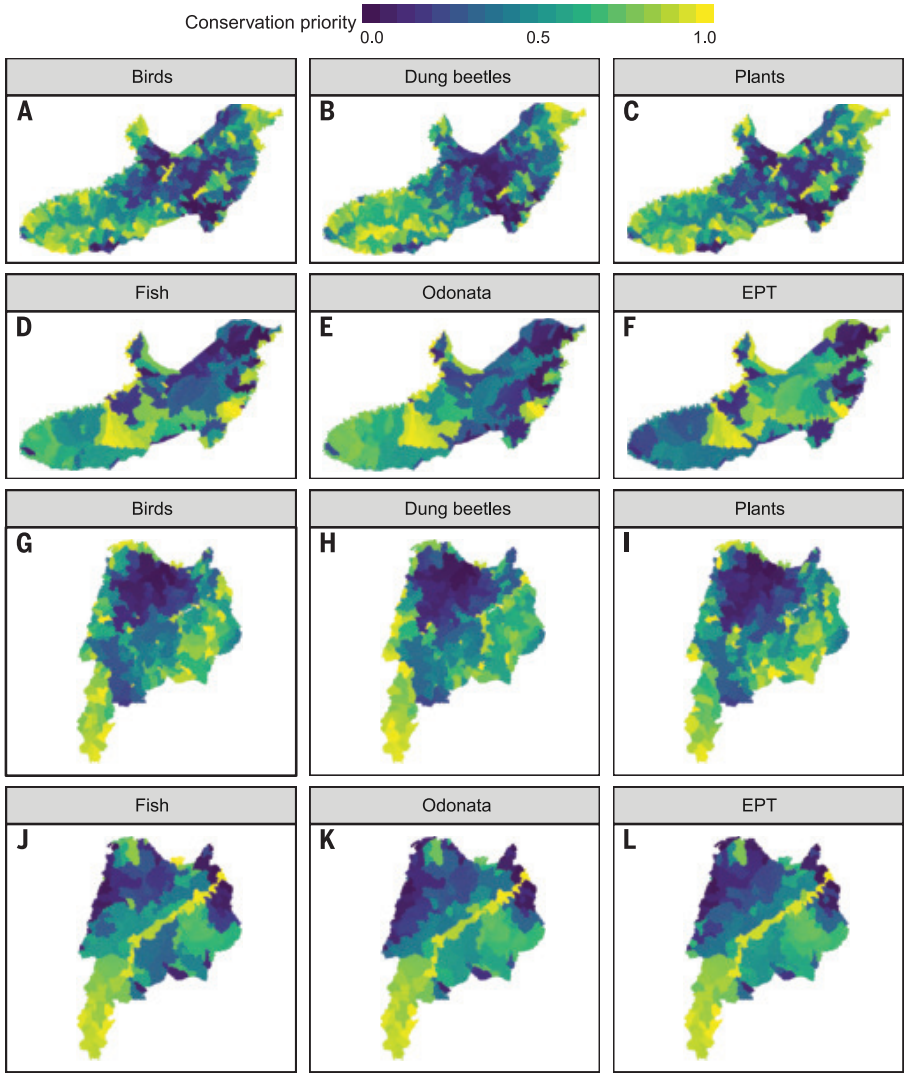


Fig. 2. Catchment prioritizations for terrestrial and freshwater biodiversity. (**A** to **L**) Catchment conservation priority rankings in [(A) to (F)] Paragominas and [(G) to (L)] Santarém for [(A) to (C) and (G) to (I)] terrestrial and [(D) to (F) and (J) to (L)] freshwater taxa. Rankings are based on catchment marginal conservation value, with 1 indicating the catchment with the highest marginal conservation value and 0 that with the lowest marginal conservation value. Results are shown for the area-constrained analysis.

financial budgets) while maintaining terrestrial benefits at their optimum. As above, we focused on area-constrained optimizations in which 20% of a landscape could be conserved.

Using the joint planning approach, freshwater benefits could be increased by on average 62 and 345% in Paragominas and Santarém, respectively, for a negligible 1% reduction in terrestrial benefits relative to their optimum (Fig. 3). A 5% reduction in terrestrial benefits resulted in an average increase in freshwater benefits of 184% in Paragominas and 365% in Santarém. The terrestrial-plus-connectivity approach generally produced lower freshwater conservation gains. Nonetheless, 1 and 5% reductions in terrestrial benefits increased freshwater benefits by 75 to 100% and 130 to 175%, respectively, in both Paragominas and Santarém. Alternatively, the freshwater gains we documented for 1 and 5% reductions in terrestrial benefits could be achieved without any terrestrial losses for, respectively, <1 and <5% increases in conservation resources (fig. S9). Trade-offs were qualitatively similar with the incorporation of opportunity costs (Fig. 3) and more and less pronounced for, respectively, lower and higher conservation resource levels (fig. S10).

Although the freshwater gains we found for negligible reductions in terrestrial protection were substantial in both Paragominas and Santarém, there were large regional differences when using the joint planning approach that incorporates both terrestrial and freshwater biodiversity data (Fig. 3). These differences arise from variation in the spatial overlap of conservation priorities between regions. In Santarém, many of the highest-priority catchments for terrestrial and freshwater groups were in the southwest (where the Tapajós National Forest is located) (Fig. 2). In Paragominas, the same spatial overlap in priorities was not apparent (Fig. 2). Thus, in Paragominas, substantial deviation from the optimal catchment prioritization for terrestrial species was required to achieve the largest increases in freshwater benefits. In Santarém, by contrast, large freshwater gains were possible simply by selecting catchments in the region of high conservation value for both realms that produced the requisite aquatic connectivity. Therefore, the realized magnitude of the freshwater gains possible from integrated planning will depend on the underlying spatial covariance in species distributions, which determines the spatial overlap in conservation priorities.

These results provide compelling evidence that the protection of freshwater species can be vastly improved without undermining terrestrial conservation goals. However, there are factors for which we did not account that could lead to substantially different terrestrial-freshwater trade-offs than we found. First, we did not incorporate the many additional socioecological benefits of freshwater conservation, meaning that our results are likely to be conservative. For example, in addition to the direct provisioning, supporting, regulating, and cultural services that freshwater ecosystems provide (*2*), by enhancing landscape connectivity freshwater conservation can also promote movement of terrestrial species, recolonization of defaunated areas, and seed dispersal and pollination services (*22*). Second and conversely, where freshwater conservation imposes external opportunity costs beyond a loss of agricultural profits—for example, by precluding the development of hydropower or imposing water-use restrictions in the surrounding landscape—the overall scope for conservation investment may be reduced, leading to fewer net benefits from integrated planning. The manifestation of these additional socioecological trade-offs that emerge when protecting freshwater ecosystems is likely to be highly dependent on local circumstances, but their consideration will be essential for designing effective and sustainable conservation projects. Last, our optimization analyses were static. Given that freshwater biodiversity data were collected in different years in Paragominas (2011) and Santarém (2010), and because the regions experienced substantially different climatic conditions during this time (*17*), some of the observed regional differences in trade-offs could result from temporal variation. Understanding and incorporating environmentally

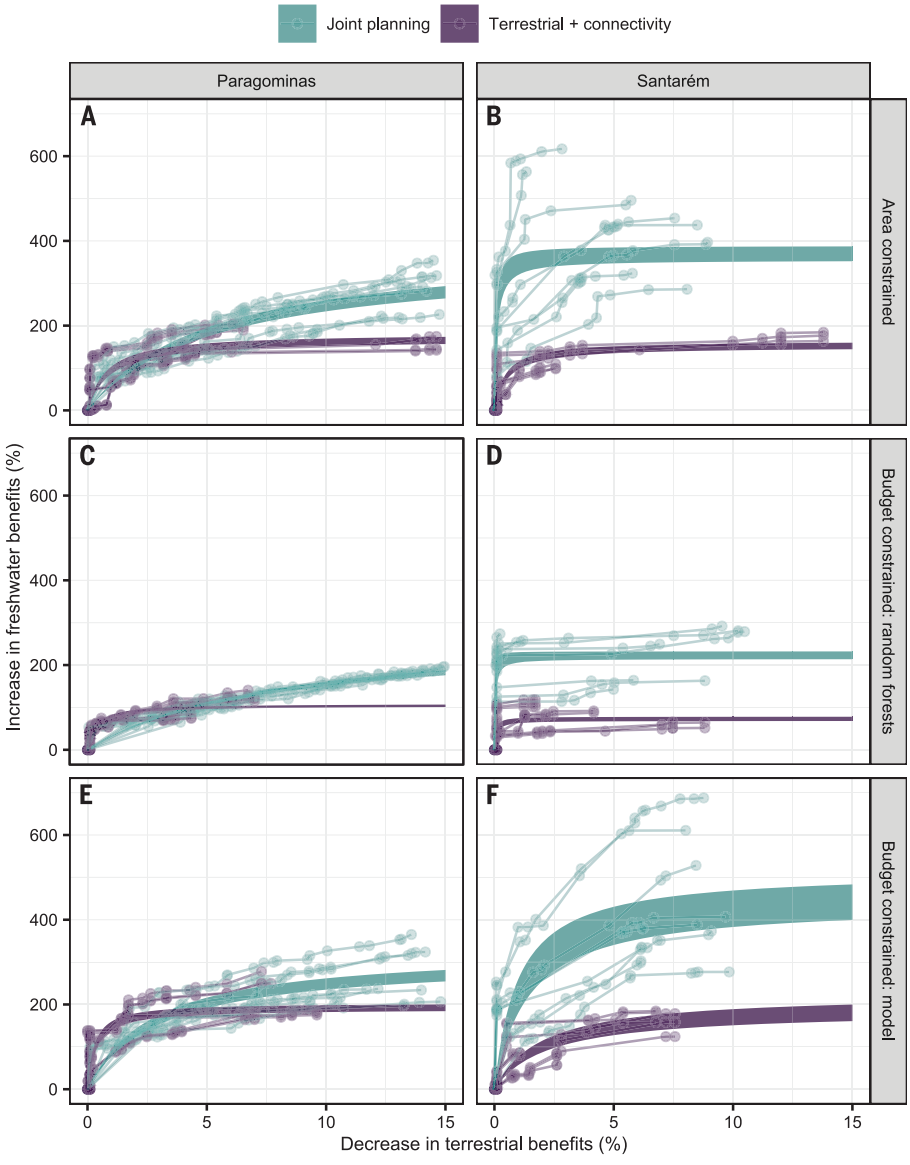


Fig. 3. Terrestrial-freshwater trade-offs. (**A** to **F**) The decrease in terrestrial benefits from their optimum required to achieve an increase in freshwater benefits through the joint-planning and the terrestrial-plus-connectivity approaches in [(A), (C), and (E)] Paragominas and [(B), (D), and (F)] Santarém. The thin lines show the results for each terrestrial-freshwater taxonomic pair. The thick lines show 1 SEM, where the mean was estimated by using Holling type-II curves, for each integrated planning approach. Results are shown for [(A) and (B)] the area-constrained analysis with the constraint that 20% of landscape could be conserved, and [(C) to (F)] the budget-constrained analyzes with two opportunity-cost estimates and with budget levels so that approximately 20% of the landscape could be conserved (*17*).

mediated changes in species distributions will be important for estimating the long-term benefits of integrated terrestrial-freshwater planning.

Identifying promising new approaches for biodiversity conservation is only the first step toward improving conservation outcomes. Given that evidence is lacking for the translation of systematic conservation planning exercises into tangible benefits (*23*), how best to turn our findings into meaningful action? First, although previous global conservation agendas—such as the United Nations' Sustainable Development Goals (SDGs) and the Convention on Biological Diversity—have recognized the need to conserve both terrestrial and freshwater ecosystems (SDG 15, Aichi target 11), recognition of their interdependence remains largely absent from conservation planning. As the world prepares to consider new, post-2020 conservation targets (*9*, *24*), we show that a truly integrated approach to conservation on land, which accounts for trade-offs and harnesses synergies among ecosystems and realms, can provide a cost-effective means to substantially improve outcomes. Understanding where such gains are highest and lowest should be a focus of future research efforts. Crucially, our findings from two biogeographically distinct regions with different biophysical drivers of species distributions (fig. S11) suggest that substantial freshwater gains ought to be attainable across the biodiverse agricultural frontier regions of the forested tropics. Also, conservation remains hampered by a severe lack of biodiversity data, especially in tropical regions (*11*, *25*). Resolving these data shortfalls will be necessary to unlock the benefits we document, and this will require more investment in large-scale ecological surveys and taxonomy (*16*, *26*). Last, to be effective and feasible, integrated terrestrial-freshwater strategies need to be aligned with or incorporated into current environmental policies and laws. In particular, freshwater-orientated planning should not come at the expense of existing protected areas, which often hold the last populations of endangered species and are coming under increasing pressure globally (*27*) and in the Amazon (*28*). Overcoming these challenges will be difficult, but the task is small compared with the enormous gains that can be made for the world's diverse and highly threatened freshwater biota.

REFERENCES AND NOTES

1. E. V. Balian, H. Segers, C. Lévèque, K. Martens, *Hydrobiologia* **595**, 627–637 (2008).
2. Millennium Ecosystem Assessment, "Ecosystems and human well-being: wetlands and water—Synthesis" (World Resources Institute, 2005).
3. A. J. Reid *et al.*, *Biol. Rev. Camb. Philos. Soc.* **94**, 849–873 (2019).
4. UNEP-WCMC and IUCN, "Protected Planet Report 2016" (UNEP-WCMC and IUCN, 2016).
5. D. Dudgeon *et al.*, *Biol. Rev. Camb. Philos. Soc.* **81**, 163–182 (2006).
6. J. S. Albert *et al.*, *Ambio* (2020).
7. T. M. Brooks *et al.*, *Science* **313**, 58–61 (2006).
8. J. E. M. Watson, N. Dudley, D. B. Segan, M. Hockings, *Nature* **515**, 67–73 (2014).
9. D. Tickner *et al.*, *Bioscience* **70**, 330–342 (2020).
10. R. Abell *et al.*, *Conserv. Lett.* **4**, 127–136 (2011).
11. A. S. L. Rodrigues, T. M. Brooks, *Annu. Rev. Ecol. Evol. Syst.* **38**, 713–737 (2007).
12. L. Castello *et al.*, *Conserv. Lett.* **6**, 217–229 (2013).
13. R. Abell, J. Allan, B. Lehner, *Biol. Conserv.* **134**, 48–63 (2007).
14. S. Linke, E. Turak, J. Nel, *Freshw. Biol.* **56**, 6–20 (2011).
15. J. Carvajal-Quintero *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **116**, 13434–13439 (2019).
16. J. Barlow *et al.*, *Nature* **559**, 517–526 (2018).
17. Materials and methods are available as supplementary materials.
18. T. A. Gardner *et al.*, *Philos. Trans. R. Soc. London B Biol. Sci.* **368**, 20120166 (2013).
19. J. L. Nessimian *et al.*, *Hydrobiologia* **614**, 117–131 (2008).
20. A. Moilanen, *Biol. Conserv.* **134**, 571–579 (2007).
21. A. Moilanen, J. Leathwick, J. Elith, *Freshw. Biol.* **53**, 577–592 (2008).
22. J. J. Tewksbury *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **99**, 12923–12926 (2002).
23. E. J. McIntosh *et al.*, *Environ. Evid.* **7**, 22 (2018).
24. G. M. Mace *et al.*, *Nat. Sustain.* **1**, 448–451 (2018).
25. W. R. T. Darwall *et al.*, *Conserv. Lett.* **4**, 474–482 (2011).
26. L. W. Drew, *Bioscience* **61**, 942–946 (2011).
27. M. B. Mascia *et al.*, *Biol. Conserv.* **169**, 355–361 (2014).
28. S. M. Pack *et al.*, *Biol. Conserv.* **197**, 32–39 (2016).

ACKNOWLEDGMENTS

We thank I. Vieira for coordinating the INCT Biodiversidade e Uso da Terra na Amazônia and the Large Scale Biosphere-Atmosphere Program (LBA) for logistical and infrastructure support during field measurements. **Funding:** This work was supported by grants from Brazil (CNPq 574008/2008-0, 458022/2013-6, and 400640/2012-0; CNPq-CAPES-PELD 441659/2016-0; Embrapa SEG:02.08.06.005.00; The Nature Conservancy–Brasil; and CAPES scholarships), the UK (Darwin Initiative 17-023; NE/F01614X/1; NE/G000816/1; NE/F015356/2; NE/I018123/1; and NE/K016431/1), Formas 2013-1571, and Australian Research Council grant DP120100797. Individual support included FAPESP funding (2017/25383-0 and 2019/05440-5, respectively) to C.G.L. and T.F.M.; H2020-MSCA-RISE (691053-ODYSSEA) funding to E.B. and J.B.; Fulbright Brasil funding to R.M.H.; NSF-Belmont Forum award (1852113) to L.C.; and CNPq funding (144452/2010-3, 307597/2016-4, 436007/2018-5, 303548/2017-7, 308970/2019-5, and 313183/2014-7, respectively) to V.C.d.O., L.J., R.P.L., P.S.P., N.H., and J.Z. Institutional support was provided by the Herbário IAN in Belém, LBA in Santarém, and FAPEMAT. This is paper no. 76 in the Sustainable Amazon Network publication series. **Ethics statement:** All biodiversity sampling was undertaken in compliance with Brazilian environmental regulations under the following licenses: (i) Sisbio license 24164 for collecting plants, issued by the Chico Mendes Institute for Biodiversity Conservation (ICMBio); (ii) Sisbio license 10061-1 for collecting dung beetles, issued by the Institute of Environment and Renewable Natural Resources (IBAMA); (iii) Sisbio licenses 10199-2 and 24355-2 for collecting fish, both issued by ICMBio (iv) Sisbio license 10873-1 for collecting insects, issued by IBAMA; and (v) Sisbio license 19102-4 for collecting Odonata, Ephemeroptera, Plecoptera and Trichoptera, issued by ICMBio. No license was required for bird sampling because the methods were observational and did not involve collecting or handling of specimens. Socioeconomic data were collected following the UK Research Integrity Office Principles for Research involving human participants, human material, and personal data and were collected with informed consent. Further approval for opportunity cost data collection was obtained from the Brazilian Agricultural Research Corporation (Embrapa) under CAAE 29054920.4.0000.5173 and Stanford University under IRB Protocol 19044. **Author contributions:** T.A.G., J.F., J.B., L.P., C.G.L., G.D.L., and S.F.B.F. designed the research with additional taxon-specific input from E.B., A.C.L., R.M.H., L.E.O.C.A., J.G.d.B., N.H., L.J., R.P.L., J.L., N.G.M., J.L.N., J.M.B.O.-J., V.H.F.O., V.C.d.O., P.S.P., R.R.C.S., and J.Z.; C.G.L., E.B., A.C.L., R.M.H., J.G.d.B., N.H., L.J., R.P.L., J.L., N.G.M., J.L.N., J.M.B.O.-J., V.H.F.O., V.C.d.O., L.P., R.R.C.S., and J.Z. collected the field data or analyzed biological samples. S.F.B.F. and T.A.G. processed the remote sensing data. G.D.L. and C.G.L. undertook the analyses with input from J.B., J.R.T., R.M., R.D.G., T.F.M., and S.F.B.F.; T.F.M., T.A.G., L.P., and R.D.G. processed the socioeconomic data; C.G.L., G.D.L., and J.B. led the writing, with all authors contributing to reviewing and editing. **Competing interests:** Authors declare no competing interests. **Data and materials availability:** The data and computer code used in these analyses are available at https://doi.org/10.6084/m9.figshare.12423881. The Zonation conservation planning software is available at: https://github.com/cbig/zonation-core.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/370/6512/117/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S13
Table S1
References (*29*–*50*)

2 January 2020; accepted 10 August 2020
10.1126/science.aba7580

SEX DETERMINATION

The mouse *Sry* locus harbors a cryptic exon that is essential for male sex determination

Shingo Miyawaki[1,2], Shunsuke Kuroki[1,2], Ryo Maeda[1,2], Naoki Okashita[1,2], Peter Koopman[3], Makoto Tachibana[1,2]*

The mammalian sex-determining gene *Sry* induces male development. Since its discovery 30 years ago, *Sry* has been believed to be a single-exon gene. Here, we identified a cryptic second exon of mouse *Sry* and a corresponding two-exon type *Sry* (*Sry-T*) transcript. XY mice lacking *Sry-T* were sex-reversed, and ectopic expression of *Sry-T* in XX mice induced male development. *Sry-T* messenger RNA is expressed similarly to that of canonical single-exon type *Sry* (*Sry-S*), but SRY-T protein is expressed predominantly because of the absence of a degron in the C terminus of SRY-S. *Sry* exon2 appears to have evolved recently in mice through acquisition of a retrotransposon-derived coding sequence to replace the degron. Our findings suggest that in nature, SRY-T, not SRY-S, is the bona fide testis-determining factor.

Sexual differentiation is essential for the survival and evolution of a species. Expression of the Y chromosomal gene *Sry* is required for male development in mammals (*1*). Since its discovery in 1990 (*2*, *3*), *Sry* has been considered a single-exon gene, containing only one open reading frame (ORF) and encoding one protein. The mouse *Sry* locus has a complex genomic structure, with 2.8 kb of specific sequence containing an ORF, flanked by ~50 kb of palindromic sequences (*4*). Using a comprehensive transcriptomics approach, we identified a sequence within the flanking palindrome that is essential for male development.

We previously developed a method to enrich somatic cell populations from mouse embryonic gonads (*5*, *6*) (fig. S1A). Using this method, NR5A1-expressing mouse pre-Sertoli cells were purified for RNA-sequencing (RNA-seq) analysis. We identified a transcribed sequence located within the palindromic sequences 3′ to the known *Sry* ORF (Fig. 1A and fig. S1B). This sequence was transcribed similarly to the known *Sry* transcript in terms of temporal and spatial specificity (fig. S1, C and D) but lacked a cap site, as evidenced by cap analysis gene expression sequencing (CAGE-seq) (Fig. 1A and fig. S1B). Long-read RNA-seq identified this sequence as the 3′ portion of a new *Sry* transcript that starts at the known *Sry* transcription start site and is spliced once at typical 5′ GT/3′ AG splicing site sequences with an internal polypyrimidine tract (Fig. 1B and fig. S1, E and F). Thus, this sequence represents a second exon of *Sry*.

Sequencing data indicated the presence of two mRNA transcripts, the known single-exon type (*Sry-S*) and a two-exon type (*Sry-T*) (Fig. 1B and fig. S2). Most (~70%) of *Sry* exon2 is occupied by one long interspersed nuclear element (LINE) L3 and three long terminal repeats (LTRs) (Fig. 1C and fig. S3). *Sry-T* expression levels were approximately half those of *Sry-S* in pre-Sertoli cells (fig. S1G) and, as evidenced by single-cell transcriptome analysis, individual pre-Sertoli cells expressed both transcripts (*7*) (fig. S1H).

The predicted SRY-T protein shares amino acids 1 to 377 with SRY-S, including the high-mobility group DNA-binding domain and poly-glutamine (poly-Q) sequences that are important for transcriptional activation (*8*, *9*). Beyond this, the C-terminal 15 amino acids of SRY-T are encoded by the second exon, whereas absence of splicing results in a different 18 amino acids at the C terminus of SRY-S (Fig. 1D). We used CRISPR/Cas9 to establish the knock-in mouse line *Sry-T*:3xFlag, in which *Sry* exon2 is targeted to produce SRY-T with a triple Flag-epitope tag at the C terminus (fig. S4). We prepared protein lysates from embryonic day 11.5 (E11.5) *Sry-T*:3xFlag gonads and enriched SRY using antibodies against the poly-Q sequence (*10*) (fig. S4A). Anti-FLAG immunoblotting identified a protein with the expected molecular weight of SRY-T-3xFLAG (~55 kDa) in the anti-SRY immune complexes (Fig. 1E). These results indicate that *Sry-T* is translated into a protein in vivo.

Next, we generated mice lacking the *Sry-T* transcript by deleting *Sry* exon2 (Fig. 2, A and B, and fig. S5, A and B). Elimination of *Sry-T* alone in XY mice resulted in complete male-to-female sex reversal (Fig. 2C and fig. S5, C to G), with a global gene expression profile of E11.5 XY *Sry-T*-deficient gonadal somatic cells similar to that of their XX counterparts (fig. S5H). *Sry-S* expression was maintained at normal levels in these mice (fig. S5I), suggesting that in wild-type (WT) XY mice, *Sry-T* and not the previously known *Sry-S* is necessary and sufficient for male sex determination.

To gain insights into this male-to-female sex reversal, we investigated total SRY protein expression in NR5A1$^+$ gonadal somatic cells in

[1]Laboratory of Epigenome Dynamics, Graduate School of Frontier Biosciences, Osaka University, 1-3 Yamadaoka, Suita, Osaka, 565-0871, Japan. [2]Division of Epigenome Dynamics, Institute of Advanced Medical Sciences, Tokushima University, 3-18-15 Kuramoto-Cho, Tokushima, 770-8503, Japan. [3]Institute for Molecular Bioscience, The University of Queensland, Brisbane, Queensland 4072, Australia.
*Corresponding author. Email: tachiban@fbs.osaka-u.ac.jp

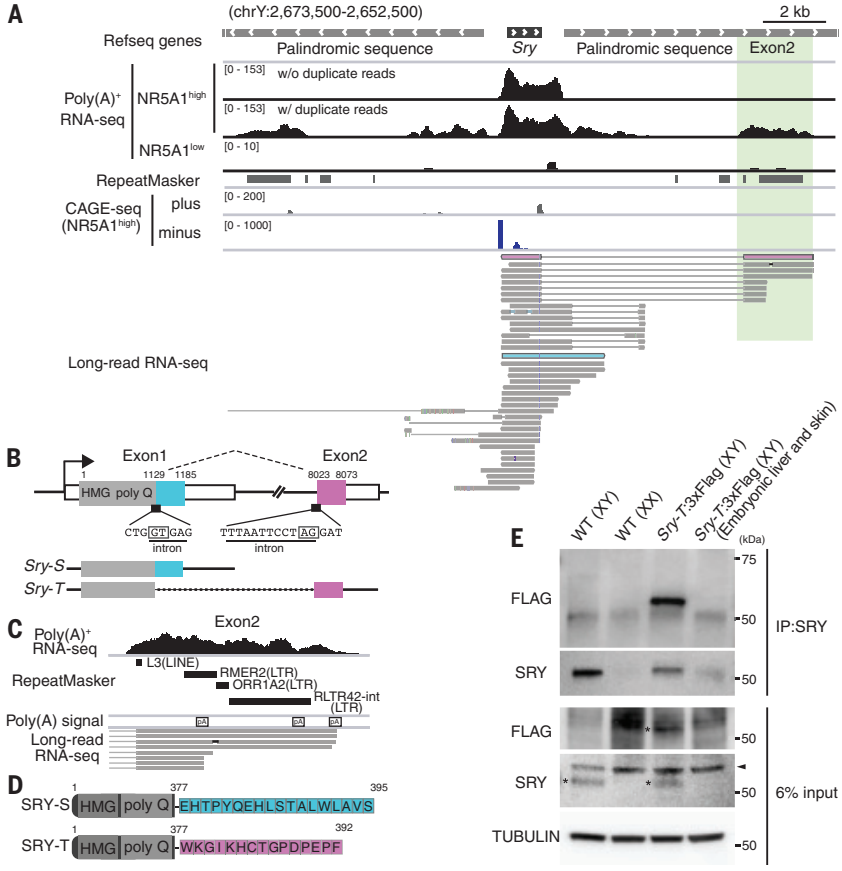


Fig. 1. A cryptic exon of mouse *Sry* generates a new transcript and protein. (**A**) *Sry* exon2 is shown shaded green. The RefSeq *Sry* (NM_011564.1) encoded in the (−) strand is presented at the top. Poly(A)$^+$ RNA-seq data were mapped with or without excluding potential duplicate reads. RepeatMasker shows transposon-derived sequences. CAGE-seq identified transcription start sites. Individual long-read RNA-seq reads (>1000 bp) were mapped. Thick and thin lines represent exonic and intronic sequences, respectively. Representative long-read RNA transcripts for the new two-exon *Sry-T* and canonical single-exon *Sry-S* are shown in red and blue, respectively. (**B**) Exon-intron structure of the *Sry* locus. Shaded and open boxes represent CDS and untranslated regions, respectively. Adenine of the *Sry* translation start codon is position 1. Sequences of *Sry* splice junctions are represented below the bold lines. 5′ GT/3′ AG splicing motifs are boxed. CDS specific for *Sry-S* (blue) and *Sry-T* (red) are shown. (**C**) Retrotransposon-derived sequences in *Sry* exon2. An L3-derived sequence accounts for exon2 CDS, whereas LTR-type retrotransposon-derived sequences provide poly(A) signals of *Sry-T*. (**D**) Predicted amino acid sequences of SRY isoforms. (**E**) Anti-SRY and anti-FLAG immunoblots using anti-SRY immune complexes prepared from E11.5 *Sry-T*:3xFlag gonads showing signals with the predicted size of SRY-T-3xFLAG (~55 kDa). Asterisks and arrowhead represent putative SRY signals and nonspecific signals, respectively.

WT and *Sry-T*–deficient XY embryos. SRY protein levels in *Sry-T*–deficient XY gonads were about one-sixth those of WT XY gonads (Fig. 2, D and E), suggesting that male-to-female sex reversal resulted from low SRY levels and that SRY-T is the predominant protein product of the *Sry* locus in XY gonads.

To compare the properties of SRY-T and SRY-S in detail, we established single-copy knock-in mouse lines conditionally expressing each isoform, which we named *Sry-T*-R26$^{KI/+}$ and *Sry-S*-R26$^{KI/+}$, respectively. In this strategy (*11*), the coding DNA sequence (CDS) of either *Sry-T* or *Sry-S* is expressed within the same context (i.e., copy number, integration site, promoter, and untranslated regions; Fig. 3A). The lines were then crossed with the *Nr5a1-Cre* transgenic line (*6*) to specifically express the *Sry* CDS in gonadal somatic cells. *Sry* mRNA levels in E11.5 XX knock-in gonads were comparable between the *Nr5a1-Cre*; *Sry-S*-R26$^{KI/+}$ and *Nr5a1-Cre*; *Sry-T*-R26$^{KI/+}$ mice (Fig. 3B). However, SRY protein levels in *Nr5a1-Cre*; *Sry-S*-R26$^{KI/+}$ mice were approximately one-seventh those in *Nr5a1-Cre*; *Sry-T*-R26$^{KI/+}$ mice (Fig. 3C), demonstrating that a posttranscriptional mechanism leads to the predominant expression of SRY-T over SRY-S.

Nr5a1-Cre; *Sry-T*-R26$^{KI/+}$ XX mice developed as males, whereas *Nr5a1-Cre*; *Sry-S*-R26$^{KI/+}$ XX mice did not (Fig. 3D and fig. S6, A and B). Immunofluorescence analysis indicated that SRY-T expression, not SRY-S expression, was sufficient to induce male development (Fig. 3, E and F). However, *Nr5a1-Cre*; *Sry-S*-R26$^{KI/KI}$ XX mice, which carry two expression cassettes of *Sry-S*, developed as males (Fig. 3D), indicating that doubling *Sry-S*-R26^{KI} expression surpassed the SRY protein-level threshold (Fig. 3, D to F) required to induce male development.

We speculated that the different C-termini of SRY-S (S18) and SRY-T (T15) (Fig. 4A) underlie the difference in posttranscriptional regulation. We generated lentiviral constructs encoding blue fluorescent protein (BFP) and enhanced green fluorescent protein (EGFP) fused to either S18 or T15, both of which were translated from the same transcript (*12*) (Fig. 4B), and transduced them into human embryonic kidney 293T cells. The effect of S18 or T15 on protein stability was evaluated by examining the EGFP/BFP ratio. Whereas EGFP-T15 expression levels were similar to those of nonfusion EGFP (Fig. 4C, right), EGFP-S18 was present at <1/10 the nonfusion EGFP levels (Fig. 4C, left). In the presence of cycloheximide, EGFP-S18 was destabilized more rapidly than EGFP and EGFP-T15, which was blocked by MG132 addition (Fig. 4D and fig. S7A). These data indicate that the C terminus of SRY-S contains a protein degradation motif, or degron.

Addition of a distal FLAG epitope counteracted the degradation of EGFP-S18 (fig. S7B), indicating that the C-terminal location of the S18 sequence is strictly required for its efficacy, consistent with a general property of C-terminal degrons (*12*). Further, a known subtype of C-terminal degrons carries valine at position -2 (V-2), and its substitution has been shown to abolish degron activity (*12*). Accordingly, we established a V-2P substitution in S18, which restored EGFP-S18 stability in transfected cells (Fig. 4E and fig. S7B). Thus, the S18 degron appears to be responsible for the differential protein expression of SRY-T and SRY-S.

To determine the significance of degron-mediated SRY-S instability in vivo, we edited the endogenous *Sry* gene using CRISPR/Cas9 to create mice expressing SRY-S with C-terminal mutations or substitutions on a *Sry-T*–deficient background and evaluated their sex development. XY *Sry-S*:T15; *Sry* exon2Δ mice, which express only SRY-S but with a degron-free T15 peptide instead of S18, developed as male (fig. S8). *Sry-S*:V394P; *Sry* exon2Δ mice, with a V-2P (amino acid position 394) substitution in SRY-S, also developed as male (Fig. 4F). Moreover, SRY protein levels were markedly elevated in *Sry-S*:V394P; *Sry* exon2Δ gonads compared with *Sry* exon2Δ gonads (Fig. 4, G and H), indicating that eliminating S18 degron activity stabilizes SRY-S to the level required for inducing male development. Finally, analysis of other mouse lines expressing SRY-S with an ordered series of

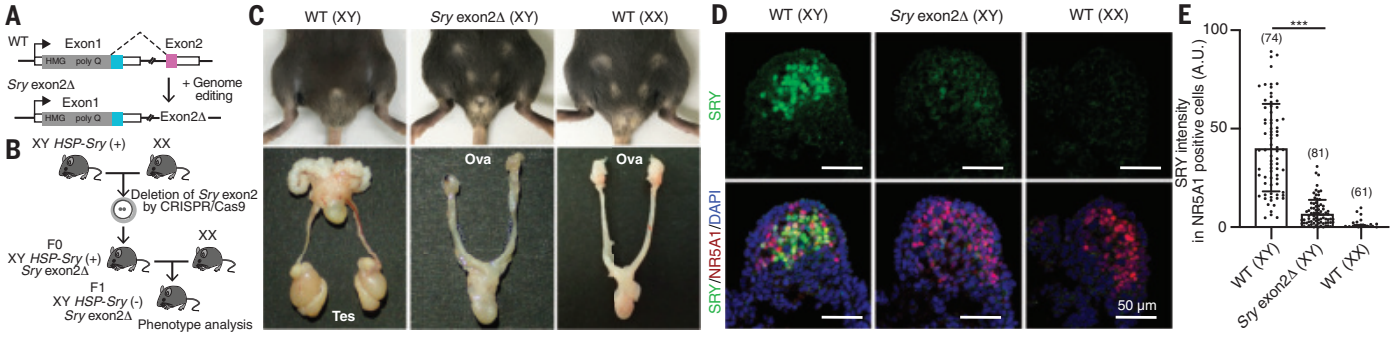
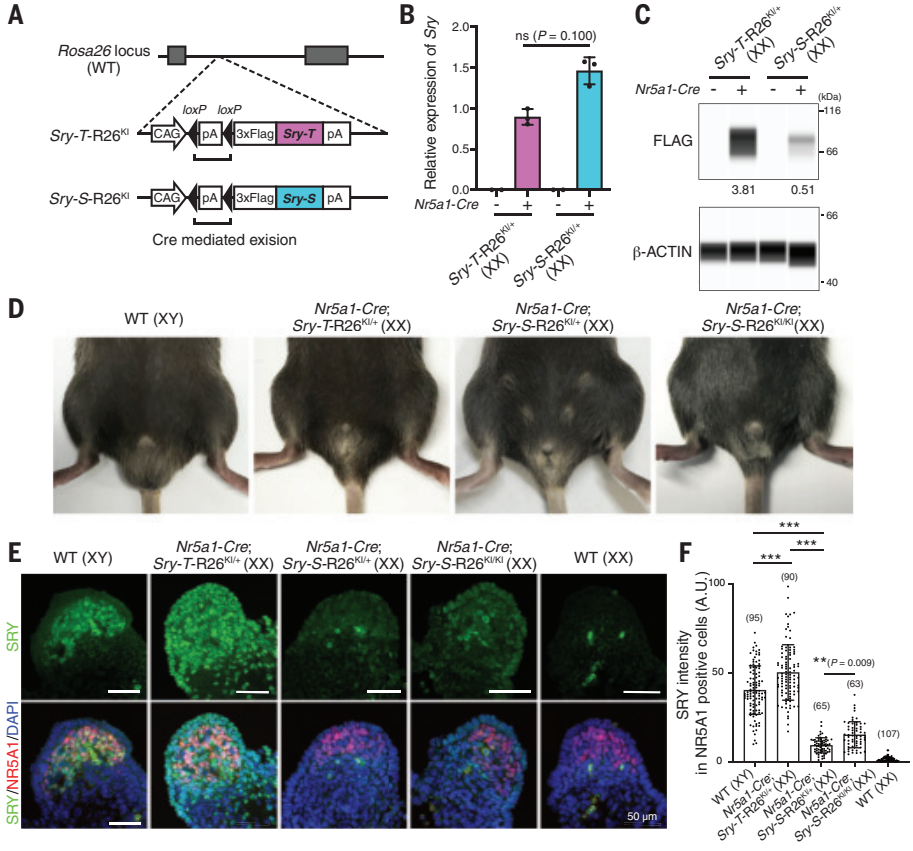


Fig. 2. *Sry-T* is necessary for male development in mice. (**A**) Deletion of *Sry* exon2 by CRISPR/Cas9. (**B**) Establishment of *Sry-T*-deficient mice. Procedures are detailed in fig. S5, A and B. To circumvent potential male infertility of *Sry-T*–deficient male mice, we used XY zygotes carrying the *HSP-Sry* transgene (*17*) for genome editing. The resultant F_1 generation without the *HSP-Sry* transgene was used for phenotypic analysis. (**C**) *Sry-T*–deficient mice exhibited male-to-female sex reversal ($n > 40$). External genitalia, nipples (top panels), and internal genitalia (bottom panels) of adult mice are shown. Tes, testis; Ova, ovary. (**D**) Coimmunostaining profiles of SRY and the gonadal somatic marker NR5A1 in E11.5 gonads of the indicated genotypes. Data are representative of biological triplicate experiments. (**E**) Quantitation of SRY in NR5A1$^+$ cells in E11.5 gonads of the indicated genotypes. A.U., arbitrary units. Data are shown as means ± SD. Numbers of analyzed cells are shown in parentheses above each bar. ***$P < 0.001$ between the indicated groups, one-way analysis of variance (ANOVA) with Tukey's post hoc test.

Fig. 3. SRY-T is the predominant SRY isoform and is sufficient to sex-reverse XX mice. (**A**) Establishment of mouse lines carrying *Sry* knocked-in *Rosa26* alleles named *Sry-T*-R26^{KI} and *Sry-S*-R26^{KI}. Cre/*lox*P-mediated genome recombination induces *Sry* expression by removing the triple-poly(A) stop cassette between the promoter and the *Sry* CDS. (**B**) Relative mRNA expression of *Sry* in E11.5 gonads of the indicated genotypes. *Sry* expression was normalized to *Gapdh*. Data are shown as means ± SD. ns, not significant, Mann–Whitney *U* test. (**C**) Capillary-based immunodetection of E11.5 gonadal cell lysates of the indicated genotypes. Expression levels of FLAG relative to β-ACTIN are indicated. (**D**) External phenotypes of 2-month-old mice of the indicated genotypes. *Nr5a1-Cre*; *Sry-T*-R26$^{KI/+}$ XX mice ($n = 6$) developed as males but *Nr5a1-Cre*; *Sry-S*-R26$^{KI/+}$ XX mice ($n = 10$) did not. *Nr5a1-Cre*; *Sry-S*-R26$^{KI/KI}$ XX mice ($n = 3$) developed as males. (**E**) Coimmunostaining profiles of SRY and NR5A1 in E11.5 gonads of the indicated genotypes. Subcellular localization of SRY-T and SRY-S are shown; for more detail, see fig. S6C. Data are representative of biological triplicate experiments. (**F**) Quantitation of SRY in NR5A1$^+$ cells in E11.5 gonads of the indicated genotypes. Data are shown as means ± SD. Numbers of analyzed cells are shown in parentheses above each bar. ***$P < 0.001$ between the indicated groups, ANOVA with Tukey's post hoc test.



C-terminal mutations on an *Sry-T*–deficient background confirmed that absence of the S18 degron is strictly required for SRY-S to activate the male pathway in vivo (fig. S8B and table S1). Thus, we conclude that the C-terminal degron prevents endogenously expressed SRY-S protein from accumulating to a level required for activating the male pathway, not only in *Sry-T*–deficient mice but also in WT XY mice.

Several previous studies showed that *Sry* transgenes encoding only SRY-S can induce testis development in XX mice under experimental conditions (*1*, *8*, *9*, *13*). The sex reversal occurred in 25% (*1*) to 75% (*8*) of mice and was shown to correlate with transgene expression levels (*9*), which was in turn influenced by integration site and/or copy number (*14*). In support of this concept, heterozygous *Nr5a1-Cre*; *Sry-S*-R26 knock-in XX mice did not develop as males but homozygotes did (Fig. 3D). Therefore, *Sry-S* transgene expression above a certain threshold can overcome the destabilizing effects of the degron. However, under physiological conditions, it is not SRY-S, but rather degron-free SRY-T, that is the bona fide testis-determining factor in the mouse.

Sry evolved from the single-exon gene *Sox3* (*15*, *16*), and all previously characterized eutherian *Sry* genes are single exon. Therefore, our findings raise the important questions of how and why an additional exon arose in mice. We found that DNA sequences similar to the mouse *Sry-S* degron are conserved in *Sry* from other

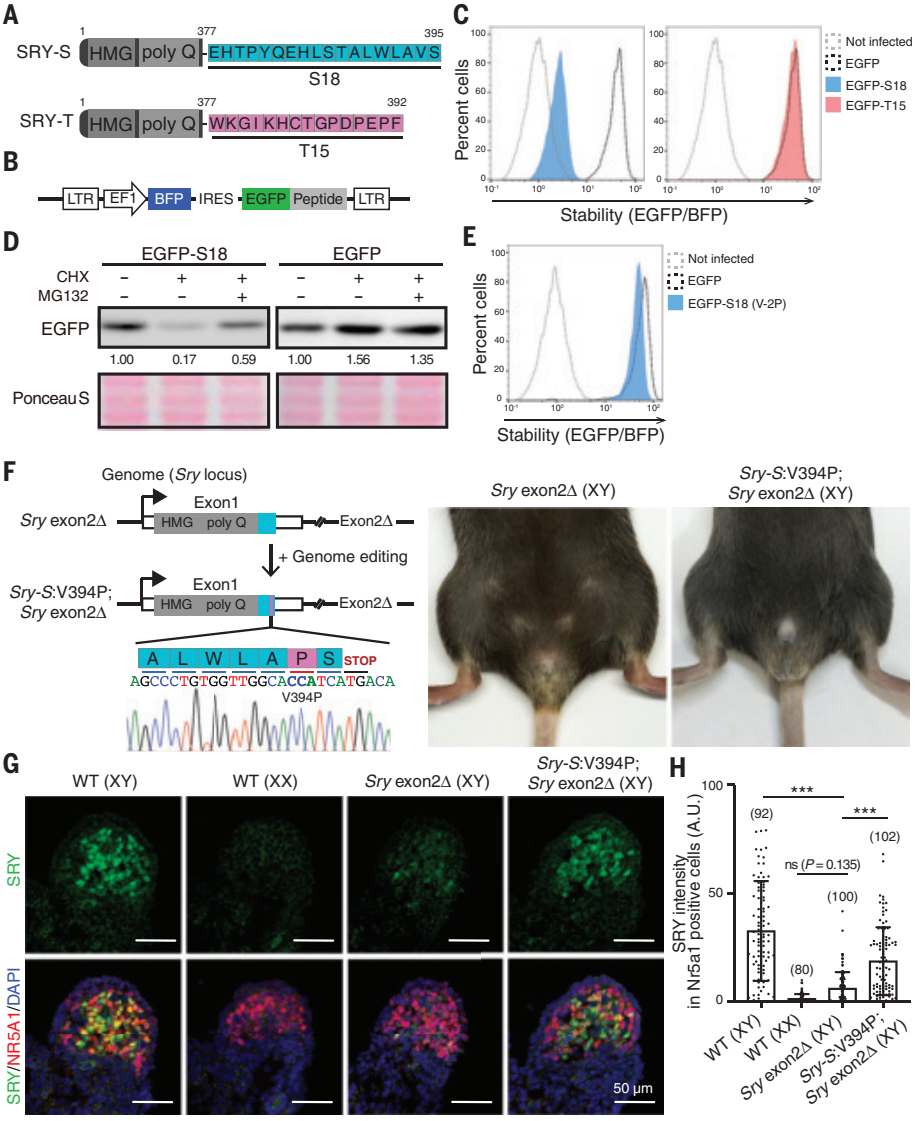


Fig. 4. SRY-S carries a C-terminal degron. (**A**) Specific C-terminal sequences of SRY-S and SRY-T were named S18 (blue) and T15 (red). (**B**) Lentivirus vector used for protein stability analysis. BFP and EGFP-fused polypeptides were translated from the same transcript. (**C**) Protein-destabilizing activity of S18 and T15 was gauged by the EGFP/BFP ratio using flow cytometry. (**D**) Western blot analysis of EGFP-S18 (left) and nonfusion EGFP (right) expressed in lentivirus-infected cells. Cells were treated with the protein synthesis inhibitor cycloheximide (CHX) for 12 hours in the presence or absence of the proteasome inhibitor MG132. Quantified protein levels are indicated. Equal protein loading was confirmed by Ponceau S staining. (**E**) V-2P substitution abolished S18 degron activity. (**F**) Editing of endogenous *Sry* to generate an *Sry-S*:V394P allele on an *Sry-T*–deficient background (left). *Sry-S*:V394P; *Sry* exon2Δ mice (n = 4 lines) developed as males (right). (**G**) Coimmunostaining profiles of SRY and NR5A1 in E11.5 gonads of the indicated genotypes. Data are representative of biological triplicate experiments. (**H**) Quantitation of SRY in NR5A1^{+} cells in E11.5 gonads of the indicated genotypes. Data are shown as means ± SD. Numbers of analyzed cells are shown in parentheses above each bar. ***P < 0.001, ns, not significant, between the indicated groups, ANOVA with Tukey's post hoc test.

rodents (fig. S9A), but SRY was not destabilized in those species (fig. S9B) because the degron was blocked from translation due to preceding stop codons (fig. S9, A to D). In mice, a different mechanism has evolved to avoid degron translation through exonization of distal retrotransposon-derived sequences (Fig. 1C and fig. S3) and splicing to excise the degron (fig. S10). The apparent replacement of SRY-S by SRY-T to overcome degron activity provides an example of Y-chromosomal gene evolution by gain of function.

Our findings overturn the conclusion that has prevailed for 30 years regarding the structure of *Sry* in the mouse, the most intensively studied mammalian model of sex determination. Further study of SRY-T will establish whether exon2 confers new functionalities on SRY in addition to protein stability and will provide a more complete understanding of how the sex-determining cascade is activated in mice compared with other mammalian species.

REFERENCES AND NOTES

1. P. Koopman, J. Gubbay, N. Vivian, P. Goodfellow, R. Lovell-Badge, *Nature* **351**, 117–121 (1991).
2. A. H. Sinclair *et al.*, *Nature* **346**, 240–244 (1990).
3. J. Gubbay *et al.*, *Nature* **346**, 245–250 (1990).
4. J. Gubbay *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **89**, 7953–7957 (1992).
5. S. Kuroki *et al.*, *Science* **341**, 1106–1109 (2013).
6. S. Kuroki *et al.*, *Genesis* **53**, 387–393 (2015).
7. I. Stévant *et al.*, *Cell Rep.* **22**, 1589–1599 (2018).
8. J. Bowles, L. Cooper, J. Berkman, P. Koopman, *Nat. Genet.* **22**, 405–408 (1999).
9. L. Zhao *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **111**, 11768–11773 (2014).
10. S. Kuroki *et al.*, *PLOS Genet.* **13**, e1007034 (2017).
11. J. M. Stenman *et al.*, *Science* **322**, 1247–1250 (2008).
12. I. Koren *et al.*, *Cell* **173**, 1622–1635.e14 (2018).
13. L. L. Washburn, K. H. Albrecht, E. M. Eicher, *Genetics* **158**, 1675–1681 (2001).
14. Y. Itoh *et al.*, *BMC Res. Notes* **8**, 69 (2015).
15. M. Stevanović, R. Lovell-Badge, J. Collignon, P. N. Goodfellow, *Hum. Mol. Genet.* **2**, 2013–2018 (1993).
16. E. Sutton *et al.*, *J. Clin. Invest.* **121**, 328–341 (2011).
17. T. Kidokoro *et al.*, *Dev. Biol.* **278**, 511–525 (2005).

ACKNOWLEDGMENTS

We thank H. Kondo and K. Ichiyanagi for critical advice on the manuscript; Y. Kanai for providing *HSP-Sry* transgenic mice; J. Kohyama for help with capillary-based immunodetection; H. Miyachi for help with animal experiments; and all members of the Tachibana laboratory for technical support. **Funding:** This work was supported by KAKENHI from the Japan Society for the Promotion of Science (JSPS) grants 19K22388 (M.T.), 18H02419 (M.T.), 17H06424 (M.T.), 17H06423 (M.T.), 20H05364 (S.K.), and 19K16490 (S.M.) and by a JSPS Research Fellowship for Young Scientists to S.M. and R.M. **Author contributions:** S.M., S.K., and M.T. designed the experiments. S.M., S.K., R.M., N.O., and M.T. performed the experiments and analyzed the data. R.M. contributed to the analyses of RNA-seq and CAGE-seq data. S.M., P.K., and M.T. interpreted the data and wrote the manuscript. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** All data are available in the main text or the supplementary materials. The accession number for the RNA-seq and CAGE-seq data generated in this study is GSE151474. The sequence of *Sry-T* was deposited in the GenBank database under accession code LC532173.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/370/6512/121/suppl/DC1
Materials and Methods
Figs. S1 to S10
Tables S1 to S3
References (*18*–*24*)
Data S1 and S2
MDAR Reproducibility Checklist

27 March 2020; accepted 6 August 2020
10.1126/science.abb6430

ELECTROCALORICS

Giant temperature span in electrocaloric regenerator

A. Torelló[1,2]*, P. Lheritier[1], T. Usui[3], Y. Nouchokgwe[1,2], M. Gérard[1], O. Bouton[1], S. Hirose[3], E. Defay[1]*

Cooling devices based on caloric materials have emerged as promising candidates to become the next generation of coolers. Several electrocaloric (EC) heat exchangers have been proposed that use different mechanisms and working principles. However, a prototype that demonstrates a competitive temperature span has been missing. We developed a parallel-plate active EC regenerator based on lead scandium tantalate multilayer capacitors. After optimizing the structural design by using finite element modeling for guidance and to considerably improve insulation, we measured a maximum temperature span of 13.0 kelvin. This temperature span breaks a crucial barrier and confirms that EC materials are promising candidates for cooling applications.

The development of highly efficient and environmentally friendly energy systems is a primary concern for mitigating global warming and promoting a sustainable use of natural resources. Approximately 20% of the world's energy consumption is used for refrigeration purposes, and the absolute amount of air conditioning units is predicted to double by 2040 (*1*). Current vapor-compression–based refrigeration systems have reached their thermodynamic limit after 100 years of advancements. They release greenhouse effect gases into the atmosphere and are typically noisy. Moreover, these systems are not suitable for the sort of miniaturization needed for several types of cooling applications.

Caloric materials undergo large entropy changes when crossing phase transitions because of stimuli such as magnetic fields (magnetocalorics), electric fields [electrocalorics (ECs)], uniaxial stress (elastocalorics), or pressure (barocalorics) (*2*–*5*). These materials have emerged as promising candidates for cooling technologies because they do not suffer from some of the disadvantages of the vapor-compression systems. Their efficiency is also presumed to be higher (*6*). In comparison with the other caloric materials, ECs are of particular interest because the coefficient of performance (COP) can be enhanced with energy recovery (*7*). Electric fields are cheaper and easier to produce than magnetic fields, which require permanent magnets for magnetocalorics. EC materials fatigue less than elastocalorics, which require stress stimuli that drives mechanical failure of these materials. Nevertheless, EC prototypes have struggled to maintain temperature spans comparable with those of magnetocaloric or elastocaloric prototypes, for which more than 10 K has been repeatedly reported (*8*–*13*). The main reasons are a modest intrinsic adiabatic EC temperature change in most EC materials but also low thermal conductivity and mandatory electrical connections, both of which hinder heat transfer.

The first EC prototype dates back to three decades ago. Sinyavsky *et al.* proposed an active EC regenerator (AER) based on lead

[1]Materials Research and Technology Department, Luxembourg Institute of Science and Technology (LIST), 41 Rue du Brill, Belvaux L-4422, Luxembourg. [2]University of Luxembourg, 2 Avenue de l'Université, Esch-sur-Alzette L-4365, Luxembourg. [3]Murata Manufacturing Co., Nagaokakyo, Kyoto 617-8555, Japan. *Corresponding author. Email: alvar.torello@list.lu; emmanuel.defay@list.lu

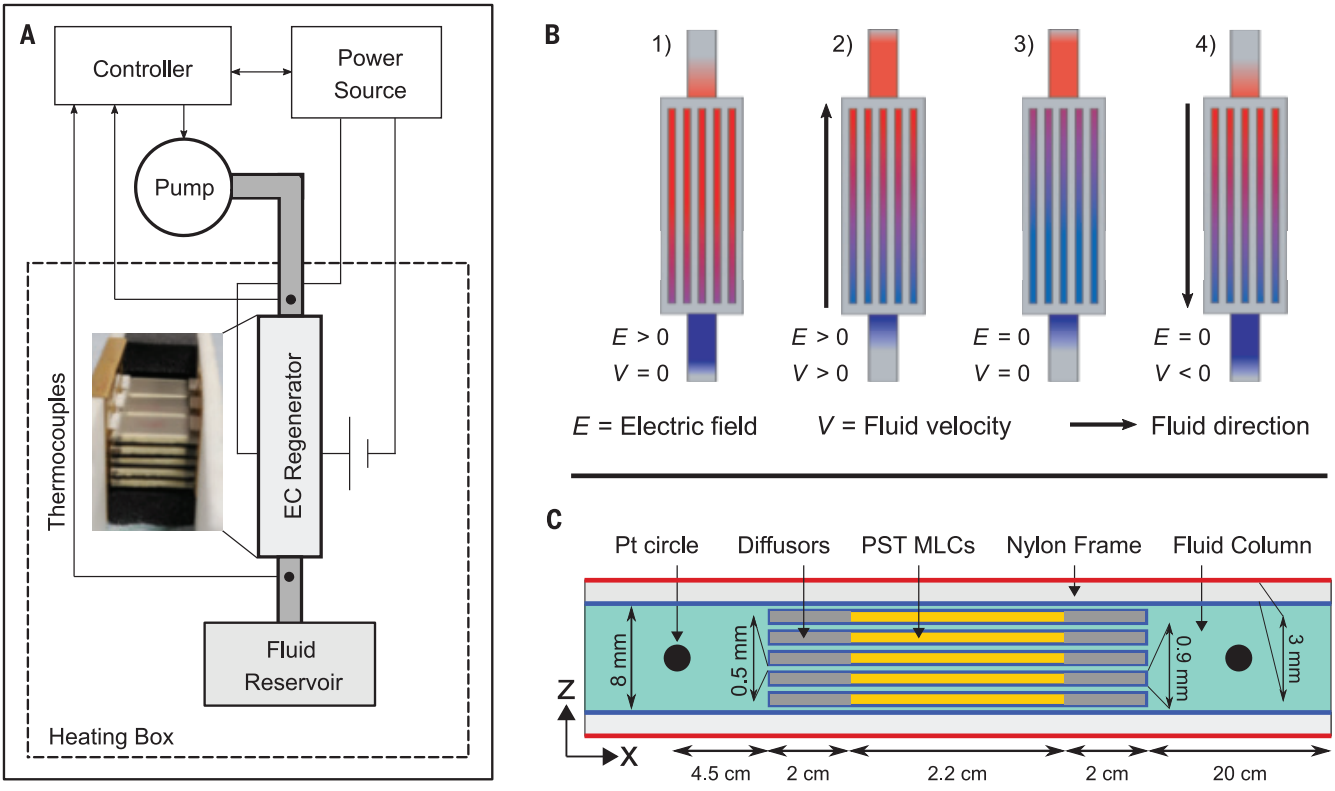


Fig. 1. Experimental design. (**A**) Schematic showing the experimental setup. (**B**) The operation principle for one cycle to stabilize the temperature gradient along the regenerator. (**C**) The geometry of our simulated model using COMSOL Multiphysics 5.2a software. We applied adiabatic conditions at the exterior walls (red lines) as a first approximation, whereas a no-slip boundary condition was set at the solid-fluid interface (blue lines).

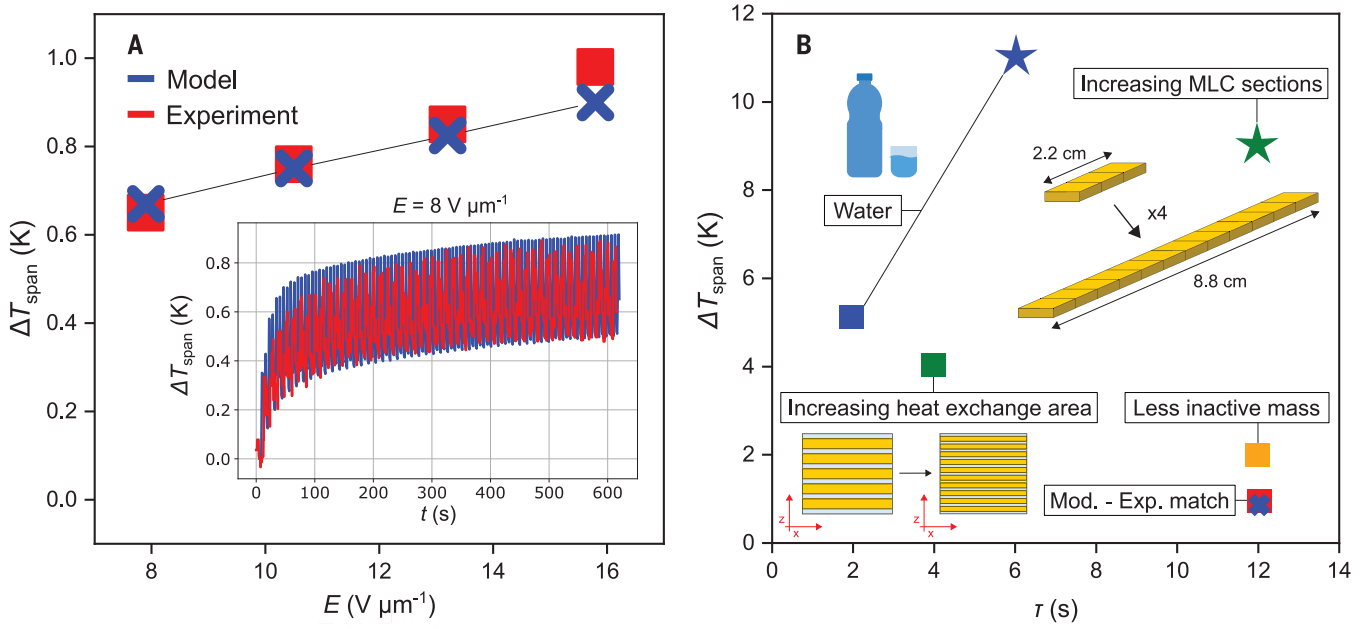


Fig. 2. Modeling results. (**A**) Temperature span (ΔT_{span}) comparison as a function of the applied electric field E and time t (inset) for REG1 (fig. S2) (*35*). (**B**) Modeled ΔT_{span} as a function of the cycle period τ. We used a 15.8 V µm^{-1} electric field, corresponding to a 2.2 K adiabatic EC temperature change of the material (ΔT_{span}).

scandium tantalate (PST) (*14*, *15*). This design was inspired by the magnetocaloric regenerator that Brown invented in 1975 (*16*), improved first by Steyert (*17*) and later on by Barclay (*18*), and reported a temperature span of 5 K. In 1995, Sinyavsky performed preliminary investigations on a larger prototype that was 30 cm long and made of 400 g of PST. Sinyavsky suggested that a temperature span of 15 K should be achievable (*19*). After the discovery in 2006 of intrinsic giant EC temperature changes in ceramic and polymer thin films (*20*, *21*), several prototypes were proposed (*22*–*28*), most of them following this same principle of regeneration. These devices are typically classified as fluid- or solid-based, depending on the substance that is used to collect the heat generated by the EC material (*29*). Examples of fluid-based coolers are given by Plaznik *et al.* (*24*) and Blumenthal *et al.* (*26*), in which lead magnesium niobate–lead titanate (PMN-PT) plates and multilayer capacitors were submerged in silicone oil. Concerning solid-state devices, Gu *et al.* (*22*) based their cooler on the oscillatory movement of stainless-steel sheets through PVDF modules, and Wang *et al.* (*23*) developed a heat switch with $BaTiO_3$ commercial multilayer capacitors (BTO MLCs). Despite the various EC materials and working principles shown, the maximum temperature span achieved by these prototypes was 6.6 K (*22*).

We followed the principle of active regeneration in designing our EC cooling device (*11*, *17*, *18*, *24*, *26*, *30*). Active regeneration is common in heat pumps that are based on materials with a low intrinsic temperature change because it permits the device to display a temperature difference between the hot and cold side of the regenerator (ΔT_{span}) larger than the adiabatic EC temperature change of the material (ΔT_{EC}). The corresponding ratio between these two temperature differences is defined as the regeneration factor. The principle of active regeneration in fluid-based coolers requires the caloric material to be porous so that a coolant fluid can be sent back and forth. The fluid movement must be synchronized with the heating and cooling steps of the corresponding caloric effect. In ECs, parallel plates are the dominating structure because applying a uniform electric field to the assembly is easy. Specifically, the structure consists of thin plates made of the caloric material that are stuck one on top of the other. In between these plates, spacers are positioned to create void channels where the fluid can flow through. We based the design of our regenerator on the circulation of a dielectric fluid through parallel plates. We built the plates from $Pb(Sc,Ta)O_3$ multilayer capacitors (PST-MLCs) (*31*) because of their sharp first-order phase transition and availability. The bespoke samples were fabricated by Murata Manufacturing. The multilayer structure consists of ceramic thin layers sandwiched with previously studied inner electrodes in between (*32*, *33*). The MLC architecture combines the advantages of bulk and thin-films samples. The inner electrodes are tens of micrometers apart, which decreases the voltage required to reach the same amount of field as in the bulk. Furthermore, the sample breakdown field increases when reducing thickness, allowing higher electric fields to be applied that favor a higher ΔT_{EC}. The overall MLC object is millimeters in size, providing more thermal mass than thin films. This prevents the heat generated with the EC effect from vanishing into the surroundings before we can use it. Moreover, the inner electrodes can also enhance heat exchange (*34*). We used two types of MLCs with 19 and 9 PST layers, each 38 µm thick. The final sample dimensions were 10.6 (L) by 7.2 (W) by 0.9 or 0.5 mm (T) (*35*).

We designed our experiment with the EC regenerator that we formed with a matrix of PST-MLCs, a syringe pump that displaces a fluid, a power supply to trigger the EC effect, and type K thermocouples to monitor the temperature (Fig. 1A). We enclosed the EC regenerator with a heating box to control the starting temperature (optimally above 25°C) of the experiment (*35*). The AER was connected to the power supply, which we chose to run in the current source mode. This supply charges (and discharges) the EC capacitors according to the compliance voltage (fig. S1) (*35*). The fluid system consists of a nonclosed single loop where one end of the regenerator is connected to the syringe pump and the other end is attached to an unsealed fluid reservoir. Our operational system (Fig. 1B) is constituted by cycles of four steps, with the first two and last two steps occurring simultaneously to emulate an Ericsson-Brayton–like cycle (*30*, *35*). We charged the EC capacitors in the first step, which increased their temperature by means of the EC effect. In the second step, we activated fluid movement to transport the heat generated to one side of the AER (hot side).

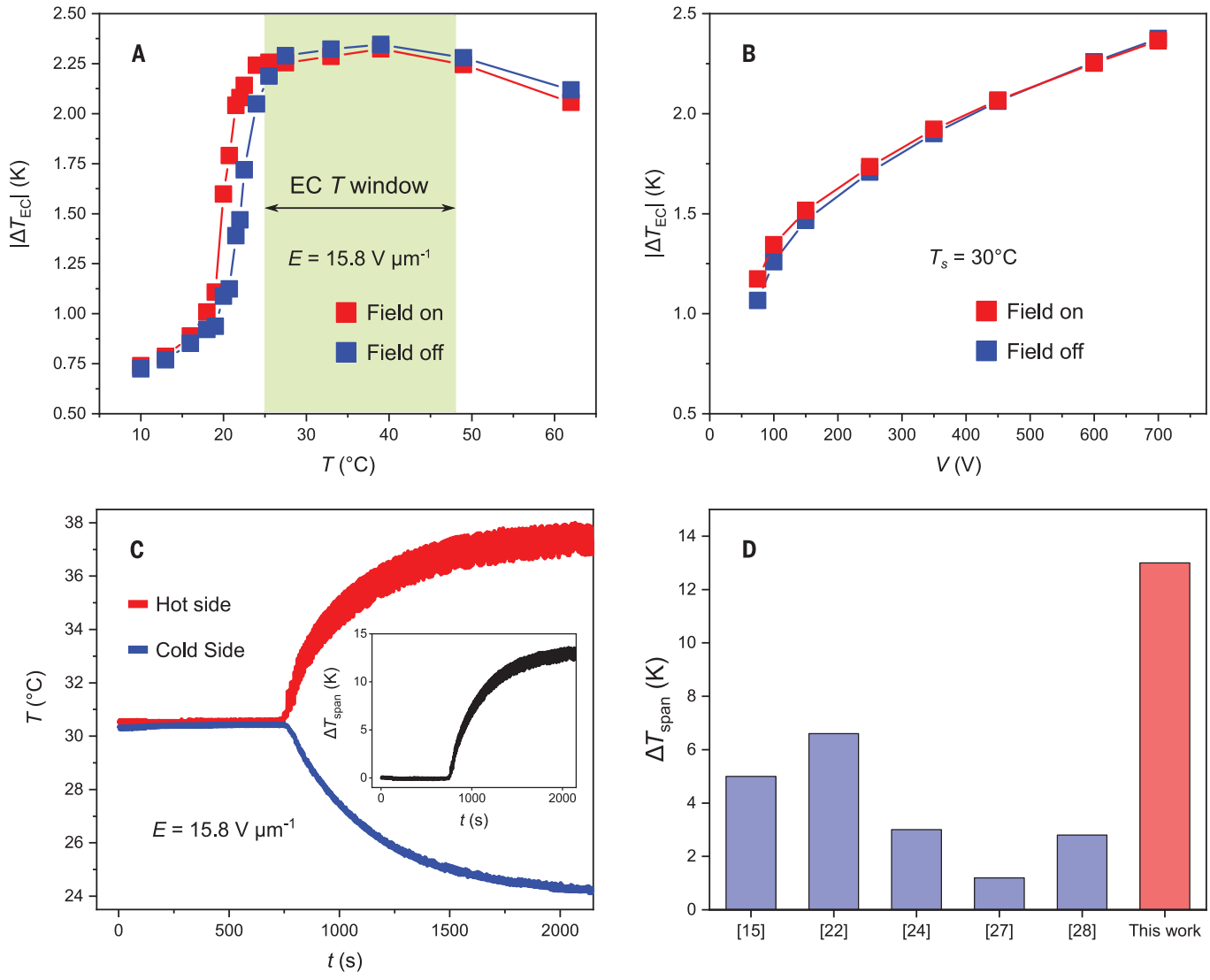


Fig. 3. Experimental results. (**A** and **B**) Adiabatic EC temperature change (ΔT_{EC}) of 0.5-mm-thick PST-MLCs (A) as a function of the starting temperature T_s for an applied electric field E = 15.8 V µm^{-1} and (B) as a function of the applied voltage V and T_s = 30°C. (**C**) Time (t) evolution T and (inset) temperature difference of the REG2's hot and cold side ΔT_{span} for an E = 15.8 V µm^{-1} and T_s = 30°C. (**D**) ΔT_{span} achieved by other EC devices (*15*, *22*, *24*, *27*, *28*).

In the third step, we discharged the EC capacitors. Their temperature thus decreases because of the EC effect. Last, we reversed the fluid direction of the fluid motion to transport the cooled fluid to the other side (cold side). After this cycle is repeated a few times, a stable temperature gradient occurs along the fluid direction in the AER. The ΔT_{span} can be several times the ΔT_{EC} of the material.

We formed our first EC regenerator, labeled as REG1 (fig. S2 and table S1) (*35*), with 15 0.9-mm-thick PST-MLCs, distributed in a matrix of three columns by five rows. We 3D-printed nylon and polyether ether ketone (PEEK) pieces to form the structure and seal it from the fluid (*24*, *26*, *35*). The fluid used was a dielectric fluid to prevent short circuits. This design generated a maximum ΔT_{span} of 1 K (Fig. 2A) from a 2.2 K ΔT_{EC}. This corresponds to a regeneration factor of 0.45.

To improve our poor experimental performance, we used finite element numerical modeling. Our simulated geometry included some of the structural bodies (Fig. 1C and tables S2 and S3) (*35*). The good match to our experiments validated the model results (Fig. 2A). This gave us confidence that we could use modeling to explore different configurations for prototypes in order to find ones that had larger ΔT_{span} (Fig. 2B). We spotted that the inactive structural parts in our first design were acting as heat traps and absorbing a non-negligible amount of heat, which compromised the overall performance (fig. S3) (*35*). Our main finding is that extra attention has to be paid to thermal insulation by decreasing as much as possible the structural parts of the regenerator.

After we removed the problematic structural elements from our model, we obtained a ΔT_{span} of 2 K (Fig. 2B, orange square). We also modified the MLCs' thickness. Thinner MLCs increases the heat exchange area. By decreasing the MLCs' thickness from 0.9 to 0.5 mm, which is possible to implement experimentally by changing the number of PST layers from 19 to 9, our model reaches a ΔT_{span} of 4 K, and the period reduces to 4 s. Increasing the length of the regenerator, by increasing the number of MLC columns, makes it more difficult to relax the temperature gradient. This allows for a larger ΔT_{span} to build. We modeled a fourfold increase of the MLCs overall length that resulted in a

ΔT_{span} of 9 K (Fig. 2B, green star). Last, we modeled the impact of using water instead of dielectric fluid, which increased the ΔT_{span} by 20% and reduced the period in half (Fig. 2B, blue square and blue star). In comparison with the dielectric fluid, water has six times more thermal conductivity, for which heat exchange is enormously enhanced, and three times more specific heat, for which more heat can be stored. In addition, water is five times less viscous, which lowers fluid self-heating and pumping pressure.

We used the modeling to guide changes to our design. We removed all structural pieces to reduce the inactive mass of the system, provided insulation with polyurethane foam, used 0.5-mm-thick PST-MLCs, and increased the regenerator length (fig. S4 and table S4) (*35*). Our final structure, labeled as REG2, consisted of 128 PST-MLCs, distributed in a matrix of 16 columns by eight rows. The dimensions were 115 (L) by 10.4 (W) by 6.25 mm (T). The total PST-MLC mass of 38.4 g was only 60% active because of the overlap of the inner MLC electrode sheets. We restricted these to avoid short-circuiting. We used the same dielectric fluid as in REG1 (*35*).

We measured the ΔT_{EC} of a 0.5-mm-thick PST-MLC as a function of the starting temperature T_s for an applied electric field (E) of 15.8 V μm^{-1} (Fig. 3A). We observed a maximum ΔT_{EC} of 2.3 K at 38°C. Below 25°C, ΔT_{EC} starts decreasing and reaches ~0.6 K at 10°C. Between 18° and 25°C, we found asymmetric behavior for ΔT_{EC} because its amplitude is different when turning the field on (heating) or off (cooling). This range corresponds to a region where the ferroelectric (FE) phase of PST is stable and applying an E cannot fully drive the desired phase transition from paraelectric (PE) to FE. The optimum working temperature lies in the range of 25° to 50°C, in which the PE phase is stable and the FE phase can be fully driven by an E. Likewise, we measured ΔT_{EC} as a function of E for a starting temperature T_s = 30°C (Fig. 3B). As we expected, ΔT_{EC} increases with the applied field and reaches a common linear regime once the field has brought the material into the FE phase. In this case, we observed no asymmetry between heating and cooling. We measured a ΔT_{EC} in the material of 2.37 K for a maximum applied voltage V of 700 V (E = 18.4 V μm^{-1}). These values are in agreement with previous results (*31*), in which better B-site–ordered 0.9-mm-thick PST-MLCs displayed a ΔT_{EC} of 2.6 K in the same temperature range for E = 15.8 V μm^{-1}. For this comparison, we have taken the value reported assuming total internal thermalization because active volume represents only 60% of the MLC body. A maximum adiabatic ΔT_{EC} > 5 K was reported at E = 29.2 V μm^{-1} near 330 K in the MLC active volume (*31*). Our samples break down at this level of an E magnitude, so we limited our field to 15.8 V μm^{-1}.

We measured a maximum ΔT_{span} of 13.0 K after 1500 s of operation in REG2 (Fig. 3C). T_s was set to 30°C to make use of the entire EC effect temperature window (Fig. 3A) and maximize the ΔT_{span}. Under these conditions, the PST-MLCs displayed a ΔT_{EC} of 2.2 K, which translates to a regenerator factor of 5.9. The ΔT_{span} we measured is more than one order of magnitude higher than our initial design and much higher than several other notable EC (Fig. 3D). The large variation of temperature we measured is very close to the most competitive elastocaloric prototypes, which have a ΔT_{span} of 15.3 K (*11*). We believe that there is still room for improvement. Most of our PST-MLCs did not have an even and flat shape. These morphologies cause irregularities in the fluid slits or even force complete closure, which compromises heat exchange and lowers heat regeneration. Moreover, the thickness of our PST-MLCs (0.5 mm) is still thicker than the EC bulk plates used in most of the literature examples, reaching values as low as 200 μm (*11*, *19*, *22*, *24*, *26*, *28*). Thus, heat regeneration could likely be further enhanced by the use of thinner and more regularly shaped PST-MLCs. Highly ordered PST-MLCs with correspondingly higher breakdown strength would allow increasing the ΔT_{EC} by more than 50% (*31*), increasing accordingly the ΔT_{span} of the device. The use of water as a working fluid instead of dielectric fluid could also increase performance, as shown by our model results. Currently, water is a challenge for use in EC prototypes because it requires electrically insulating the entire EC material.

Last, a third regenerator, REG3, with 32 0.9-mm-thick PST-MLCs and an electrical heater in the cold side, was built to report cooling power values following the same design as REG2. We measured 3 K of no-load temperature span and 0.26 W (12 W kg^{-1} per mass of PST-MLCs) of maximum cooling power. These data were used to fit our numerical model to predict the performances of other configurations. For the REG2 device, simulations displayed a no-load temperature span of 12.8 K, which is very close to the 13 K in the experiment, and a maximum cooling power of 1.22 W (32 W kg^{-1}). In addition, the hypothetical configuration of 0.3- and 0.2-mm-thick PST-MLCs in (*31*), with ΔT_{EC} of 5.5 K, and water as a coolant fluid, was modeled. Simulations reported a no-load temperature span of 33 and 47.5 K and a maximum cooling power of 10.6 W (550 W kg^{-1}) and 16.3 W (850 W kg^{-1}), respectively (fig. S5) (*35*), which is similar to the best magnetocaloric and elastocaloric coolers.

Our work should raise interest not only in the scientific community, but also in industry; now that the 10 K barrier has been crossed, large cooling powers are predicted, and enhanced efficiency can be obtained through energy recovery (*7*). We hope that this will promote the development of EC prototypes in the years to come as an alternative to the currently dominating vapor-compression systems.

REFERENCES AND NOTES

1. IEA, "The Future of Cooling" (IEA, 2018); www.iea.org/reports/the-future-of-cooling.
2. S. Fähler *et al.*, *Adv. Eng. Mater.* **14**, 10–19 (2012).
3. L. Mañosa, A. Planes, M. Acet, *J. Mater. Chem. A Mater. Energy Sustain.* **1**, 4925–4936 (2013).
4. X. Moya, S. Kar-Narayan, N. D. Mathur, *Nat. Mater.* **13**, 439–450 (2014).
5. I. Takeuchi, K. Sandeman, *Phys. Today* **68**, 48–54 (2015).
6. X. Moya, E. Defay, V. Heine, N. D. Mathur, *Nat. Phys.* **11**, 202–205 (2015).
7. E. Defay *et al.*, *Nat. Commun.* **9**, 1827 (2018).
8. S. Qian *et al.*, *Sci. Technol. Built Environ.* **22**, 500–506 (2016).
9. M. Schmidt, A. Schütze, S. Seelecke, *Int. J. Refrig.* **54**, 88–97 (2015).
10. H. Ossmer, C. Chluba, S. Kauffmann-Weiss, E. Quandt, M. Kohl, *APL Mater.* **4**, 064102 (2016).
11. J. Tušek, K. *et al.*, *Nat. Energy* **1**, 16134 (2016).
12. H. Ossmer *et al.*, *Smart Mater. Struct.* **25**, 085037 (2016).
13. B. Yu, M. Liu, P. W. Egolf, A. Kitanovski, *Int. J. Refrig.* **33**, 1029–1060 (2010).
14. Y. V. Sinyavsky, N. D. Pashkov, Y. M. Gorovoy, G. E. Lugansky, L. Shebanov, *Ferroelectrics* **90**, 213–217 (1989).
15. Y. V. Sinyavsky, V. M. Brodyansky, *Ferroelectrics* **131**, 321–325 (1992).
16. G. V. Brown, *J. Appl. Phys.* **47**, 3673–3680 (1976).
17. W. A. Steyert, *J. Appl. Phys.* **49**, 1216–1226 (1978).
18. J. A. Barclay, W. A. Steyert, Active magnetic regenerator. U.S. patent no. 4.332.135 (1982).
19. Y. V. Sinyavskii, *Chem. Petrol. Eng.* **31**, 295–306 (1995).
20. A. S. Mischenko, Q. Zhang, J. F. Scott, R. W. Whatmore, N. D. Mathur, *Science* **311**, 1270–1271 (2006).
21. B. Neese *et al.*, *Science* **321**, 821–823 (2008).
22. H. Gu *et al.*, *Appl. Phys. Lett.* **102**, 122904 (2013).
23. Y. D. Wang *et al.*, *Appl. Phys. Lett.* **107**, 134103 (2015).
24. U. Plaznik *et al.*, *Appl. Phys. Lett.* **106**, 043903 (2015).
25. D. Sette *et al.*, *APL Mater.* **4**, 091101 (2016).
26. P. Blumenthal, C. Molin, S. Gebhardt, A. Raatz, *Ferroelectrics* **497**, 1–8 (2016).
27. T. Zhang, X.-S. Qian, H. Gu, Y. Hou, Q. M. Zhang, *Appl. Phys. Lett.* **110**, 243503 (2017).
28. R. Ma *et al.*, *Science* **357**, 1130–1134 (2017).
29. P. Blumenthal, A. Raatz, *EPL* **115**, 17004 (2016).
30. A. Kitanovski *et al.*, "Active Magnetic Refrigeration" in *Magnetocaloric Energy Conversion* (Springer, 2015), pp 97–166.
31. B. Nair *et al.*, *Nature* **575**, 468–472 (2019).
32. S. Kar-Narayan, N. D. Mathur, *J. Phys. D Appl. Phys.* **43**, 032002 (2010).
33. X. Moya, E. Defay, N. D. Mathur, S. Hirose, *MRS Bull.* **43**, 291–294 (2018).
34. Y. Liu, B. Dkhil, E. Defay, *ACS Energy Lett.* **1**, 521–528 (2016).
35. Materials and methods are available as supplementary materials.

ACKNOWLEDGMENTS

We thank X. Moya, N. Mathur, and B. Nair for discussions. We thank R. Faye, H. Strozyk, and S. Nicolau for their help in the development of REG1. We also thank N. Furusawa, Y. Inoue, and K. Honda for their kind assistance in fabricating the MLCs. **Funding:** A.T., P.L., Y.N., M.G., O.B., and E.D. acknowledge the Fonds National de la Recherche (FNR) of Luxembourg for supporting this work through the projects MASSENA PRIDE/15/10935404/Defay-Siebentritt, CAMELHEAT C17/MS/11703691/Defay, and COFERMAT FNR/P12/4853155/Kreisel. **Author contributions:** E.D. suggested the experimental study. S.H and T.U. prepared the PST-MLC samples. M.G. designed REG1 with E.D. and A.T.; A.T. developed and investigated the numerical modeling that led to Fig. 2, A and B. E.D. and A.T. designed REG2 with M.G.; A.T. and Y.N. ran the set of experiments that led to Fig. 3, A and B. A.T. and P.L. ran the experiments with the prototype that led to Fig. 3C. A.T. ran the experiments and simulations that led to fig. S5. E.D., A.T., and P.L. interpreted the key findings. P.L. prepared

and automated the experimental set up with O.B.; A.T. prepared the figures. A.T. wrote the manuscript with E.D. **Competing interests:** Three authors (A.T., M.G., and E.D.) have filed a patent in December 2019 about an EC regenerator design under the number LU101559 (Luxembourg). The remaining authors declare no competing interests; **Data and materials availability:** All data are available in the manuscript or in the supplementary material. Correspondence and requests for materials should be addressed to A.T. or E.D.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/370/6512/125/suppl/DC1
Materials and Methods
Figs. S1 to S5
Tables S1 to S9
Reference (*36*)

20 March 2020; accepted 12 August 2020
10.1126/science.abb8045

ELECTROCALORICS

A high-performance solid-state electrocaloric cooling system

Yunda Wang[1]*, Ziyang Zhang[1], Tomoyasu Usui[2], Michael Benedict[1], Sakyo Hirose[2], Joseph Lee[1], Jamie Kalb[1], David Schwartz[1]*

Electrocaloric (EC) cooling is an emerging technology that has broad potential to disrupt conventional air conditioning and refrigeration as well as electronics cooling applications. EC coolers can be highly efficient, solid state, and compact; have few moving parts; and contain no environmentally harmful or combustible refrigerants. We report a scalable, high-performance system architecture, demonstrated in a device that uses $PbSc_{0.5}Ta_{0.5}O_3$ EC multilayer ceramic capacitors fabricated in a manufacturing-compatible process. We obtained a system temperature span of 5.2°C and a maximum heat flux of 135 milliwatts per square centimeter. This measured heat flux is more than four times higher than other EC cooling demonstrations, and the temperature lift is among the highest for EC systems that use ceramic multilayer capacitors.

Over the two centuries since the invention of vapor compression, heat pumps have become an increasingly essential technology, with applications ranging from air conditioning and refrigeration to the stabilization of precision electronic components. Space cooling currently accounts for about 20% of the total electricity used in buildings and 10% of total electricity consumption around the world (*1*). As large countries continue to develop, the demand for air conditioning will increase. The global installation of room air conditioners is estimated to reach 4.5 billion units by 2050 (*2*). Applications extend beyond large-scale cooling, with efficient heat pumps being critical for thermal management of many electronic devices and sensors (*3*), including night-vision infrared sensors (*4*) and laser diodes (*5*).

Nevertheless, cooling technologies have only seen incremental changes in the past few decades. Vapor compression refrigeration, patented in 1803 (*6*), remains the predominant cooling technology in use. Vapor compression cooling has been difficult to displace because many decades of development have led to high efficiency, scalability, reliability, and relatively compact size (*7*). However, several issues suggest a need to move beyond this technology. The high-performance, nonflammable refrigerants commonly in use are hydrofluorocarbons (HFCs)—global-warming forcers typically thousands of times more potent than carbon dioxide (*8*). Phase-down of HFCs may be one part of a response to climate change. Replacement

[1]PARC, A Xerox Company, 3333 Coyote Hill Rd., Palo Alto, CA 94304, USA. [2]Murata Manufacturing Co., Ltd., 1-10-1, Higashikotari, Nagaokakyo, Kyoto 617-8555, Japan.
***Corresponding author. Email: yundaw@parc.com (Y.W.); dschwart@parc.com (D.S.)**

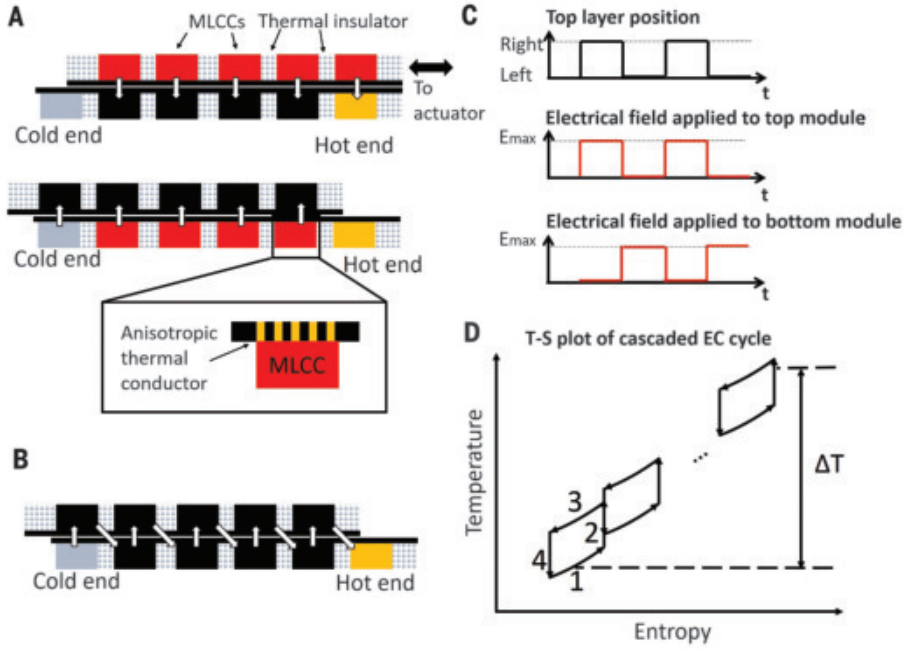


Fig. 1. Operation of the EC system. (**A**) Relative positions of the EC modules in the two heat transfer stages. (**B**) The heat flow path. (**C**) Time-domain illustration of the cyclic application of electric field and actuation. E_{max}, The maximum electric field applied to the module; t, time. (**D**) Schematic of the Brayton cycle as observed for each MLCC. Stage 1 is a cold isoelectric stage. The electric field across the MLCC is low, the MLCC absorbs heat from an adjacent MLCC in the opposite module, and both temperature and entropy increase. Stage 2 is a nominally adiabatic stage. The electric field across the MLCC is increased, resulting in an ECE-induced temperature increase of the MLCC; at the same time, the actuator voltage is switched so that the MLCC module alignment shifts. This stage is ideally isentropic. Stage 3 is a hot isoelectric stage. With the electric field still high, the MLCC rejects heat to the adjacent MLCC of the opposite module; both temperature and entropy decrease. Stage 4 is a nominally adiabatic stage. The electric field is switched back to zero, leading to a temperature decrease in the MLCC; the actuator is switched so that the MLCCs return to their original alignment. This stage is ideally isentropic. T-S, temperature-entropy.

refrigerants require trade-offs between cost, efficiency, and safety (*9*). Trans-critical CO_2 systems have favorable thermodynamic properties but are incompatible with existing compressors and other equipment (*10*).

The inveteracy of vapor compression has hindered the development of zero–global warming potential (GWP) alternative technologies (*11*). High efficiency, characterized by the coefficient of performance (*COP*) (the heat pumping power divided by input power at a set operating temperature), is essential for addressing indirect CO_2 emissions associated with electricity consumption. Of non-vapor compression space-cooling technologies, only absorption coolers and evaporative coolers have achieved measurable market penetration, yet these are typically appropriate for specific-use cases or environments. Other technologies in development include those based on the thermoelectric effect (TE) (*12*), the magnetocaloric effect (*13*), the electrocaloric effect (ECE) (*14*, *15*), the elasto- and barocaloric effects (*16*), adsorption cooling (*17*), and enhanced radiative cooling (*18*). Gas-phase technologies, including Stirling and thermoacoustic coolers, have found niche applications, for example, in cryogenic pulse-tube refrigerators. The only technology that has been widely commercialized is based on the TE, primarily as Peltier devices for electronics cooling. The success of TE technology is directly attributable to its compact form factor and simple operation. Despite decades of intensive research, the low efficiency of TE commercial devices arises from fundamental material and form-factor limitations, which have not been substantially improved. There appears to be no clear path to achieving vapor compression–equivalent efficiencies, especially for larger-scale systems (*19*).

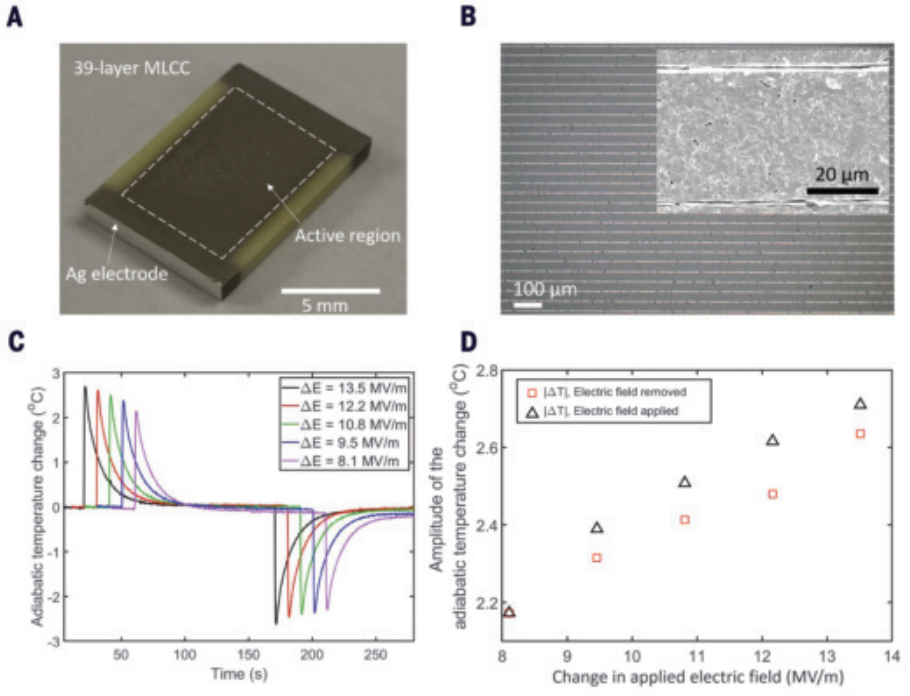


Fig. 2. MLCC images and characterization results. (**A**) Photograph of an MLCC. (**B**) Photograph and SEM (inset) of a cross section of an MLCC. (**C**) Induced temperature changes of the MLCC over time with the application and removal of polarizing electric fields. Measurements are performed at room temperature. ΔE, change in applied electric field. (**D**) Amplitude of the temperature changes during heating and cooling as a function of electric field magnitude.

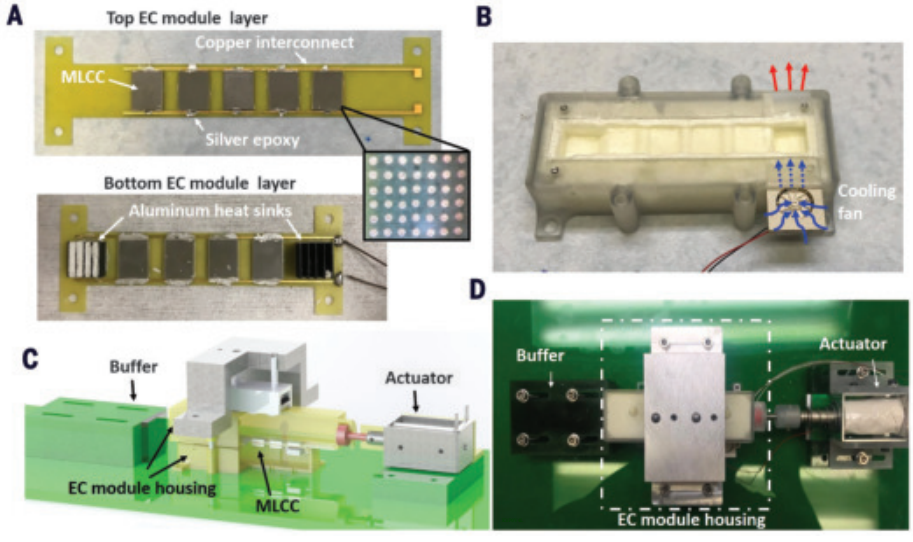


Fig. 3. EC cooler and key components. (**A**) Top and bottom EC modules assembled on ATC plates. The inset shows the copper through-vias behind the MLCCs. (**B**) Bottom housing structure, including miniature fan and air flow path. (**C**) Solid model of the cooler assembly. (**D**) Photograph of the cooler assembly.

Among emerging technologies, electrocaloric (EC) cooling has the potential for high efficiency and near solid-state operation. The ECE is a material property characterized by an adiabatic temperature change with an applied electric field (*20*). A heat pump can be realized by thermally coupling EC materials alternately to a heat sink and source synchronously with the application and removal of a polarizing field. The ECE is driven by electric fields, eliminating the need for large compressors, pumps, or magnets. The strategy is a scalable replacement, both for vapor compression systems and for thermoelectric devices (*21*–*24*). In recent decades, tremendous progress has been made in material development, yielding extremely promising materials that evince a "giant" ECE associated with a first- or second-order ferroelectric phase change. Both high-performing relaxor ferroelectric ceramic (*14*, *25*–*30*) and polymer EC materials (*15*, *27*, *28*) have been developed. Predictions of adiabatic temperature changes greater than 40°C were based on extrapolations from laboratory measurements (*27*). Nevertheless, demonstrations of cooling systems based on advanced materials have lagged. The primary limitation has been the inability to realize these materials in scalable form factors at quantities compatible with incorporation into systems. Material properties have been characterized in thin films and small bulk samples in laboratories. Achieving similar properties

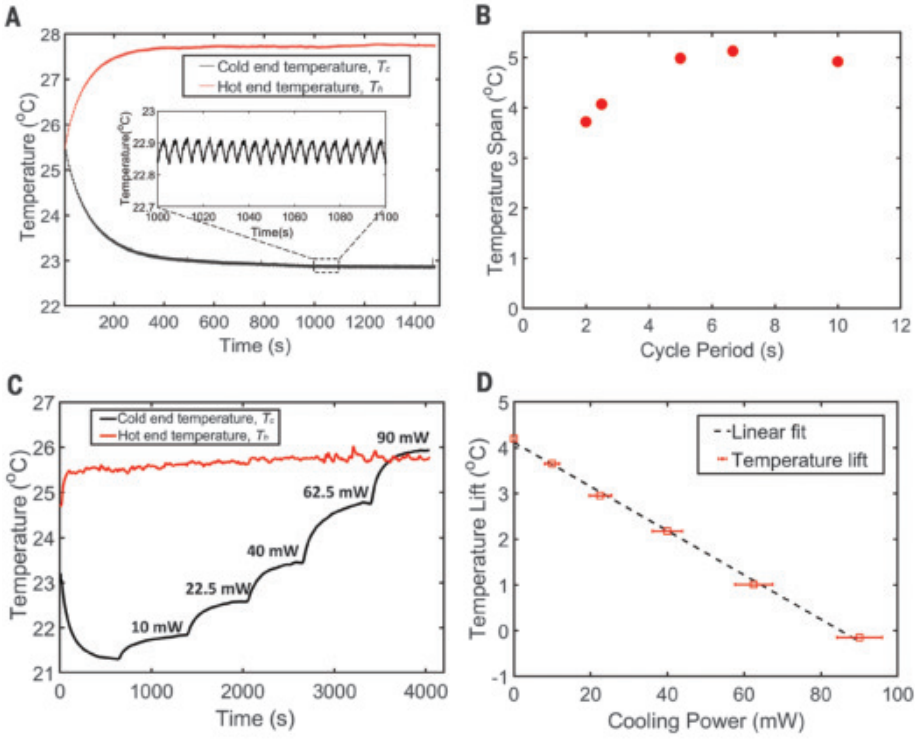


Fig. 4. Characterization result of the EC cooler. (**A**) Measured hot- and cold-end temperatures with a 400-V polarizing voltage and 5-s cycle period. Data are sampled at 20 Hz. (**B**) The maximum temperature span as a function of operating cycle period with 400-V operation. (**C**) Measured hot- and cold-end temperatures at different applied heater powers. Data are sampled at 20 Hz and filtered by a 200-point moving average. (**D**) System temperature lift, $T_h - T_c$, as a function of cooling power. Error bars show the maximum value range based on the calculated uncertainty in the heater power measurement.

in components with high thermal mass has remained elusive. Most system demonstrations (*31*–*38*) have low temperature lifts, and only a few reported measured cooling power (*31*, *32*). Polymer and ceramic materials have similar volumetric EC energy densities, and systems with both material types have been demonstrated. The highest temperature span of a polymer system reported in a peer-reviewed publication is 5 K (*34*). A temperature span up to 8.4°C of a polymer system was claimed in a government project report (*39*). However, the low thermal conductivity of polymers necessitates the use of thin films of not more than several tens of micrometers for adequate heat transfer for a practical heat pump [for example, (*31*)]. This makes scaling to high-capacity and high-performance systems very challenging. Ceramic materials have higher thermal conductivity, enabling thicker, higher–thermal mass devices.

We developed a scalable, high-performance system based on ceramic EC materials in a modular cascaded self-regenerating architecture with low thermal loss. We demonstrated the system in a device that uses partially ordered $PbSc_{0.5}Ta_{0.5}O_3$ (PST)–based EC multilayer ceramic capacitors (MLCCs) as the working elements. These MLCCs have an adiabatic temperature change of 2.5°C for a polarizing field of 10.8 MV/m at room temperature. The system achieves a temperature span of 5.2°C and a heat flux of 135 mW/cm^2 separately, using the same polarizing field. This temperature span is one of the highest reported in a ceramic EC system since the foundational work at the Moscow Power Engineering Institute in the early 1990s (*29*, *30*). The fluid-regenerated systems built there used 0.3-mm-thick PST plates as the working body and reported a temperature lift as high as 12.7°C (*30*). Unlike that early work, which was designed as a laboratory experiment, we designed our system to be scalable, using multilayer capacitors fabricated in an electronics manufacturing process, a modular configuration, and solid-state regeneration. With a few exceptions (*31*, *32*), most EC device reports only include measurements of maximum temperature span and do not measure cooling capacity. This is a critical quantity for technology comparison. For EC systems, heat flux gives insight into material utilization and overall system size. Our measured heat flux is ~4.5 times the 29.7 mW/cm^2 that was reported for a system that achieved a maximum temperature span of 2.8°C at zero heat flux in a separate experiment (*31*). Scalable materials, together with mechanical simplicity and modular design, can ultimately enable a system whose size, efficiency, and cost are competitive with those of vapor compression.

Our system comprises a pair of stacked modules, each containing EC MLCCs distanced with thermal insulating materials between them (Fig. 1, A and B). The inclusion of the insulation is a critical improvement of this design over continuous active regenerators because it interrupts the thermal shunting along the temperature gradient, a substantial source of loss. The modules are thermally coupled such that heat is easily transferred from one to the other. They are moved laterally relative to one another as the polarizing electric field is switched synchronously (Fig. 1C). In this way, we generated a temperature lift between the two ends of the device that is greater than the MLCC adiabatic temperature change. Another key design innovation is the use of anisotropically thermally conductive (ATC) plates to enhance heat exchange between layers while maintaining low lateral thermal leakage. The ATC plates are designed to have high through-plane and low in-plane thermal conductivity.

We prepared the EC MLCCs using a solid-state reaction and tape casting process (*40*, *41*), a large-volume capacitor manufacturing method commonly used in the electronics industry. We formed the dielectric layer from PST EC material with a B-site cation order of ~0.70 to 0.78. The Curie temperature of this material is 13°C and can be readily modified with chemical substitution and by controlling B-site cation ordering (*42*). The inner electrodes are Pt. The dielectric and electrode layers are ~37- and ~1.5-μm thick, respectively. We used Ag paste to form the external terminals. We photographed and obtained a scanning electron microscopy (SEM) image of an MLCC cross section (Fig. 2, A and B). We characterized the adiabatic temperature change of the MLCC by direct measurement. We measured the dynamic EC temperature as the polarizing fields are applied and removed (Fig. 2C). We also measured the heating and cooling adiabatic temperature changes of the MLCC for several magnitudes of the polarizing field (Fig. 2D) (*43*).

We optimized the system to maximize heat transfer between the MLCC modules while minimizing thermal leakage (Fig. 3, C and D). We fabricated each ATC plate from glass-reinforced epoxy (FR-4) laminates and copper in a standard commercial printed circuit board (PCB) process. In the regions of the plate where the MLCCs are attached, plated-copper through-vias serve as thermal shunts for high through-plane thermal conductivity. The diameter of the vias and the pitch of the

array are 150 and 300 μm, respectively. The effective through-plane thermal conductivity in the region of the vias is about 85.7 W/(m·K) based on a finite-element simulation in COMSOL, assuming 400 W/(m·K) as the thermal conductivity of electroplated copper and 0.29 and 0.81 W/(m·K), respectively, as the through-plane and in-plane thermal conductivities of FR-4 (*44*). The MLCCs are assembled onto the ATC plates to form the top and bottom EC modules (Fig. 3A), which contained five and four MLCCs, respectively. The bottom module has an aluminum plate-fin heat sink at each end to facilitate achieving stable temperatures when required. The modules are seated in a 3D-printed VeroClear housing filled with polyurethane foam insulation with a thermal conductivity of 0.04 W/(m·K). The housing (Fig. 3B) provides structural support and minimizes heat leakage. We attached a miniature fan to one end to flow air across the hot-end heat sink. We connected the top housing to a spring-returned single-acting linear solenoid actuator with a maximum stroke length of 12.7 mm and a contraction force of 7.2 N at 50% stroke. Application of 12-V dc to the actuator draws the top module toward the actuator. When the voltage is returned to zero, the solenoid relaxes and a spring returns it to its original position. The extents of actuation in each direction are set by adjusting the position of the actuator and an adjustable buffer. The EC module housing provides vertical pressure through a set of wheels to enhance the thermal contact between the two modules without hindering sliding. Control software synchronizes the actuation of the top module layer with the application of electric fields to the EC capacitors (*43*).

We operated the MLCCs under an EC Brayton cycle, with two isoelectric and two isentropic stages (Fig. 1D). To measure the maximum temperature span, we fully insulated both ends of the device. The heat sinks are filled with silicon paste to prevent air flow, the fan is removed, and the hot end is enclosed in foam insulation. Thermistors for measurement of the hot and cold end temperatures, T_h and T_c, respectively, were fitted into bore holes in the aluminum heat sinks. The variation in offset between the two thermistors was measured to be less than 0.1°C, allowing accurate measurement of small temperature lifts. We measured hot-end and cold-end temperatures with the cooler operated at a 400-V polarizing voltage (equivalent to ~10.5-MV/m electric field) and a 0.2-Hz switching frequency (5-s period) with a 50% duty cycle and negligible transition time (Fig. 4A). We also determined the maximum temperature lift $\Delta T = T_h - T_c$ for several different operating cycle periods, all with a polarizing voltage of 400 V (Fig. 4B). We achieved a maximum temperature span of 5.2°C at an operation frequency of 0.15 Hz (6.7-s period). This is equivalent to a ~0.09°C/mm temperature gradient along the heat pumping axis.

To measure the cooling power, we attached a 1-mm–by–2-mm 100-ohm chip resistor to the cold end and used it as an electric heater. The fan provides air flow across the uninsulated hot-end heat sink to maintain T_h near ambient temperature. We characterized the system by varying the current through the heater resistor and allowing the temperature lift ΔT to stabilize (Fig. 4, C and D). Because the coupling to ambient temperature is imperfect, over the course of the experiments, the heat sink temperature, T_h, increases slightly as heat is pumped to the hot end. The cold end temperature, T_c, changes in discrete steps as the heater power is varied. Because the temperature difference between the heater and the environment is small and the insulation is very good, heat leakage is negligible, and the heat pumped by the cooler is well approximated by the applied heater power. The measurement results show a linear relationship between the cooling power and temperature lift. We achieved a maximum heat power of ~85 mW with no temperature lift. This value is equivalent to a heat flux of ~135 mW/cm^2, given the 0.63-cm^2 active cooling area of the MLCC. The specific cooling power normalized to the volume of the active electrocaloric material is ~116 mW/cm^3. We also calculated the cooling power normalized to the total volume and to the total mass—each including the inactive material of the MLCC, the ATC plate, and the insulating material—to be ~29.2 mW/cm^3 and ~6.8 mW/g, respectively. The equivalent heat flux in the heat pumping direction is ~156 mW/cm^2, given the 0.54-cm^2 cross-sectional area normal to the heat flow direction.

A key determinant of the efficiency of a heat switch–based EC cooling system is the effective thermal contrast ratio K', defined as the ratio of "effective" on and off conductivity, k_h' and k_l' (*45*). We can model our system as a cascaded heat switch–based architecture (*43*). In an otherwise ideal heat switch–based system operating under a Carnot cycle, the *COP* is given by

$$COP_{K'} = \frac{Q_c}{W_{ECE}} = COP_{r,K'} \cdot COP_{Car} = \left(\frac{\sqrt{K'}-1}{\sqrt{K'}+1}\right)^2 COP_{Car} \quad (1)$$

where W_{ECE} is the work required to move heat through actuation of the ECE, K' is the effective heat-switch contrast ratio, Q_c is the heat collected from the cold side of the device, and $COP_{Car} = T_c/(T_h - T_c)$ is the maximum thermodynamic (Carnot) *COP*. $COP_{r,K'}$ is the maximum *COP* relative to COP_{Car} for a heat switch–based system with effective thermal contrast ratio K' (*33*). Assuming a necessarily imperfect charge recovery, system efficiency is given by

$$\begin{aligned} COP_C &= Q_c/W_{tot} \\ &= Q_c/[W_{ECE} + W_{elec}(1-\eta_{ECR})] \\ &= \eta_{elec} \cdot COP_{K'} \end{aligned} \quad (2)$$

where W_{elec} is the additional electrical work associated with charging and discharging the EC capacitor, and η_{ECR} is the electrical charge recovery (ECR) efficiency, or the portion of electrical energy recovered each cycle. This work is often disregarded in thermodynamic-cycle efficiency calculations that assume that input work and net cycle work are equivalent, leading to unrealistically high efficiency estimates. We did not observe measurable frictional or Joule heating in the system. Additionally, the mechanical work required to move the reciprocating system is orders of magnitude smaller than other work terms (*43*). Thus, we can neglect this term as well as the electrical work required to drive the actuator, assuming a reasonably efficient actuator. The electrical efficiency factor η_{elec} is given by

$$\eta_{elec} = \frac{W_{ECE}}{W_{ECE} + W_{elec}(1-\eta_{ECR})} \quad (3)$$

and can be approximated as

$$\eta_{elec} = \frac{1}{1 + \frac{T\varepsilon}{\rho c_E}\left(\frac{\Delta E}{\Delta T}\right)^2 (1-\eta_{ECR})} \quad (4)$$

for given electrical permittivity ε, material density ρ, and heat capacity at constant electric field c_E (*43*). As detailed in (*43*), this approximation is based on a number of simplifying assumptions to allow analytical insight. A more complete thermodynamic model is required for an accurate calculation of η_{elec} and system *COP* (*46*). Operating on a Brayton cycle instead of a Carnot cycle, the *COP* is reduced by the ratio of the temperature difference between the heat source and heat sink to the maximum temperature difference internal to the device, Δ/δ (*45*), giving the system relative *COP* (sometimes termed efficiency):

$$\begin{aligned} COP_r &= \frac{\Delta}{\delta} \cdot \frac{COP_C}{COP_{Car}} \\ &= \frac{\Delta}{\delta} \frac{1}{1 + \frac{T\varepsilon}{\rho c_E}\left(\frac{\Delta E}{\Delta T}\right)^2 (1-\eta_{ECR})} \left(\frac{\sqrt{K'}-1}{\sqrt{K'}+1}\right)^2 \end{aligned} \quad (5)$$

We determined η_{elec} as a function of η_{ECR} for the partially ordered PST material used in this work (fig. S5) and the material parameters (table S2). As the material ECE improves relative to the permittivity, the impact of ECR

decreases. A simple tank circuit can achieve η_{ECR} as high as 98% in the best case (*47*). This value would yield an electrical efficiency factor of 77.3%. With a more conservative η_{ECR} of 95%, the electrical efficiency factor is 57.7%.

In our system, the thermal contrast ratio K' was calculated to be 6.7 (*43*), corresponding to a $COP_{K'}$ of 20% and a potential COP_r of 11.5% in the limit of $\Delta/\delta = 1$, assuming 95% effective charge recovery, and a potential COP_r of 15.5% with 98% charge recovery efficiency, based on a simple, linearized efficiency model. This value is itself competitive with thermoelectric cooling devices, which typically are limited to $COP_r < 15\%$ (*48*). Although the thermal contrast ratio of 6.7 is notably smaller than the 27 previously reported by our team (*32*), the previous result was for a single-stage silicon heat switch–based system that is not readily scalable to high temperature lifts. The low thermal contrast ratio in the present system is largely attributable to low "off" conductance associated with the high thermal conductance of the thick copper traces used. We anticipate being able to achieve a contrast ratio of 23 (table S3) through improved metallization design. Much higher efficiencies are possible with increased K' through straightforward system improvements. These include the following: (i) Reducing the thickness of the copper traces, t_m, on the PCB from 36 to 1 μm. Sputtered ~1-μm-thick copper traces have been experimentally verified to carry the transient switching current. (ii) Replacing the PCB with a polymer material like acrylic. Polymers have much lower thermal conductivity [~0.2 W/(m·K)] than the in-plane thermal conductivity of FR-4 [0.81 W/(m·K)]. (iii) Improving the MLCC form factor. MLCCs of the type used in this device have higher thermal conductivity in the direction parallel to the metal inner electrodes (*49*), yet this value is constrained to be normal to the preferred heat transfer direction in the reported device because of the relatively few layers and large area of the MLCCs used. If the PST MLCCs are fabricated with the same form factor as a commercial MLCC (1210ZG226ZAT2A, AVX, Northern Ireland) and assembled into the system (see fig. S6), the performance can be substantially enhanced. Further improvements are also possible, for example, by using a higher-conductivity metal for electrodes. Our estimate of the COP_r of the current design, along with assumptions of these improvements for η_{ECR} of 95 and 98%, shows that up to 56.4% may be achievable with the existing PST material (table S3). These values could make our system competitive with vapor compression cooling.

REFERENCES AND NOTES

1. International Energy Agency (IEA), "The future of cooling: Opportunities for energy efficient air conditioning" (Report, IEA, 2018).
2. S. Sneha, I. Campbell, A Kalanki, "Solving the global cooling challenge: How to counter the climate threat from room air conditioners" (Report, Rocky Mountain Institute, 2018).
3. S. Trutassanawin, A. Eckhard, paper presented at the International Refrigeration and Air Conditioning Conference, paper R172, Purdue University, West Lafayette, IN, 2004.
4. J. S. Lewis, N. K. Dhar, L. A. Elizondo, R. Dat, *Proc. SPIE* **9609**, 960902 (2015).
5. H. J. Lee, J. S. Yoon, C.-J. Kim, *Heat Transf. Asian Res.* **30**, 357–370 (2001).
6. T. Moore, *An Essay on the Most Eligible Construction of Ice-houses: Also, a Description of the Newly Invented Machine Called the Refrigerator* (Bonsal & Niles, 1803).
7. W. Goetzler, R. Zogg, J. Young, C. Johnson, "Energy savings potential and RD&D opportunities for non-vapor-compression HVAC technologies" (Report, U.S. Department of Energy, 2014).
8. The United Nations Environment Programme (UNEP), "2010 Report of the Refrigeration, Air Conditioning and Heat Pumps Technical Options Committee (RTOC). Montreal protocol on substances that deplete the ozone layer" (UNEP, 2011).
9. W. Goetzler, R. Zogg, J. Young, C. Johnson, *ASHRAE J.* **56**, 12–23 (2014).
10. P. Nekså, H. T. Walnum, A. Hafner, paper presented at the Ninth IIR Gustav Lorentzen Conference on Natural Working Fluids (GL 2010), Sydney, Australia, 12 to 14 April 2010.
11. J. Shi *et al*., *Joule* **3**, 1200–1225 (2019).
12. B. Russ, A. Glaudell, J. J. Urban, M. L. Chabinyc, R. A. Segalman, *Nat. Rev. Mater.* **1**, 16050 (2016).
13. V. Franco *et al*., *Prog. Mater. Sci.* **93**, 112–232 (2018).
14. A. S. Mischenko, Q. Zhang, J. F. Scott, R. W. Whatmore, N. D. Mathur, *Science* **311**, 1270–1271 (2006).
15. B. Neese *et al*., *Science* **321**, 821–823 (2008).
16. J. Tušek *et al*., *Nat. Energy* **1**, 16134 (2016).
17. Y. Tu, R. Wang, Y. Zhang, J. Wang, *Joule* **2**, 1452–1475 (2018).
18. Y. Zhai *et al*., *Science* **355**, 1062–1066 (2017).
19. J. W. Peeples, *Electronics Cooling Magazine* **7**, 16–24 (2001).
20. I. Kriaa, N. Abdelmoula, A. Maalej, H. Khemakhem, *J. Electron. Mater.* **44**, 4852–4856 (2015).
21. S. Pandya *et al*., *Phys. Rev. Appl.* **7**, 034025 (2017).
22. M. Valant, *Prog. Mater. Sci.* **57**, 980–1009 (2012).
23. S. Lu, Q. Zhang, *Adv. Mater.* **21**, 1983–1987 (2009).
24. S. Patel, A. Chauhan, R. Vaish, *Energy Technol.* **4**, 1097–1105 (2016).
25. Y. Hou, L. Yang, X. Qian, T. Zhang, Q. M. Zhang, *Philos. Trans. R. Soc. London Ser. A* **374**, 20160055 (2016).
26. Y. Hou, L. Yang, X. Qian, T. Zhang, Q. M. Zhang, *Appl. Phys. Lett.* **108**, 133501 (2016).
27. S. G. Lu *et al*., *Appl. Phys. Lett.* **97**, 162904 (2010).
28. X. Li *et al*., *J. Mater. Chem. C* **1**, 23–37 (2013).
29. Y. V. Sinyavsky, V. Brodyansky, *Ferroelectrics* **131**, 321–325 (1992).
30. Y. V. Sinyavskii, *Chem. Petrol. Eng.* **31**, 295–306 (1995).
31. R. Ma *et al*., *Science* **357**, 1130–1134 (2017).
32. Y. D. Wang *et al*., *Appl. Phys. Lett.* **107**, 134103 (2015).
33. R. Epstein, K. Malloy, *J. Appl. Phys.* **106**, 064509 (2009).
34. H. Gu *et al*., *Appl. Phys. Lett.* **102**, 122904 (2013).
35. T. Zhang, X.-S. Qian, H. Gu, Y. Hou, Q. M. Zhang, *Appl. Phys. Lett.* **110**, 243503 (2017).
36. H. Gu, X. S. Qian, H. J. Ye, Q. M. Zhang, *Appl. Phys. Lett.* **105**, 162905 (2014).
37. U. Plaznik *et al*., *Appl. Phys. Lett.* **106**, 043903 (2015).
38. Y. D. Wang *et al*., paper presented at the Eighth International Conference on Caloric Cooling (Thermag VIII), Darmstadt, Germany, 16 to 20 September 2018.
39. S. R. Annapragada, P. Verma, A. Sur, W. Xie, "High-efficiency solid-state heat pump module" (Project Final Report, U.S. Department of Energy, 2018).
40. S. Hirose *et al*., *APL Mater.* **4**, 064105 (2016).
41. B. Nair *et al*., *Nature* **575**, 468–472 (2019).
42. N. Setter, L. E. Cross, *J. Appl. Phys.* **51**, 4356–4360 (1980).
43. Materials and methods are available as supplementary materials.
44. K. Azar, J. E. Graebner, in *Proceedings of the 12th IEEE SEMI-THERM Symposium* (IEEE, 1996), pp. 169–182.
45. S. J. Smullin, Y. D. Wang, D. E. Schwartz, *Appl. Phys. Lett.* **107**, 093903 (2015).
46. J. Gong, A. J. H. McGaughey, *Int. J. Energy Res.* **44**, 5343–5359 (2020).
47. E. Defay *et al*., *Nat. Commun.* **9**, 1827 (2018).
48. M. K. Rawat, H. Chattopadhyay, S. Neogi, *Int. J. Emerg. Technol. Adv. Eng.* **3**, 362–367 (2013).
49. S. Kar-Narayan, N. D. Mathur, *Appl. Phys. Lett.* **95**, 242903 (2009).

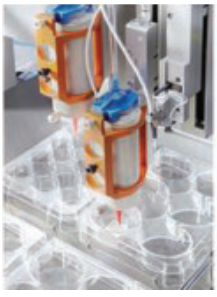
ACKNOWLEDGMENTS

We thank Q. Wang at PARC for his support on microfabrication and also C. Minami, N. Furusawa, Y. Inoue, and K. Honda for their kind assistance in fabricating MLCCs. **Funding:** PARC and Murata internal research funding. **Author contributions:** Y.W., D.S., and M.B. conceived and designed the system. Z.Z., M.B., J.K., J.L., and Y.W. fabricated the prototype. Z.Z. performed the experiments. T.U. and S.H. fabricated and tested the MLCCs. D.S. performed the thermodynamic analysis. Y.W., D.S., Z.Z., and M.B. organized the data and wrote the manuscript. All authors reviewed the manuscript. **Competing interests**: D.S. and Y.W. are inventors of U.S. patent application serial number 15/375,713, which claims the major features of the cascaded self-regenerating design. **Data and materials availability:** All data are available in the manuscript or the supplementary materials.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/370/6512/129/suppl/DC1
Materials and Methods
Figs. S1 to S6
Tables S1 to S3
References

8 December 2019; resubmitted 7 May 2020
Accepted 12 August 2020
10.1126/science.aba2648



3D Bioprinter
Analytik's BioScaffolder Prime is an affordable, high-performance 3D bioprinter that delivers precision engineering in an advanced, customizable platform. The rapidly expanding field of 3D bioprinting for tissue engineering and regenerative medicine combines biocompatible/biodegradable polymers with living cells. This attractively priced yet complete bioprinter package enables researchers to create bioscaffolds for cell growth and to deposit layers of bioinks on implants or microfluidic objects. This multiheaded 3D bioprinter can be equipped with multiple dispensing tools, including unique core/shell tools for simultaneous dispensing of different materials. Decentralized units for printing, media control, and computing save precious space in your biosafety cabinet and ensure superb heat dissipation. Silent but smart XYZ-drives deliver micrometer precision. In addition, the system comes with a Peltier heater/cooler cartridge for temperature-controlled bioprinting and a built-in UV source UV-LED pen. Designed to fit and operate in a standard biosafety cabinet, BioScaffolder Prime allows you to undertake your 3D printing applications quickly, safely, and in a sterile environment.
Analytik
For info: +44-(0)-1954-232776
analytik.co.uk/product/3d-bioprinter-bioscaffolder

SARS-CoV-2 Testing Workstations
With the ability to process up to 10,000 samples per day on a single system, the scalable and modular explorer automated workstations for SARS-CoV-2 testing improve the efficiency of laboratories 24/7 while eliminating errors and reducing labor requirements, by automating and standardizing disparate workflow steps. Different versions are configured for either reverse transcription PCR or ELISA testing, with varying degrees of automation based on workload and staffing resources. The scalable and modular nature of these workstations ensure that as your needs change, your automation solutions change with you. After the coronavirus pandemic subsides, your explorer workstations can be used to automate other laboratory processes.
PerkinElmer
For info: www.perkinelmer.com/contactus
www.perkinelmer.com

Protein Purification Workstation
Hamilton Company and Integrated Micro-Chromatography Systems (IMCS) introduce the affinityPure STAR assay ready workstation for the automation of high-throughput, small-scale affinity purification workflows, such as the Protein A purification of antibodies while screening biotherapeutics against SARS-CoV-2, the novel coronavirus that causes COVID-19. The affinityPure STAR is also ideal for other large-molecule research and manufacturing applications. Specialized affinity purification IMCStips used in the workflow contain resins such as Protein A, with custom formulations available upon request. The resins are loosely packaged in Hamilton's automation-ready Compressed O-Ring Expansion (CO-RE) pipette tips, which feature a tight seal to ensure precision and accuracy during liquid-handling steps. Patented dispersive mixing technology in the IMCStips facilitates maximum contact between each resin and protein of interest and efficient binding of the target analyte for high recovery during elution. A predefined hardware configuration processes up to 96 samples in 10 min–30 min compared to throughput-limiting and lengthy manual methods such as traditional spin columns.
Hamilton
For info: 800-648-5950
www.hamiltoncompany.com

Microplate Handler
BioTek announces that Agilent's BenchCel Microplate Handler is now added to its portfolio of walkaway automated solutions. The modular BenchCel Microplate Handler is a high-speed robot with an open, flexible, and scalable format to bring efficiency to diverse applications such as ELISA, endpoint add-and-read assays, and cell fix–stain–image processes. By integrating BenchCel between a BioTek liquid-handling device and detection or imaging system, sample throughput is increased while manual intervention is decreased. BenchCel is powered by VWorks software and incorporates an easy-to-use interface for dynamic scheduling of workflows. Additionally, user-friendly software interfaces may be created for a streamlined experience. Compatible BioTek devices include the EL406 Washer/Dispenser, MultiFlo FX Multi-Mode Dispenser, Synergy Neo2 Hybrid Multi-Mode Reader, Epoch 2 Microplate Spectrophotometer, and Cytation 5 Cell Imaging Multi-Mode Reader. With its small footprint, BenchCel may be used on the benchtop or in a biosafety cabinet, and for added flexibility it is compatible with a variety of microplates, including deepwell plates.
BioTek
For info: 888-451-5171
www.biotek.com/products/software-robotics-robotics/benchcel

Automated Analysis of Fluorescent Biomarkers
AMS Biotechnology (AMSBIO) announces a service designed to provide high-content cytometry data on cells and tissues, enabling the automated analysis of fluorescent biomarkers on living cells or long-term, stored biobanked cells and tissues. Leveraging the innovative ChipCytometry platform, AMSBIO can offer a complete workflow solution (products and services) for high-content cytometry on cells and tissues, from sample preparation and biobanking to biomarker analysis and data mining. The ChipCytometry platform operates by immobilizing cells in microfluidic chips, which are later stained with antibodies for multiplex biomarker detection. The platform allows quantitative analysis of a virtually unlimited number of protein biomarkers in a single sample. AMSBIO also provides custom assays facilitating analysis of almost any antibody/biomarker target. This technology not only delivers quantitative biomarker analysis but also facilitates both pre- and posttreatment immune-cell profiling—a key part of characterizing the complexities of cancer biology. AMSBIO offers stable sample storage and reinterrogation for up to 2 years. All customers using this service receive a detailed analysis report as well as the full .fcs raw data file.
AMS Biotechnology
For info: 617-945-5033
www.amsbio.com/cytometry-service

Electronically submit your new product description or product literature information! Go to www.sciencemag.org/about/new-products-section for more information.

Newly offered instrumentation, apparatus, and laboratory materials of interest to researchers in all disciplines in academic, industrial, and governmental organizations are featured in this space. Emphasis is given to purpose, chief characteristics, and availability of products and materials. Endorsement by *Science* or AAAS of any products or materials mentioned is not implied. Additional information may be obtained from the manufacturer or supplier.

CityU
香港城市大學
City University of Hong Kong

48th World University
4th World's top 50 Universities under age 50
1st World's Most International Universities
1st Engineering/Technology/Computer Sciences in Hong Kong
2nd Business School in Asia

Worldwide Search for Talent

City University of Hong Kong is a dynamic, fast-growing university that is pursuing excellence in research and professional education. As a publicly funded institution, the University is committed to nurturing and developing students' talents and creating applicable knowledge to support social and economic advancement. The University has nine Colleges/Schools. As part of its pursuit of excellence, the University aims to recruit **outstanding scholars** from all over the world in various disciplines, including **business, creative media, data science, energy and environment, engineering, humanities and social sciences, law, science, veterinary medicine and life sciences.**

The University welcomes applications and nominations for all faculty positions of Chair Professor, Professor, Associate Professor and Assistant Professor. The remuneration package will be highly competitive, commensurate with qualifications and experience. Interested parties are invited to submit an online application with current curriculum vitae to apply for current openings at http://go.cityu.hk/hrojobus or by email to "hrojob@cityu.edu.hk".

City University of Hong Kong is an equal opportunity employer and we are committed to the principle of diversity. Personal data provided by applicants will be used for recruitment and other employment-related purposes.

Worldwide recognition ranking 48th, and 4th among top 50 universities under age 50 (QS survey 2021); 1st in the World's Most International Universities (THE survey 2020); 1st in Engineering/Technology/Computer Sciences in Hong Kong (ARWU survey 2016); and 2nd Business School in Asia-Pacific region (UT Dallas survey 2017).

Perelman School of Medicine
University of Pennsylvania

Department of Physiology Faculty Positions

The Department of Physiology at the Perelman School of Medicine at the University of Pennsylvania seeks highly qualified candidates for faculty positions in the tenure track at the Assistant and Associate Professor ranks. Applicants must have a Ph.D. and/or M.D. or equivalent degree. Teaching responsibilities may include graduate and medical student instruction and training graduate students and postdoctoral fellows. Research responsibilities include developing a successful extramurally-supported research program. We seek investigators with demonstrated excellence in research grounded in fundamental mechanistic investigation that leverages molecular, biophysical, structural and cell-biological insights to inform physiological functions, particularly those emphasizing emerging areas, novel technologies, and interdisciplinary research.

Assistant Professor: Applicants in the early stage of career development are encouraged to apply. Apply online: **http://apply.interfolio.com/79065**

Associate Professor: Applicants are expected to have funding and an internationally recognized reputation of innovative research excellence and productivity, including high impact publications. Apply online: **http://apply.interfolio.com/79068**

The Perelman School of Medicine, one of the top ranked medical schools in the country for NIH funding, is a highly collaborative environment with state-of-the-art core facilities (https://www.med.upenn.edu/cores/), There are opportunities for appointments in numerous thematic institutes and centers (https://www.med.upenn.edu/psom/centers.html/), including the Cardiovascular Institute, Mahoney Institute for Neurosciences, Pennsylvania Muscle Institute, and Abramson Cancer Center. The School of Medicine is integrated within the campus of the University of Pennsylvania, a world-class institution with an easily walkable campus located near central Philadelphia. We seek candidates who embrace and reflect diversity in the broadest sense. The University of Pennsylvania is an EOE. Minorities/Women/Individuals with disabilities/Protected Veterans are encouraged to apply.

香港中文大學(深圳)
The Chinese University of Hong Kong, Shenzhen

Faculty Positions in Biomedical Engineering
The Chinese University of Hong Kong, Shenzhen

The Chinese University of Hong Kong, Shenzhen (CUHK-Shenzhen) invites applications for full-time faculty positions in Biomedical Engineering (BME). Interested candidates in biomaterials and regenerative medicine, biomechanics, biomedical imaging, biomolecular engineering and nanomedicine, biosensors and bioelectronics, immunoengineering, medical devices and robotics, neuroengineering, systems and synthetic biology, tissue engineering and regenerative medicine are welcome to apply at academic ranks of professors/associate professors/assistant professors/lecturers. Qualified candidates must have an earned doctoral degree in biomedical sciences, bioengineering and/or related disciplines, postdoctoral training and work experience. Applicants for Professors and Associate Professors should have an established research program with funding history, teaching experiences and an excellent publication record. Lecturer and Senior Lecturer positions are also available for candidates with teaching experience in biomedical engineering.

Established in 2014, CUHK-Shenzhen is a research-intensive university that inherits the fine academic traditions of The Chinese University of Hong Kong. The University adopts a tenure-track system for Assistant Professors and above. English is the main language for classroom teaching, and graduates receive degrees of The Chinese University of Hong Kong. The BME program was jointly established between the School of Life and Health Sciences (LHS) and the School of Science and Engineering (SSE), with LHS hosting the program and taking part in basic medical teaching/research in the School of Medicine. Qualified applicants may receive primary appointment at either LHS or SSE. Please visit **lhs.cuhk.edu.cn** for additional descriptions of the program. Interested applicants should submit their curriculum vitae, teaching statements and research descriptions to: http://academicrecruit.cuhk.edu.cn/lhs. Applications will be reviewed on a rolling basis until the positions are filled.

Sign up for Chinese Online Academic Forums | Overseas Scholars' Visit to Top Chinese Universities

Check the Details from www.edu.cn/zgx

- **10,000+ academic job vacancies in China**
- **Free one-to-one consultation service**

Send your CV to consultant@acabridge.edu.cn

- Taishan International Forum for Young Scholars
- Face to Face with NUIST Presidents
- CJLU & Yiwu City Government 2020 "B&R" Forum for Young Scholars
- Southeast University Youth Forum

HIGH-LEVEL GLOBAL TALENTS RECRUITMENT

2020 Global Online Job Fair (www.edu.cn/cv)

October 16, High-level Global Talents Recruitment
00:30-14:30 GMT (08:30-22:30 Beijing time)

November 06, China High-level Talents Recruitment
05:30-09:30 GMT (13:30-17:30 Beijing time)

November 27, High-level Global Talents Recruitment
00:30-14:30 GMT (08:30-22:30 Beijing time)

December 18, High-level Global Talents Recruitment
00:30-14:30 GMT (08:30-22:30 Beijing time)

Qualification for Applicants
Global scholars, Doctor and Post-doctor, Doctoral Candidates

Key Disciplines
Art & Science, Business, Economics, Computer Science and Information Technology, Chemistry, Agriculture, Fisheries and Food Science, Law, Life sciences, Mathematics, Medicine, Physics and Engineering, Psychology

Participating Universities
Beijing Jiaotong University, Harbin Engineering University, Hebei University of Technology, Zhengzhou University, National University of Defense Technology, Northeastern University(continuously updating)

Job Vacancies in China's Universities and Institutes
Please visit ***www.acabridge.edu.cn***
Contact ***consultant@acabridge.edu.cn***

Scan to check the sessions

EVEN SEPARATED BY MOUNTAINS AND RIVERS, WE HAVE OUR BLOOD TIGHTLY CONNECTED

The Coronavirus pandemic has forced many of us to consider different ways of working and communicating. Chinese universities and colleges are now holding online job fairs to help overseas scholars explore career options in China.

By **Rachel Mason**

A leap of faith

I spent a lot of *Free Solo* peeking through my fingers, waiting for disaster to strike. I was on a flight to visit family and had decided to watch the documentary about climber Alex Honnold's extraordinary, rope-free ascent of the 900-meter sheer cliff face of El Capitan. I have no head for heights, and maneuvering up a rock face is the last thing you'd find me doing. Yet one moment gave me an odd jolt of recognition. In it, Honnold has to let go of one set of handholds, step out over the void, and land on the next hold—just knowing that it will work out. It's a true leap of faith that made me think of the step into the unknown I had just made in my own career.

A few years earlier, I was an astronomer with a tenured position at a major international observatory, but something wasn't right. I just wasn't that into galaxies anymore. I had been living off the grid in rural Hawaii, raising animals, growing vegetables, and becoming increasingly curious about the science of sustainable agriculture—a topic that had been at the back of my mind since my childhood in the English countryside. I had never studied it formally, instead falling into chemistry for my bachelor's degree and then—after hearing a fascinating lecture about nuclear fusion in stars—being drawn into astronomy for my Ph.D. But now I just couldn't let go of this question: How can we have abundant food, happy people and animals, and a healthy environment?

"I was suspended in the void, not yet safely on the other side of my career transition."

After much thought—and some sessions with a counselor to make sure I wasn't about to make a huge mistake—I decided that returning to school was the extended "sabbatical" I needed to figure out what to do next. I was fortunate enough to be able to make that radical change. I don't have any dependents; my employer agreed that I could return to my old job if I had second thoughts; and my partner fully supported my decision, even though it would mean living apart. So, I left my secure career in Hawaii and moved to Vermont to begin a master's program in sustainable agriculture.

Strangely enough, this transition didn't seem so scary. It was a well-defined move: I would spend 2 years working toward a credential, and I knew I could thrive in an academic setting. I was looking forward to expanding my intellectual horizons, and the question of what I would do next—well, that could wait.

My El Capitan moment came as I was finishing my master's degree. I had no clue how I would put a new career together, which was more than a little unnerving. I had new skills and knowledge and I felt like a much better scientist than ever before, but I didn't know where I could put all this to use. The only things I was sure of were that I liked doing detective work, analyzing data, and figuring out how to handle complex projects, and I wanted to be in a position where I could make a difference.

I looked at all kinds of possibilities, including project management at a startup that uses blockchain technology to enforce food sustainability standards, a role at a company developing cultured meat, and a few postdoc positions. But I felt both over- and underqualified, not sure where my education and experience really fit. Judging by the number of rejection letters I received, others felt the same way. I was suspended in the void, not yet safely on the other side of my career transition but holding on to the belief that, somehow, I would get there.

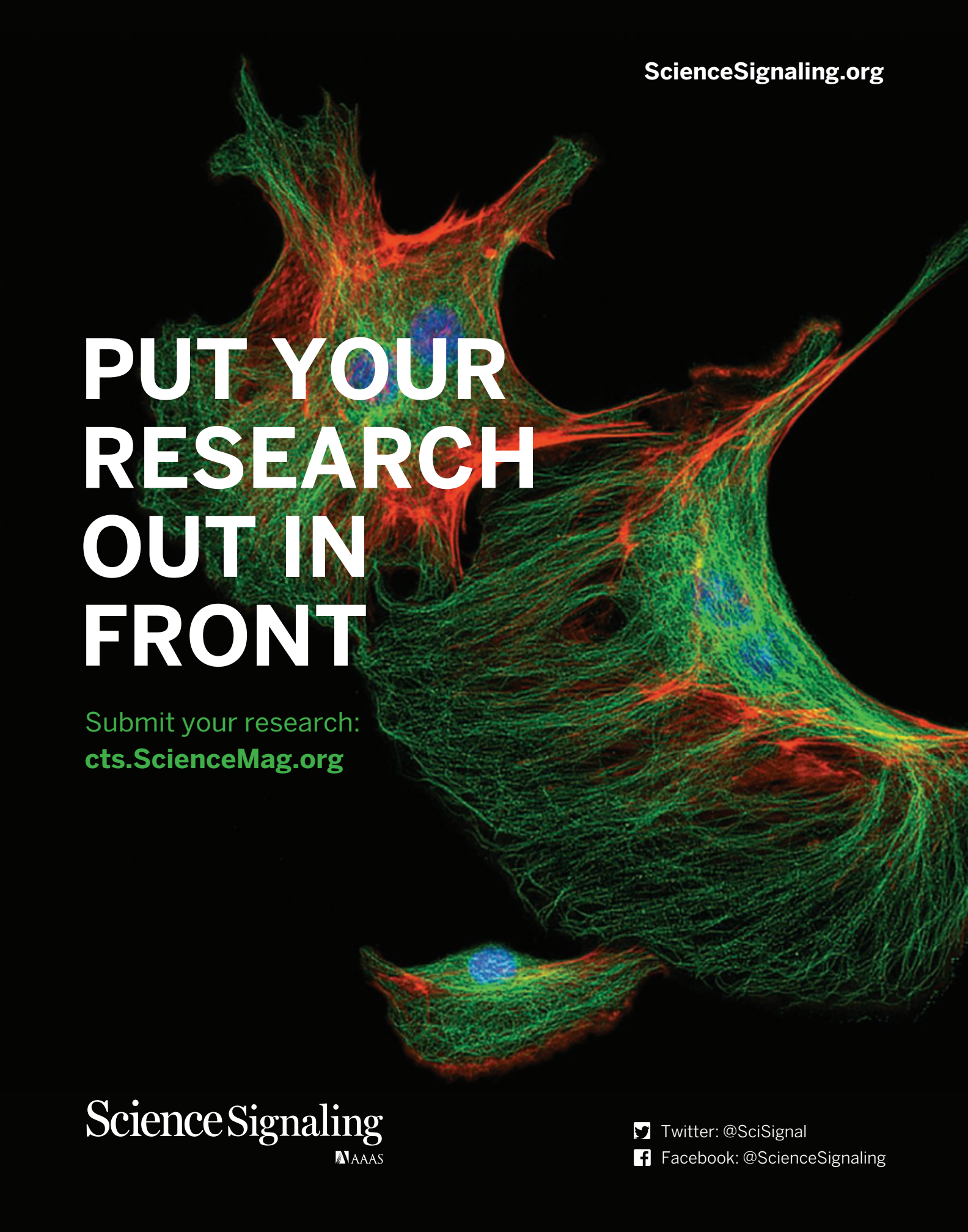
Fortunately, there are people who see potential in unconventional career paths, often because they have followed them. One of those people is the physicist-turned-ecologist who helped me land by offering me an academic research position. He is now my mentor as I study how global changes such as rising carbon dioxide levels affect the nitrogen content of plants and what that means for people and animals who depend on these plants for protein.

My leap of faith was a lot less elegant than Honnold's precise maneuver, and the route upward from here is less clear than his was. I'm 8 months into a 2-year appointment, and when that time is up, I may find myself facing another El Capitan moment. But I've made the leap before. It won't be as scary the next time around. ■

Rachel Mason is an assistant research scientist at the National Socio-Environmental Synthesis Center in Annapolis, Maryland.

ILLUSTRATION: ROBERT NEUBECKER

ScienceSignaling.org



PUT YOUR RESEARCH OUT IN FRONT

Submit your research:
cts.ScienceMag.org

Science Signaling
AAAS

Twitter: @SciSignal
Facebook: @ScienceSignaling

counter intelligence

The Scepter™ 3.0 Automated Handheld Cell Counter

Upgrade your tissue culture room by getting accurate cell counts quickly, and recording results digitally. Our Scepter™ 3.0 handheld cell counter combines smart capabilities with the Coulter impedance principal to hand you precise, consistent counts in seconds.

Easy-to-read cell count histogram display

Mounts anywhere for storage and charging

Ergonomic design for sustained, fatigue-free use

Wireless capability for instant printing or transfer of count data

Precise Coulter technology in a compact, microfluidic sensor

Accuracy you can count on, smarter than ever.

To learn more, visit
SigmaAldrich.com/Scepter

© 2019 Merck KGaA, Darmstadt, Germany and/or its affiliates. All Rights Reserved. MilliporeSigma and the vibrant M are trademarks of Merck KGaA, Darmstadt, Germany or its affiliates. All other trademarks are the property of their respective owners. Detailed information on trademarks is available via publicly accessible resources. 2019-21675

The life science business of Merck KGaA, Darmstadt, Germany operates as MilliporeSigma in the U.S. and Canada.

Millipore®

Preparation, Separation, Filtration & Monitoring Products