'Murder nurses,' or victims of shoddy statistics? p. 228

Light pollution is rapidly increasing pp. 234 & 265

Guessing errors and fitness in red squirrels p. 269

Science

$15
20 JANUARY 2023
science.org

AAAS



TIGHT SQUEEZE

Cramming an allene into a ring unlocks a chemical synthesis pp. 237 & 261





SCIENCE FOR HUMANITY

AAAS | ANNUAL MEETING

The 2023 AAAS Annual Meeting will convene in-person in Washington, D.C. for the first time in two years. Virtual access to the meeting will be available as well.

Register today and join us March 2-5, 2023!

aaas.org/meetings | #AAASmtg

BE A PART OF THE CONVERSATION

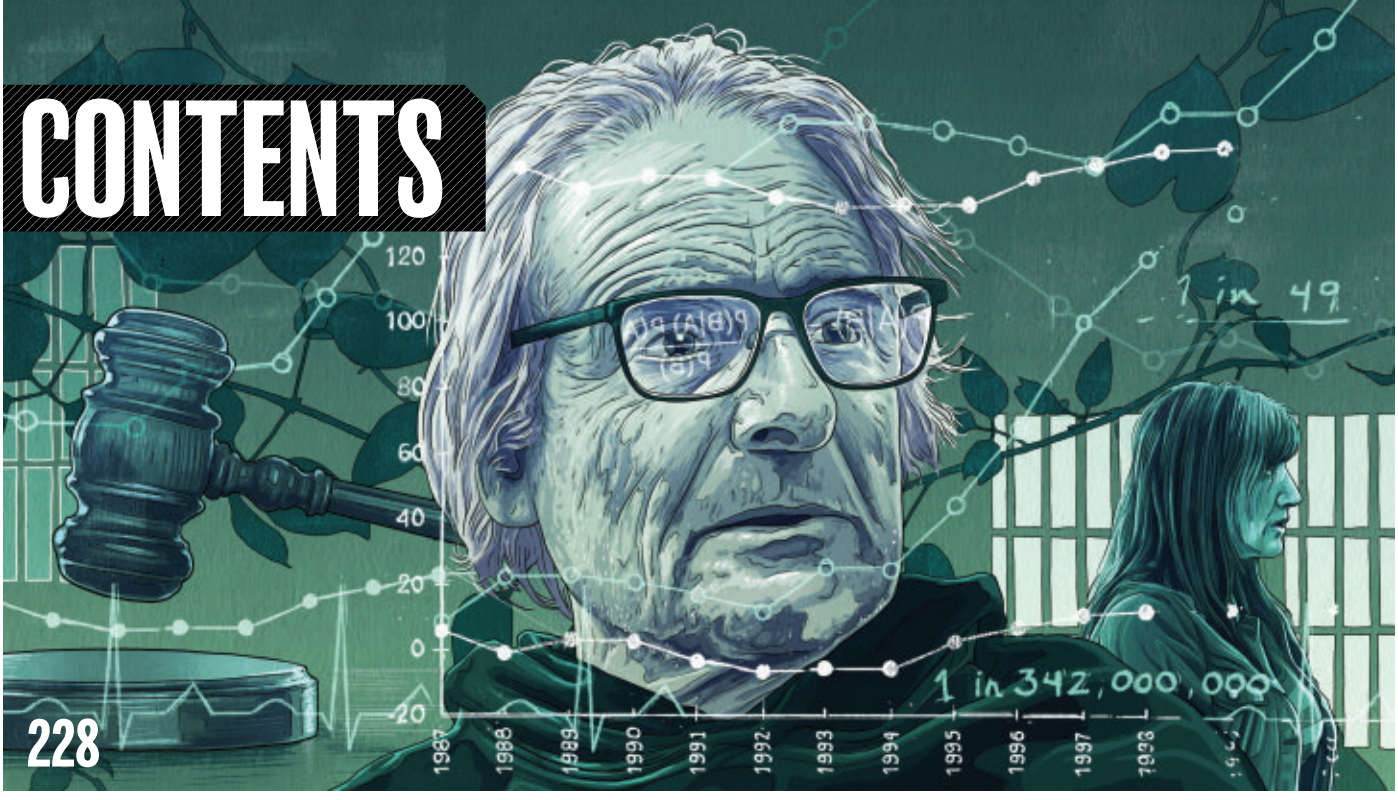
- Diversity, Equity and Inclusion
- Research Funding and Infrastructure
- Science Communication
- Artificial Intelligence
- STEM Education
- Information Technology and Security
- International S&T Collaboration
- Science and Values
- Social and Behavioral Sciences
- Global and Public Health
- Science Policy
- Climate Change

and more...





CONTENTS



228

20 JANUARY 2023 • VOLUME 379 • ISSUE 6629

NEWS

IN BRIEF

220 News at a glance

IN DEPTH

222 Taliban decree deepens Afghanistan's health crisis
Ban on female NGO staff hampers critical aid work and jeopardizes polio eradication campaign *By L. Roberts*

223 To reduce 'reputational bias,' NIH may revamp grant scoring
Proposed rating system seeks to minimize influence of applicant's resume and workplace *By J. Kaiser*

224 Stem cell factors reverse signs of aging in mice
Will reprogramming technique one day help people? *By C. Offord*

225 NSF under fire for not tracking sexual orientation
In major demographic survey of scientists, agency will only query gender identity
By K. Langin

226 A cell-by-cell view of a worm's metamorphosis
Single-cell analysis reveals shifts in identity as shapeless larva becomes wriggling adult
By E. Pennisi

227 Now in charge, House Republicans ramp up probes
New investigative panels will delve into pandemic origins and research ties with China *By J. Mervis and J. Cohen*

FEATURES

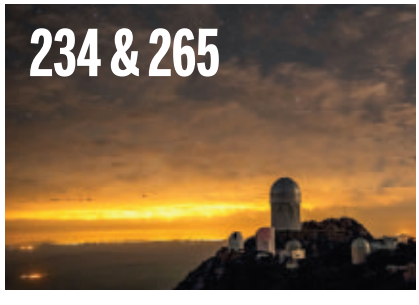
228 Unlucky numbers
Richard Gill is fighting the shoddy statistics that put nurses in prison for serial murder
By C. O'Grady

232 Mothers damned by statistics
By C. O'Grady
PODCAST

INSIGHTS

PERSPECTIVES

234 Light pollution is skyrocketing
Data from citizen scientists reveal a worrying growth in light pollution over the past decade
By F. Falchi and S. Bará
RESEARCH ARTICLE p. 265



236 Clicking zeolites together
A new mechanism to synthesize catalytic zeolites gives a previously unknown topology
By R. E. Morris
RESEARCH ARTICLE p. 283

237 A strained ring for stereoselective synthesis
The spring-loaded tension of a cyclic allene expedites synthesis of a natural product
By C. L. Jankovic and F. G. West
RESEARCH ARTICLE p. 261

238 Germline study points to sarcoma pathways
Pathogenic variants related to mitosis and telomere integrity are enriched in sarcoma patients *By D. Mandelker and M. Ladanyi*
RESEARCH ARTICLE p. 253

POLICY FORUM

240 Remaking the Chinese Academy of Sciences
Under pressure to reinvent itself, the CAS should concentrate on managing large-scale research infrastructures *By X. Yang* et al.

BOOKS *ET AL.*

244 Making modern science
Victorian values permeate contemporary scientific culture, maintains a historian
By B. Lightman

245 Pirates and politics
An anthropologist argues that experimental communities in Madagascar influenced the European Enlightenment
By J. Dua

CREDITS: (ILLUSTRATION) TAVO MONTAÑEZ; (PHOTO) HARUN MEHMEDINOVIC/500PX/GETTY IMAGES

YOUR RESEARCH HAS ITS REWARDS.



Submit an essay by 3/15/23 to enter to win $25K.

The prestigious international *Science* & PINS Prize is awarded annually to outstanding early-career researchers in the emerging field of neuromodulation. Findings already show improved outcomes in several neurological disorders. To enter, just write a 1000-word essay about your research performed in the last three years. The Grand Prize winner will be presented a US$25,000 award, as well as have their essay published in *Science* magazine and on *Science* Online. A runner-up will have their essay published online. For additional inquiries, email SciencePINSPrize@aaas.org.

SCIENCE-PINS PRIZE FOR NEUROMODULATION

SCIENCE-PINS

SCIENCE.ORG/PINS

LETTERS

246 Avian flu threatens Neotropical birds
By V. Gamarra-Toledo et al.

246 Atmospheric goals for sustainable development
By M. Keywood et al.

247 Evidence required for ethical social science
By R. Littman et al.

RESEARCH

IN BRIEF

248 From *Science* and other journals

REVIEW

251 CRISPR
CRISPR technology: A decade of genome editing is only the beginning
J. Y. Wang and J. A. Doudna
REVIEW SUMMARY; FOR FULL TEXT: DOI.ORG/10.1126/SCIENCE.ADD8643

RESEARCH ARTICLES

252 Lymphoma
BTG1 mutation yields supercompetitive B cells primed for malignant transformation
C. Mlynarczyk et al.
RESEARCH ARTICLE SUMMARY; FOR FULL TEXT: DOI.ORG/10.1126/SCIENCE.ABJ7412

253 Cancer genomics
Heritable defects in telomere and mitotic function selectively predispose to sarcomas
M. L. Ballinger et al.
PERSPECTIVE p. 238

261 Organic chemistry
Total synthesis of lissodendoric acid A via stereospecific trapping of a strained cyclic allene *F. M. Ippoliti* et al.
PERSPECTIVE p. 237

265 Light pollution
Citizen scientists report global rapid reductions in the visibility of stars from 2011 to 2022 *C. C. M. Kyba* et al.
PERSPECTIVE p. 234

269 Behavioral adaptation
Phenotype–environment mismatch errors enhance lifetime fitness in wild red squirrels
L. Petrullo et al.

272 Plant science
Nanobody-driven signaling reveals the core receptor complex in root nodule symbiosis
H. Rübsam et al.

278 Quantum simulation
A superconducting quantum simulator based on a photonic-bandgap metamaterial
X. Zhang et al.

283 Zeolites
A 3D extra-large-pore zeolite enabled by 1D-to-3D topotactic condensation of a chain silicate *J. Li* et al.
PERSPECTIVE p. 236



269
A juvenile North American red squirrel in the Yukon, Canada

CREDITS: (PHOTO) RYAN TAYLOR; (ILLUSTRATION) ROBERT NEUBECKER



288 Solar cells
Radical polymeric p-doping and grain modulation for stable, efficient perovskite solar modules *S. You* et al.

294 Optics
Breaking the limitation of polarization multiplexing in optical metasurfaces with engineered noise *B. Xiong* et al.

300 Stellar astrophysics
Spin-down by dynamo action in simulated radiative stellar layers *L. Petitdemange* et al.

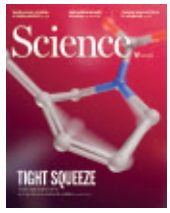
DEPARTMENTS

219 Editorial
A funding mosaic for loss and damage
By K. Warner and M. Weisberg

306 Working Life
Changing the climate
By S. Dee

ON THE COVER

This image illustrates the unusual chemical structure of a cyclic allene. Allenes typically contain two consecutive double bonds in a linear arrangement of three carbon atoms. Confinement in a ring instead enforces a bent geometry that produces strain and thus high reactivity. This enhanced reactivity enabled an efficient total synthesis of lissodendoric acid A, a complex natural product of pharmaceutical interest for its antioxidant properties. See pages 237 and 261. *Illustration: V. Altounian/*Science*; Data: Neil Garg/University of California, Los Angeles*



Science Staff 218
Science Careers 304

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 © 2023 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): $2411; Foreign postage extra: Air assist delivery: $107. 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.science.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 Valda Vinson

Editor, Research Jake S. Yeston **Editor, Insights** Lisa D. Chong **Managing Editor** Lauren Kmec

DEPUTY EDITORS Gemma Alderton (UK),Stella M. Hurtley (UK), Phillip D. Szuromi, Sacha Vignieri **SR. EDITORS** Caroline Ash (UK), Michael A. Funk, Brent Grocholski, Di Jiang, Priscilla N. Kelly, Marc S. Lavine (Canada), Mattia Maroso, Yevgeniya Nusinovich, Ian S. Osborne (UK), L. Bryan Ray, Seth Thomas Scanlon (UK), H. Jesse Smith, Keith T. Smith (UK), Jelena Stajic, Peter Stern (UK), Valerie B. Thompson, Brad Wible **ASSOCIATE EDITORS** Bianca Lopez, Madeleine Seale (UK), Corinne Simonti, Yury V. Suleymanov, Ekeoma Uzogara **LETTERS EDITOR** Jennifer Sills **LEAD CONTENT PRODUCTION EDITORS** Chris Filiatreau, Harry Jach **SR. CONTENT PRODUCTION EDITOR** Amelia Beyna **CONTENT PRODUCTION EDITORS** Robert French, Julia Haber-Katris, Nida Masiulis, Abigail Shashikanth, Suzanne M. White **SR. EDITORIAL MANAGERS** Carolyn Kyle, Beverly Shields **SR. PROGRAM ASSOCIATE** Maryrose Madrid **EDITORIAL ASSOCIATE** Joi S. Granger **SR. EDITORIAL COORDINATORS** Aneera Dobbins, Jeffrey Hearn, Lisa Johnson, Jerry Richardson, Alice Whaley (UK), Anita Wynn **EDITORIAL COORDINATORS** Maura Byrne, Clair Goodhead (UK), Alexander Kief, Ronmel Navas, Isabel Schnaidt, Qiyam Stewart, Brian White **RESEARCH & DATA ANALYST** Jessica L. Slater **ADMINISTRATIVE COORDINATOR** Karalee P. Rogers **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** Shraddha Chakradhar, 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, Michael Price, Kelly Servick **NEWS REPORTERS** Adrian Cho, Jennifer Couzin-Frankel, Jocelyn Kaiser, Rodrigo Pérez Ortega (Mexico City), Robert F. Service, Erik Stokstad, Paul Voosen, Meredith Wadman **INTERNS** Zack Savitsky, Viviana Flores, Katherine Irving **CONTRIBUTING CORRESPONDENTS** Warren Cornwall, Andrew Curry (Berlin), Ann Gibbons, Sam Kean, Eli Kintisch, Kai Kupferschmidt (Berlin), Andrew Lawler, Mitch Leslie, Eliot Marshall, Virginia Morell, Dennis Normile (Tokyo), Elisabeth Pain (Careers), Charles Piller, Gabriel Popkin, Michael Price, Joshua Sokol, Richard Stone, Emily Underwood, Gretchen Vogel (Berlin), Lizzie Wade (Mexico City) **CAREERS** Rachel Bernstein (Editor), Katie Langin (Associate Editor) **COPY EDITORS** Julia Cole (Senior Copy Editor), Morgan Everett, Cyra Master (Copy Chief) **ADMINISTRATIVE SUPPORT** Meagan Weiland

Creative Director Beth Rakouskas

DESIGN MANAGING EDITOR Chrystal Smith **GRAPHICS MANAGING EDITOR** Chris Bickel **PHOTOGRAPHY MANAGING EDITOR** Emily Petersen **MULTIMEDIA MANAGING PRODUCER** Kevin McLean **WEB CONTENT STRATEGY MANAGER** Kara Estelle-Powers **DESIGN EDITOR** Marcy Atarod **DESIGNER** Christina Aycock **SENIOR SCIENTIFIC ILLUSTRATOR** Valerie Altounian **SCIENTIFIC ILLUSTRATORS** Austin Fisher, Kellie Holoski, Ashley Mastin **INTERACTIVE GRAPHICS EDITOR** Kelly Franklin **SENIOR GRAPHICS SPECIALISTS** Holly Bishop, Nathalie Cary **SENIOR PHOTO EDITOR** Charles Borst **SENIOR PODCAST PRODUCER** Sarah Crespi **VIDEO PRODUCER** Meagan Cantwell **SOCIAL MEDIA STRATEGIST** Jessica Hubbard **SOCIAL MEDIA PRODUCER** Sabrina Jenkins **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 **MANAGER, BUSINESS ANALYSIS** Cory Lipman **BUSINESS ANALYSTS** Kurt Ennis, Maggie Clark **FINANCIAL ANALYST** Isacco Fusi **BUSINESS OPERATIONS ADMINISTRATOR** Taylor Fisher **SENIOR PRODUCTION MANAGER** Jason Hillman **SENIOR MANAGER, PUBLISHING AND CONTENT SYSTEMS** Marcus Spiegler **CONTENT OPERATIONS MANAGER** Rebecca Doshi **SENIOR CONTENT & PUBLISHING SYSTEMS SPECIALIST** Jacob Hedrick **SENIOR PRODUCTION SPECIALIST** Kristin Wowk **PRODUCTION SPECIALISTS** Kelsey Cartelli, Audrey Diggs **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** Aaron Helmbrecht, Ashley Hylton, Mike Romano, Lorena Chirinos Rodriguez, Jenna Voris **SENIOR DESIGNER** Kim Huynh

DIRECTOR AND SENIOR EDITOR, CUSTOM PUBLISHING Sean Sanders **ASSISTANT EDITOR, CUSTOM PUBLISHING** Jackie Oberst **PROJECT MANAGER** Melissa Collins

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 **PRODUCT ASSOCIATE** Caroline Breul, Anne Mason **SPJ ASSOCIATE MANAGER** Samantha Bruno Fuller **SPJ ASSOCIATE** Casey Buchta

MARKETING MANAGER Kess Knight **BUSINESS DEVELOPMENT MANAGER** Rasmus Andersen **SENIOR INSTITUTIONAL LICENSING MANAGER** Ryan Rexroth **INSTITUTIONAL LICENSING MANAGER** Marco Castellan, Claudia Paulsen-Young **SENIOR MANAGER, INSTITUTIONAL LICENSING OPERATIONS** Judy Lillibridge **SENIOR OPERATIONS ANALYST** Lana Guz **SYSTEMS & OPERATIONS ANALYST** Ben Teincuff **FULFILLMENT ANALYST** Aminta Reyes

DIRECTOR, GLOBAL SALES Tracy Holmes **US EAST COAST AND MID WEST SALES** Stephanie O'Connor **US MID WEST, MID ATLANTIC AND SOUTH EAST SALES** Chris Hoag **US WEST COAST SALES** Lynne Stickrod **ASSOCIATE DIRECTOR, ROW** Roger Goncalves **SALES REP, ROW** Sarah Lelarge **SALES ADMIN ASSISTANT, ROW** Victoria Glasbey **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** Rie Rambelli (Tokyo), Miyuki Tani (Osaka)

DIRECTOR, COPYRIGHT, LICENSING AND SPECIAL PROJECTS Emilie David **RIGHTS AND PERMISSIONS ASSOCIATE** Elizabeth Sandler **LICENSING ASSOCIATE** Virginia Warren **CONTRACT SUPPORT SPECIALIST** Michael Wheeler

---

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
science.org/authors/
science-information-authors

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

MEDIA CONTACTS
scipak@aaas.org

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

INSTITUTIONAL SALES AND SITE LICENSES
science.org/librarian

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

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

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

MEMBERSHIP AND INDIVIDUAL SUBSCRIPTIONS
science.org/subscriptions

MEMBER BENEFITS
aaas.org/membership/benefits

AAAS BOARD OF DIRECTORS
CHAIR Susan G. Amara
PRESIDENT Gilda A. Barabino
PRESIDENT-ELECT Keith Yamamoto
TREASURER Carolyn N. Ainslie
CHIEF EXECUTIVE OFFICER
Sudip Parikh
BOARD Cynthia M. Beall
Ann Bostrom
Janine Austin Clayton
Kaye Husbands Fealing
Maria M. Klawe
Jane Maienschein
Robert B. Millard
Babak Parviz
William D. Provine
Juan S. Ramírez Lugo
Susan M. Rosenberg

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.

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

Erin Adams, *U. of Chicago*
Takuzo Aida, *U. of Tokyo*
Leslie Aiello, *Wenner-Gren Fdn.*
Deji Akinwande, *UT Austin*
Judith Allen, *U. of Manchester*
Marcella Alsan, *Harvard U.*
James Analytis, *UC Berkeley*
Paola Arlotta, *Harvard U.*
Delia Baldassarri, *NYU*
Nenad Ban, *ETH Zürich*
Christopher Barratt, *U. of Dundee*
Nandita Basu, *U. of Waterloo*
Franz Bauer, *Pontificia U. Católica de Chile*
Ray H. Baughman, *UT Dallas*
Carlo Beenakker, *Leiden U.*
Yasmine Belkaid, *NIAID, NIH*
Philip Benfey, *Duke U.*
Kiros T. Berhane, *Columbia U.*
Joseph J. Berry, *NREL*
Alessandra Biffi, *Harvard Med.*
Chris Bowler, *École Normale Supérieure*
Ian Boyd, *U. of St. Andrews*
Malcolm Brenner, *Baylor Coll. of Med.*
Emily Brodsky, *UC Santa Cruz*
Ron Brookmeyer, *UCLA* **(S)**
Christian Büchel, *UKE Hamburg*
Dennis Burton, *Scripps Res.*
Carter Tribley Butts, *UC Irvine*
Johannes Buchner, *TUM*
György Buzsáki, *NYU School of Med.*
Mariana Byndloss, *Vanderbilt U. Med. Ctr.*
Annmarie Carlton, *UC Irvine*
Simon Cauchemez, *Inst. Pasteur*
Ling-Ling Chen, *SIBCB, CAS*
Wendy Cho, *UIUC*
Ib Chorkendorff, *Denmark TU*
Chunaram Choudhary, *Københavns U.*
Karlene Cimprich, *Stanford U.*
Laura Colgin, *UT Austin*
James J. Collins, *MIT*
Robert Cook-Deegan, *Arizona State U.*
Virginia Cornish, *Columbia U.*
Carolyn Coyne, *Duke U.*
Roberta Croce, *VU Amsterdam*
Christina Curtis, *Stanford U.*
Ismaila Dabo, *Penn State U.*
Jeff L. Dangl, *UNC*
Nicolas Dauphas, *U. of Chicago*
Frans de Waal, *Emory U.*
Claude Desplan, *NYU*
Sandra Díaz, *U. Nacional de Córdoba*
Samuel Díaz-Muñoz, *UC Davis*
Ulrike Diebold, *TU Wien*
Stefanie Dimmeler, *Goethe-U. Frankfurt*
Hong Ding, *Inst. of Physics, CAS*
Dennis Discher, *UPenn*
Jennifer A. Doudna, *UC Berkeley*
Ruth Drdla-Schutting, *Med. U. Vienna*
Raissa M. D'Souza, *UC Davis*
Bruce Dunn, *UCLA*
William Dunphy, *Caltech*
Scott Edwards, *Harvard U.*
Todd Ehlers, *U. of Tübingen*
Nader Engheta, *UPenn*
Karen Ersche, *U. of Cambridge*
Beate Escher, *UFZ & U. of Tübingen*
Barry Everitt, *U. of Cambridge*
Vanessa Ezenwa, *U. of Georgia*
Toren Finkel, *U. of Pitt. Med. Ctr.*
Gwenn Flowers, *Simon Fraser U.*
Natascha Förster Schreiber, *MPI Extraterrestrial Phys.*
Peter Fratzl, *MPI Potsdam*
Elaine Fuchs, *Rockefeller U.*
Daniel Geschwind, *UCLA*
Ramon Gonzalez, *U. of South Florida*
Gillian Griffiths, *U. of Cambridge*
Nicolas Gruber, *ETH Zürich*
Hua Guo, *U. of New Mexico*
Taekjip Ha, *Johns Hopkins U.*
Daniel Haber, *Mass. General Hos.*
Sharon Hammes-Schiffer, *Yale U.*
Wolf-Dietrich Hardt, *ETH Zürich*
Louise Harra, *UCL*
Carl-Philipp Heisenberg, *IST Austria*
Janet G. Hering, *Eawag*
Christoph Hess, *U. of Basel & U. of Cambridge*
Heather Hickman, *NIAID, NIH*
Hans Hilgenkamp, *U. of Twente*
Janneke Hille Ris Lambers, *ETH Zürich*
Kai-Uwe Hinrichs, *U. of Bremen*
Deirdre Hollingsworth, *U. of Oxford*
Randall Hulet, *Rice U.*
Auke Ijspeert, *EPFL*
Gwyneth Ingram, *ENS Lyon*
Darrell Irvine, *MIT*
Akiko Iwasaki, *Yale U.*
Stephen Jackson, *USGS & U. of Arizona*
Erich Jarvis, *Rockefeller U.*
Peter Jonas, *IST Austria*
Johanna Joyce, *U. de Lausanne*
Matt Kaeberlein, *U. of Wash.*
William Kaelin Jr., *Dana-Farber*
Daniel Kammen, *UC Berkeley*
Kisuk Kang, *Seoul Nat. U.*
V. Narry Kim, *Seoul Nat. U.*
Robert Kingston, *Harvard Med.*
Nancy Knowlton, *Smithsonian*
Etienne Koechlin, *École Normale Supérieure*
Alex L. Kolodkin, *Johns Hopkins U.*
LaShanda Korley, *U. of Delaware*
Paul Kubes, *U. of Calgary*
Chris Kuzawa, *Northwestern U.*
Laura Lackner, *Northwestern U.*
Gabriel Lander, *Scripps Res.* **(S)**
Mitchell A. Lazar, *UPenn*
Hedwig Lee, *Duke U.*
Ryan Lively, *Georgia Tech*
Luis Liz-Marzán, *CIC biomaGUNE*
Omar Lizardo, *UCLA*
Jonathan Losos, *WUSTL*
Ke Lu, *Inst. of Metal Res., CAS*
Christian Lüscher, *U. of Geneva*
Jean Lynch-Stieglitz, *Georgia Tech*
David Lyons, *U. of Edinburgh*
Fabienne Mackay, *QIMR Berghofer*
Zeynep Madak-Erdogan, *UIUC*
Anne Magurran, *U. of St. Andrews*
Ari Pekka Mähönen, *U. of Helsinki*
Asifa Majid, *U. of Oxford*
Oscar Marín, *King's Coll. London*
Charles Marshall, *UC Berkeley*
Christopher Marx, *U. of Idaho*
David Masopust, *U. of Minnesota*
Geraldine Masson, *CNRS*
C. Robertson McClung, *Dartmouth*
Rodrigo Medellín, *U. Nacional Autónoma de México*
C. Jessica Metcalf, *Princeton U.*
Tom Misteli, *NCI, NIH*
Alison Motsinger-Reif, *NIEHS, NIH* **(S)**
Danielle Navarro, *U. of New South Wales*
Daniel Nettle, *Newcastle U.*
Daniel Neumark, *UC Berkeley*
Thi Hoang Duong Nguyen, *MRC LMB*
Beatriz Noheda, *U. of Groningen*
Helga Nowotny, *Vienna Sci. & Tech. Fund*
Pilar Ossorio, *U. of Wisconsin*
Andrew Oswald, *U. of Warwick*
Isabella Pagano, *Istituto Nazionale di Astrofisica*
Elizabeth Levy Paluck, *Princeton U.*
Jane Parker, *MPI Cologne*
Giovanni Parmigiani, *Dana-Farber* **(S)**
Daniel Pauly, *U. of British Columbia*
Ana Pêgo, *U. do Porto*
Samuel Pfaff, *Salk Inst.*
Julie Pfeiffer, *UT Southwestern Med. Ctr.*
Philip Phillips, *UIUC*
Matthieu Piel, *Inst. Curie*
Kathrin Plath, *UCLA*
Martin Plenio, *Ulm U.*
Katherine Pollard, *UCSF*
Elvira Poloczanska, *Alfred-Wegener-Inst.*
Julia Pongratz, *Ludwig Maximilians U.*
Philippe Poulin, *CNRS*
Lei Stanley Qi, *Stanford U.*
Trevor Robbins, *U. of Cambridge*
Joeri Rogelj, *Imperial Coll. London*
Amy Rosenzweig, *Northwestern U.*
John Rubenstein, *SickKids*
Mike Ryan, *UT Austin*
Miquel Salmeron, *Lawrence Berkeley Nat. Lab*
Nitin Samarth, *Penn State U.*
Erica Ollmann Saphire, *La Jolla Inst.*
Joachim Saur, *U. zu Köln*
Alexander Schier, *Harvard U.*
Wolfram Schlenker, *Columbia U.*
Susannah Scott, *UC Santa Barbara*
Anuj Shah, *U. of Chicago*
Vladimir Shalaev, *Purdue U.*
Jie Shan, *Cornell U.*
Beth Shapiro, *UC Santa Cruz*
Jay Shendure, *U. of Wash.*
Steve Sherwood, *U. of New South Wales*
Brian Shoichet, *UCSF*
Robert Siliciano, *JHU School of Med.*
Lucia Silvotti, *UCL*
Emma Slack, *ETH Zürich & U. of Oxford*
Richard Smith, *UNC* **(S)**
John Speakman, *U. of Aberdeen*
Tara Spires-Jones, *U. of Edinburgh*
Allan C. Spradling, *Carnegie Institution for Sci.*
V. S. Subrahmanian, *Northwestern U.*
Ira Tabas, *Columbia U.*
Eriko Takano, *U. of Manchester*
A. Alec Talin, *Sandia Natl. Labs*
Patrick Tan, *Duke-NUS Med. School*
Sarah Teichmann, *Wellcome Sanger Inst.*
Rocio Titiunik, *Princeton U.*
Shubha Tole, *Tata Inst. of Fundamental Res.*
Maria-Elena Torres Padilla, *Helmholtz Zentrum München*
Kimani Toussaint, *Brown U.*
Barbara Treutlein, *ETH Zürich*
Jason Tylianakis, *U. of Canterbury*
Wim van der Putten, *Netherlands Inst. of Ecology*
Matthew Vander Heiden, *MIT*
Ivo Vankelecom, *KU Leuven*
Judith Varner, *UC San Diego*
Henrique Veiga-Fernandes, *Champalimaud Fdn.*
Reinhilde Veugelers, *KU Leuven*
Bert Vogelstein, *Johns Hopkins U.*
Julia Von Blume, *Yale School of Med.*
David Wallach, *Weizmann Inst.*
Jane-Ling Wang, *UC Davis* **(S)**
Jessica Ware, *Amer. Mus. of Natural Hist.*
David Waxman, *Fudan U.*
Chris Wikle, *U. of Missouri* **(S)**
Terrie Williams, *UC Santa Cruz*
Ian A. Wilson, *Scripps Res.* **(S)**
Hao Wu, *Harvard U.*
Li Wu, *Tsinghua U.*
Wei Xie, *Tsinghua U.*
Benjamin Youngblood, *St. Jude*
Yu Xie, *Princeton U.*
Jan Zaanen, *Leiden U.*
Kenneth Zaret, *UPenn School of Med.*
Lidong Zhao, *Beihang U.*
Bing Zhu, *Inst. of Biophysics, CAS*
Xiaowei Zhuang, *Harvard U.*
Maria Zuber, *MIT*

EDITORIAL

A funding mosaic for loss and damage

In the context of climate policy, "loss and damage" refers to unavoided climate change impacts, including those from extreme weather events and slow-onset events such as sea-level rise, increasing temperatures, and loss of biodiversity. Some argue that focusing on loss and damage distracts industrialized countries from attempts to avoid losses and damages in the first place through decarbonization and adaptation. Others argue that insufficient mitigation by industrialized countries has unleashed an unfair burden on developing countries, so developed countries alone must address loss and damage financially. The impasse was broken at the 27th Conference of the Parties (COP27)—the international climate negotiations—in November 2022, through a decision on new funding arrangements for loss and damage, including a new fund. Reflecting on the funding needs, Maldives' Environment Minister Aminath Shauna argued at COP27 that the world requires a "mosaic of solutions." As a transitional committee begins to assemble recommendations to operationalize this mosaic of funding arrangements, several critical questions require insights from the scientific community.

> "... improve peoples' situation, not keep them at the brink of survival."

To address major needs, what solutions can resolve gaps? Climate change is anticipated to have enduring effects on the real economy, manifesting as depressed production, decreased employment, diminishing social coherence, and cultural loss. All of these stress societies and public and private budgets. Climate-related emergencies could involve displacement or relocation (planned or unplanned) and may necessitate climate-resilient reconstruction and recovery. Slow-onset events may involve redistribution of populations through migration. This will require insights from social, economic, and climate data. Addressing these needs should improve peoples' situation, not keep them at the brink of survival.

What steps will improve coherence and coordination across the landscape of global, regional, and national institutions that fund activities to address loss and damage? Support in response to climate-related emergencies is currently channeled through humanitarian funding; emergency payouts inject cash and support for temporary shelter, food, and basic life support. Debt instruments such as loans, bonds, and budget reallocations currently finance reconstruction and recovery. But future climate impacts are needed that foster demand for public and private investments in ways that buoy people's ability to navigate disruptions across their life cycle, such as investments in health, vocational education, social safety nets, and restoration of ecosystems.

Regional and global monetary coordination will increasingly be invoked to provide fiscal space for governments when disruptions affect their finances. For example, regional financing arrangements allow groups of countries to pledge mutual financial support to members experiencing financial difficulties in their regions. Special drawing rights from the International Monetary Fund can help stabilize exchange rate volatility, which can be a measure of a country's economic health. Considerable research is needed on economic tools that provide stability in spite of disruptions.

How can current financial modalities better address needs and promote stability? New and additional capital from a new fund alone will not solve the loss and damage finance issue. The decision recognizes that a variety of mechanisms—including through international financial institutions, multilateral development banks, central banks, and social safety net programs—are required to shape and channel support to address needs. Research is necessary for the design of systems and structures to provide stability and enhance well-being in the face of progressive upheaval and uncertainty. Operationalization of funding arrangements raises questions about how to adjust timing, eligibility requirements, and access to promote stability. Information is needed about funding gaps, and creative and effective solutions for filling those gaps to help vulnerable populations and ecosystems on which they depend.

The scientific community can bolster operationalization of loss and damage funding arrangements through insights about these three questions. Analysis about the real economy will deepen understanding of the full dimensions of escalating, nonlinear impacts. Developing this mosaic of solutions means harnessing the strength of the global economy to deliver stability in robust, flexible ways that can anticipate and address societal needs in the face of advancing adverse climate impacts in every region of the world.

–Koko Warner and Michael Weisberg

Koko Warner is a visiting fellow at Perry World House, University of Pennsylvania, Philadelphia, PA, USA, and director of Global Data Institute at the International Organization for Migration, Berlin, Germany. kkwarner@iom.int

Michael Weisberg is Bess W. Heyman President's Distinguished Professor of Philosophy and director of Postgraduate Programs at Perry World House, University of Pennsylvania, Philadelphia, PA, USA. He is also nonresident senior adviser at the International Peace Institute, New York, NY, USA. weisberg@phil.upenn.edu

10.1126/science.adg5740



NEWS

IN BRIEF

Edited by **Martin Enserink** *and* **Jeffrey Brainard**

CLIMATE SCIENCE

In 2022, Earth set new records for warming

Temperatures continued to rise at an alarming pace in 2022, which became the fifth- or sixth-hottest year in modern history, U.S. and European science agencies reported last week. Earth's average recorded surface temperatures were some 1.2°C warmer than preindustrial times. Nearly 30 countries set individual all-time heat records, and some 850 million people experienced the warmest temperatures of their lives last year. As in 2021, the warming was suppressed by a persistent, multiyear La Niña cooling pattern in the tropical Pacific Ocean, the agencies said. But La Niña is expected to wane this year, setting the stage for even higher temperatures. Meanwhile, the world's oceans, which capture 90% of the excess heat from global warming and are less prone to short-term temperature fluctuations, again had their hottest year on record in 2022—as they have nearly every year since the 1990s.

China reports COVID-19 deaths

PUBLIC HEALTH | China's government said last week that nearly 60,000 people have died after contracting COVID-19 since it abandoned its zero-COVID policy on 7 December 2022—a major departure from previous assertions, deemed not credible by outsiders, that fewer than 10 people per day died since the policy ended. The new tally includes hospitalized patients for whom COVID-19 was either the direct cause of death or a contributing factor, a National Health Commission official said at a 14 January press briefing. The average age of the deceased was 80, and more than 90% had underlying conditions. The official added that 300,000 COVID-19 patients have been hospitalized and that infections are now tapering off.

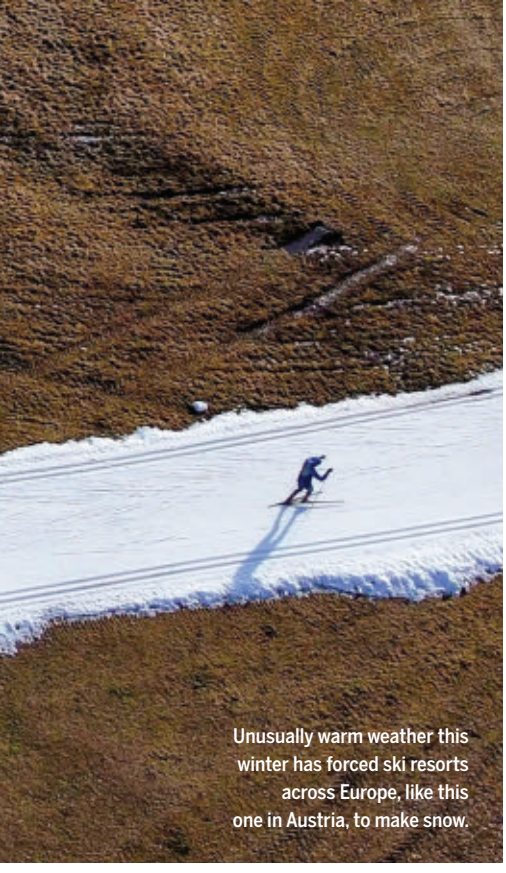
Hungary protests funding freeze

POLITICS | Hungary has vowed to fight an EU decision to suspend millions of euros for research and higher education as part of an ongoing dispute about control of the country's universities. In December 2022, the European Union notified the Hungarian government that it would exclude 21 of the country's three dozen universities from the Horizon Europe research funding system and from the Erasmus program, which funds international student exchanges. The European Union says the universities' oversight system, introduced by the current government, fails to meet EU standards because it allows politicians to sit on the schools' governing bodies. Since 2021, Hungary has received roughly €60 million from Horizon Europe to fund nearly 200 projects, and in 2020, more than 20,000 Hungarian nationals received roughly €40 million in Erasmus grants. A government spokesperson said last week that Hungary would sue to reinstate the funding if ongoing negotiations fail.

Japan aims to boost Ph.D. jobs

WORKFORCE | Alarmed by the country's sagging industrial prowess and a dearth of jobs for Ph.D. holders, the Japanese government will offer tax breaks for corporations

PHOTO: AP PHOTO/MATTHIAS SCHRADER



Unusually warm weather this winter has forced ski resorts across Europe, like this one in Austria, to make snow.

PHOTO: TRUMPF/MARTIN STOLLBERG

hiring recent doctoral graduates. Starting in April, companies will be able to claim a corporate tax credit worth 20% of salaries and other costs associated with hiring researchers who've earned a Ph.D. within the past 5 years. To qualify, companies also have to boost the share of R&D salaries going to staff with a doctoral degree by at least 3% annually. Ph.D. holders face bleak job prospects in Japan. Companies prefer hiring people with master's degrees and training them in-house, and the number of research positions at Japan's publicly supported universities is effectively capped.

Reef sharks face extinction

CONSERVATION | Sharks living near coral reefs are nearly twice as likely to be threatened with extinction as sharks in general, according to a new analysis. Many shark species are overfished for their fins and meat, which led the Convention on International Trade in Endangered Species of Wild Fauna and Flora in November 2022 to regulate the exports of about 10% of all 1199 species of sharks and rays. Fifty-nine percent of the 134 species of reef-dwelling sharks and rays are in danger of extinction, primarily from overfishing but also from climate change and habitat degradation, researchers report this week in *Nature Communications*. They call for better enforcement of fisheries regulations and an increase in marine protected areas.

Court nixes defamation verdict

#METOO | An appeals court in Peru has nullified a lower court's verdict against archaeologist Marcela Poirier, who was convicted of defamation after publicly accusing prominent archaeologist Luis Jaime Castillo Butters of sexual harassment. Castillo Butters had sued Poirier after she reported on Facebook that Castillo Butters's employer, the Pontifical Catholic University of Peru, had found evidence of harassment. In May 2021, Poirier was given a $48,400 fine and a 20-month suspended jail sentence, which she appealed. The appeals court ordered a retrial after finding that the proceedings were "plagued with irregularities," according to Poirier's lawyer. Castillo Butters did not respond to requests for comment.

Laser guides lightning

PHYSICS | Like the Greek god Zeus aiming thunderbolts, physicists working on a mountaintop in Switzerland have used a high-power laser to steer lightning toward a 124-meter-tall radio tower. The advance marks the culmination of decades of efforts to show that a laser can drill an ionized channel through the atmosphere that provides a path of least resistance for lightning to follow. The research team steered the bolt over the final 50 meters before impact. The finding, reported this week in *Nature Photonics*, could open the way to using lasers to create "virtual lightning rods" that protect rocket launchpads and other sensitive infrastructure by preemptively draining away the threat. The technique could shield an area larger than that covered by conventional lightning rods, the researchers say. For most applications, however, the multimillion-dollar laser system likely won't replace those humble and far less costly devices.



A mountaintop laser was fired 1000 times per second to create a channel that guided a lightning bolt.

THEY SAID IT

> "We also need Congress to pass a bill that codifies these protections into law, so they don't fade away under a future presidency."

Jacob Carter of the Union of Concerned Scientists, about a scientific integrity policy unveiled last week by the Biden administration to safeguard U.S. government scientists from political meddling.

JWST glimpses its first exoplanet

ASTRONOMY | NASA's JWST orbiting observatory has made its first discovery of a previously unknown planet around another star and is searching for its atmosphere. The newly discovered world, dubbed LHS 475 b, is almost exactly Earth-size. Other scopes can parse the atmospheres of gas giants from starlight passing through, but the atmospheres of small rocky exoplanets present a much harder challenge; researchers hope JWST will help crack it. The discoverers of LHS 475 b told the American Astronomical Society last week they will need more observations to pinpoint whether it has an atmosphere, and if so, its composition.

China's population starts to fall

DEMOGRAPHY | It's official: China's population has started to shrink, a turning point that could herald major economic challenges. In 2022, the number of Chinese residents—excluding those of Hong Kong, Macau, and Taiwan—fell by 850,000 people to 1.4 billion, China's National Bureau of Statistics announced this week. The drop, the country's first since the 1960s, was computed before the current wave of COVID-19 deaths began and may have been accelerated by couples deciding against having children during the pandemic. Demographers had long predicted the decline and expect it to be a long-term trend. China's birth rate has fallen for years, even after the government ended its one-child policy in 2016, and is now among the world's lowest. One reason is the large-scale migration to cities, where raising children is expensive, demographers say.



A hospital ward for severe malnutrition in Kabul. Six million people in Afghanistan are on the brink of famine.

PHOTO: OLIVER WEIKEN/PICTURE-ALLIANCE/DPA/AP IMAGES

GLOBAL HEALTH

Taliban decree deepens Afghanistan's health crisis

Ban on female NGO staff hampers critical aid work and jeopardizes polio eradication campaign

By **Leslie Roberts**

A decree issued by the Taliban last month that bars women from working in national and international nongovernmental organizations (NGOs) is another devastating blow to women's rights in Afghanistan. It also threatens to plunge a country beset by hunger and natural disasters even deeper into a public health crisis.

The United Nations and its humanitarian partners are engaged in intense negotiations to persuade the Taliban to reverse the edict. But for now, many NGOs, which depend heavily on female staff, have made the agonizing decision to suspend their operations, which provide vital food, hygiene, and medicines. The decree also jeopardizes the global campaign to eradicate polio, in which women play crucial roles. Afghanistan and its neighbor Pakistan are the last two countries where the wild poliovirus is endemic, and the campaign is going full bore to wipe it out by the end of this year.

Since August 2021, when the extremists took over, Afghanistan's economy has collapsed and the country has been hit with a punishing drought, an earthquake, floods, and a brutal winter. More than 28 million people are in need of humanitarian aid, such as food, shelter, clean water, hygiene, warm clothes, child protection, education, cash transfers, immunizations, and essential health services. Some 6 million are on the brink of famine.

Afghanistan's Ministry of Economy pronounced the edict in a 24 December letter to the Agency Coordinating Body for Afghan Relief and Development (ACBAR), which has 183 national and international NGOs as members. The letter claimed some female NGO staff were not wearing the hijab properly, says ACBAR Director Fiona Gall, and said it would revoke the operating licenses of organizations that did not comply with the ban on female staff. (The letter came after the Taliban barred women from universities and girls from secondary schools.) The Taliban now visits NGO offices regularly to see whether any women are at work; if so, they may get arrested. Women-led NGOs are under particular scrutiny.

For cultural reasons, Afghan women can't interact with male aid workers. By issuing the decree, "The Taliban has in effect suspended aid for half the population of Afghanistan," Sima Bahous of UN Women said in a 25 December statement, adding that "11.6 million women and girls are no longer receiving vital assistance." And many national and international NGOs say they had no choice but to temporarily suspend some or all operations. A quick survey of 87 NGOs conducted by the Humanitarian Access Working Group on 12 January found that 83% had done so.

"We are in a really difficult spot. We have no intention of abandoning the communities we work with," says Keyan Salarkia of Save the Children Afghanistan, which has more than 5000 staff, about half of them women. But, "We can't reach women or children and can't keep our staff safe. That is not a compromise we could make."

Kochay Hassan, executive director of the Afghan Women's Educational Center, says her organization continues to operate, but with most women working at home. The decree "will not stop us … if women can still work from home and engage with our male colleagues," Hassan says. The Vision Development Organization, which works on issues such as women's empowerment, girls' education, and health promotion, is also persevering, says founder and Executive Director Madina Mahboobi.

The United Nations's own agencies are technically exempt from the ban but could still see their work crippled by it. The UN World Food Programme (WFP), for example, which is aiming to feed 15 million people through the winter, works with about

100 local partners. "The whole humanitarian community is affected by the decision in one way or another," a WFP spokesperson says.

Afghanistan's Ministry of Public Health has said female NGO workers in the health sector are exempt from the decree as well, but exactly what that statement means is unclear, Gall says, and many organizations worry about the safety of their female staff. It appears that female health staff can work in hospitals and clinics, but the status of mobile teams is less clear, Gall says.

As a workaround, Hassan and others have adopted new terminology for their NGOs' work. In a number of provinces, "safe spaces" for women and children, which provide health interventions among other services, are now "health centers" that continue to operate with female staff, who carefully follow the Taliban's dictates on dress code and male chaperones.

Leaders of the Global Polio Eradication Initiative (GPEI) are trying to clarify what the edict means for their work. With just two cases caused by the wild poliovirus in Afghanistan in 2022, they say they have no intention of letting up. Two days after the edict came out, GPEI went ahead with a campaign in four eastern provinces, using female vaccinators. But UNICEF decided not to send its female staff, who inform women about upcoming campaigns and their benefits. "That is not sustainable going forward," says Hamid Jafari of the World Health Organization, who directs polio operations in the region.

GPEI is planning for a large campaign targeting 5.3 million Afghan children in late January, Jafari says. "We are working closely with the ministry of health to ensure that women can stay in campaigns." If not, campaigns will proceed with male vaccinators, as they already do in parts of the country. Fewer children will be vaccinated, especially babies who can't be taken out of the house, Jafari says, "But we will go ahead and try our best."

The U.N. and humanitarian organizations are still hoping the decree will be reversed. The U.N. Resident Coordinator in Afghanistan, Ramiz Alakbarov, met with Qari Din Mohammad Hanif, the Taliban minister of economy, on 26 December, and discussions have continued. Observers note dissent among the Taliban; some ministers have said they don't agree with the decree. And the Taliban seems to have left the door open, at least partially, to more discussions. "That means there is room for compromise," Gall says. "But it might not happen as soon as we would like." ■

Leslie Roberts is a journalist in Washington, D.C.

PEER REVIEW

To reduce 'reputational bias,' NIH may revamp grant scoring

Proposed rating system seeks to minimize influence of applicant's resume and workplace

By **Jocelyn Kaiser**

Researchers familiar with the National Institutes of Health's (NIH's) peer-review process can cite many cases where the reputation of a scientist or their workplace seemed to count for more than the strength of their ideas. There was the scientist who moved from the Ivy League to a large public university and saw scores on his grant applications drop. The investigators from historically Black institutions who always seemed to do worse than applicants from majority white schools. And the big names from research powerhouses whose proposals won stellar ratings without much scrutiny.

"Anyone who actually attends a study section sees [reputational bias] happening," says Noni Byrnes, director of NIH's Center for Scientific Review (CSR), referring to NIH peer-review panels. Now, in a bid to reduce that bias, her center is proposing the first overhaul of NIH's scoring system in 14 years. But the idea is getting a mixed reception.

> **"There is value in some objective score ... to help to assess the investigator."**
> **Shelley Berger**, University of Pennsylvania

The proposed reform was prompted, in part, by a striking pattern in NIH awards: About 70% of grants go to just 10% of all NIH-funded institutions. NIH says both anecdotes and data suggest reputational bias is partly responsible. So, in 2020, a working group of CSR's advisory council began to examine ways to reduce this bias, as well as ease the workload for reviewers.

Currently, reviewers rate proposals on significance, investigator, innovation, approach, and environment. The reform proposal, released last month, would reorganize those five criteria into three factors. Two would be scored: importance, and rigor and feasibility. Reviewers would also consider, but not score, a third factor: expertise and resources. If the reviewer sees problems in these areas—because the applicant lacks relevant experience, for example, or access to key equipment—the reviewer could adjust the overall score and write a brief explanation for NIH staff.

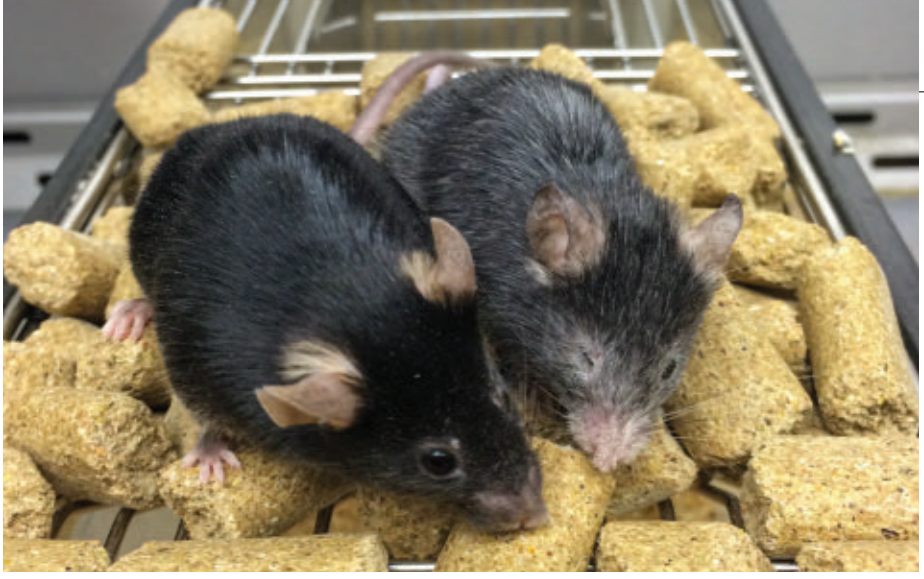
Some researchers worry these changes would go too far. To Shelley Berger, an epigenetics researcher at the University of Pennsylvania and a member of the NIH director's advisory panel, de-emphasizing an applicant's background is a mistake. "There is value in some objective score to try to help to assess the investigator," she said at the panel's December 2022 meeting. Rather than remake the scores, NIH could use "other methods to try to help underrepresented investigators and institutions to get into the NIH funding system," she wrote in an email.

In contrast, others think the changes don't go far enough. They say NIH should simply blind reviewers by preventing them from seeing an applicant's identifying information, at least initially. NIH is already experimenting with such blinded reviews for a limited number of grants. In the first year, the approach appeared to result in a more diverse pool of applicants for those grants and greater racial diversity among awardees. Some other U.S. research agencies, such as NASA, are also testing blinded reviews.

At NIH, Byrnes says implementing blinded reviews on a large scale would be impractical, because anonymizing proposals is "unwieldy" and reviews take longer. In addition, the agency has found in some tests of blinding that reviewers could still identify about 20% of the applicants.

NIH is now collecting comments on the proposal, with a 10 March deadline. CSR expects to release the final review criteria this summer.

If NIH adopts the new system, it would launch in 2025. The agency should monitor whether it is having any impact, says drug abuse researcher Michael Taffe of the University of California, San Diego, who has long criticized NIH for not doing enough to narrow a funding success gap between Black and white scientists. "NIH has a tendency to put these policy changes in place and not ... determine if they are producing the desired outcome," he says. ■



A 16-month-old mouse with induced DNA breaks (right) looks much older than a normal counterpart.

PHOTO: D. SINCLAIR/HMS

BIOLOGY

Stem cell factors reverse signs of aging in mice

Will reprogramming technique one day help people?

By **Catherine Offord**

A decade after Kyoto University biologist Shinya Yamanaka won a share of a Nobel Prize for discovering a cocktail of proteins that reprogram adult cells into versatile stem cells, two teams argue the proteins can turn back the clock for entire organisms—perhaps one day humans. One group at a biotech used gene therapy to deliver some of the so-called Yamanaka factors into old mice, and modestly extended their life span. A separate team followed a similar strategy to reverse aging-like changes in genetically engineered mice.

In both cases, the Yamanaka factors appear to have rejuvenated part of the animals' epigenome, chemical modifications on DNA and proteins that help regulate gene activity. But other scientists say suggestions of age reversal are premature. "These studies use reprogramming factors to reverse epigenetic changes that happen during aging," says Matt Kaeberlein, a geroscientist at the University of Washington, Seattle, but that's a far cry from making an old animal young again.

Several groups had already found that mice genetically engineered to begin expressing Yamanaka factors in adulthood show reversal of certain aging symptoms. To explore an approach that might lead to a more practical treatment for people, San Diego–based company Rejuvenate Bio injected elderly (124-week-old) mice with adeno-associated viruses (AAVs) carrying genes for three of the factors, collectively known as OSK.

These animals lived another 18 weeks on average, compared with 9 weeks for a control group, the company reported in a preprint on bioRxiv this month. They also partially regained patterns of DNA methylation—a type of epigenetic mark—typical of younger animals. Although some studies have suggested Yamanaka factors can promote cancer, Noah Davidsohn, Rejuvenate's chief scientific officer and co-founder, says the company has so far seen no obvious negative effects.

"I would say it is provocative—possibly a breakthrough," says Steven Austad of the University of Alabama, Birmingham, who studies the biology of aging. "But it will need to be replicated and the mechanism explored before we can say for sure."

The second study, published in *Cell* last week, is from a team led by Harvard Medical School geneticist David Sinclair, who has backed several controversial "anti-aging" interventions. (Rejuvenate's approach grew from an earlier collaboration between Sinclair and Davidsohn, but Sinclair isn't involved in the company's research, Davidsohn says.) Sinclair's team set out to test his "information theory of aging," which posits that our bodies get old because of the cumulative loss of epigenetic marks. Cells' DNA repair mechanisms, operating throughout a lifetime to fix DNA cuts and other damage, are what degrade these marks, he argues.

To test the theory in mammals, the team genetically engineered mice that, when given a drug, make an enzyme that cuts their DNA at 20 sites in the genome. The animals' cells faithfully repaired the damage. But widespread changes in cells' DNA methylation patterns and gene expression followed, consistent with Sinclair's theory. The mice ended up with an epigenetic signature more like that of older animals, and their health deteriorated. Within weeks, they lost hair and pigment; within months, they showed multiple signs of frailty and tissue aging.

To see whether the epigenetic degradation was reversible, the researchers injected some of these seemingly aged mice with AAVs carrying OSK genes, which Sinclair's group recently reported could reverse loss of vision in aging rodents. Analyses of the mice's muscles, kidneys, and retinas suggest the cocktail reversed some of the epigenetic changes induced by the DNA breaks. The findings point to a way to drive an animal's age "forwards and backwards at will," Sinclair says, and support the idea of epigenome-targeting treatments for aging in humans.

Molecular biologist Wolf Reik, director of the rejuvenation-focused Altos Cambridge Institute of Science, praised the sophistication of the Harvard team's study, but says the team's indirect way of inducing epigenetic changes with dramatic DNA breaks makes it hard to prove those changes are what's causing aging. The DNA breaks could be taking their toll through a different route. It's also unclear how well mice with induced DNA breaks mimic naturally aging animals, says Jan Vijg, a geneticist at the Albert Einstein College of Medicine.

He and others stress that aging is a complex process with multiple contributing factors, and that in both papers, the effects of OSK treatment were moderate: a small extension of life span in one, and a partial reversal of artificially induced symptoms in the other. "The jump that now aging is a program" that can be wound backward isn't justified by the research, Vijg says.

Still, both groups want to move their work toward the clinic. Rejuvenate is examining the mechanisms underlying the treatment's action and tweaking its delivery and composition, Davidsohn says. "OSK might not be the final set" of factors, he adds. Sinclair says his team is already testing AAV-delivered OSK in the eyes of monkeys. "If those studies in monkeys go well and everything looks safe enough for humans, the plan is to immediately apply to the FDA [Food and Drug Administration] to do a study in one or more [age-related] diseases of blindness." ■

Catherine Offord is a science journalist in Barcelona.

WORKFORCE

NSF under fire for not tracking sexual orientation

In major demographic survey of scientists, agency will only query gender identity

By **Katie Langin**

The U.S. National Science Foundation (NSF) says it does not plan to include a question about sexual orientation in a major national workforce survey, prompting hundreds of researchers to send a letter of protest.

Last month, the agency submitted its plans for the 2023 National Survey of College Graduates (NSCG), a biennial survey of more than 160,000 U.S. bachelor's degree holders with a focus on the science and engineering workforce, to the Office of Management and Budget (OMB) for approval. Many LGBTQ scientists were pleased that the survey will, for the first time, include a question about gender identity for all respondents. But the absence of a sexual orientation question is "incredibly disappointing," says Ramón Barthelemy, an assistant professor at the University of Utah who has studied the experiences of LGBTQ scientists in physics. Speaking as a gay man, he says, "We have fought so hard for so long to try to get representation in the scientific community, and what NSF is communicating to us is, they don't want us to have that representation."

The agency had raised hopes in 2021 when it pilot tested questions about gender identity and sexual orientation on the NSCG. Advocates took that as a sign that in coming years the agency would fully implement those questions in its suite of workforce surveys, including the Survey of Earned Doctorates and the Survey of Doctorate Recipients, which focus on Ph.D.-educated scientists and engineers. Those surveys provide crucial data about which demographic groups are underrepresented in the U.S. STEM community.

Up to now, respondents have had the option of identifying their sex only as "male" or "female." If NSF's plans are approved, this year's NSCG will split that question into two, first asking respondents to specify their sex assigned at birth and then their current gender identity. For the latter, respondents will have the option of selecting all that apply among male, female, and transgender; typing their own descriptor as free text; or both. This approach is consistent with OMB recommendations released this month detailing how to collect gender identity data, though some think the survey should offer more options, such as nonbinary and gender-fluid. Gender is "quite an expansive spectrum," says Abby Ray, a microbiology Ph.D. student at the University of California, Davis, and vice president of communications for oSTEM, an organization that represents LGBTQ+ people in STEM. For oSTEM's own data collection efforts, "We try to give as many options as possible so people have a chance to see themselves represented," they say.

The addition of a gender identity question is a "small step in the right direction," says Kaitlin Rasmussen, a postdoc in astrophysics at the University of Washington, Seattle, who has written on collecting data about marginalized genders. But it's frustrating that sexual minorities won't be counted as well, they add. "Gender and sexuality are really deeply intertwined and you might say you can't consider one without the other."



ILLUSTRATION: DAVIDE BONAZZI C/O THEISPOT

NSF told OMB it decided to omit the question about sexual orientation because it didn't perform well on the pilot test, taking longer to complete than the gender identity question and resulting in more changed answers and respondents exiting the survey. But the agency provided no data to back up those assertions or show how the results compare with those for questions about, for instance, race, disability, or income. "If they really want to stick to what they're claiming, then they should release the data," says Jon Freeman, an associate professor of psychology at Columbia University who has spent the past 5 years advocating that NSF collect data on LGBTQ scientists. An NSF spokesperson wrote in a statement to *Science* that the agency "will make those and other methodological findings available to the public later this year."

It's "perplexing," says Nancy Bates, a retired federal statistician. She points out that questions about sexual orientation have been asked in other federal surveys for years and have generally performed well. "There's less debate and sort of worry about measurement error for that particular question, as opposed to gender identity," says Bates, who co-chaired a National Academies of Sciences, Engineering, and Medicine working group that issued recommendations in 2022 detailing best practices for measuring sexual orientation and gender identity.

The NSF spokesperson said the agency will continue to explore whether to add a sexual orientation question to future workforce surveys. "The findings [from the 2021 pilot test] are already helping inform questions for future surveys. However, the results indicated that additional research and testing are needed to ensure it results in accurate data that protects privacy," the spokesperson added.

Freeman hopes OMB will force NSF's hand. In an effort to sway OMB's decision, due next month, he drafted a 16-page letter detailing his concerns about NSF's plans, which has been cosigned by 1700 other researchers. The letter, submitted to OMB last week, argues that NSF hasn't been transparent about its decision-making process and that the agency could still move forward with a sexual orientation question by borrowing methodological approaches that have already been extensively tested and vetted by other federal agencies, such as the Department of Education, the Census Bureau, and the Centers for Disease Control and Prevention.

"[OMB] could say, there's no reason that you're not also doing sexual orientation when that's what all these other agencies have done," Freeman says. ■

DEVELOPMENTAL BIOLOGY

A cell-by-cell view of a worm's metamorphosis

Single-cell analysis reveals shifts in identity as shapeless larva becomes wriggling adult

By **Elizabeth Pennisi**, *in Austin, Texas*

A newborn infant looks unmistakably human, with legs, mouth, ears, and bottom all in place. The same can't be said about the youngest sea stars, worms, or butterflies: Many invertebrates start out looking nothing like the adults they will become. Now, researchers have monitored one worm's larval cells during the transformation to adulthood, spying on their fates and how their identities changed. The work, reported earlier this month at the annual meeting of the Society for Integrative and Comparative Biology here, provides some of the first clear cell-by-cell clues about what happens during these radical changes in body form.

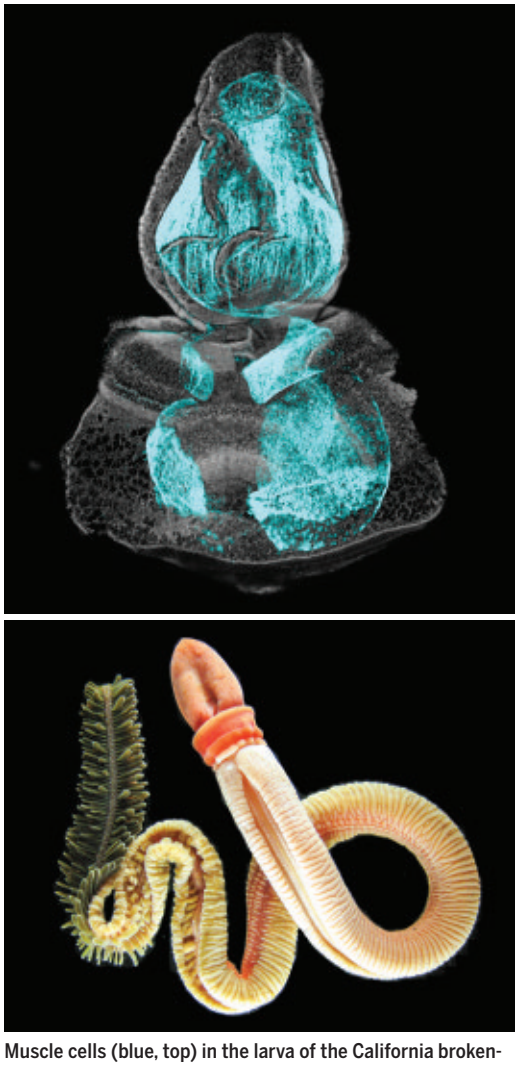
"We are beginning to get at how a larva produces an adult body plan," says Christina Zakas, an evolutionary biologist at North Carolina State University (NCSU) who was not involved with the work. The study's reliance on a sophisticated technique for tracking gene activity in single cells is also a milestone. To date, most single-cell analyses have focused on well-studied "model" animals such as the fruit fly, mouse, or zebrafish. "That [the author] got this new technology to work on something that was not a model organism was just stunning," says Greg Rouse, a marine biologist at the Scripps Institution of Oceanography who was not involved with the work. The work could inspire other biologists to study their own favorite metamorphosis at the level of single cells, he says.

Invertebrates undergo several kinds of larva-to-adult transformations. Caterpillars become butterflies through a well-understood process called catastrophic metamorphosis, in which a small group of cells that were dormant in the larva multiply and diversify into the adult, leaving the dead larval body behind. But Paul Bump, now a postdoctoral fellow at Harvard University, wanted to know what happens in species lacking such a group of cells.

Bump, who was then a Stanford University graduate student working with evolutionary biologist Christopher Lowe, focused on the California broken-hearted worm (*Schizocardium californicum*), a marine worm found in California mud flats. The 30-centimeter-long adult grows out of a larva that is little more than a sesame seed–size blob with an eyespot at one end and a band of cilia around its body. Several months after starting life, in just 48 hours or so, the larva transforms into a tiny juvenile worm with bulging mouthparts, a thickened midsection, and a long, sinuous body.

To track what happens to the larva's cells as it transitions to adulthood, Bump and his colleagues first sequenced the worm's genome and figured out how to tease out individual cells. He measured gene activity in cells isolated from larvae at different stages of metamorphosis and from the juvenile worm. Using those results, he grouped the larval and juvenile cells into different types—nerve cells, muscle cells, and the like. He also attached labels to DNA at different time points so he could see where it—and the cell containing it—wound up as metamorphosis proceeded.

The method could even identify cells that formed during metamorphosis. The protocol is "a new way of asking questions about cell identities," says Caroline Albertin, a developmental biologist at the Marine Biological Laboratory who was not involved with the work.



Muscle cells (blue, top) in the larva of the California broken-hearted worm are one of many cell types rearranged during the metamorphosis that produces an adult worm (bottom).

Depending on the cell type, one of three things happened. Muscle cells and some other cell types survived the transformation with little change. Other larval cells, including a few nerve cells, died off and disappeared. But to Bump's surprise, many cells—perhaps almost half, including other nerve cells and gut cells—seemed to hang around but switched their repertoire of active genes, taking on different roles in the adult. "What a cell is, and is capable of, is more flexible than we previously appreciated," he said at the meeting. "It seems that there are different cell types at different developmental stages," says Jose Aguilar, an evolutionary biologist at NCSU who hopes to use the single-cell technique in his own work on other worms.

Bump also noted clues to the evolutionary roots of metamorphosis. Around the time the worm's larva began its transformation, many cells started to produce an enzyme that, in insects, helps activate a molecule called juvenile hormone. The worm appears to make a related hormone, Bump reported, and the timing suggests the hormone plays a role in the worm's metamorphosis. Because juvenile hormone is known to regulate metamorphosis in insects, the finding could mean that some of the molecular mechanisms controlling the process evolved in a common ancestor of worms and insects.

Elaine Seaver, a developmental biologist at the University of Florida, suggests more insights may be forthcoming. "We're starting to apply single-cell sequencing to all kinds of animals. It's a powerful tool to learn more about evolution and about life histories." ■

PHOTOS: (TOP TO BOTTOM) PAUL BUMP; P. GONZALEZ *ET AL.*, *CURRENT BIOLOGY* (2016) 10.10.16/J.CUB.2016.10.047



Representative James Comer (R–KY) will play a key role in the COVID-19 pandemic probe.

PHOTO: GRAEME SLOAN/SIPA USA/AP IMAGES

U.S. SCIENCE POLICY

Now in charge, House Republicans ramp up probes

New investigative panels will delve into pandemic origins and research ties with China

By **Jeffrey Mervis** *and* **Jon Cohen**

The new Republican majority in the U.S. House of Representatives is preparing to shine a bright light on science—and scientists. Last week, it created two investigative panels that will scrutinize the country's relationship with China and its response to the COVID-19 pandemic. Both committees are expected to grill many prominent scientists and federal research officials on their actions over the past several years.

In approach and style, however, they are likely to be very different. The China panel, officially the Select Committee on Strategic Competition between the United States and the Chinese Communist Party, is fueled by bipartisan concerns that China has used economic espionage to help it become a global leader in key high-tech fields, including energy storage, artificial intelligence, and quantum computing. "It's time to understand the urgency of the threat [posed by China] … by working on a bipartisan basis," says its new chair, Representative Mike Gallagher (R–WI), a former intelligence officer.

In contrast, the Select Subcommittee on the Coronavirus Pandemic is widely regarded as an opportunity for House Republicans to criticize President Joe Biden's administration's response to the global crisis. Approved on a party line vote of 220 to 213 as part of a package of new House rules, the panel is also expected to focus on claims, unsupported by evidence, that a laboratory in Wuhan, China, either bioengineered or accidentally released SARS-CoV-2 obtained from bats.

The China committee will build on previous congressional reports that have questioned the value of continued scientific collaboration with China despite the many students and scientists born in China who conduct research on U.S. campuses and run high-tech companies. The 365-to-65 House vote to create the panel points to a bipartisan consensus that U.S. higher education can do more to protect federally funded research from foreign influences, without sacrificing the traditional free flow of talent and ideas across national borders.

Although a majority of House Democrats voted to create the panel, some worry it may be a way to revive the now-canceled China Initiative. Under former President Donald Trump's administration that law enforcement initiative led to the criminal prosecutions of some two dozen academic scientists of Chinese ancestry.

"We have always recognized that there are legitimate concerns with the actions of the [Chinese] government," said Representative Judy Chu (D–CA), chair of the Congressional Asian Pacific American Caucus, who opposed setting up the special panel. "But this committee should not be used as an open invitation to engage in blatantly xenophobic anti-China rhetoric … and to promote policies that result in the racial profiling of our communities."

Ties with China will be a focus of the pandemic panel as well. The Wuhan lab received some U.S. government funding, leading to charges that U.S. officials squelched concerns about the lab and attempted to mislead the public about the pandemic's origins. Representative James Comer (R–KY), chair of the House oversight panel that will administer the select committee, has said it will investigate "what the U.S. government knew regarding the origins of COVID-19 and when the government knew it."

The pandemic panel's chair has yet to be named, but it is expected to aggressively pursue requests, sent last month by Comer and Representative Jim Jordan (R–OH), for detailed information from more than 40 academic scientists and government officials. It also is likely to examine how the U.S. government regulates and funds "gain-of-function" research, which can make viruses more dangerous to humans in order to better understand the risks they pose and devise better vaccines and treatments.

One high-profile target will be immunologist Anthony Fauci, former head of the National Institute of Allergy and Infectious Diseases (NIAID). Other targets include conservation biologist Peter Daszak, president of the EcoHealth Alliance, a nonprofit that subcontracted a grant from NIAID to the Wuhan laboratory, and four scientists involved in producing a 2020 *Nature Medicine* paper that concluded SARS-CoV-2 was not engineered in a laboratory. They are Kristian Andersen of Scripps Research, Michael Farzan of UF Scripps Biomedical Research, Robert Garry of Tulane University, and Ian Lipkin of Columbia University.

Comer has alleged that "Fauci was warned early on that the virus appeared manmade and pointed to a lab leak and instead of blowing the whistle may have attempted to cover it up." Fauci has rejected allegations of any cover-up and says he is willing to answer questions from Congress. EcoHealth issued a statement saying it has "cooperated with every bipartisan initiative that has contacted us" and told Comer and Jordan it will "assist" lawmakers in their effort "to understand the scientific evidence behind COVID-19's origins."

Other researchers contacted by Comer and Jordan said they are reviewing the requests to provide a wide array of "documents and communications" related to the *Nature Medicine* paper and other issues.

Neither select panel has the power to write or approve legislation. And any policy recommendations would have to win enough votes in the Democrat-controlled Senate and avoid a presidential veto to become law. ■

FEATURES

UNLUCKY NUMBERS

Richard Gill is fighting the shoddy statistics that put nurses in prison for serial murder

By **Cathleen O'Grady**, *in Leiden, the Netherlands*



When a Dutch nurse named Lucia de Berk stood trial for serial murder in 2003, statistician Richard Gill was aware of the case. But he saw no reason to stick his nose into it.

De Berk was a pediatric nurse at Juliana Children's Hospital in The Hague. In 2001, after a baby died while she was on duty, a colleague told superiors that De Berk had been present at a suspiciously high number of deaths and resuscitations. Hospital staff immediately informed the police. When investigators reexamined records from De Berk's shifts, they found 10 suspicious incidents. Three other hospitals where De Berk had previously worked added another 10. The probability of such a pattern happening by chance was one in 7 billion, the police said. De Berk was arrested on 13 December 2001, suspected of murdering five children. Newspapers called her a "murder nurse" and an "angel of death."

Gill, then working as a statistics professor at Leiden University, remembers his wife telling him about a "witch trial" and saying, "They're using statistics; you should get involved, do something useful." But Gill knew the statistician working on the case and considered him a decent, careful person. "So I thought I didn't have to. And anyway, I was obsessed with quantum mechanics," he says. In 2003, De Berk was found guilty of four murders and three attempted murders and sentenced to life in prison. An appeals court convicted her again in 2004. The Dutch Supreme Court upheld the conviction 2 years later.

It wasn't until late 2006, when Gill read two whistleblowers' account of the trial, that he started to look into the case—and became incandescent. Tunnel vision, bad statistics, and poor human intuitions about coincidence had marred the investigation. When Gill ran the numbers himself, he found the string of deaths on De Berk's watch might well be entirely due to coincidence. Along with fellow statisticians, whistleblowers, and others, Gill campaigned for a retrial that eventually led to De Berk's exoneration in 2010. Her case is now considered one of the worst miscarriages of justice in the Netherlands.

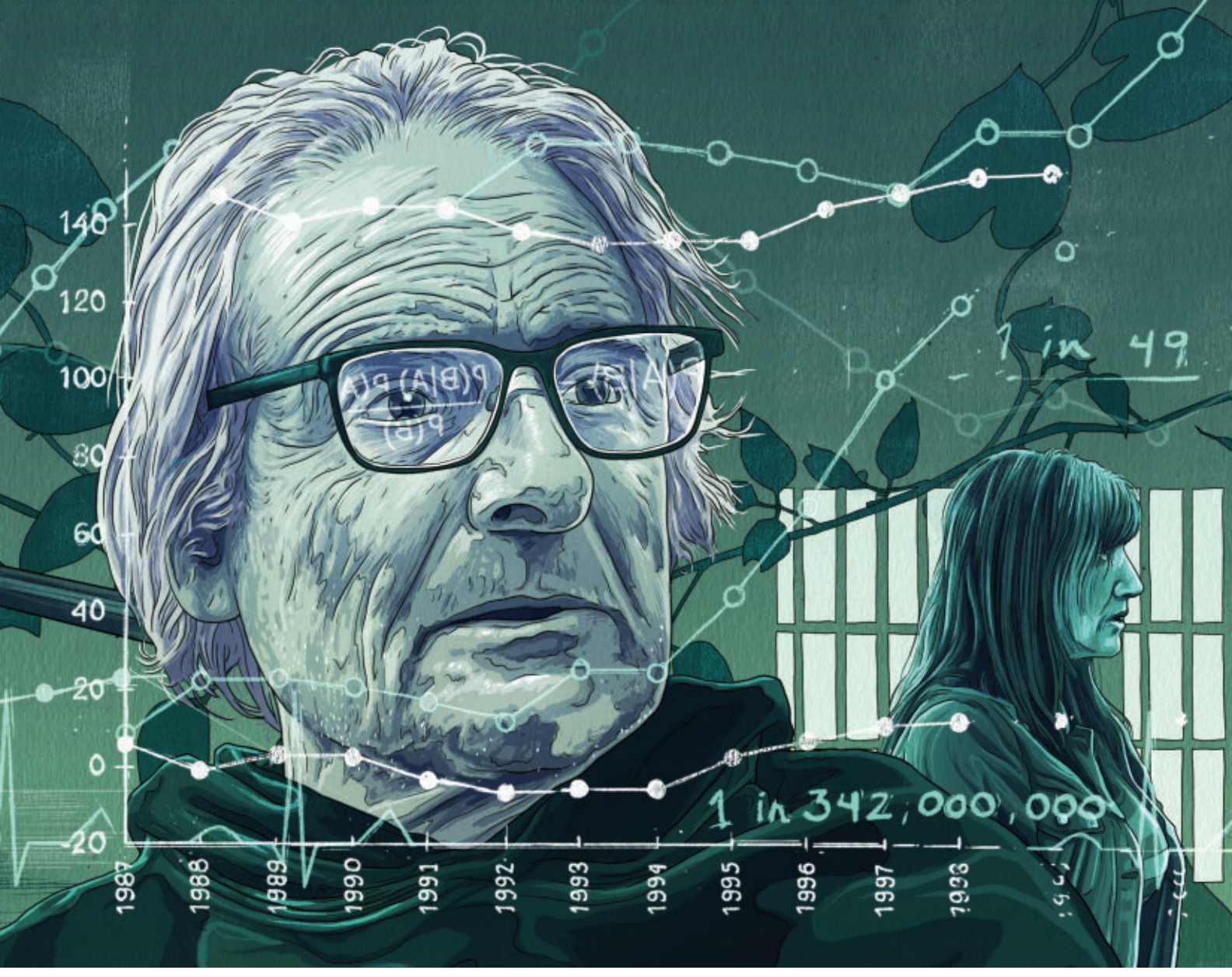
It also opened a new chapter in Gill's professional life: He became a leading expert on the statistics of medical murder cases similar to De Berk's—and a loud, persistent voice warning of the shoddy statistics that are sometimes central to prosecutors' arguments. "In a normal murder case, you actually have a body which has clearly been murdered," he says. When there's only a suspicious cluster of deaths, investigators may assume a murderer is at work and selectively focus on evidence that supports that assumption. People's intuition of an "impossible coincidence" joins the dots in the evidence.

Gill worked with defense lawyers and campaigned—in vain—to overturn the conviction of British nurse Ben Geen, found guilty in 2006 of two murders and 15 counts of grievous bodily harm. He also helped secure the October 2021 acquittal of nurse Daniela Poggiali, accused of two murders in a high-profile case in Italy. By now, the misuse of statistics has drawn enough attention that prosecutors sometimes insist their evidence is not statistical, Gill says, but often, "hidden statistics" seep into the cracks.

In a report peer reviewed and distributed by the Royal Statistical Society (RSS) in September 2022, Gill and colleagues detailed the statistical missteps in past medical murder trials and made recommendations for how legal systems can do better. Gill hopes the report will help with the case of another British nurse, Lucy Letby, who is now on trial for the alleged murder of seven babies and attempted murder of 10 more in a neonatal unit at the Countess of Chester Hospital.

"Similar issues have arisen across many, many different jurisdictions," says criminologist William Thompson, professor emeritus at the University of California, Irvine, and a co-author of the RSS report. "The same investigative dynamics play out … the same cognitive biases, and the tunnel vision." Gill likes to point out such errors with an outspokenness that frequently ruffles feathers, says statistician Peter Grünwald of the Center for Mathematics and Informatics (CWI) in

ILLUSTRATION: TAVO MONTAÑEZ



Amsterdam, a friend and colleague who also campaigned for De Berk's retrial. "He will give very radical opinions. … But somehow he's a very pleasant person to disagree with."

GILL HAS NOT ALWAYS been a troublemaker. His career has been defined by long detours into the bowels of arcane mathematical problems. Injustice bothers him—but so does error. He spends a lot of time debating quantum mechanics "crackpots" on the internet.

Gill had a serene childhood in the English countryside. His father, a physicist, spent his career in industry. His mother, Gill discovered after World War II intelligence was declassified in 1974, had been one of the human "computers" who helped crack Germany's Enigma code at an outstation of Bletchley Park. "I wasn't surprised," he says. "I always thought I got my brains as much from her as from my father."

At the University of Cambridge, where he studied math, it was statistics that most captured his attention. It had "weighty ethical and philosophical implications," he says. "It was a branch of mathematics that really means something."

As a student, he was not much of an activist. He says he feels guilty about not speaking up more about injustice when he was young. One incident in particular haunts him: his role as a statistician in a 1970s experiment that severed the front legs of rats to investigate whether bipedalism reshaped their skulls. "What upset me most is that I didn't have the strength of character to refuse to do that job."

The study was one of Gill's first assignments as a statistical consultant at a predecessor to CWI. Newly married, with three children born in quick succession, his most pressing concern was finding and keeping a good job, and the variety of consulting projects at the center fit neatly with his desire to do something practical. He obtained a Ph.D. on the mathematical underpinnings of "survival analysis," the study of the expected time until an event—such as a mechanical failure or a death in a clinical trial—occurs. Later, statistical problems in quantum mechanics were his main focus.

And then he started to look into the story of Lucia de Berk.

DE BERK'S CASE became famous for a number: one in 342 million. That was the probability that the many "incidents" on her shifts were due to random bad luck, according to Henk Elffers, a law psychologist then at the Netherlands Institute for Crime and Law Enforcement and an expert witness for the prosecution. His figure was less stark than the police's one in 7 billion, but still very damning.

Elffers's reasoning was controversial and came under fire from statistical experts



Lucia de Berk (facing camera) embraces Metta de Noo, who played a key role in her defense, after her exoneration in 2010. De Berk spent more than 6 years in prison.

during De Berk's appeals. He had multiplied the probability of De Berk's pattern of death across multiple wards, which would make any nurse look guiltier with each job change. For example, even a mundane one in 20 chance at one hospital, and the same chance at the next, would transform into a more suspicious one in 400 chance.

But prosecutors had additional evidence: Investigators had found traces of the heart medication digoxin in the body of one alleged victim and an overdose of the sedative chloral hydrate in another. At De Berk's first appeal—essentially a retrial—the Court of Appeal in The Hague ruled this evidence of foul play meant other deaths could be safely attributed to De Berk based on weaker evidence, including the overall "pattern" of incidents, and her diary, which spoke of her "very great secret" and "compulsion." The appeals court convicted De Berk again, adding three additional murders to her count.

De Berk, who suffered a stroke 5 days after the Supreme Court rejected her second appeal in 2006, maintained her innocence throughout.

That might have been the end of the case if it hadn't been for Metta de Noo, a geriatrician who had inside information. De Noo's sister-in-law was the head pediatrician at Juliana, where De Berk worked, and had aided the police investigation. But when De Noo examined documents from the case, she found what she believed were flaws in the medical evidence. The infant who had allegedly died of digoxin poisoning had been declining for days after heart surgery. And the hospital had prescribed the maximum dose of chloral hydrate for the other child, allowing additional doses if needed. De Berk had been agitating for doctors to pay attention to the child's deteriorating condition.

When De Noo asked specialists for support, she met with hostility and ridicule. Her doggedness destroyed her good relationship with her brother and his wife. She eventually turned to Ton Derksen, her other brother and a philosopher of science who had spent his career writing about flaws in reasoning of the type that permeated the De Berk investigation.

With De Noo's help, Derksen published a bombshell book in 2006: *Lucia de B.: Reconstruction of a Miscarriage of Justice*. (In the Netherlands, suspects' last names are commonly withheld to protect their privacy.) Derksen dismantled the figure of one in 342 million, giving a meticulous account of statistical errors, weak medical evidence, and bias in the investigation. For example, investigators examining the "incidents" connected with De Berk had classified deaths and resuscitations as suspicious when she was on duty, and not suspicious when she was off.

> **"We humans are terribly good at seeing patterns when they're not there."**
>
> **Peter Green**, University of Bristol

The prosecution had also argued that De Berk's ward had seen a total of five deaths between 1996 and 2001, and all had occurred after De Berk had started working in 1999. But the ward had a different name until 1999, and earlier deaths were excluded, Derksen found. In reality, there were seven deaths in the 3 years before De Berk joined and six in the 3 years after. (De Noo published her own account of the case—and the way it tore her family apart—in 2010.)

Grünwald, then a young assistant professor, brought Derksen's book to Gill's attention and asked whether he would join a campaign for De Berk's case to be reopened. Gill says reading the book made him "absolutely furious" with himself for trusting Elffers and not getting involved earlier. And he was angry that the appeals court had claimed its verdict did not rely on statistics: "Ton Derksen showed that it was soaked in statistics."

Gill quickly reanalyzed the data himself. In a write-up posted online in January 2007, he reported a much less outlandish probability of one in 100,000—even before removing biases in the data. Gill has refined his analysis over the years, building in complexities such as the fact that nurses could be expected to have different mortality rates based on their skill, choices, and work patterns. In a paper in *Chance* in 2018, he and colleagues calculated a probability of one in 49.

In 2007, convinced of De Berk's innocence, Gill organized a petition to reopen the case. His quantum mechanics work was "useful after all," he says, because he persuaded Nobel Prize–winning physicist Gerard 't Hooft to sign, which generated headlines. But in other ways Gill was less diplomatic. He called some doctors "criminals" and said "outrageous things" to journalists, Grünwald says: "Metta, Ton, and I basically had to hold him back." Haga Hospital even threatened to sue him after he posted previously unpublished details about the case on his website.

Yet Grünwald says Gill's cheerful fearlessness was crucial. Many Dutch statisticians knew and liked Elffers, he says. "People … were afraid to say out loud that he was doing something stupid and nonsensical. Richard had no problems with that at all." (Elffers declined to comment for this story.)

PHOTO: ROB VOSS/ANP/REDUX

The efforts paid off. In 2006, the Commission for the Evaluation of Closed Criminal Cases decided to reconsider the case and appointed a subcommittee to investigate. In a "drab government building" in The Hague, Gill helped explain how bungled statistics had put De Berk in prison. In 2007, the commission recommended reopening the case; in 2008, the Dutch Supreme Court agreed. That same year, the Dutch government suspended De Berk's sentence and she was released from prison, pending a retrial.

THE MISTAKES in De Berk's case were far from unique, Gill and others say. "We humans are terribly good at seeing patterns when they're not there," says statistician Peter Green, a professor emeritus at the University of Bristol and one of the RSS report's authors.

Investigators sometimes enhance those patterns by only tallying the evidence that confirms their theory, discarding or not even noticing data that don't. Even investigators who aim to be unbiased can make minor choices that add up to a skewed picture, Thompson says (see graphic, right). "You end up with a piece of evidence that looks extraordinarily unlikely to have occurred by chance. And of course, the problem is it didn't exactly occur by chance—you kind of helped it along."

Gill worries this is what led to the 2006 conviction of Geen, who was given 17 life sentences, with a minimum term of 30 years. Prosecutors argued there was a high rate of unexplained respiratory arrests—which are typically rarer than cardiorespiratory arrests—on Geen's shifts, although they did not try to quantify the probability that this "unusual pattern" occurred by chance. As in De Berk's case, there was other evidence, including the fact that Geen had a syringe containing muscle relaxant in his pocket when he was arrested. The prosecution argued that he had injected patients with the drug in order to cause respiratory arrest and then play the hero by resuscitating them.

Geen's defense lawyers challenged the "unusual pattern" in a 2009 appeal, submitting a report by University of Warwick medical statistician Jane Hutton. The appeals court upheld the conviction. "The judges seemed to be very overconfident that they could detect an unusual pattern without putting in some of the most basic information that you need as a comparison," Hutton says.

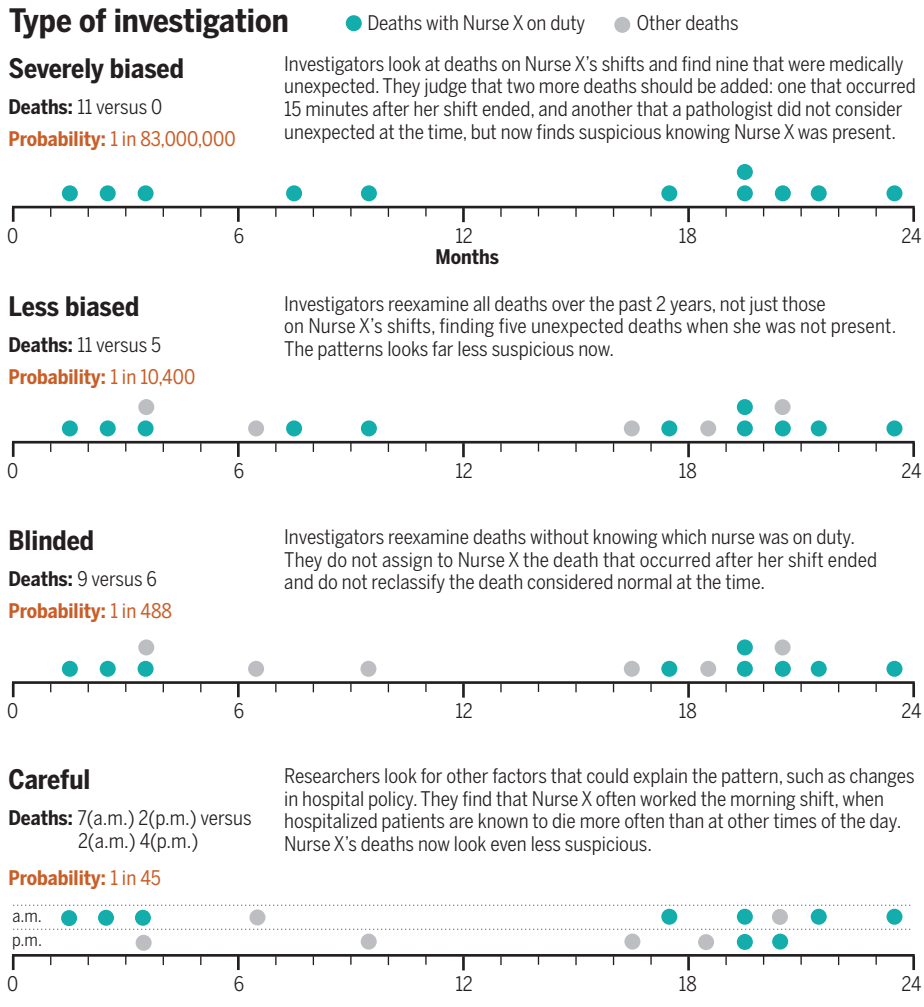
In a 2022 paper published in *Laws*, Gill and colleagues argued that blinded investigators might have reached different conclusions about Geen's case. The high rate of respiratory arrests on his shifts was accompanied by a drop in cardiorespiratory arrests, suggesting a bias in how these cases were classified. Compared with data from the same hospital over a wider time period, the deaths and resuscitations on Geen's shifts do not seem extraordinary, Gill and his co-authors said. He and other statisticians wrote letters of support in 2015 when Geen asked the Criminal Cases Review Commission to look into his case. The request was denied; Geen remains in prison.

EVEN WHEN statistical experts do get involved in a case, they may not be immune to errors of reasoning, as Elffers's work showed. In the case of Poggiali, the Italian nurse, statisticians wrote that a very high level of statistical significance is a "guarantee" that "there is a causal effect"—in this case between Poggiali being on duty and the deaths. But this is a well-known error of reasoning: "Correlation is not causation," Green says. Thompson says clusters may have surprising causes that are difficult or impossible to uncover. He points to cases where chemicals leached from equipment or changes in baby formula were at fault.

Gill and his colleagues found that Poggiali's death rate was higher than her colleagues', even after various controls, but argued this could be at least partly explained by Poggiali's long hours—she arrived very early and left late from her shifts—which meant she was present at more death certifications

A murderous nurse at work—or just coincidence?

Cognitive biases can easily lead an investigation astray and have drastic effects on how suspicious a cluster of deaths seems. In this imaginary example drawn from real-world errors, a doctor reports that many deaths seem to occur while Nurse X is on duty. The hospital launches an investigation, reexamining deaths at Nurse X's ward over the past 2 years. A simple statistical test* compares the rate of suspicious deaths when Nurse X is on duty with the rate when she is off. It then calculates the probability of seeing this pattern purely as a result of random chance.** The outcome depends greatly on the type of investigation.



* Log-linear model and chi-sq test. Method adapted from Appendix 5 in "Healthcare serial killer or coincidence? Statistical issues in investigation of suspected medical misconduct," published by the Royal Statistical Society in 2022. Script available at https://fal.cn/UnluckyNumbers.
** The p-value gives the probability of finding these results—or more extreme ones—if the hypothesis that the nurse killed the patients is not true.

GRAPHIC: C. BICKEL/*SCIENCE*

Mothers damned by statistics

Next month, a judge in Sydney will hear new expert testimony in a criminal case that has fascinated Australia for 2 decades: that of Kathleen Folbigg, who in 2003 was convicted of the murder of three of her infant children and manslaughter in the death of the fourth.

There is no medical evidence that Folbigg's children were murdered. Her case rests partly on the vanishingly small chance that unexplained medical tragedy would strike the same family four times. Like some other infanticide cases, it parallels the murder convictions of doctors and nurses based on suspicious clusters of patient deaths (see main story, p. 228). As those cases show, seemingly common-sense statistical assumptions can mislead—with horrifying consequences.

Folbigg's children all died between 1989 and 1999, at ages between 19 days and 19 months. Her husband reported Folbigg to the police after discovering her diary, in which she had described anger and frustration with her children, and a sense of responsibility for their deaths: "With Sarah all I wanted was her to shut up. And one day she did."

For each child, doctors found possible, but not definitive, evidence for natural causes of death. Yet taken together, expert witnesses said, the deaths were suspicious, because multiple cases of sudden infant death syndrome (SIDS) within a single family are extremely rare—let alone four of them. The New South Wales Supreme Court sentenced Folbigg to 40 years in prison, reduced to 30 years by a 2005 appeal. A 2019 inquiry upheld her conviction, and a 2021 appeal was dismissed.

Critics say the case rested heavily on the reasoning popularized in the 1990s by British pediatrician Roy Meadow, who asserted that with respect to child deaths, "one is a tragedy, two is suspicious, and three is murder unless there is proof to the contrary." Pediatrician Susan Beal cited a variation of "Meadow's law" during a 2003 hearing on what evidence could be admitted in Folbigg's trial.

Meadow testified in court cases himself as well. But his reputation fell apart after the case of British solicitor Sally Clark, who in 1999 was convicted of murdering her two infant sons. Meadow testified that the chance of two SIDS deaths in a low-risk family like Clark's was one in 73 million. That calculation assumed SIDS could not have inherited risk factors, statistician Phil Dawid of the University of Cambridge wrote in a report for Clark's first appeal in 2000. He put the chance of the two deaths at a less outlandish one in 1 million, "or even much higher," and pointed out that double infanticide is also vanishingly rare. The court should weigh both rare possibilities against each other, he says, along with all the other evidence.

Clark lost the appeal, but she was exonerated at a second appeal in 2003, partly because it came to light that pathologist Alan Williams had failed to disclose evidence that one of the babies had *Staphylococcus aureus* in his spinal fluid, a possible natural cause of death. The appeal judges said Meadow's statistical evidence—which could have had "a major effect" on the jury—should not have been admitted.

Williams was barred from working for the U.K. Home Office for 3 years and Meadow lost his medical license, a decision later overturned by the U.K. High Court. After the scandal, the attorney general ordered a review of 297 infanticide cases, and decided to drop charges in three cases and review the convictions in 28 others.

There may be exculpatory medical evidence in Folbigg's case as well. In 2020, a group of researchers led by Peter Schwartz at the Italian Auxological Institute published a paper showing Folbigg's two daughters both had a newly discovered genetic variant that impairs cells' ability to regulate calcium, leading to a greatly increased risk of cardiac arrhythmia and sudden death. The paper led the Australian Academy of Science and Folbigg's lawyers to launch a petition in March 2021, signed by 90 scientists, asking New South Wales Governor Margaret Beazley to pardon Folbigg. Beazley ordered a new inquiry; hearings are due to begin in February.

Clark, despite her vindication, never recovered and died of acute alcohol poisoning in 2007. Her family and the coroner's office attributed the death to severe distress from the "catastrophic experience." —*C.O.G.*

during shift handovers. They also pointed out a statistical flaw in the medical evidence: A toxicologist had said the potassium concentration found in one of the victim's eyes was unexpectedly high, suggesting potassium chloride poisoning. But this did not take into account any statistical uncertainty in the data on expected levels of potassium, Gill and colleagues wrote in a 2021 paper in *Law, Probability and Risk* summarizing the findings that had helped secure Poggiali's acquittal.

The Letby case now in court shows many of the same troubling features as earlier cases, Gill and others say. Letby was moved to clerical duties in 2016 after a series of deaths and resuscitations on her shifts, and first arrested in 2018. She is accused of murdering seven babies and attempting to murder 10 more, using methods such as insulin poisoning and injection of air bubbles.

The similarities go beyond statistics to the way Letby has been vilified. Social media commentary will "make your stomach turn," Gill says. "People are saying we should bring back hanging, shoot the bitch." The media have portrayed her as an "evil creature," says Neil Mackenzie, a lawyer based in Edinburgh, Scotland, who specializes in medical negligence cases and co-authored the RSS report. "I think there's possibly misogyny in there," Mackenzie says. "The press loves bad women."

The RSS report Gill and others published in September does not claim Letby is innocent, in part because public comment on the guilt or innocence of a person standing trial may be considered contempt of court in U.K. legal systems. "We've got to have no opinion on this case," Green says, but "there's potential here for miscarriage of justice."

Gill says a deep cognitive bias works against defendants like Letby. People "don't believe in chance, actually," he says. "Quantum mechanics has been shouting at us for 100 years that the physical universe is built on randomness. … But we don't understand this. It upsets us deeply. When a succession of bad things happens, we know there must have been an agent responsible. And so we naturally believe in devils and witches, gods and angels."

NOT ALL MEDICAL MURDER cases are witch hunts, however. "This is an instance where there actually are some witches," Thompson says.

In 2000, for example, a British physician named Harold Shipman was convicted of murdering 15 patients over a period of 3 years after an investigation yielded evidence that he had given overdoses of diamorphine—heroin, used in the United

Kingdom for severe pain—and falsified the medical records of numerous patients, suggesting they had been sicker than they were to make their deaths appear less suspicious. (Shipman was in one patient's will but his motives have not become clear.) Shipman, who is suspected of killing hundreds more, was sentenced to life in prison and died by suicide in 2004.

A 5-year government inquiry in the wake of the case identified ways to better protect patients, such as more oversight of death certificates. The case also led statisticians to explore a new question: Could statistics detect real murderers, based on a suspicious pattern of deaths alone?

Cambridge statistician David Spiegelhalter, who gave advice to the panel, believes so. He and his colleagues adapted a method from industrial quality control to compare the rate of death certificates signed by Shipman over time with deaths at other local doctors' practices. They found they could have identified a worrying pattern in Shipman's patients 13 years before he was arrested.

Such a system would produce false alarms; more people might die under the care of a doctor or nurse who works with particularly difficult cases, for example. But a robust method would prevent too many misfires, Spiegelhalter says, and a "ping" in the system should never be taken as anything more than a sign that a human should look at the data.

But implementing this kind of routine monitoring would be very complicated, says Bruce Guthrie, professor of general practice at the University of Edinburgh. The kind of data Spiegelhalter and his colleagues used is not routinely collected—it was pieced together as part of the Shipman investigation. And Shipman worked alone, which few family doctors do; many patients are likely to see multiple doctors. Only the "most horrendously prolific murderers" would be likely to show up, Guthrie wrote in an email to *Science*.

Meanwhile Thompson, Gill, and others are calling for cultural and institutional fixes to prevent unjust convictions. Many lawyers find statistics challenging, Mackenzie says. "This is one of the evils of the 'two cultures' myth," he says: Some students are channeled into scientific subjects, and others into humanities, and "never the twain shall meet." Niamh Nic Daeid, a forensic science researcher at the University of Dundee, says she routinely encounters anxiety and resistance about statistics. Nic Daeid, Spiegelhalter, and others have produced a range of statistics training materials, including an RSS "primer" and a free online course for lawyers.

But training is not enough, Thompson says, because the biases that underlie errors are "built into our perceptual processes." Instead, he says, it's crucial to change investigative procedures. The RSS report recommends that investigators be blinded. For example, pathologists should classify deaths as suspicious or not without knowing which medical personnel were in attendance, adapting the standardized blinding methods used in epidemiology to study disease outbreaks.

But blinding has proved to be a hard sell among forensic scientists, in part because it's often more challenging than it seems, says Peter Stout, CEO of the Houston Forensic Science Center and a strong advocate of blinding and other measures to improve forensic science. It can mean, for example, that a forensics lab—already strapped for funding and time—needs an extra person to serve as a case manager who screens possibly biasing information from a blinded analyst. And the line between relevant and irrelevant information is not always clear.

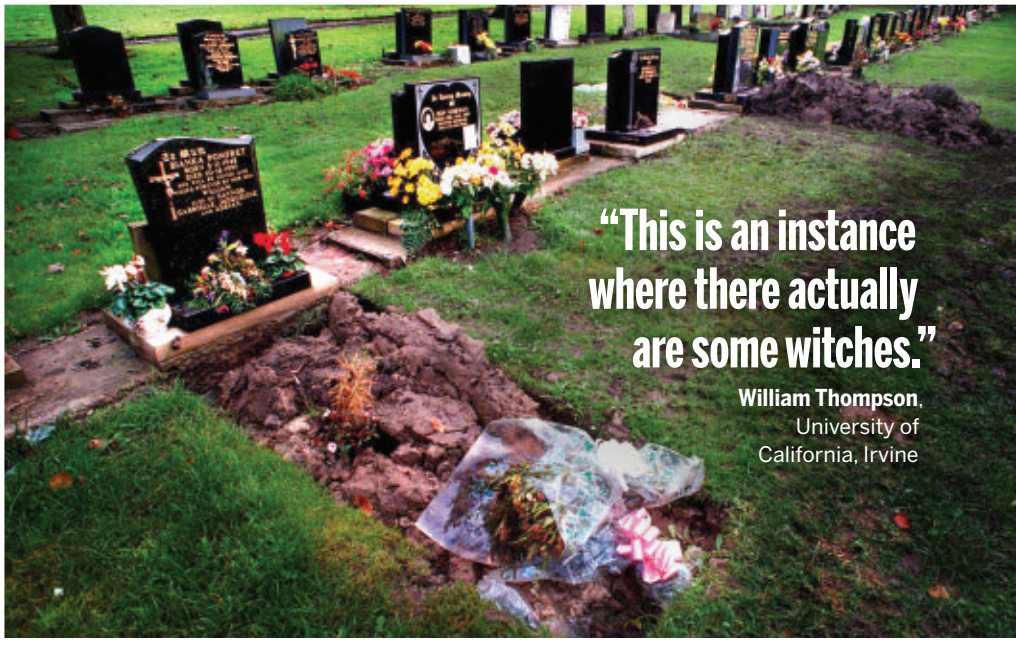
Adele Quigley-McBride, a cognitive bias researcher at Duke University, trains analysts in a technique called sequential unmasking. The method gets around the blurred line between relevant and irrelevant information by giving investigators access to increasing amounts of information with each round of analysis. Analysts note their observations and conclusions in each round; if new information changes their opinion, they have to explain why.

GILL WAS AT THE COURT of Appeal in Arnhem in 2010 when De Berk's exoneration was announced. "It was one of the biggest events of my life," he says. "It was really joyful." De Berk was immediately rushed off by her lawyers and journalists swarmed Derksen and De Noo. "I bought a marijuana cigarette," Gill says, and then he took a train to The Hague and went to the beach. "I smoked my joint, and I ate a dish of oysters, and drank some white wine."

De Berk later received a written apology from the Dutch minister of justice and an undisclosed financial compensation for the 6.5 years she spent in prison. Gill stays in contact with her; she likes his posts on Facebook sometimes. She told Gill she did not want to be interviewed for this story. "She's managed to put it all far away, and she needs to keep it that way," he says.

Thirteen years on, Gill, now retired, is watching the Letby case closely, but his obsession with forensic statistics has begun to subside. His retirement projects include a range of statistical kerfuffles with lower stakes, such as the rating of Dutch herring sellers. He has plenty of other things to occupy his attention—winemaking, an amateur distillery, grandchildren. "I think I've reached the point where I want to spend more time in the forest picking mushrooms, actually," he says.

He hopes younger statisticians will feel compelled to help when bad statistics lead to injustice, as he did. "I sensed that in the Lucia case, I could make a difference," Gill says. "And that therefore I must." ■



Graves were exhumed in Hyde, United Kingdom in 1998 to search for evidence against physician Harold Shipman, who was convicted of killing 15 patients in 2000.

PHOTO: INDEPENDENT/ALAMY STOCK PHOTO

INSIGHTS

PERSPECTIVES

ASTRONOMY

Light pollution is skyrocketing

Data from citizen scientists reveal a worrying growth in light pollution over the past decade

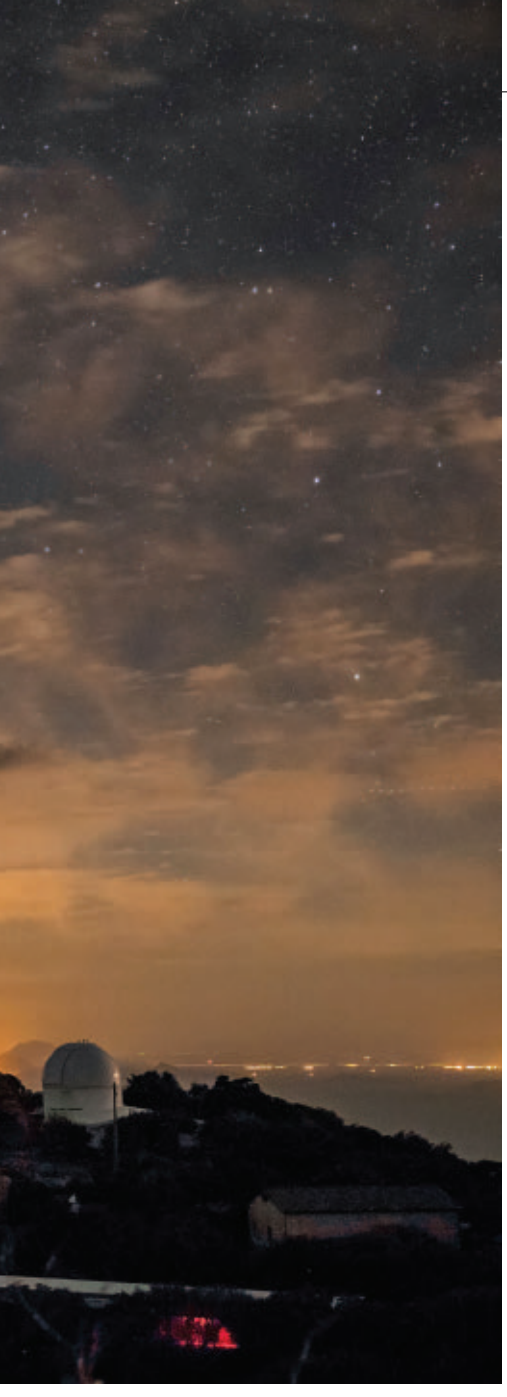


By **Fabio Falchi**[1,2] *and* **Salvador Bará**[1,3]

Artificial lighting at night is contributing to a rapid increase in light pollution, as reported by Kyba *et al.* (*1*) on page 265 of this issue. By analyzing tens of thousands of observations made by citizen scientists around the world over the past 12 years, the authors found that the dimmest stars in the night sky are being hidden by a 10% yearly increase in the sky background as a result of artificial light. This growth is difficult to discern with satellites now in operation because their detectors are blind to the blue light emitted by light-emitting diodes (LEDs), which are progressively replacing older lamps (such as high-intensity discharge). The results of the study point to the need for satellites that can detect the visible light in different color bands and make a strong case for reversing light pollution and its wide-ranging negative effects, from ecological to cultural to health-related.

Kyba *et al.* studied the collection of observations of the Globe at Night project (https://www.globeatnight.org). More than 50,000 observations were submitted by citizens—using the naked human eye—to detect the dimmest stars visible in a site. The data allowed the authors to determine the background brightness of the night sky. The brighter the night sky is, the brighter a star must be to be seen by the unassisted human eye. Analysis of data from the past dozen years (from 2011 to 2022) show that the sky brightness resulting from artificial light is increasing exponentially in the world with an alarming average of 10% each year—i.e., doubling in less than 8 years. This increase is much higher than estimates of the evolution of artificial light emissions (~2% yearly) based on radiance measurements taken by the Defense Meteorological Satellite Program and Suomi National Polar-orbiting Partnership satellites (*2, 3*). Part of this discrepancy could be explained by the inability of these satellites to detect blue light emitted by LEDs, which started to be used outdoors ~10 years ago. These satellites are also unable to detect light emitted almost horizontally, such as that from ultrabright LED billboards and build-

PHOTO: HARUN MEHMEDINOVIC/500PX/GETTY IMAGES



Light pollution intrudes on the Kitt Peak National Observatory in Arizona, USA.

ing façades. These weak points must be overcome by next-generation satellites equipped with both multiwavelength and multiangle monitoring capabilities (*4, 5*) that this global environmental problem (*6*) deserves.

The study of Kyba *et al.* is biased in that most of the data from citizen scientists are from Europe and North America. This does not subtract value from the findings; on the contrary, they are reinforced because these continents show a constant or slightly decreasing light pollution trend, as determined by satellites (*2*). The other continents show an increasing light pollution trend with satellites, so a greater increase would be expected for those continents when citizen science data become available.

Perhaps the most important message that the scientific community should glean from the Kyba *et al.* study is that light pollution is increasing, notwithstanding the countermeasures purportedly put in operation to limit it. Looking at the International Space Station's images and videos of Earth's night hemisphere, people generally are only struck by the "beauty" of the city lights, as if they were lights on a Christmas tree. They do not perceive that these are images of pollution. It is like admiring the beauty of the rainbow colors that gasoline produces in water and not recognizing that it is chemical pollution. Awareness must greatly increase for artificial light at night to be perceived not as an always-positive thing, but as the pollutant it really is.

Most people associate artificial light with road safety and personal security, connections that are not well supported by evidence (*7, 8*). Thus, year after year, more lights are installed to brighten the night. However, there are notable negative consequences that are on the rise. Astronomers were quick to realize that night lights affected their research, and hence they moved observatories farther from cities. Unfortunately, as the use of outdoor lights spreads, this strategy has become less effective. In addition, as the efficiency of light technology improves, more lighting is possible for the same amount of money. This is translating into a flourishing of new lights to illuminate roads, buildings, and natural landscapes; for digital billboards; and for garden "beautification," among other applications. To make things worse, thousands of satellites are being launched by a few corporations into low Earth orbits. This will prevent all humankind from admiring a pristine sky anywhere on the planet. The wilderness experience one can have in a national park at night, for example, may be increasingly affected by the glows of cities and satellite megaconstellations. The loss of the starry night is an unprecedented loss for all cultures.

Light pollution is also, and mainly, an environmental problem (*9*). Life on Earth evolved with sunlight during the day and starlight and the light from the Moon during the night. It is expected that introducing artificial light into ecosystems at levels that surpass the amount of natural light experienced in natural conditions will trigger changes in animal behavior. Some species, including predators, may take advantage of more light. Other species will be in danger of local or global extinction, which will reduce biodiversity (*10*). Moreover, animals, including humans, will diminish their production of melatonin in response to artificial light at night (*11*). Melatonin is a fundamental hormone that controls the internal circadian clock of animals, and thus animal physiology. A lack of this hormone could have numerous negative consequences on health (*12*). It is also expected that processes needed to produce, install, and operate outdoor artificial lighting will release greenhouse gases into the atmosphere and contribute to global warming. An estimated 200 billion kg of carbon dioxide is generated to produce the ~400 terawatt-hours of electric energy needed for outdoor lights (*13*).

Attempts to control light pollution have been carried out in the past decades, including aiming light below the horizon plane. This approach is not sufficient because any new light, even if shielded, will add pollution to the night environment after being reflected off the surfaces being lit. New approaches that treat light as a true pollutant could include introducing total caps and red lines on its production that are based on the control of light pollution indicators, such as the night sky zenith luminance or the illuminance on the ground coming from the entire sky (*14, 15*). This would be analogous to the reduction of most air contaminants that was achieved through legislative interventions, such as the Clean Air Act in the United States.

Light pollution is an environmental problem and as such should be confronted and solved. This may restore the need for an uncontaminated night environment by animals and humankind. ■

[1]Departamento de Física Aplicada, Universidade de Santiago de Compostela, Santiago de Compostela, Galicia, Spain. [2]Light Pollution Science and Technology Institute (ISTIL), Thiene, Italy. [3]Independent Scholar, Santiago de Compostela, Galicia, Spain. Email: falchi@istil.it

REFERENCES AND NOTES

1. C. C. M. Kyba, Y. Ö. Altıntaş, C. E. Walker, M. Newhouse, *Science* **379**, 265 (2023).
2. C. C. M. Kyba *et al.*, *Sci. Adv.* **3**, e1701528 (2017).
3. A. Sánchez de Miguel, J. Bennie, E. Rosenfeld, S. Dzurjak, K. J. Gaston, *Remote Sens.* **13**, 3311 (2021).
4. C. D. Elvidge *et al.*, *Int. J. Remote Sens.* **28**, 2645 (2007).
5. C. C. M. Kyba *et al.*, *J. Geophys. Res. Atmos.* **127**, e2021JD036382 (2022).
6. F. Falchi *et al.*, *Sci. Adv.* **2**, e1600377 (2016).
7. P. R. Marchant, P. D. Norman, *Appl. Spat. Anal. Policy* **15**, 1583 (2022).
8. R. Steinbach *et al.*, *J. Epidemiol. Community Health* **69**, 1118 (2015).
9. K. J. Gaston, A. Sánchez de Miguel, *Annu. Rev. Environ. Resour.* **47**, 373 (2022).
10. R. Sordello *et al.*, *Landsc. Urban Plan.* **219**, 104332 (2022).
11. M. Grubisic *et al.*, *Sustainability* **11**, 6400 (2019).
12. R. G. Stevens, G. C. Brainard, D. E. Blask, S. W. Lockley, M. E. Motta, *CA Cancer J. Clin.* **64**, 207 (2014).
13. C. Walker, P. Benvenuti, Eds., Dark and Quiet Skies II Working Group Reports, Zenodo (2022); https://doi.org/10.5281/zenodo.5874725.
14. S. Bará, F. Falchi, R. C. Lima, M. Pawley, *Environ. Chall.* **5**, 100212 (2021).
15. F. Falchi, S. Bará, *R. Soc. Open Sci.* **7**, 201501 (2020).

10.1126/science.adf4952

CATALYSIS

Clicking zeolites together

A new mechanism to synthesize catalytic zeolites gives a previously unknown topology

By **Russell E. Morris**

Zeolites are some of the most successful catalysts used in industry, particularly in petrochemicals processing. The topology of zeolites is important because this controls access to the catalytic sites inside these porous solids. There is a growing need for new catalytic materials that will accommodate larger substrates that are common in, for example, bio-based fuel refining. On page 283 of this issue, Li *et al.* (*1*) describe a mechanism by which an extra-large-pore zeolite is made by "clicking" together one-dimensional (1D) silicate chains that are present in a precursor solid. The synthesis hinges on the precise arrangement of the silicate chains in the precursor, which are aligned to connect to a highly crystalline and almost defect-free product. The new zeolite also features structural units that offer hitherto unforeseen potential for distinct applications such that the size of the zeolite pores matches the substrate better than what is possible now.

Zeolites are porous, silicate-based solids whose catalytic function comes from the addition of other elements into the framework structure. The nature of the catalytic functionality depends on the specific element that is added: For example, aluminum and titanium enable acid and redox catalysis, respectively. The zeolite topology controls access to this catalytic functionality, limiting which molecules can react according to those that can diffuse into or out of the pores. Zeolites are traditionally prepared using crystallization techniques under hydrothermal conditions (*2*). This is a dynamic process in which the successive making and breaking of chemical bonds leads to the growth of highly ordered crystal structures.

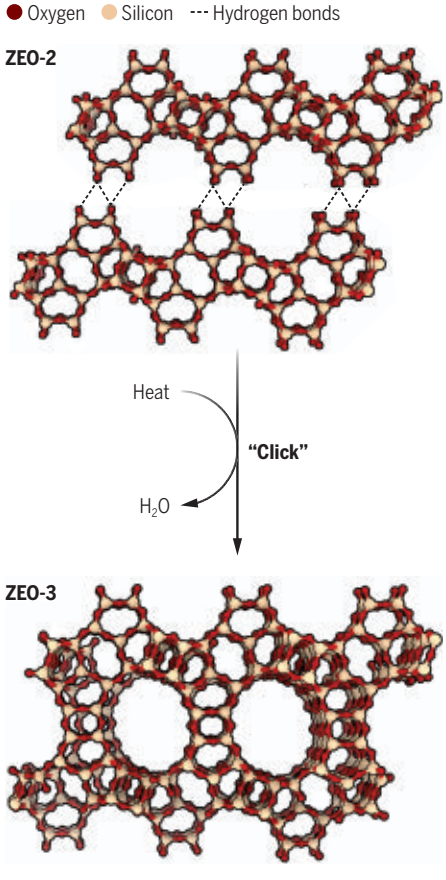
Over the years, there have been many advances based on modification of hydrothermal crystallization that have led to improved control over zeolite topology. The development of organic structure–directing agents (OSDAs) has been especially fruitful (*3*), but other approaches to controlling zeolite topology, such as using solvents (*4*) and chemical compositions that have structure-directing properties (*5*), have expanded the range of possible zeolites that are available. An important goal of this work is the generation of so-called extra-large-pore materials that can handle larger substrates (*6*). As a result of extensive research, understanding of the features of hydrothermal crystallization that promote larger pores in the final material has improved. These features include the correlation between the density of the zeolite and the ring sizes that are present in the material—the smaller the average ring size, the lower the density (*7*). In the study of Li *et al.*, the synthesis of a zeolite they call ZEO-3 seems to break this correlation and produces a material with unexpectedly high porosity for its large average ring size.

The difference between the traditional approach to zeolite synthesis and that for the preparation of ZEO-3 is the final, framework-forming step in the process (see the figure). Rather than a reversible crystallization, the last step is an irreversible condensation reaction of a precursor material called ZEO-2. ZEO-2 contains chains of silicate that are held together by hydrogen bonds that undergo a condensation reaction upon heating to produce ZEO-3, which contains chains that are then connected by stronger covalent bonding. The reaction is topotactic in that the structure of the chains themselves remains untouched by the process: Only the interchain bonding is affected. The condensation reaction is reminiscent in many ways of the click chemistry approach that has been so successful in organic chemistry and that won the 2022 Nobel Prize in Chemistry. The difference between the two processes (apart from the difference in the chemistry involved) is that for the formation of a crystalline solid based on a fully connected lattice, the clicking together needs to be a concerted process over a long range of chain length. The key to this being possible is the relative arrangement of the chains in ZEO-2. The use of an unusual phosphorus-containing OSDA in the synthesis of ZEO-2 leads to the chains being almost perfectly arranged such that they click seamlessly into place in the final material. Topotactic condensations of 2D (layered) precursors have been carried out previously (*8*) to provide new zeolites, but a 1D-to-3D transformation had not been demonstrated before. Applying a similar approach to the condensation of chains results in a surprisingly high degree of crystallinity and defect-free final material.

What really makes this work intriguing, however, is the unexpectedly high porosity of the final material. It has been noted before that changing the mechanism of zeolite synthesis, such as in the so-called assembly-disassembly-organization-reassembly (ADOR) method, can lead to unexpected materials (*9*, *10*). This route has been demonstrated for small- and medium-pore zeolites, but the preparation of the extra-large-pore zeolite, ZEO-3, opens up the possibility

From one dimension to three

Precise organization of one-dimensional units, such as the chains present in ZEO-2, leads to the possibility of an irreversible condensation reaction to form ZEO-3, a zeolite with unusual and surprising three-dimensional topology. The success of this approach relies on the relative organization of the silicate chains in ZEO-2: They are almost perfectly arranged and held together by hydrogen bonding, which allows them to "click" together in a condensation reaction to form a covalently bonded extra-large-pore zeolite.



GRAPHIC: A. MASTIN/*SCIENCE*

EaStCHEM School of Chemistry, University of St Andrews, St Andrews, UK. Email: rem1@st-andrews.ac.uk

of targeting materials that can be used to catalyze reactions of large substrates. That the preparation of ZEO-3 is even possible implies that the previous "rules" for targeting high-porosity zeolites are mechanism dependent and that finding new mechanisms will inevitably lead to new topologies. This is particularly attractive given the large number of hypothetical zeolites that have been computationally predicted (*11*) but not yet realized in the laboratory. In addition, Li *et al.* show that elements such as titanium can be incorporated into ZEO-3, which is beneficial because such elements doped into the zeolites provide the active sites where catalysis takes place. In the case of titanium, this opens up the possibility of new redox catalysis. Developing such active zeolites is vital work because transferring their utility to new technologies, such as the formation of biodegradable plastics from renewable sources (*12*), is key to developing a society that is less reliant on petroleum products.

The next question is whether the same type of chemistry can be done to form zeolites with different, even larger, pore structures. The real key in the work of Li *et al.* is the almost perfect organization of the precursor, ZEO-2. Can the same level of organization be produced with other low-dimensional silicate materials (layers, chains, or even molecules)? In ZEO-2, the organization came about through a crystallization process, but an expansion of the potential of this mechanism would come from more active control of organization. For example, there have been advances in the high-yield preparations of exfoliated zeolitic nanolayers (*13*)—can such materials be organized so that they can be clicked together in the same way as ZEO-3? It is a major challenge to have such control over organization, but if it is possible, the resulting zeolites could also break the current "rules" and have structural properties different from those prepared using traditional methods, in turn opening up new applications. ■

REFERENCES AND NOTES

1. J. Li *et al.*, *Science* **379**, 283 (2023).
2. C. S. Cundy, P. A. Cox, *Chem. Rev.* **103**, 663 (2003).
3. A. Corma, F. Rey, J. Rius, M. J. Sabater, S. Valencia, *Nature* **431**, 287 (2004).
4. E. R. Cooper *et al.*, *Nature* **430**, 1012 (2004).
5. J. L. Paillaud, B. Harbuzaru, J. Patarin, N. Bats, *Science* **304**, 990 (2004).
6. Q. F. Lin *et al.*, *Science* **374**, 1605 (2021).
7. G. O. Brunner, W. M. Meier, *Nature* **337**, 146 (1989).
8. L. Schreyeck, P. Caullet, J. C. Mougenel, J. L. Guth, B. Marler, *Microporous Mater.* **6**, 259 (1996).
9. W. J. Roth *et al.*, *Nat. Chem.* **5**, 628 (2013).
10. M. Mazur *et al.*, *Nat. Chem.* **8**, 58 (2016).
11. V. A. Blatov, G. D. Ilyushin, D. M. Proserpio, *Chem. Mater.* **25**, 412 (2013).
12. M. Dusselier, P. Van Wouwe, A. Dewaele, P. A. Jacobs, B. F. Sels, *Science* **349**, 78 (2015).
13. W. J. Roth *et al.*, *Sci. Adv.* **6**, eaay8163 (2020).

10.1126/science.adf3961

CHEMISTRY

A strained ring for stereoselective synthesis

The spring-loaded tension of a cyclic allene expedites synthesis of a natural product

By **Christian L. Jankovic** *and* **F. G. West**

Effective chemical synthesis of a complex natural product in a laboratory flask requires careful planning so that the smallest possible number of reaction steps is used and they all occur at high chemical yields. These efforts can be thwarted when key strategic steps (forming important chemical bonds or establishing stereogenic centers) are unselective—that is, they furnish undesired reaction products instead of or in addition to the one needed to complete the synthetic sequence. To overcome these challenges, new types of transformations are sought that proceed with high selectivity under the mildest possible conditions. On page 261 of this issue, Ippoliti *et al.* (*1*) describe the use of a new class of reactive intermediates, cyclic allenes, in the enantioselective synthesis of the alkaloid lissodendoric acid A. This synthesis highlights how cyclic allene chemistry can work in tandem with other methods to obtain daunting products in a selective fashion.

Allenes are three-carbon structures containing two alkene π bonds that typically exist with the carbon atoms arranged linearly and with the groups attached to the terminal carbons in a 90° orientation to each other (see the figure). However, when this functional group is embedded in a six-membered ring, the linear geometry is no longer possible, causing a deformation that bends the ends toward each other and rotates the terminal groups closer to planarity. This distortion imparts considerable potential energy—ring strain—into the molecule, and the release of this energy through a reaction that relieves the ring strain makes cyclic allenes highly reactive (*2*, *3*). Indeed, six-membered cyclic allenes are so reactive that they exist only transiently and cannot be isolated (*2*). This high reactivity has been demonstrated by the observation of many types of trapping reactions that occur at or below room temperature (*4*–*8*).

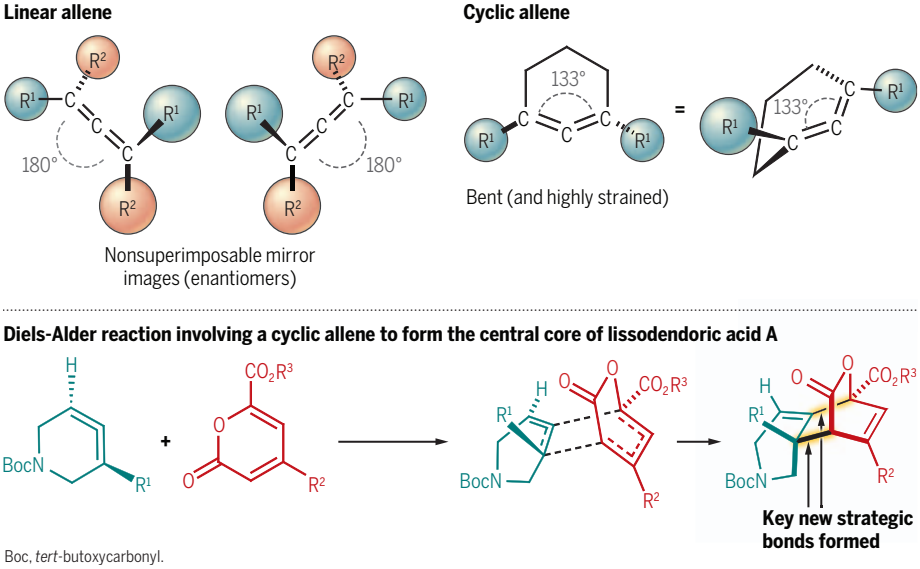
Department of Chemistry, University of Alberta, E3-43 Gunning-Lemieux Chemistry Centre, Edmonton, AB, Canada. Email: fwest@ualberta.ca

Reactive strained intermediates have been used widely as building blocks for complex and functional molecules. Benzynes, aromatic rings that possess an endocyclic carbon-carbon triple bond, have long been the premier example of this concept, with many total syntheses using these fascinating molecules in crucial bond-forming reactions (*9*). Another class of strained molecules, cyclic alkynes, played an essential role in the development of biorthogonal chemistry, which was awarded with the 2022 Nobel Prize in Chemistry (*10*). The potential energy embedded in strained cyclic alkynes allows for the facile covalent linkage of biologically relevant moieties within living systems. However, in contrast to these strained ring molecules, cyclic allenes—particularly those of smaller ring sizes—have thus far found limited application. Only now are cyclic allenes emerging as candidates for use in the construction of biologically and synthetically useful scaffolds. For example, the high reactivity of cyclic allenes was harnessed in the preparation of DNA-encoded libraries (*11*), and the total synthesis of lissodendoric acid A reported by Ippoliti *et al.* provides another important illustration of their potential value. In particular, the synthesis of lissodendoric acid A exploits a critical difference between cyclic allenes and other strained intermediates: their intrinsic chirality.

Chirality usually results from the presence of one or more atoms with four different groups attached. A distinct feature of cyclic allenes is their inherent axial chirality, in which each enantiomer can furnish a different product. Although typically generated as 50:50 mixtures of both enantiomers, methods exist to enable the selective generation of one cyclic allene enantiomer (*5*, *12*). A reaction that uses this single enantiomer would be expected to furnish a predominant product with one three-dimensional arrangement, which is needed if the goal involves laboratory synthesis of a naturally occurring compound. Alkaloids such as lissodendoric acid A are usually formed in nature as a single enantiomer. For natural products with therapeutic properties, generally a single enantiomer will be biologi-

Adding to the synthesis toolbox

Allenes are three-carbon structures that typically exist in a linear arrangement (top left). However, when this functional group is embedded in a six-membered ring, the geometry becomes distorted, and this imparts ring strain into the molecule, making cyclic allenes highly reactive (top right). The enantioselective generation of a cyclic allene intermediate can be used to synthesize the central core of the natural product, lissodendoric acid A (bottom).



Boc, *tert*-butoxycarbonyl.

cally functional, whereas the others may be less effective or even harmful to living cells.

Ippoliti *et al*. observed that the central core of lissodendoric acid A could be constructed with a Diels-Alder reaction that uses a cyclic allene, a reaction that would give access to the central fused ring scaffold. After careful validation of this hypothesis, the authors found a set of conditions that allowed for the high-yielding synthesis of a cyclic allene intermediate, using a fluoride ion to cause an elimination reaction of a silicon-containing precursor. This reaction proceeded with a high degree of stereospecificity for the generation of the required enantiomer. A remarkably concise synthetic route was then applied to obtain lissodendoric acid A. This sequence of steps included the extrusion of carbon dioxide at a key point to introduce a 1,3-diene present in lissodendoric acid A, along with a highly effective ring-closing metathesis reaction to form a challenging 14-membered ring. The elegance of this approach to a complex alkaloid natural product was made possible through the use of an enantiomerically enriched and highly reactive cyclic allene intermediate. Furthermore, this laboratory synthesis also confirmed the stereochemical assignment of the naturally occurring functional enantiomer, which had previously been surmised but unproven.

Although the existence of cyclic allenes has been known for many decades, the study by Ippoliti *et al.* details the advantageous use of cyclic allenes in the chemical synthesis of a complex natural product. This achievement demonstrates that cyclic allenes can function as efficient building blocks en route to biologically intriguing molecules previously found only in minute quantities from natural sources. Access to a robust and selective reaction expedites development of a streamlined synthetic sequence with greatly reduced costs of time and reagents. At the same time, valuable insights into the distinct reactivity of these unusual and transient molecules have been uncovered. This new process is a welcome addition to the synthetic chemist's toolbox, and this work is likely to be one of many future applications of cyclic allenes in complex molecule synthesis. ■

REFERENCES AND NOTES

1. F. M. Ippoliti *et al.*, *Science* **379**, 261 (2023).
2. G. Wittig, P. Fritze, *Angew. Chem. Int. Ed.* **5**, 846 (1966).
3. W. R. Moore, W. R. Moser, *J. Am. Chem. Soc.* **92**, 5469 (1970).
4. J. S. Barber *et al.*, *J. Am. Chem. Soc.* **138**, 2512 (2016).
5. J. S. Barber *et al.*, *Nat. Chem.* **10**, 953 (2018).
6. V. A. Lofstrand, F. G. West, *Chemistry* **22**, 10763 (2016).
7. V. A. Lofstrand, K. C. McIntosh, Y. A. Almehmadi, F. G. West, *Org. Lett.* **21**, 6231 (2019).
8. M. M. Yamano *et al.*, *Angew. Chem. Int. Ed.* **58**, 5653 (2019).
9. P. M. Tadross, B. M. Stoltz, *Chem. Rev.* **112**, 3550 (2012).
10. N. J. Agard, J. A. Prescher, C. R. Bertozzi, *J. Am. Chem. Soc.* **126**, 15046 (2004).
11. M. V. Westphal *et al.*, *J. Am. Chem. Soc.* **142**, 7776 (2020).
12. M. Christl, H. Fischer, M. Arnone, B. Engels, *Chemistry* **15**, 11266 (2009).

ACKNOWLEDGMENTS

The authors thank the Natural Science and Engineering Research Council of Canada (NSERC) for support of our studies in this field.

10.1126/science.ade7122

CANCER

Germline study points to sarcoma pathways

Pathogenic variants related to mitosis and telomere integrity are enriched in sarcoma patients

By **Diana Mandelker**[1] *and* **Marc Ladanyi**[1,2]

Germline cancer predisposition is an important cause of premature mortality. Traditionally, gene discovery in hereditary cancer involved either familial linkage analysis to segregate a variant in a family with several instances of the same cancer type, case-control studies for a limited number of genes or variants to establish relative risks, or genomewide association studies (GWASs) that look for the association of common variants to cancer development. On page 253 of this issue, Ballinger *et al.* (*1*) identify genes and pathways associated with sarcoma predisposition, based on a case-control study design using whole-genome sequencing (WGS) of 1644 individuals with sarcoma and 3205 matched controls without sarcoma paired with genetic ontology analysis. They found pathogenic germline variants in genes related to mitosis and telomere integrity in 44 sarcoma patients, representing a significant enrichment relative to controls and, more broadly, highlighting the importance of these processes to sarcomagenesis.

The known spectrum of genes associated with sarcoma predisposition is broad, but largely limited to rare genetic syndromes with other clinical features (*2*). Ballinger *et al.* identified 14 new candidate sarcoma-predisposition genes in seemingly sporadic sarcoma cases (see the table). These genes clustered functionally in the telomeric and mitotic pathways, specifically genes that were related to or encoded components of the shelterin complex (which consists of six

[1]Department of Pathology and Laboratory Medicine, Memorial Sloan Kettering Cancer Center, New York, NY, USA. [2]Human Oncology and Pathogenesis Program, Memorial Sloan Kettering Cancer Center, New York, NY, USA. Email: mandelkd@mskcc.org; ladanyim@mskcc.org

GRAPHIC: V. ALTOUNIAN/*SCIENCE*

proteins involved in protecting telomere ends from being misrecognized as DNA damage by DNA repair mechanisms) and the centrosome complex (whose proper functioning in mitosis prevents chromosome mis-segregation and aneuploidy). This suggests that alterations in these biological processes are key to sarcoma predisposition.

Although one of these telomere shelterin complex genes, protection of telomeres 1 (*POT1*), had already been implicated in predisposition to angiosarcoma (*3, 4*), the study of Ballinger *et al*. extends its role in susceptibility to a wider spectrum of sarcoma types. Moreover, the identification of pathogenic variants in other shelterin complex genes reinforces the idea that abnormally long telomeres and increased telomere fragility predispose to sarcomas. The pathogenic variants in shelterin complex genes and centrosome complex genes were observed in 44 patients with both diverse complex karyotype sarcomas such as undifferentiated pleomorphic sarcoma (UPS) and angiosarcoma as well as, perhaps unexpectedly, simple karyotype, gene fusion–driven sarcomas such as synovial sarcoma, myxoid liposarcoma (MLPS), and alveolar rhabdomyosarcoma (ARMS), representing the two broad genomic classes of sarcomas (*5, 6*).

The study by Ballinger *et al*. should prompt further work to more firmly establish these candidates as sarcoma predisposition genes and to determine penetrance and the spectrum of cancer predisposition they confer. Although the study design did not allow for estimates of penetrance, the pedigrees suggest a low penetrance for sarcoma but a possible shared predisposition to melanoma. Specifically, the authors find an overrepresentation of shelterin complex variants in families with melanoma and thyroid cancers, in addition to sarcoma. However, these family members were not available for familial studies to prove that the variant segregates with disease to define a new cancer susceptibility syndrome. Furthermore, the study cohort was composed of a diverse mix of sarcoma types. Cancer predispositions can be specific to sarcoma type. For example, germline pathogenic variants in *TP53* (which encodes p53) in people with Li-Fraumeni syndrome are associated with predisposition to osteosarcomas and complex karyotype soft tissue sarcomas, but simple karyotype, gene fusion–driven sarcomas are not a part of Li-Fraumeni syndrome. Therefore, studies of larger cohorts of specific sarcoma types may refine the cancer susceptibility spectrum conferred by these newly identified genes.

Additional analyses of somatic alterations in sarcomas with these germline pathogenic variants may provide more insight into the biology of these cancers (*7*). Some of the somatic analyses presented by Ballinger *et al*. may argue against these germline variants having a classic loss-of-function role in tumorigenesis. For example, for germline pathogenic variants in centrosome genes, only 2 of 24 of the tumors had evidence of loss of heterozygosity (LOH) of the wild-type allele. Traditionally, hereditary cancer susceptibility genes require a second somatic hit (such as LOH) to fully inactivate the gene and drive tumorigenesis. Thus, ~80 to 90% of ovarian cancers arising in carriers of *BRCA1* or *BRCA2* germline variants demonstrate LOH of the wild-type allele (*8*). Additionally, four of five gastrointestinal stromal tumors (GISTs) in carriers of centrosome gene variants had somatic *KIT* mutations, whereas germline pathogenic variants in other GIST-associated syndromes usually have wild-type *KIT* (*9*, *10*). Therefore, further work is needed to determine whether some of the variants identified may act in a haplo-insufficient or dominant-negative manner, where loss of the wild-type allele is not required for tumorigenesis. Most cancer predisposition genes are also sporadically inactivated by somatic mutations, as has been shown for *POT1* in angiosarcomas (*4*), so it will be germane to screen the other genes nominated by Ballinger *et al*. for somatic mutations in sporadic sarcomas.

Over the past few years, genetic testing of broad cohorts of cancer patients has revealed that ~10 to 20% harbor pathogenic germline variants in cancer susceptibility genes (*11–13*), despite the absence of clear family histories of the cancer types that would have triggered germline testing. Increasingly, this germline genetic testing is becoming integral to the care of cancer patients and their families because patients can receive targeted therapies [for example, germline *BRCA1* or *BRCA2* mutations sensitize ovarian, breast, and pancreatic cancers to poly(ADP-ribose) polymerase (PARP) inhibitors] and at-risk family members can undergo enhanced cancer surveillance or risk-reducing surgery. Notably, however, these genetic studies on large cancer cohorts have mostly focused on genes with well-established associations to cancer susceptibility and therefore do not account for cancer hereditability that may result from genes not previously linked to cancer susceptibility, such as those identified by Ballinger *et al*.—hence the importance of the unbiased WGS approach in this study. By using a nuanced experimental approach involving WGS, an extreme phenotype design, and genetic ontology analysis, the authors have discovered 14 candidate genes associated with sarcoma predisposition. If confirmed, this will be critical to the management of individuals and families with sarcoma because it will enable more at-risk genetic testing and cancer surveillance. Furthermore, this experimental design has the potential to reveal more genes that confer susceptibility to a wide variety of diseases and enhance the ability to practice individualized genetic medicine in the future. ■

Candidate sarcoma predisposition genes

Pathogenic variants in 14 genes belonging to two broad functional classes were identified in 44 individuals with diverse sarcomas.

Shelterin complex and related genes

GENE	LOCATION	SARCOMA TYPES*
POT1	7q31	WDLPS, UPS (2), angiosarcoma, myxofibrosarcoma, MPNST
SMARCAL1	2q35	Osteosarcoma (2), PNET, UPS (2), GIST
STAG3	7q22	Osteosarcoma, rhabdomyosarcoma NOS, UPS
TERF1	8q21	UPS, chondrosarcoma
TERF2IP	16q23	GIST, WDLPS
TIMELESS	12q13	Chondrosarcoma (2), synovial sarcoma
TINF2	14q12	Liposarcoma NOS, myxoid liposarcoma, UPS

Centrosome complex and related genes

CEP72	5p15	UPS, synovial sarcoma (2), GIST (2), PNET, ARMS
CEP63	3q22	UPS, GIST, MPNST
CEP89	19q13	UPS (2)
HAUS4	14q11	GIST**, chordoma
HAUS5	19q13	GIST (3)**, leiomyosarcoma
PCM1	8p22	GIST
SSNA1	9q34	Leiomyosarcoma

*Each sarcoma type represents one individual unless otherwise specified.
**One patient with GIST harbored truncating variants in both *HAUS4* and *HAUS5*.
ARMS, alveolar rhabdomyosarcoma; *CEP*, centrosomal protein; GIST, gastrointestinal stromal tumor; *HAUS*, HAUS augmin like complex subunit; MPNST, malignant peripheral nerve sheath tumor; NOS, not otherwise specified; *PCM1*, pericentriolar material 1; PNET, primitive neuroectodermal tumor; *POT1*, protection of telomeres 1; *SMARCAL1*, SWI/SNF-related matrix-associated actin-dependent regulator of chromatin subfamily A–like protein 1; *SSNA1*, SS nuclear autoantigen 1; *STAG3*, stromal antigen 3; *TERF1*, telomeric repeat binding factor 1; *TERF2IP*, TERF2 interacting protein; *TINF2*, TERF1 interacting nuclear factor 2; UPS, undifferentiated pleomorphic sarcoma; WDLPS, well-/dedifferentiated liposarcoma.

REFERENCES AND NOTES

1. M. L. Ballinger *et al*., *Science* **379**, 253 (2023).
2. M. Farid, *Oncologist* **21**, 1002 (2016).
3. O. Calvete *et al*., *Nat. Commun.* **6**, 8383 (2015).
4. E. Shen *et al*., *J. Med. Genet.* **57**, 664 (2020).
5. B. S. Taylor *et al*., *Nat. Rev. Cancer* **11**, 541 (2011).
6. Cancer Genome Atlas Research Network, *Cell* **171**, 950 (2017).
7. D. Mandelker, O. Ceyhan-Birsoy, *Trends Cancer* **6**, 31 (2020).
8. K. N. Maxwell *et al*., *Nat. Commun.* **8**, 319 (2017).
9. S. A. Boikos *et al*., *JAMA Oncol.* **2**, 922 (2016).
10. D. Mandelker *et al*., *NPJ Precis. Oncol.* **7**, 1 (2023).
11. D. Mandelker *et al*., *JAMA* **318**, 825 (2017).
12. N. J. Samadder *et al*., *JAMA Oncol.* **7**, 230 (2021).
13. O. Ceyhan-Birsoy *et al*., *Genome Med.* **14**, 92 (2022).

10.1126/science.adf8572

POLICY FORUM

SCIENCE POLICY

Remaking the Chinese Academy of Sciences

Under pressure to reinvent itself, the CAS should concentrate on managing large-scale research infrastructures

By **Xiyi Yang[1], Xiaoyu Zhou[1], Cong Cao[2]**

The decades-long reform of China's science and technology system has repeatedly ignited a critical question about the position of the Chinese Academy of Sciences (CAS). The academy's early contributions to China's strategic weapons programs, and indeed, to China's nation-building efforts, have left a legacy that partly mirrors the development of Chinese science but is almost impossible to replicate. But the CAS has been under enormous pressure to reinvent itself by producing original and more visible outcomes to justify its existence as the most prominent research institution in China. We discuss major challenges faced by the CAS and the antecedents of its current reform, a reshuffling of its 100-plus research institutes under a new Pioneer Initiative (*1*). By concentrating resources on big-science centers managing large-scale research infrastructures (RIs), CAS may be able to enhance its uniqueness and strategic importance among China's many research institutions and to potentially bolster its future.

ONGOING S&T SYSTEM REFORM

After some 30 years' intensive investments, China has achieved remarkable progress in science and technology (S&T), at least in quantitative terms (*2*). But scratching beneath that surface reveals a creeping suspicion that the country's innovation system is still underperforming (*3*, *4*). The long-standing criticism of China's lack of originality, along with the intensified global technological competition, has further prompted the country to reform its S&T system to become a major S&T power in the world by 2050.

The Chinese government's most recent national 5-year plan, announced in 2020 with objectives for 2035, sets three major targets for S&T reform. First, China will concentrate resources and especially accelerate investments to strengthen its strategic S&T force. National laboratories, an important fixture of the strategic S&T force, would be formed to tackle key technological problems faced by the country.

Second and related, China will continuously invest in the construction of big-science RIs. Globally, big-science RIs have been the foundation for some of the most successful national laboratories, including the Oak Ridge National Laboratory in the US, the European Organization for Nuclear Research (CERN) in Switzerland, and the High Energy Accelerator Research Organization (KEK) in Japan. By using big-science RIs, such laboratories undertake not only basic research but also research tasks with strategic goals. Furthermore, the construction and operation of these laboratories and the associated big-science RIs are often coordinated by governmental institutions or transnational research organizations that go beyond the capability of a single university or research institution.

Third, China's ongoing efforts of forming a market-oriented and enterprise-centered innovation system require a more delicate division of labor in its different actors. Market-oriented research and innovation should gradually be transferred to enterprises. Universities should turn out the next generation of talents who are able to handle increasingly complex tasks. And national laboratories should focus their efforts on cutting-edge research that fits into national agendas, or the research in Pasteur's quadrant (*5*, *6*).

In this context, as China's largest and premier research institution, the CAS needs to refine its role. Headquartered in Beijing and with an employment of 69,000 and a postgraduate student body of 79,000, the academy currently comprises 11 branches and more than 100 institutes throughout the country. During his visit to the CAS in July 2013, China's President Xi Jinping, while affirming the academy's past contributions, also challenged it to become a pioneer in four aspects: achieving leap-forward development in S&T, nurturing a national innovative talent pool, enhancing its role as the nation's premier S&T think tank, and building world-class research institutes. Apparently, the political leadership is still unsatisfied with the academy's most recent performance, especially given its consumption of a substantial amount of state funding (in 2021, it operated on an S&T expenditure of USD$13.26 billion in 2022 dollars, including USD$5.77 billion, or 43.51%, from government appropriation, and its R&D expenditure reached USD$12.52 billion, considerably more than any other Chinese institutions of learning).

MAJOR CHALLENGES FACING THE CAS

Early on, the division of labor between universities focused on education, and industrial R&D institutes focused on problems of their own sectors, justified the existence of the CAS as a national academy. From the 1950s to the 1970s, the CAS demonstrated its value, especially by its contributions to the success of China's strategic weapons programs, through which it also leveraged to achieve substantial disciplinary development.



The Large High Altitude Air Shower Observatory (LHAASO), shown in June 2022, was built in Daocheng County, Sichuan Province, to detect cosmic rays.

PHOTO: LIU ZHONGJUN/CHINA NEWS SERVICE/GETTY IMAGES

[1]School of Entrepreneurship and Management, ShanghaiTech University, Shanghai, China. [2]Nottingham University Business School China, University of Nottingham Ningbo China, Ningbo, China. Email: zhouxy@shanghaitech.edu.cn; cong.cao@nottingham.edu.cn



In the mid-1980s when China started to reorganize its S&T system, the CAS adopted a "one academy, two systems" strategy that required the majority of its staff to engage in commercialization and projects of direct relevance to the economy while concentrating a small number of scientists on basic research and on following the global trend in high and new technology. Consequently, the academy's overall research quality and ability to tackle fundamental research questions were undermined.

The launch of the Knowledge Innovation Program (KIP) in 1998 was to directly address these problems and the academy's uncertain future. The CAS initially retrenched its staff, aiming to make its institutes nimble and its workforce mobile to better respond to new challenges (*7*). However, the downsizing of the academy, along with the strengthening of research capability in universities and the national defense sector, had threatened its existence. In response, in the late stage of the KIP, the CAS not only reversed its earlier approach but also swung to the other extreme by substantially expanding its reach. It established research institutes in emerging scientific disciplines and in cities where it did not have a presence. It also formed alliances with provincial and local governments and industrial organizations to help raise technological content of the local economy, thus deviating from its mission as a national academy.

Entering the 21st century, the CAS was further challenged for its vitality in China's S&T system. The research areas of many CAS institutes overlapped with each other and with those of universities, and a number of projects repeated one another unnecessarily. The quality of research was not satisfactory (*8*). The academy also faced an increasing challenge from top universities in recruiting talents. For instance, among the 667 "Young Thousand Talents" recruited in 2015, 94 went to CAS-affiliated organizations while 43 and 42 went to Tsinghua and Peking University, respectively.

Furthermore, there was a lack of interest among CAS scientists in seeking opportunities to apply their research, for two primary reasons. First, the Chinese government has been using its wealth of funds and political will to push innovation from the top. A substantial proportion of research projects, including those done at the CAS, were state led instead of market oriented, which limited the opportunity of commercialization (*4*). Second, performance evaluation in most CAS institutes determined that scientists had little incentive for commercialization. Following the global trend to conduct research and being the first to publish remained the primary goal of CAS scientists.

In addition, the academy's multiple functions and missions—from research, talent training, strategic high-tech development, commercialization and local engagement to the provision of policy advice as a think tank and through its honorific members—had made both the overall organizational management and the evaluation of individual institutes and scientists across the CAS extremely difficult. Indeed, the crux is more than that of funding or politics, but rather is related to its institutional setting with its unclearly defined tasks overwhelming and unsettling the CAS and its scientists (*9*).

In August 2014, just 1 year after President Xi's visit, the CAS proposed the Pioneer Initiative to reform itself in another

unprecedented way. This time, it proposed to reorganize all its institutes into four categories: centers of excellence focused on basic research, innovation institutes doing applied research, big-science centers managing large-scale RIs, and specialized institutes that address problems specific to a region of China (*1*). Those institutes that did not fit into the restructuring would be merged, moved to the three universities under its jurisdiction, or dissolved. The emphasis on basic research, commercialization, and regional development is like putting old wine in a new bottle as they were the central themes of earlier reforms like the KIP, and doing so would not differentiate the CAS much from China's many other R&D organizations and universities. By contrast, centers of big-science RIs may bring real opportunities for the CAS to prove its uniqueness and strategic importance in China's S&T system.

BIG-SCIENCE INFRASTRUCTURE

"Big science" is a style of scientific research characterized by the usage of large-scale RIs by groups of scientists to address grand challenges of human society. Big-science RIs have become indispensable for scientific breakthroughs in physics, astronomy, materials science, biomedical science, and many other areas. For instance, with synchrotron radiation sources, imaging facilities, relativistic heavy-ion colliders, and free-electron lasers, the Brookhaven National Laboratory has spawned at least seven Nobel Prizes since 1947. Similarly, high-energy physics and cosmology discoveries have brought numerous Nobel Prizes to CERN, where some of the world's largest and most complex RIs are located. International collaboration has also been strengthened as the construction and operations of big-science RIs involve substantial resources and knowledge transfer from other countries.

Initially, the US, European Union, and Japan were key players in hosting big-science RIs. Major facilities include Stanford Linear Accelerator Center (SLAC) in the US, Large Hadron Collider (LHC) in CERN, and the SuperKEKB in KEK of Japan for research on particle physics, as well as major synchrotron light sources such as Advanced Photon Source (APS) in the US, Super Photon Ring-8 GeV (SPring-8) in Japan, and European Synchrotron Radiation Facility (ESRF) in France. Later on, some developing economies also started to invest in RIs. For instance, the Synchrotron-light for Experimental Science and Applications in the Middle East (SESAME) and Sirius, the new synchrotron light source in Brazil.

As a latecomer in scientific research, China began to lay out its own big-science RIs in the 1980s (*10*). In the 21st century, mainly through the Medium and Long-Term Plan for the Development of Science and Technology (2006–2020), the country further initiated a series of mega-programs that involved generous government investment in big-science RIs. Consequently, China is now home to a few of the world's largest RIs, such as the Five-hundred-meter Aperture Spherical Telescope and the Shanghai High Repetition-rate X-ray Free Electron Laser and Extreme Light Facility (SHINE) (see the table for a list of major big-science RIs in China).

REMAKING THE CAS

According to its self-assessment, by 2020, the CAS had accomplished the initial step of the Pioneer Initiative by restructuring most of its research units. From 2020 to 2030, it will continue reforming its personnel system and mobilizing resources across the four categories of institutes to become the country's main sources of innovative ideas, talents, and scientific achievements, as challenged by President Xi. Compared to the other three categories, concentrating resources on the big-science centers may bring more distinctive competitiveness to the CAS to meet challenges from top universities and other research institutes as well as industrial R&D endeavors.

This conclusion is derived from analysis of two critical attributes of the four categories of institutes: internal resourcefulness and external uniqueness. Resourcefulness refers to how many resources the CAS has already owned and can access, and it determines the performance potential of the reformed institutes. Uniqueness, taking an external perspective, considers how the reformed institutes can distinguish or stand out from other actors of the S&T system, and defines CAS's strategic role in China's S&T landscape.

The centers of excellence face fierce competition as the research capability and achievements of many universities and ministry-sponsored research institutes have

Major big-science research infrastructure in China

RESEARCH INFRASTRUCTURE NAME	YEAR OF CONSTRUCTION	CONSTRUCTION COST (IN MILLION USD)	RESPONSIBLE ORGANIZATION	LOCATION
Beijing Electron Positron Collider (BEPC)	1984	165.14	CAS	Beijing
Hefei Light Source	1984	55.32	CAS	Hefei
Shenguang-II High Power Laser	1994	N/A	CAS	Shanghai
Heavy Ion Research Facility (HIRFL)	2000	69.67	CAS	Lanzhou
Experimental Advanced Superconducting Tokamak (EAST)	2000	47.47	CAS	Hefei
Large Sky Area Multi-Object Fiber Spectroscopic Telescope (LAMOST)	2001	55.38	CAS	Beijing
Shanghai Synchrotron Radiation Facility (SSRF)	2004	271.34	CAS	Shanghai
China Spallation Neutron Source (CSNC)	2007	250.23	CAS	Dongguan
Daya Bay Reactor Neutrino Experiment	2007	N/A	CAS	Shenzhen
Steady High Magnetic Field Facility (SHMFF)	2008	49.21	CAS and University of Science and Technology	Hefei
Five-hundred-meters Aperture Spherical Telescope (FAST)	2009	N/A	CAS	Guizhou
National Center for Protein Science	2010	145.30	CAS	Shanghai
Soft X-ray Free Electron Laser	2011	36.42	CAS and Peking University	Shanghai
Dalian Coherent Light Source (DCLS)	2014	17.47	CAS	Dalian
Large High Altitude Air Shower Observatory (LHAASO)	2016	N/A	CAS	Daocheng
Earth System Numerical Simulation Facility (EarthLab)	2018	198.43	CAS and Tsinghua University	Beijing
Shanghai High Repetition-rate X-ray Free Electron Laser and Extreme Light Facility (SHINE)	2018	1581.11	ShanghaiTech University	Shanghai
High Energy Photon Source (HEPS)	2019	N/A	CAS	Beijing

The original costs were first converted to RMB as of 2022 to account for inflation and then converted from RMB to USD based on the currency market exchange rate in December 2022. Sources: https://lssf.cas.cn/ CAS, Chinese Academy of Sciences; N/A, no publicly available information of construction cost.

DATA: CHINESE ACADEMY OF SCIENCES

risen rapidly, with many research areas and programs often overlapping with those of the CAS. Attracting top talents and applying for basic research funding have already been difficult for the CAS. It is also uncertain whether the new centers of excellence can produce more scientific breakthroughs than universities or other research organizations without exhausting too many resources. The success of the innovation institutes depends on whether the CAS can resolve the discrepancy between the needs of its scientists and the industrial sector or whether it can spin off new ventures as it did before. Although scientists care more about being the first in their findings, integrating science with cost effectiveness is more vital to the commercialization of research and high-tech entrepreneurship. It is therefore questionable whether institutes in this category can outperform industrial R&D institutes where staff have more incentive and market knowledge for commercialization. Moreover, the specialized institutes focusing on regional issues may also be too fragmented for the CAS to generate a national impact. Overall, these three types of reformed institutes may not warrant their affiliation with the national academy as measured by their resourcefulness and uniqueness.

The big-science centers, however, satisfy both conditions. The CAS has built and administered about 80% of all big-science RIs currently operated in China (*11*), reflecting its distinctive feature and unmatched advantage. It has cumulated experience in building and managing these facilities, which are fully government financed. For instance, the construction of the High Energy Photon Source, the world's brightest fourth-generation synchrotron facility, is supported by special funding from the central government and Beijing municipal government; funding for constructing SHINE, China's most expensive RI, has come from the National Development and Reform Commission and Shanghai municipal government (*12*). Despite not having a substantial presence within the CAS, in 2019, 65% of the funding from the nine special basic science projects secured by the CAS from the National Key R&D Program, a program administered by the Ministry of Science and Technology, went to the big-science centers. Pivoting the CAS to these big-science infrastructure projects also provides flexibility for the academy to establish and extend its competitive advantage. By nature, big science's most ambitious projects—satellites and space probes, particle accelerators, and gravitational-wave observatory—rival those of universities and industrial organizations in their size and complexity.

Furthermore, many existing big-science RIs have already generated world-class scientific breakthroughs. Examples include the discovery of a four-quark matter from the Beijing Electron Positron Collider (*13*) and a new form of neutrino oscillations from the Daya Bay Reactor Neutrino Experiment (*14*), with the latter involving a multinational team with researchers from China, as well as Chile, the US, Russia, and the Czech Republic. The RIs have also facilitated technological innovations that meet the strategic demands of China's industrial and economic development. For instance, the development of the BeiDou Navigation Satellite System has benefited from the BPL/BPM Time Service System, a research facility that provides reliable high-precision time service to industrial developments in China since the 1970s. Some preliminary work suggests that the construction of the Shanghai Synchrotron Radiation Facility (SSRF) enhanced China's internal innovation capability (*15*). Investing in big-science centers may thus have the potential to meet China's imperative for scientific excellence and technological competitiveness.

Prioritizing these big-science centers also helps settle the recent discussion on the establishment of national laboratories. Rather than diverting and potentially wasting China's increasing but still scarce resources to emergent but untested organizations, turning part of the CAS into national laboratories may be most cost-effective and feasible. By doing so, the CAS may evolve into a national management agency responsible for the construction and operation of all big-science RIs in China. The strategic human capital and tacit knowledge that the CAS has accumulated will help it to become the leading big-science training center in China, providing professional training for research scientists and architectural engineers in disciplines related to large-scale RIs. Hosting big-science RIs at the CAS and turning it into national laboratories will reshape the academy's position in China's S&T system, avoiding disorderly competition with universities and industrial research institutes. Doing so also fulfills a major task of the reform of China's S&T system.

The CAS may continue to learn from its international peers regarding how to build and operate national laboratories around big-science RIs. Indeed, since 2017, the Chinese government has already been experimenting with a few national laboratories in Shanghai, Beijing, Hefei, and Guangzhou, with extensive involvement of the CAS. However, China has seen a recent rush to host new RIs by local governments, but not all of them have a good grasp of the necessity and advanced nature of the facilities. Given the huge costs, the construction of big-science RIs must anticipate the scientific needs of the future instead of merely considering the existing demands of a mature body of users. Hence, there is a need for coordination of efforts as well as for cultivation of a growing user community for the RIs.

The process of turning the CAS into national laboratories oriented toward administering and supervising China's big-science RIs is complicated and even radical and painful, especially given its storied history. It will be less a scientific and more a political, economic, and administrative decision. The CAS leadership may start considering shifting resources toward the big-science centers, within its "Pioneer Initiative" framework. Of course, we do not mean to suggest an immediate separation of other categories of institutes from the academy. However, by orienting resources toward big-science–related institutes, the CAS may be able to gradually downsize other categories of institutes and reduce its redundant or ineffective functions. The questions become whether the government would support and facilitate such a transformation after carefully weighing the costs and benefits and what the CAS should do to streamline the process, especially regarding the other parts of the institution. ■

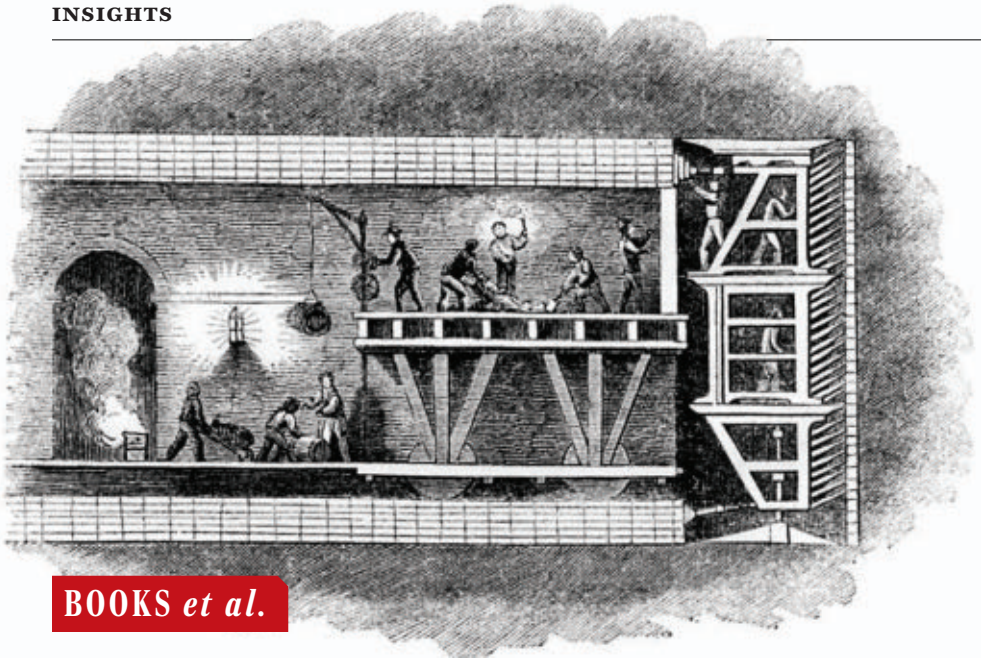
REFERENCES AND NOTES

1. D. Cyranoski, *Nature* **513**, 468 (2014).
2. J. Tollefson, *Nature* **553**, 390 (2018).
3. C. Cao, N. Li, X. Li, L. Liu, *Science* **341**, 460 (2013).
4. A. Regina, W. Kirby, W. McFarlan, *Harv. Bus. Rev.* **92**, 107 (2014).
5. D. E. Stokes, *Pasteur's Quadrant: Basic Science and Technological Innovation* (Brookings Institution Press, 1997).
6. R. P. Suttmeier, B. Shi, *Greater China's Quest for Innovation* (Shorenstein Asia-Pacific Research Center, Stanford, 2008), pp. 35–56.
7. R. P. Suttmeier, C. Cao, D. F. Simon, *Science* **312**, 58 (2006).
8. M. Poo, L. Wang, *Natl. Sci. Rev.* **1**, 618 (2014).
9. M. Cao *et al.*, *Int. J.* **506**, 120 (2021).
10. H. Chen, *Large Research Infrastructures Development in China: A Roadmap to 2050* (Springer, 2011).
11. Chinese Academy of Sciences, Introduction, https://english.cas.cn/about_us/introduction/201501/t20150114_135284.shtml (accessed 5 June 2022).
12. ShanghaiTech University, Shanghai-XFEL Beamline Project (SBP) and Shanghai Soft X-ray Free Electron Laser Facility (SXEFL) reach another milestone: First light and first data, https://www.shanghaitech.edu.cn/eng/2021/0626/c1260a66742/page.htm (accessed 5 June 2022).
13. E. Swanson, *Physics* (College Park Md.) **6**, 69 (2013).
14. F. P. An *et al.*, *Phys. Rev. Lett.* **108**, 171803 (2012).
15. X. Yang, X. Zhou, C. Cao, SSRN Working Paper (2022).

ACKNOWLEDGMENTS

We are grateful to G. Chen, H. Jiang, H. Li, N. Li, R. Li, Zhi Liu, Zhongkai Liu, Y. Sun, R. P. Suttmeier, H. Wang, and three anonymous reviewers for their comments. This research is supported in part by the National Natural Science Foundation of China (grants 71774091 and 72102145) and the Double First-Class Initiative Fund of ShanghaiTech University. The views expressed in the paper are those of the authors and do not necessarily represent those of their employers and the funder.

10.1126/science.add3428



Workers dig the Thames Tunnel in this engraving from the *Illustrated London News*.

BOOKS *et al.*

HISTORY OF SCIENCE

Making modern science

Victorian values permeate contemporary scientific culture, maintains a historian

By **Bernard Lightman**

In *How the Victorians Took Us to the Moon*, Iwan Rhys Morus has drawn on his vast knowledge of the history of the physical sciences in 19th-century Britain to depict the emergence of "new ways of thinking about and organising science that were directed at the future in a wholly new and unprecedented way, and some of the key consequences of that reorientation." Understanding this, Morus argues compellingly, helps us understand our current situation, for we moderns think about and organize science in the same way.



How the Victorians Took Us to the Moon
Iwan Rhys Morus
Pegasus, 2022. 400 pp.

Although Morus develops a complex argument throughout the book, his elegant and accessible writing style will appeal to a variety of audiences, including historians of science, scientists, and casual readers. His ability to synthesize recent scholarship to present a novel, coherent story is truly impressive.

Each of the book's chapters introduces the reader to a different group of scientists and engineers who were central to 19th-century discoveries in the physical sciences, including the transatlantic cable, the railways, the telegraph and telephone, calculating machines, and flying machines. In each case, Morus shows that development of the new science and technology led Victorians to believe that they were transforming the world and that, as a result, the future would be very different from the past. In other words, scientists and engineers were not just inventing new scientific theories and technologies, they were also "invent[ing] the future as we know it." To illuminate this point, Morus frequently references contemporaneous works of science fiction, a literary genre born in this period.

Morus's explanation for how the Victorians remade the future is tied to how science itself was reinvented in the early 19th century. The story therefore begins with a new breed of experts, led by John Herschel, William Whewell, and Charles Babbage, who came on the scene in the 1820s. These individuals attempted (unsuccessfully, at first) to reform the old Royal Society and later, in 1831, created the British Association for the Advancement of Science. Through it, they shaped the era's science.

Victorians considered self-discipline to be a masculine virtue, and their approach to science—then considered an exclusively male pursuit—reflected this. Even a man who came from a humble working-class background, like the natural philosopher Michael Faraday, could become an exemplary man of science, they believed, because men knew how to discipline themselves (unlike women, who were ever "at the mercy of their uncontrollable bodies").

Morus also draws a connection between this image of the disciplined scientist and the importance of showmanship to the business of invention. He pays considerable attention to scientific and industrial exhibitions such as the Adelaide Gallery, the Great Exhibition of 1851, and later exhibitions and museums in which the inventor was as much on show as their invention.

Men of invention, such as Isambard Kingdom Brunel, John Henry Pepper, and Thomas Edison, were held up as powerful, charismatic, and heroic figures during this era. As Morus points out, however, this image of the scientific loner was an illusion. Scientific discovery and invention were often the product of the efforts of a multitude of expert workers.

Another strand of Morus's argument brings in the imperial dimensions of 19th-century science. "All the technological innovations—actual and imagined—described in this book had imperial entanglements," he declares. By this, he means that they were dependent on the resources of an imperial power and that they themselves provided ways of maintaining and expanding the reach of empire.

Morus has given the history of 19th-century science something that it has desperately needed: a new big picture that integrates the insights of recent scholarship. Of course, there are gaps. His emphasis on the physical sciences excludes, for example, more engagement with recent scholarship on the history of the life sciences, including the theory of evolution. And despite the discussion of imperialism, Morus spends most of his time analyzing British scientists and engineers. But one book can only do so much.

Morus ends with a sobering thought: Unreflectingly following the Victorian recipe for future-making sets limitations on how we might remake our own futures. There is no escape from the fact that the science that governs our lives is the product of an imperial, masculine culture. If, Morus insists, we want "to change science and its culture for the better, we need to start by remembering this history." ■

The reviewer is at the Department of Humanities, York University, Toronto, ON M3J 1P3, Canada.
Email: lightman@yorku.ca

10.1126/science.adf4691

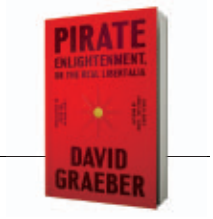
CREDIT: ENGRAVING FROM THE *ILLUSTRATED LONDON NEWS*

ANTHROPOLOGY

Pirates and politics

An anthropologist argues that experimental communities in Madagascar influenced the European Enlightenment

Pirate Enlightenment, or the Real Libertalia
David Graeber
Farrar, Straus and Giroux, 2023. 208 pp.



By **Jatin Dua**

In 1695, as it made its way through the Red Sea, the *Ganj-i-Sawai,* a trading vessel belonging to the Mughal emperor Aurangzeb, was attacked and ransacked by a ragtag group led by Englishman Henry Avery, the notorious "king of Pirates." Although European pirates had been crucial to the making of the British Empire (and other European empires), the British sought to distance themselves from Avery and his comrades, launching a worldwide manhunt against him, the first such global operation in recorded history. Avery, however, eluded capture and was rumored to have sailed to Madagascar, where he found a home amid the burgeoning pirate communities on the northeast coast.

Avery's high jinks, along with other daring stories of piracy on the high seas, captured the public imagination in England and throughout Europe. Through fictional and fictionalized accounts and biographies, a proliferation of pirate stories—including many that featured a legendary pirate colony in Madagascar known as Libertalia—circulated throughout Europe, emphasizing the fantasy and romance of piracy, a lore that lives on in contemporary renderings.

In his final posthumously published book, *Pirate Enlightenment, or the Real Libertalia*, anthropologist and political activist David Graeber revisits this historical moment, alongside his earliest ethnographic fieldwork. This elegantly breezy treatise takes readers on a journey to the monsoonal waters of the Indian Ocean and the verdant landscape of Madagascar.

Piracy, as Graeber compellingly notes, is not just myth and romance but also an actual political experiment in forms of social organization and living that are nonhierarchical and antiauthoritarian. Divided into three parts, *Pirate Enlightenment* draws from a number of contemporaneous accounts of pirate exploits, colonial archives, historiography on Madagascar, and Graeber's own fieldwork in the region to create a vivid narrative that seeks to move beyond Eurocentric understandings of the expansion of piracy from the Atlantic to the Indian Ocean in the 17th and 18th centuries.

European merchants and colonists were part of a long history of encounters that shaped Malagasy social life. However, unlike other travelers who formed complex political alliances or incorporated themselves within local society through marriage, European settlers frequently remained aloof or used violence as a mode of establishing presence in Madagascar, reveals Graeber. The pirates who arrived on the northeast coast were far more amenable to incorporating themselves into local systems, he maintains, and were open to creating new forms of political, economic, and social organization. While acknowledging the "meager" and often "sensationalistic" nature of sources, Graeber cites examples drawn from contemporaneous popular accounts where newly settled pirates played important roles as advisers and mediators in establishing new settlements and polities infused with an egalitarian ethos. Specifically, Graeber notes that pirates successfully "convert[ed] the democratic institutions first developed on board ships into forms that would be viable on land."



Saint Mary's Island, off the coast of Madagascar, was once a pirate haven.

In the second part of the book, Graeber switches vantage points and examines the arrival of pirates from the lens of Malagasy society. Here, he includes a fascinating discussion on the agency and importance of Malagasy women traders in creating alternative loci of political, economic, and social authority.

Graeber concludes by expanding on the provocation that Madagascar's pirate colonies can be understood as an Enlightenment political experiment that resonated far beyond the region and shaped emerging notions of liberty, freedom, and egalitarianism in Europe at the turn of the 18th century. Absent a direct line of connection between these experimental communities and figures such as Montesquieu, Locke, and Hobbes, Graeber bases his thesis on both the physical presence of numerous "pirate envoys" at European courts and the proliferation of pirate tales across Europe, arguing that pirates were a major topic of conversation and "the Enlightenment was an intellectual movement uniquely tied to conversational forms."

Pirate Enlightenment pluralizes and globalizes our understanding of whose ideas and actions are considered impactful and whose vision shapes the world, a framing that still resonates in contemporary times. Many of today's self-proclaimed pirates, for example, would claim that they—as coastal villagers, fishermen, and traders—are legitimate protectors of the sea. Piracy was and is many things, including a demand to be taken seriously as political actors. Yet global hierarchies continue to divide up worlds into the West and the rest.

In his academic writing and political commitments, David Graeber exemplified an ethos of action and conversation. There is a certain bittersweetness to this text, one that ends with an exhortation toward the arts of speaking and conversation. Graeber himself is no longer around to speak, to debate, or to inspire protest and action.

As anthropologists have noted, gifts are inalienable—they contain within them something of the giver. Graeber's final book is certainly such a gift. ■

The reviewer is at the Department of Anthropology, University of Michigan, Ann Arbor, MI 48109, USA. Email: jdua@umich.edu

10.1126/science.adf6028

PHOTO: JAVARMAN3/ISTOCKPHOTO

LETTERS

Edited by **Jennifer Sills**

Avian flu threatens Neotropical birds

A highly pathogenic avian influenza virus (HPAIv) that spread through the Holarctic region in 2022, affecting millions of birds (*1*), has reached South America. The disease threatens marine and terrestrial birds, including endangered species. Real-time information gathering and steps to prevent the spread of disease will be required to mitigate this outbreak.

Along the Peruvian coastline alone, HPAIv killed more than 22,000 wild birds in just 4 weeks in 2022, mainly Peruvian pelicans (*Pelecanus thagus*) and Peruvian boobies (*Sula variegata*) (*2*), both of which are categorized as Endangered in Peru (*3*). The extant Peruvian population of pelicans in coastal protected areas is around 124,000 (*4*). The virus has also killed Near Threatened Guanay cormorants (*Leucocarbo bougainvilliorum*) (*2*, *3*). Beyond their conservation status and intrinsic value, these species are of economic interest. The guano produced by the Peruvian pelican, the Peruvian booby, and the Guanay cormorant is used as fertilizer and constitutes one of the most valuable economic resources in Peru (*5*).

HPAIv is spreading across the continent. By December 2022, the virus had been found in birds in Ecuador, Colombia, Venezuela, and Chile (*6*). In addition to marine birds, the virus could threaten scavengers, as it has in other regions of the world (*7*), including the emblematic Andean condor (*Vultur gryphus*) (*8*). Viral transmission could occur if those scavengers consume infected carcasses while traveling from the Andean region to the coastline (*9*).

We urge Peruvian and other South American authorities and conservation managers to take steps to track and mitigate the spread of this disease. Up-to-date information about the wild bird species affected and the number of birds remaining in the region will be necessary to understand the extent of HPAIv's effects. Managers should perform active and passive epidemiological surveillance and reduce the availability of infected carcasses in the environment to prevent transmission to marine birds, obligate scavenger species, and humans. Given the difficulty of controlling emerging pathogens, reducing anthropogenic threats such as environmental pollution, resource overexploitation, and habitat loss should also be an urgent priority. Minimizing other risks will give bird populations a better chance against emerging pathogens. Finally, we recommend continuing these efforts even if this flu event passes; the epidemiological behavior of this virus indicates that recurrent outbreaks are likely (*10*).



An outreak of highly pathogenic avian flu threatens the Peruvian pelican (*Pelecanus thagus*) population.

PHOTO: VÍCTOR GAMARRA-TOLEDO

Víctor Gamarra-Toledo[1,2]*, Pablo I. Plaza[2], Roberto Gutiérrez[1,3], Paola Luyo[3], Lizbeth Hernani[3], Fernando Angulo[4], Sergio A. Lambertucci[2]
[1]Museo de Historia Natural, Universidad Nacional de San Agustín de Arequipa, Arequipa, Peru. [2]Grupo de Investigaciones en Biología de la Conservación, Instituto de Investigaciones en Biodiversidad y Medio ambiente (Consejo Nacional de Investigaciones Científicas y Técnicas–Universidad Nacional del Comahue), Bariloche, Argentina. [3]Servicio Nacional de Áreas Naturales Protegidas por el Estado, Lima, Peru. [4]Centro de Ornitología y Biodiversidad, Lima, Peru.
*Corresponding author.
Email: vgamarrat@unsa.edu.pe

REFERENCES AND NOTES

1. M. Wille, I. G. Barr, *Science* **376**, 459 (2022).
2. Ministerio de Salud, "Alerta epidemiológica: Epizootia de influenza aviar, tipo A, subtipo H5 en aves silvestres y aves de traspatio en el país. Código: AE 029 - 2022" (2022); https://www.dge.gob.pe/epipublic/uploads/alertas/alertas_202229_07_152617.pdf [in Spanish].
3. Ministerio de Agricultura y Riego, "Decreto Supremo N° 004-2014-MINAGRI," *El Peruano* (2014), pp. 520497–520504 [in Spanish].
4. Agrorural, "Reporte Mensual de conservación en Islas y Puntas Guaneras" (2021); https://www.gob.pe/institucion/agrorural/informes-publicaciones/2313145-reporte-mensual-de-conservacion-en-islas-y-puntas-guaneras-octubre-2021 [in Spanish].
5. D. Plazas-Jiménez, M. V. Cianciaruso, *Trends Ecol. Evol.* **35**, 757 (2020).
6. Pan American Health Organization, "Epidemiological update: Outbreaks of avian influenza and public health implications in the region of the Americas" (2022); https://www.paho.org/en/documents/epidemiological-update-outbreaks-avian-influenza-and-public-health-implications-region.
7. M. F. Ducatez *et al.*, *Emerg. Infect. Dis.* **13**, 611 (2007).
8. P. I. Plaza *et al.*, *Ibis* **162**, 1109 (2020).
9. V. Gamarra-Toledo *et al.*, *Environ. Pollut.* **317**, 120742 (2023).
10. World Animal Health Information System, "Reports: Animal disease events" (2023); https://wahis.woah.org/#/event-management.

10.1126/science.adg2271

Atmospheric goals for sustainable development

Atmospheric chemistry and composition underlie the existential threats of climate change and ozone depletion (*1*), and poor air quality represents the greatest environmental health issue in the modern world (*2*). However, the only targets and indicators related to atmospheric health and clean air in the United Nations Sustainable Development Goals (SDGs) focus on reducing fine particulate pollution levels and the associated mortality rates (*3*, *4*). By creating targets for a broad range of specific pollutants, the SDGs could facilitate more effective actions, monitoring, and funding.

The SDGs consider air quality only in terms of the mass density of particulate matter of 2.5 μm or less (PM2.5). However, PM2.5 is not a single pollutant but rather a measurement that contains many different compounds, including sulfate, black carbon, metals, and organic compounds of all particle sizes below 2.5 μm. Measuring PM2.5 does not provide specific information about the concentrations of the components that make up PM2.5, nor does it account for the impact of other air pollutants known to be harmful, such as ozone and nitrogen dioxide.

Given that we still cannot fully explain the cause of the statistical link between PM2.5 and adverse health outcomes (*5*, *6*), this measurement alone does not provide adequate information. The SDGs would be more valuable if they also included targets to reduce deaths from specific pollutants. Particle number concentration and chemical composition may be more important than fine particulate matter mass density for health outcomes (*7*, *8*). Moreover, instead of tracking excess mortality alone, targets should include nonfatal health impacts, as well as impacts on ecosystems and agriculture.

Metrics for reduction of other atmospheric pollutants could be added to future developments of the SDGs. For example, in SDG 10 (Reduced Inequalities), reducing pollutants could contribute to access to clean air to breathe, a major contributor to health inequalities (*9*). Targets that address stratospheric ozone chemistry could be added to SDG 15 (Life on Land), given that life is only possible because of a healthy atmosphere, including the ozone layer (*10*). Inclusion of targets for tropospheric ozone would increase the pressure to control atmospheric pollution holistically, which—if ozone levels were successfully reduced—would also increase agricultural yields (SDG 2, Zero Hunger) (*11*). SDG 13 (Climate Action) could improve transparency by providing explicit targets for the peak concentrations of all major climate forcers, including greenhouse gases and particulates, which would help to track our progress in meeting global climate change commitments.

Updating the SDGs to incorporate atmospheric targets would improve data, policies, and incentives. Better regional policies on air pollution control and monitoring would also be supported in accordance with the broader guidance. In addition, a more holistic strategy would provide incentives for funders to support scientists researching pollutants that are not categorized as PM2.5.

Melita Keywood[1]*, Clare Paton-Walsh[2], Mark Lawrence[3], Christian George[4], Paola Formenti[5], Roybn Schofield[6], Helen Cleugh[1,7], Nathan Borgford-Parnell[8], Anthony Capon[9]
[1]Commonwealth Scientific and Industrial Research Organisation Environment, Melbourne, Australia. [2]University of Wollongong, Centre Atmospheric Chemistry, Wollongong, Australia. [3]Research Institute for Sustainability, Potsdam, Germany [4]University of Lyon, Université Claude Bernard Lyon 1, Centre National de la Recherche Scientifique (CNRS), Lyon, France. [5]Laboratoire Interuniversitaire des Systèmes Atmosphériques, Université Paris Cité Paris and Université Paris Est Creteil, CNRS, Paris, France. [6]School of Geography, Earth, and Atmospheric Sciences, University of Melbourne, Melbourne, Australia. [7]Australian National University, Canberra, Australia. [8]United Nations Environment Programme, Climate & Clean Air Coalition, Paris, France. [9]Monash University, Monash Sustainable Development Institute, Melbourne, VIC, Australia.
*Corresponding author.
Email: melita.keywood@csiro.au

REFERENCES AND NOTES

1. "Climate change 2021: The physical science basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change," V. Masson-Delmotte *et al.*, Eds. (2021).
2. J. Lelieveld *et al.*, *Nature* **525**, 367 (2015).
3. M. Elder, E. Zusman, "Strengthening the linkages between air pollution and the Sustainable Development Goals," Institute for Global Environmental Strategies, Policy Brief 35 (2016).
4. B. Lode *et al.*, *Rev. Eur. Compar. Int. Environ. Law* **25**, 27 (2016).
5. L. A. Cox Jr., *Risk Anal.* **33**, 2111 (2013).
6. M. Franklin *et al.*, *Epidemiology* **19**, 680 (2008).
7. M. Strak *et al.*, *Environ. Health Perspect.* **120**, 1183 (2012).
8. F. R. Cassee *et al.*, *Inhalation Toxicol.* **25**, 802 (2013).
9. A. Hajat *et al.*, *Curr. Environ. Health Rep.* **2**, 440 (2015).
10. J. C. Van der Leun, *Photodermatol. Photoimmunol. Photomed.* **20**, 159 (2004).
11. R. Van Dingenen *et al.*, *Atmos. Environ.* **43**, 604 (2009).

COMPETING INTERESTS

M.K. is the president of the International Commission for Atmospheric Chemistry and Global Pollution. C.P.-W. is co-chair of the International Global Atmospheric Chemistry Project. M.L. is co-chair of the German National Science Platform for Sustainability in the Agenda 2030. R.S. is president of the Australian Council of Environmental Deans and Directors and a board member of the Australian Community Climate and Earth System Simulator–National Research Infrastructure. H.C. is the vice chair of the World Climate Research Program Joint Scientific Committee.

10.1126/science.adg2495

Evidence required for ethical social science

The history of science is littered with violations of the human rights of research participants (*1*, *2*). One response has been regulation. In the United States, the Belmont Report lays out ethical principles for human subjects research (*3*), which form the foundation of modern-day institutional review boards. According to the Belmont Report, ethics committees and investigators need to determine whether the benefits of the research outweigh the costs and whether the costs unevenly burden particular groups. However, the assumptions underlying such assessments can be flawed (*4*, *5*). Without evidence about the causal effects of social science research on participant welfare, researchers risk causing unexpected harms or missing unexpected benefits.

Randomized experiments can provide data to replace assumptions about the potential benefits and costs of research on study participants. By randomly assigning some people to participate in a study and others not to participate, researchers can estimate the causal effects of participation on welfare (*6*). Researchers can also randomly assign and compare multiple versions of a protocol to learn about the marginal costs and benefits to participants, or they could randomize the sequence of a study component and participant welfare questions (*7*). Finally, researchers can add participant welfare questions to the end of an existing experimental study to understand the costs and benefits of a treatment. In each case, the experiments should be designed to assess whether costs and benefits are distributed equally among different subsets of participants.

Research ethics has come a long way since the Belmont Report, but researchers who work with human subjects still lack the causal evidence to make fully informed ethical choices. Researchers know they can do better, as calls for ethics appendices or ethics sections in preregistrations illustrate (*8*, *9*). To facilitate the adoption of these strategies, funders could require more transparency on the ethics of the work they support. The experimental designs we propose can be embedded into ongoing research efforts, adding little expense. Researchers can also coordinate efforts by using open access forums to post study instruments, data, and results, and when appropriate can implement harmonized experiments to build generalizable evidence across cases (*10*). Randomized experiments that assess the effects of participating in research can augment existing practices and build credible evidence to inform future risk-benefit calculations and improve ethical decision making.

Rebecca Littman[1], Rebecca Wolfe[2]*, Graeme Blair[3], Sarah Ryan[4]
[1]Department of Psychology, University of Illinois-Chicago, Chicago, IL, USA. [2]Harris School of Public Policy, University of Chicago, Chicago, IL, USA. [3]Department of Political Science, University of California Los Angeles, Los Angeles, CA, USA. [4]Department of Psychology, University of California Berkeley, Berkeley, CA, USA.
*Corresponding author.
Email: rebeccawolfe@uchicago.edu

REFERENCES AND NOTES

1. A. M. Brandt, "Racism and research: The case of the Tuskegee Syphilis Study," Hastings Center Report (1978), pp. 21–29.
2. V. Spitz, *Doctors from Hell: The Horrific Account of Nazi Experiments on Humans* (Sentient Publications, 2005).
3. US National Commission for the Protection of Human Subjects of Biomedical, and Behavioral Research, "The Belmont report: Ethical principles and guidelines for the protection of human subjects of research, vol. 2" (Department of Health, Education, and Welfare, 1978).
4. D. Kahneman, O. Sibony, C. R. Sunstein, *Noise: A Flaw in Human Judgment* (Little, Brown, 2021).
5. T. Eyal *et al.*, *J. Person. Soc. Psych.* **114**, 547 (2018).
6. J. J. Muehlenkamp *et al.*, *Clin. Psych. Sci.* **3**, 26 (2015).
7. C. Grady *et al.*, *PLOS One* **26**, e0172607 (2017).
8. E. Asiedu, D. Karlan, M. Lambo Quayefio, C. Udry, *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2024570118 (2021).
9. H. Baron, L. E. Young, *Polit. Sci. Res. Meth.* **10**, 1 (2021).
10. G. Blair, G. McClendon, in *Advances in Experimental Political Science* (Cambridge University Press, 2021), ch. 22.

10.1226/science.adf8329

RESEARCH

IN *SCIENCE* JOURNALS

Edited by **Michael Funk**



BEHAVIORAL ADAPTATION

Risky payoffs

In nature, environmental fluctuation is common, and some years are good and some are bad for species trying to survive and reproduce. Some animals take bets on what the upcoming environment might be like, and different strategies can affect overall fitness. Petrullo *et al.* studied different strategies in a population of red squirrels that has been monitored for decades. They found that mother squirrels that bet on a good environment, and produced more offspring, had higher overall fitness than mothers who bet that the environment would be poor. This was the case even when the optimistic mothers were wrong. —SNV *Science*, abn0665, this issue p. 269

Over a lifetime, American red squirrel (*Tamiasciurus dsonicus*) mothers successfully raise more young when they assume that their environment will be favorable.

PLANT SCIENCE

Parsing nodulation pathways

Legumes benefit from symbiotic microbes resident in root nodules that fix nitrogen, whereas most other plants depend on externally supplied or gathered nitrogen. Rübsam *et al*. identified a bifunctional receptor complex in the model legume *Lotus japonicus* that initiates development and infection of the root nodules. Development of root nodules could be driven solely by the intracellular domains containing the kinases once the receptor complex was formed, but the Nod factor–recognizing ectodomains were required to support rhizobial infection. Receptors identified in the cereal barley, which does not form symbiotic nitrogen-fixing root nodules, were able to support *Lotus* root nodule organogenesis, suggesting that the path to nitrogen fixation in cereals might be shorter than we thought. —PJH

Science, ade9204, this issue p. 272

SOLAR CELLS

Two additives for perovskite modules

The performance of large-area perovskite solar cells and modules fabricated by the vacuum flash solution-processing method has been enhanced with two different additives. You *et al*. synthesized a fullerene derivative with a phosphonate group that accelerated electron transfer across grain boundaries in the perovskite film and suppressed ion migration. They also prepared a polypropylene oxide polymer substituted with a redox active species that could facilitate p-doping of the hole transport layer quickly, unlike the conventional air-doping approach. Large-area modules (17 square centimeters) had power-conversion efficiencies of up to 21.4%. The modules retained 95% of their efficiency after more than 3000 hours of continuous illumination at elevated temperature. —PDS

Science, add8786, this issue p. 288

QUANTUM SIMULATION

A hybrid platform for quantum simulation

Quantum simulators are typically constructed from a set of quantum particles that are controllably placed on a lattice and then allowed to interact with each other, but there are limitations. Simulators based on neutral atoms lack the flexibility to independently control and read out single atoms, trapped-ion based quantum systems are difficult to scale beyond tens of ions, and superconducting quantum circuits are limited to local interactions between qubits. Zhang *et al*. constructed a many-body quantum simulator by interfacing superconducting qubits with

PHOTO: SUMIO HARADA/MINDEN PICTURES

a microwave photonic band-gap metamaterial waveguide. This hybrid superconducting qubit-metamaterial approach represents a route toward developing a large-scale quantum simulator platform, extending the lattice to two dimensions and hosting a larger number of quantum particles. —ISO

Science, ade7651, this issue p. 278

OPTICS

For more information, make some noise

Holograms can be considered large-capacity memory storage media that are capable of holding information about three-dimensional scenes and compressing it into a two-dimensional pattern (metasurface). Attempting to store more images onto the one hologram pattern tends to result in cross talk and corruption of the stored data. Xiong *et al*. found that introducing engineered noise into the process enabled an increase in storage capacity. This approach should work for metasurfaces in applications such as high-capacity optical displays, information encryption, and data storage. —ISO

Science, ade5140, this issue p. 294

STELLAR ASTROPHYSICS

A simulated dynamo in radiative stars

Strong stellar magnetic fields are generated by dynamos, which typically require convection of material inside the star. Petitdemange *et al.* used numerical simulations to show how a stellar dynamo can form in layers that are purely radiative (meaning that they have no convection). Their results matched most of the properties of a theoretical prediction known as the Tayler-Spruit dynamo. The resulting magnetic field remains trapped inside the star, so it would not be observable on the surface but could potentially be inferred using asteroseismology. Because magnetic fields can transport angular momentum, the mechanism slows the spin of stellar cores and increases the rotation of the outer layers, enhancing the mixing of chemical elements. —KTS

Science, abk2169, this issue p. 300

CANCER

SCD-ing into home for GBMs

Glioblastomas (GBMs) are highly enriched with GBM stem–like cells (GSCs) that cause resistance to standard therapy. These cells rely on deregulated de novo lipid synthesis, which represents a potential target for treatment; however, known inhibitors are not brain penetrant and susceptibility is unknown. Eyme *et al*. found that the stearoyl CoA desaturase inhibitor YTX-7739 was able to trigger lipotoxicity in patient-derived GSCs and showed therapeutic efficacy in mouse models. Mechanistically aberrant MEK/ERK signaling endowed sensitivity to this enzyme inhibitor, whereas AMPK activation rendered the GSCs resistant. —DLH

Sci. Transl. Med. **15**, eabq6288 (2023).

PHYSIOLOGY

A hepatokine elevates lipolysis

Increased circulating triglyceride levels are a risk factor for the development of cardiovascular and fatty liver diseases. Kim *et al.* identified a hepatokine called CREBH-C that stimulates the lipolysis and clearance of circulating triglycerides. CREBH-C is secreted from the liver in response to increased energy demands (such as those induced by fasting) or hepatic stress (such as that induced by obesity or a high-fat diet) in mice and humans. CREBH-C treatment of mice fed a high-fat diet reduced circulating triglyceride amounts and increased triglyceride uptake by nonhepatic tissues. —WW

Sci. Signal. **16**, eadd6702 (2023).

IN OTHER JOURNALS

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



NEUROSCIENCE

Dynamic cortical Down states

When asleep, under anesthesia, or in some damaged states, the brain exhibits slow oscillations. These alternate between active periods with neuronal firing called Up states and quiescent periods associated with unconsciousness called Down states. The mechanisms underlying alternations between Up and Down states are still not fully understood. Caramassa *et al.* analyzed Up and Down state transitions and found that they contain two complementary periods that determine relative responsiveness: A highly synchronized period after the Up-to-Down transition with a time course that is defined by the exponential decay of the phase and another stochastic period. The work offers some insight into how the brain slips in and out of unconsciousness. —PRS *J. Neurosci.* **42**, 9387 (2022).

SIGNAL TRANSDUCTION

Designer anti-inflammatory receptors

The modular nature of protein interactions that mediate cellular signal transduction offers the opportunity to design therapeutics. In the human body, hundreds of billions of cells are estimated to die each day, and they are mostly disposed of by a process called efferocytosis. If efferocytosis fails, then damaging inflammation can ensue. Morioka *et al*. constructed receptors that combined a sensor domain recognizing phosphatidylserine on the surface of dead cells with an adapter domain from a cytoplasmic protein called ELMO that helps to propagate signals for efferocytosis. Expression of such chimeras in

PHOTO: YIPENGGE/ISTOCK PHOTO



BIOSENSORS

Spying on the microbiome

Gut microbiota play important roles in animal digestion and immune response, and many microbes are capable of metabolic transformations not performed by the host. Vaaben *et al.* generated and characterized eight biosensor plasmids that, when transformed into a suitable bacterial host, could report on the presence of oxygen, by-products of metabolism, and host-derived molecules in vitro and in a model worm. This collection of sensors could in theory be adapted to report on multiple stimuli at once to build a more complete picture of the gut environment in model animals. —MAF *ACS Synth. Biol.* **11**, 4184 (2022).

Bioengineered biosensors can detect changes in lactate concentrations in the gut lumen of individual *Caenorhabditis elegans* nematodes (shown in this false-color image).

phagocytic cells enhanced efferocytosis and was protective in mouse models of inflammatory diseases. —LBR

Cell **185**, 4887 (2022).

ABORTION

Risk of suicide

Abortion access remains a divisive issue in the United States, and its media coverage can affect women's mental health. Although abortion legislation has been linked to women's anxiety and depression rates, suicide risk has not been addressed. Zandberg *et al.* examined enforcement of state-level restrictions on reproductive care, including abortions, with state-level suicide rates between 1974 and 2016. The authors used the Targeted Regulation of Abortion Providers (TRAP) laws index, a measure that increases when a TRAP law is enforced and decreases when the law is blocked. Their analyses compared the impact of TRAP laws among reproductive-aged women (20 to 34 years old) with a control group of post-reproductive-age women (45 to 64 years old). Enforcement of TRAP laws increased the annual suicide rate by 5.81% among the younger, reproductive-aged women. The findings have implications for suicide prevention interventions for younger women in light of the recent reversal of legislation that made access to an abortion a federal right in the United States. —EEU

JAMA Psychiatry 10.1001/jamapsychiatry.2022.4394 (2022).

COSMOLOGY

Supernovae support a cosmological constant

Type Ia supernovae (SN Ias) can be used as standard candles, astronomical objects with intrinsic brightness that can be determined from observations, allowing their distance to be inferred independently of their redshift. Analysis of SN Ias can therefore be used to constrain cosmological models, which predict the redshift at a given distance. Brout *et al.* performed a joint analysis of more than 1500 SN Ias, spanning redshifts from 0.001 to 2.3, and measured cosmological parameters with greater precision than previous supernova studies. Their results agree with standard cosmology and constrain the dark energy equation of state, finding that it is consistent with a cosmological constant. —KTS

Astrophys. J. 10.3847/1538-4357/ac8e04 (2022).

METABOLISM

A new player in obesity

Drugs that target serotonin have broad uses in treating neuropsychiatric disorders, but there can be limiting side effects such as increased hunger and weight gain. The appetite-suppressing actions of serotonin are mediated through the 5-hydroxytryptamine receptor 2C ($5HT_{2C}R$) expressed by hypothalamic proopiomelanocortin neurons. He *et al*. performed whole-exome sequencing in 2548 people with severe obesity and found rare loss-of-function variants in the *HTR2C* gene, which encodes $5HT_{2C}R$. Mice expressing one of the *HTR2C* variants exhibited increased appetite and developed obesity when fed a high-fat diet despite increased physical activity. Thus, *HTR2C* variants likely contribute to some cases of severe childhood-onset obesity, and $5HT_{2C}R$ signaling may offer a therapeutic target. —GKA

Nat. Med. **28**, 2537 (2022).

FULLERENE CHEMISTRY

Decorating molecular soccer balls

The extraordinary finale of last year's World Cup treated fans to the marvelous ball-handling skills of Messi and Mbappé. The molecular realm also features a soccer ball, the icosahedral C_{60} fullerene, and chemists too strive to manipulate it skillfully. Lu *et al.* report chiral pyramidal capsules that can trap fullerene inside and adhere substituents to it selectively in just one of two mirror-image orientations. Specifically, the C_{60} reacts with anthracene groups on the capsule faces through Diels-Alder cycloaddition, after which the capsule can be disassembled by imine exchange to liberate the decorated ball. —JSY

Nat. Chem. 10.1038/s41557-022-01103-y (2021).

PHOTO: HEITI PAVES/GETTY IMAGES

ALSO IN *SCIENCE* JOURNALS

Edited by **Michael Funk**

LYMPHOMA

How lymphomas outcompete

Lymphomas are cancers of the immune system that arise from B cells undergoing a strict natural selection process required for immunity. These highly mutating and dividing B cells vigorously compete against each other for T cell help to survive. Mutations affecting *B cell translocation gene 1* (*BTG1*) are exclusive to B cell lymphoma and associated with poor clinical outcomes. Mlynarczyk *et al*. found that mutant *BTG1* effects were limited to conferring B cells with only subtle acceleration of their T cell help response. This effect occurred at the checkpoint that governs natural selection of B cells, so these cells became "supercompetitors" that outpaced and replaced their normal counterparts. This behavior mirrors embryonic-specific supercompetition processes, pointing to BTG1 as an evolutionary "gatekeeper" of natural selection during the adaptive immune response. —SMH and PNK

Science, abj7412, this issue p. 252

CRISPR

A decade of CRISPR

In the decade since the publication of CRISPR-Cas9 as a genome-editing technology, the CRISPR toolbox and its applications have profoundly changed basic and applied biological research. Wang and Doudna now review the origins and utility of CRISPR-based genome editing, the successes and current limitations of the technology, and where innovation and engineering are needed. The authors describe important advances in the development of CRISPR genome-editing technology and make predictions about where the field is headed. They also highlight specific examples in medicine and agriculture that show how CRISPR is already affecting society, with exciting opportunities for the future. —DJ

Science, add8643, this issue p. 251

CANCER GENOMICS

Sarcoma genes in the limelight

Sarcomas are cancers of muscles, bones, and other connective tissues that tend to develop in younger patients and are often aggressive and difficult to treat. Because of their relative rarity, the biology of sarcomas is not nearly as well understood as that of more common cancers. In a massive, multinational genomic study involving thousands of patients, their families, and controls, Ballinger *et al*. identified distinct biological pathways where mutations increase the inherited risk for developing sarcoma through alterations of telomere biology and mitotic function (see the Perspective by Mandelker and Ladanyi). More work is needed before these findings can lead to therapeutic advances, but the study provides much needed biological insight into a deadly disease. —YN

Science, abj4784, this issue p. 253; see also adf8572, p. 238

ORGANIC CHEMISTRY

Synthesis by strain

Chemists often rely on strained intermediates to drive reactions that relieve the strain. In this context, however, cyclic allenes that cram adjacent double bonds into tight carbon rings have been underexploited. Ippoliti *et al*. now report a synthetic route to lissodendoric acid A, a marine natural product that relies on transient generation of a cyclic allene to prepare the fused ring core (see the Perspective by Jankovic and West). They accessed a single mirror-image form of the allene from a chiral precursor by fluoride attack on a silicon substituent and displacement of an adjacent bromide, setting up a stereospecific Diels-Alder cycloaddition with a pyrone reagent. —JSY

Science, ade0032, this issue p. 261; see also ade7122, p. 237

ZEOLITES

Condensing chains into extra-large pores

Zeolites with extra-large pores could absorb and process larger molecules, but strategies for synthesizing these materials are limited. Li *et al*. show that a silicate precursor with one-dimensional chains can undergo condensation reactions upon heating to remove the organic template to form Si-O-Si bridges that maintain the crystal topology (see the Perspective by Morris). The resulting low-density, pure silica zeolite has 14- and 16-membered silicate rings and highly thermal and hydrothermal stability. —PDS

Science, ade1771, this issue p. 283; see also adf3961, p. 236

IMMUNOTHERAPY

Activating the STING in tumors

Agents that activate the DNA-sensing cGAS-STING pathway have been explored for cancer immunotherapy in both preclinical and clinical settings, but how to optimally activate the pathway remains unclear. Jneid *et al.* found that incorporation of the naturally occurring STING ligand cyclic guanosine monophosphate-adenosine monophosphate (cGAMP) into noninfectious enveloped virus-like particles (VLPs) selectively activated STING in antigen-presenting cells, including dendritic cells, when administered intratumorally. Compared with a nontargeted synthetic STING agonist, VLPs delivering cGAMP systemically enhanced tumor-specific T cell responses and antitumor effects during immune checkpoint blockade in mice. These results demonstrate that selectively targeting STING agonists to antigen-presenting cells may improve their therapeutic effects, particularly in poorly immunogenic tumors. —CO

Sci. Immunol. **8**, eabn6612 (2023).

LIGHT POLLUTION

The night sky is rapidly getting brighter

Artificial lighting that escapes into the sky causes it to glow, preventing humans and animals from seeing the stars. Satellites can measure the light emitted upward, but they are not sensitive to all wavelengths produced by LED lighting or to light emitted horizontally. Kyba *et al*. used data from citizen scientists to measure how light pollution is affecting human views of the stars worldwide (see the Perspective by Falchi and Bará). Participants were shown maps of the sky at different levels of light pollution and asked which most closely matched their view. Trends in the data showed that the average night sky got brighter by 9.6% per year from 2011 to 2022, which is equivalent to doubling the sky brightness every 8 years. —KTS

Science, abq7781, this issue p. 265; see also adf4952, p. 234

Science
JOURNALS AAAS



Publish your research in the *Science* family of journals

The *Science* family of journals (*Science, Science Advances, Science Immunology, Science Robotics, Science Signaling*, and *Science Translational Medicine)* are among the most highly-regarded journals in the world for quality and selectivity. Our peer-reviewed journals are committed to publishing cutting-edge research, incisive scientific commentary, and insights on what's important to the scientific world at the highest standards.

Submit your research today!

Learn more at **Science.org/journals**

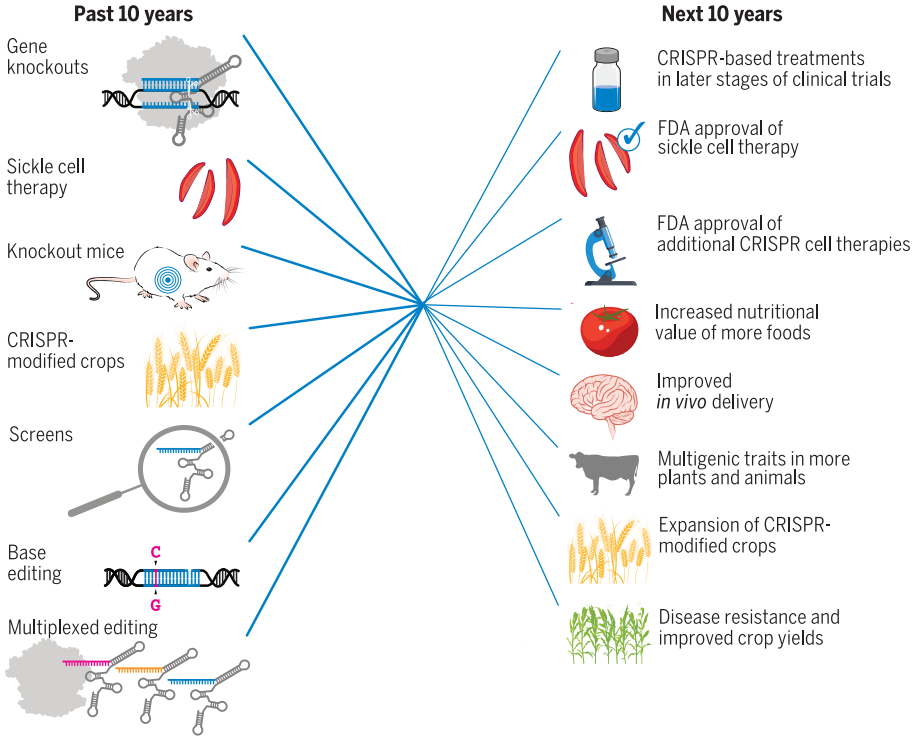
REVIEW SUMMARY

CRISPR

CRISPR technology: A decade of genome editing is only the beginning

Joy Y. Wang and Jennifer A. Doudna*

BACKGROUND: The fields of molecular biology, genetics, and genomics are at a critical juncture—a moment in history when a convergence of knowledge and methods has made it both technically possible and incredibly useful to edit specific base pairs or segments of DNA in cells and living organisms. The advent of clustered regularly interspaced short palindromic repeat (CRISPR) genome editing, coupled with advances in computing and imaging capabilities, has initiated a new era in which we can not only diagnose human diseases and even predict individual susceptibility based on personal genetics but also act on that information. Likewise, we can both identify and rapidly alter genes responsible for plant traits, transforming the pace of agricultural research and plant breeding. The applications of this technology convergence are profound and far reaching—and they are happening now. In the decade since the publication of CRISPR-Cas9 as a genome editing technology, the CRISPR toolbox and its applications have profoundly changed biological research, impacting not only patients with genetic diseases but also agricultural practices and products. As a specific example from the field of genomic medicine, it has become feasible to obtain a complete sequence of the human genome in less than 24 hours—a staggering advance considering the first such sequence took 5 years to generate. Notably, designing and putting to use a potent CRISPR genome editor to obtain clinically actionable information from that genome—previously a near-intractable challenge—now takes only a matter of days. For additional background and related topics, we refer readers to in-depth reviews of the microbiology and structural biology of CRISPR systems and to articles about the considerable ethical and societal challenges of this technology.



CRISPR: past, present, and future. The past decade of CRISPR technology has focused on building the platforms for generating gene knockouts, creating knockout mice and other animal models, genetic screening, and multiplexed editing. CRISPR's applications in medicine and agriculture are already beginning and will serve as the focus for the next decade as society's demands drive further innovation in CRISPR technology.

ADVANCES: The past decade has witnessed the discovery, engineering, and deployment of RNA-programmed genome editors across many applications. By leveraging CRISPR-Cas9's most fundamental activity to create a targeted genetic disruption in a gene or gene regulatory element, scientists have built successful platforms for the rapid creation of knockout mice and other animal models, genetic screening, and multiplexed editing. Beyond traditional CRISPR-Cas9–induced knockouts, base editing—a technology utilizing engineered Cas9's fused to enzymes that alter the chemical nature of DNA bases—has also provided a highly useful strategy to generate site-specific and precise point mutations. Over the past decade, scientists have utilized CRISPR technology as a readily adaptable tool to probe biological function, dissect genetic interactions, and inform strategies to combat human diseases and engineer crops. This Review covers the origins and successes of CRISPR-based genome editing and discusses the most pressing challenges, which include improving editing accuracy and precision, implementing strategies for precise programmable genetic sequence insertions, improving targeted delivery of CRISPR editors, and increasing access and affordability. We examine current efforts addressing these challenges, including emerging gene insertion technologies and new delivery modalities, and describe where further innovation and engineering are needed. CRISPR genome editors are already being deployed in medicine and agriculture, and this Review highlights key examples, including a CRISPR-based therapy treating sickle cell disease, a more nutritious CRISPR-edited tomato, and a high-yield, disease-resistant CRISPR-edited wheat, to illustrate CRISPR's current and potential future impacts in society.

OUTLOOK: In the decade ahead, genome editing research and applications will continue to expand and will intersect with advances in technologies, such as machine learning, live cell imaging, and sequencing. A combination of discovery and engineering will diversify and refine the CRISPR toolbox to combat current challenges and enable more wide-ranging applications in both fundamental and applied research. Just as during the advent of CRISPR genome editing, a combination of scientific curiosity and the desire to benefit society will drive the next decade of innovation in CRISPR technology. ■

The list of author affiliations is available in the full article online.
*Corresponding author. Email: doudna@berkeley.edu
Cite this article as J. Y. Wang, J. A. Doudna, *Science* 379, eadd8643 (2023). DOI: 10.1126/science.add8643

READ THE FULL ARTICLE AT
https://doi.org/10.1126/science.add8643

REVIEW

CRISPR

CRISPR technology: A decade of genome editing is only the beginning

Joy Y. Wang[1,2] and Jennifer A. Doudna[1,2,3,4,5,6,7,8]*

The advent of clustered regularly interspaced short palindromic repeat (CRISPR) genome editing, coupled with advances in computing and imaging capabilities, has initiated a new era in which genetic diseases and individual disease susceptibilities are both predictable and actionable. Likewise, genes responsible for plant traits can be identified and altered quickly, transforming the pace of agricultural research and plant breeding. In this Review, we discuss the current state of CRISPR-mediated genetic manipulation in human cells, animals, and plants along with relevant successes and challenges and present a roadmap for the future of this technology.

Starting with a 1987 report about repetitive DNA sequences in a bacterial genome (*1*), a small set of researchers working in the fields of microbiology and food science began studying mysterious DNA sequence arrays known as clustered regularly interspaced short palindromic repeats (CRISPRs), commonly found in microbial genomes together with genes encoding CRISPR-associated (Cas) proteins. The presence of short DNA sequences within CRISPRs matching those in viruses hinted at the function of these systems as adaptive immunity pathways used to prevent viral infection (Fig. 1A) (*2*). Curiosity-driven research ultimately showed how CRISPR systems use RNA molecules transcribed from the sequence arrays to guide Cas proteins to cut, and thereby destroy, viral DNA or RNA (*3*, *4*). Furthermore, this line of research showed how CRISPR's RNA-programmed cutting action (*5*, *6*) could be used to alter DNA sequences in any cell with unprecedented ease [reviewed in (*7*)]. Over the past decade, scientists around the world have rapidly adapted CRISPR to enable both fundamental research and wide-ranging applications in animals, plants, and humans.

The most widely used genome editor is the CRISPR-Cas9 protein complexed with its partner RNA (Fig. 1B). The power of CRISPR as a genome editing technology stems from its chemical mechanism of DNA cutting at a site dictated by RNA-determined sequence recognition. Because Cas proteins use RNA-DNA base pairings for DNA recognition, the same protein, such as Cas9, can target a wide range of DNA sequences by simply swapping guide RNAs (Fig. 1B). In eukaryotic cells, DNA breaks are efficiently repaired, enabling targeted changes to DNA sequences at will (Fig. 1B). Mutating the amino acids required for cleavage activity in the Cas9 active sites allows targeted DNA nicking (introducing a single-stranded DNA cut) or DNA binding by a catalytically inactive Cas9. As a result, the first examples of engineered CRISPR-Cas involved transcriptional repression or activation to silence or up-regulate specific genes (*8*, *9*). Other forms of engineered Cas9 are fused to enzymes that enable individual nucleobase editing, chromatin modification, or sequence insertion (*10*–*13*). Other Cas proteins, including RNA-targeting proteins, have been explored as genome-modifying tools, enabled by discovery efforts and extensive biochemical and structural characterization [reviewed in (*14*–*17*)]. Some of these enzymes have also been harnessed for the development of imaging methods (*18*–*20*) and diagnostic approaches (*21*, *22*).

Together, applications of CRISPR technology have provided the foundation for clinical trials of therapies to treat sickle cell disease, beta-thalassemia, the degenerative disease transthyretin (TTR) amyloidosis, and congenital eye disease, as well as planned clinical trials for both rare (progeria, severe combined immunodeficiency, familial hypercholesterolemia) and common (cancer, HIV infection) diseases. CRISPR technology has enabled agricultural advances including slick-coat cattle and a more nutritious tomato. It has spurred research across fields of molecular and cell biology, fueling the publication of thousands of research articles and providing a tool base for many companies focused on therapeutics, agriculture, and synthetic biology. However, CRISPR technology and its potential impact are still in their early stages. As we discuss in the next section, some genome editing applications have now become routine whereas others remain difficult due to limitations of today's tools. These genome editing challenges provide opportunities for new discoveries and engineering to advance the field by offering a more complete toolbox for genetic manipulation.

CRISPR-induced gene knockouts

The past decade has witnessed the astounding success of CRISPR-induced gene knockouts, which have transformed basic and translational research and demonstrate tremendous potential in agriculture and therapeutic development. Traditional CRISPR-induced knockout methods in eukaryotic cells involve the CRISPR-Cas9 ribonucleoprotein (RNP), composed of the Cas9 nuclease and an engineered single-guide RNA molecule (sgRNA) (*23*–*26*). The sgRNA directs Cas9 to the target site, where it creates a double-stranded DNA break (DSB) that is repaired by endogenous repair pathways including the nonhomologous end joining (NHEJ) and the microhomology-mediated end joining pathways and the more precise homology-directed repair (HDR) pathway that uses a repair template [reviewed in (*27*–*29*)] (Fig. 2A). Because of the high targeting specificity and efficacy of CRISPR-Cas9, such gene knockouts are now routine in research applications, providing a streamlined process to disrupt genes for functional study.

CRISPR-Cas9 has proven to be successful in enabling the rapid creation of knockout (KO) mice and other animal models (*30*, *31*) (Fig. 2B). Traditional gene targeting methods used inefficient homologous recombination in embryonic stem (ES) cells, followed by laborious screening of modified ES cells for the desired sequence change and injection into WT embryos (*32*, *33*). CRISPR-Cas9 provides a way to introduce DSBs in a one-cell-stage embryo, bypassing the screening stage of suitable targeted ES cells and greatly simplifying the production of gene-edited animals (*34*, *35*). This has reduced the time needed for generating genetically modified mice, from 1 year to as few as 4 weeks (*36*). As a result, the production of KO and transgenic mice has now become routine for research applications. Additionally, because most mammalian species lack established ES cell lines, CRISPR-Cas9 editing has facilitated the development of genetically engineered animal models in new species (*37*, *38*). The generation of KO and transgenic animal models has become even more efficient with advancements in strategies for introducing CRISPR-Cas9 components into zygotes, including CRISPR RNP electroporation of zygotes (CRISPR-EZ) (*39*), CRISPR RNP electroporation and AAV donor infection (CRISPR-READI) (*40*), and improved genome editing through oviductal delivery of

[1]Department of Chemistry, University of California, Berkeley, Berkeley, CA, USA. [2]Innovative Genomics Institute, University of California, Berkeley, Berkeley, CA, USA. [3]Howard Hughes Medical Institute, University of California, Berkeley, Berkeley, CA, USA. [4]Department of Molecular and Cell Biology, University of California, Berkeley, Berkeley, CA, USA. [5]California Institute for Quantitative Biosciences (QB3), University of California, Berkeley, Berkeley, CA, USA. [6]Molecular Biophysics & Integrated Bioimaging Division, Lawrence Berkeley National Laboratory, Berkeley, CA, USA. [7]Gladstone Institutes, University of California, San Francisco, San Francisco, CA, USA. [8]Gladstone-UCSF Institute of Genomic Immunology, San Francisco, CA, USA.
*Corresponding author. Email: doudna@berkeley.edu

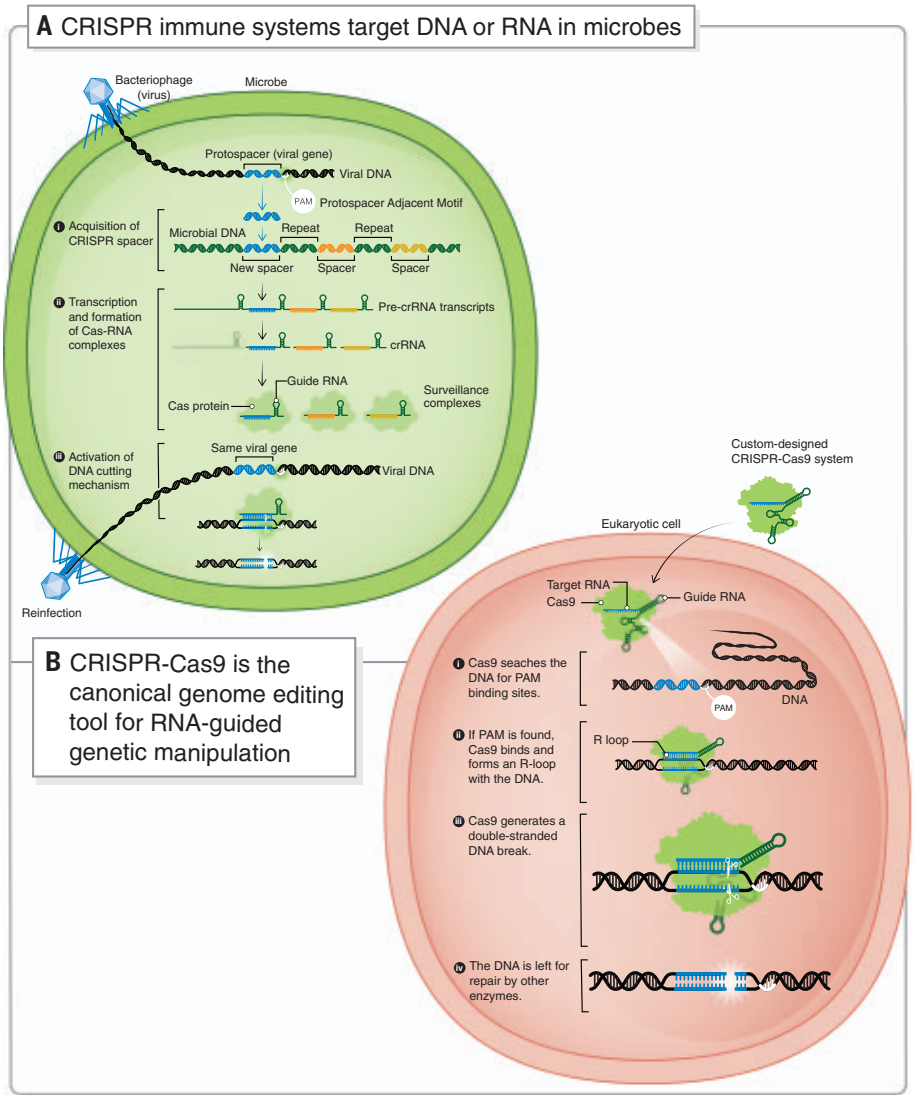


Fig. 1. CRISPR-based adaptive immunity provides programmable genome editing tools. (**A**) CRISPR immune systems target DNA or RNA in microbes (illustration depicts DNA targeting). Three steps to immunity include: (i) acquisition of CRISPR spacer sequence matching an infectious agent; (ii) transcription and formation of Cas-RNA complexes; (iii) seek-and-destroy surveillance mechanisms. (**B**) CRISPR-Cas9 is the canonical genome editing tool for RNA-guided genetic manipulation. Cas9 searches for target sites in a genome by engaging with PAM sequences, forming an R-loop with complementary DNA, generating a double-strand DNA (dsDNA) break, and finally releasing DNA for repair.

nucleic acids (i-GONAD) (*41*). Beyond germline editing, CRISPR-Cas9 is also used in somatic editing, which is useful in situations where whole-body knockouts are lethal for embryos and in many cases can more accurately model cancer progression and realistic modes of therapeutic treatment (*42*). Advancements in in vivo delivery strategies (discussed later in this review) have expanded the types of somatic animal models that can be created. Through CRISPR-genome engineering methods, animal models have been developed for many diseases, including tyrosinemia, Duchenne muscular dystrophy, cancer, osteoporosis, Huntington's disease, amyotrophic lateral sclerosis, Alzheimer's disease, and HIV-1/AIDS as just a few examples [reviewed in (*38*)]. The ability to rapidly create animal models has advanced and will continue to advance the study of genetic diseases, allowing researchers to study the causal relationships of specific genetic variations and disease and to develop and test new treatments for these diseases.

CRISPR screens

With the ease of introducing CRISPR-induced gene knockouts, researchers have successfully applied this technology to genetic screens, i.e., the systematic, targeted genetic alteration of a number of genes in parallel. Such CRISPR screens have become a powerful approach for understanding genetic interactions and dissecting biological pathways and have given rise to major advances in target discovery and drug development. The capabilities of CRISPR screens are continuing to expand, especially when combined with advancements in single-cell multiomics technologies. In general, genetic screens involve one or multiple gene perturbations, a model system such as engineered human cells, and a selection assay or readout to evaluate the effects of the perturbation(s) (*43*, *44*) (Fig. 2C). Because of its efficiency and flexibility, CRISPR editing is a powerful strategy for introducing perturbations that can be used for close study of how a single gene disruption affects a cell of interest, as well as high-throughput testing of thousands of perturbations in pooled screens (*45*, *46*). The ease of designing and cloning guide RNA (gRNAs) has enabled the development of gRNA libraries up to genome-wide scales, allowing researchers to perturb every gene in the human genome (*47*, *48*). CRISPR technology advancements have also expanded the types of CRISPR screens that researchers can use for different applications. Beyond CRISPR KO screens, CRISPR interference (CRISPRi) and CRISPR activation (CRISPRa) screens have also become popular approaches that use reversible gene expression control [reviewed in (*49*, *50*)]. Saturation genome editing utilizing Cas9-mediated HDR enables the generation of all possible single-nucleotide polymorphisms (SNPs) for functional screening (*51*–*53*). More recently, as alternatives to Cas9-mediated HDR, researchers have also started applying CRISPR base and prime editing (discussed later in the Review) for genetic screens (*54*–*56*). Base editing, which can introduce point mutations more efficiently than Cas9-mediated HDR with minimal indel formation, may serve as an improved strategy for functional variant screening (*54*). Prime editing allows for the introduction of small insertions and deletions in addition to point mutations, which essentially enables saturation mutagenesis across residues. This technology is still relatively new, and it remains to be seen whether prime editing screens can achieve similar levels of flexibility as HDR with regard to target selection and targetable window size (*56*).

The successful editing of diverse cells and organisms using CRISPR technologies provides flexibility for choosing a model system for the genetic screen to best answer the relevant biological question. Beyond primary cells, CRISPR screens have been developed in more complex model systems including organoids, animals, and plants (*57*–*60*). CRISPR screens are now a common method for probing gene function in cancer and have allowed for identification of a variety of cancer drivers and regulators [reviewed in (*45*, *46*)].

Following the introduction of CRISPR components into the model, a variety of techniques

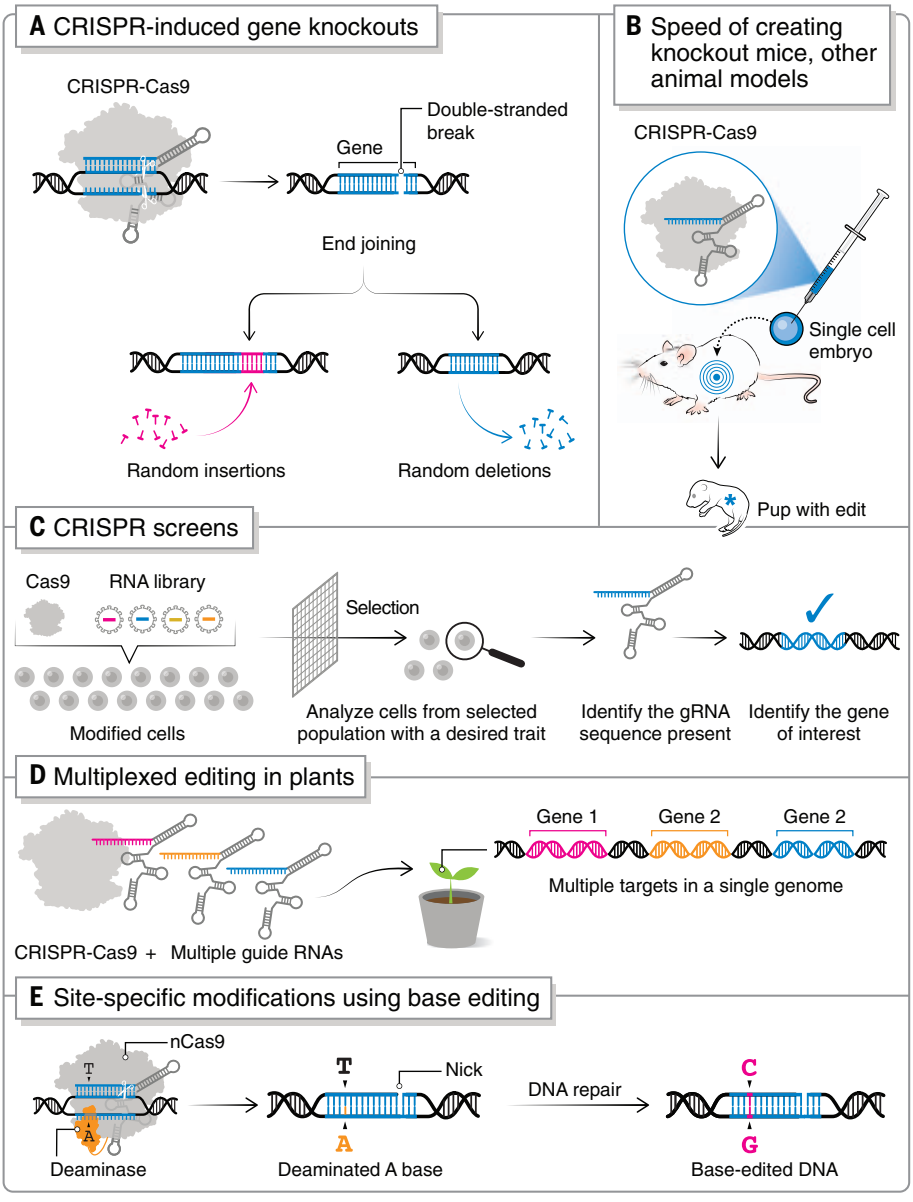


Fig. 2. The genome editing toolbox, part 1: what works well. (**A**) CRISPR-induced gene knockouts in eukaryotic cells result from the DNA DSB, created by the Cas9 RNP and then commonly repaired by endogenous end joining repair pathways. (**B**) The speed of creating KO mice and other animal models has greatly increased with the development of CRISPR-Cas9 editing technology, which can edit single-cell embryos to generate gene-modified mice. (**C**) CRISPR screens are utilized for functional genetic screening. CRISPR-Cas9 editors introduce genetic perturbations into a model, which is subjected to a selection assay, followed by a readout to evaluate the effects of the perturbations. (**D**) CRISPR-Cas9 is a platform for multiplexed editing in plants. Multiple gRNAs can be used with Cas9 to simultaneously edit multiple targets in a genome. (**E**) CRISPR base editors, typically composed of nCas9 or dCas9 fused to a deaminase, enable site-specific modifications without DSBs.

can be used for the selection assay and readout. Common selection strategies include viability- or proliferation-based screens, fluorescence-activated cell sorting or microfluidics-assisted cell screening-based screens using cell surface proteins as markers (i.e., PD1, PDL1, MHC), and in vivo screens assaying phenotypes, such as tumor growth or sensitivity/resistance to immunotherapy (*45*, *46*, *61*). One exciting area of development in the readout of CRISPR screens is new methods that provide simultaneous proteomic, epigenetic, and/or transcriptomic analyses, such as Procode (*62*), Perturb-ATAC (*63*), Perturb-seq (*64*), and ECCITE-seq (*65*), which can provide a wealth of information. The successes of CRISPR screens will continue to accelerate with new technologies that improve the sensitivity of these assays and readouts. We are only beginning to see the impact of combining CRISPR screens and single-cell multiomics modalities with the rapidly advancing infrastructure for big data collection and analysis (*64*, *66*, *67*). Further discovery and engineering, including the enhancement of orthologous Cas9 enzymes or other RNA-guided nucleases such as Cas12a [reviewed in (*7*)], can also greatly increase the potential of combinatorial or multiplex CRISPR screens, which can reveal novel and complex genetic interactions.

Multiplexed editing in plants and beyond

Multiplex genome editing, or the simultaneous targeting of multiple specific DNA loci in a genome, represents another area where CRISPR-induced gene knockout technology has been scaled up and adapted into a successful platform—particularly in the plant science fields [reviewed in (*68*, *69*)] (Fig. 2D). Over the past decade, CRISPR-Cas9 has become a popular tool for plant editing. Traditional crop trait engineering methods involved random mutagenesis (e.g., with radiation) or transgenesis with Agrobacterium followed by laborious crossing and screening to identify a plant with a new trait of interest. These processes are lengthy and hard to control in addition to facing substantial regulatory hurdles. By contrast, CRISPR-induced modifications are targeted, can be made rapidly, and generally represent small indels, i.e., insertions and/or deletions, or point mutations at locations specified by the trait engineer [reviewed in (*70*)]. CRISPR-Cas9 has been adapted for simultaneous multilocus editing, which is especially useful for editing crop species that can carry multiple copies of the target gene (e.g., hexaploid wheat) (*71*) and for crop domestication, which involves targeting multiple different genes [reviewed in (*72*)]. One advantage of CRISPR-Cas9 in regards to multiplexed editing is separation of the nuclease and gRNA, such that multiple gRNAs can be used with one Cas protein to edit different targets (*23*). The gRNAs can be provided as multiple individual expression cassettes, each transcribed from their own promoters (*73*–*75*), or as a single polycistronic cassette that is processed posttranscriptionally (*76*–*78*). In the past few years, CRISPR multiplex genome editing has become a successful strategy for creating new crop genotypes and agriculturally useful traits in a single generation. One area where multiplexed CRISPR-Cas9 editing has achieved successful results is in crop domestication and improvement. Examples include the use of multiplexed editing to disrupt domestication genes, introduce characteristics such as herbicide resistance, and increase crop yield and quality (*68*, *79*–*81*).

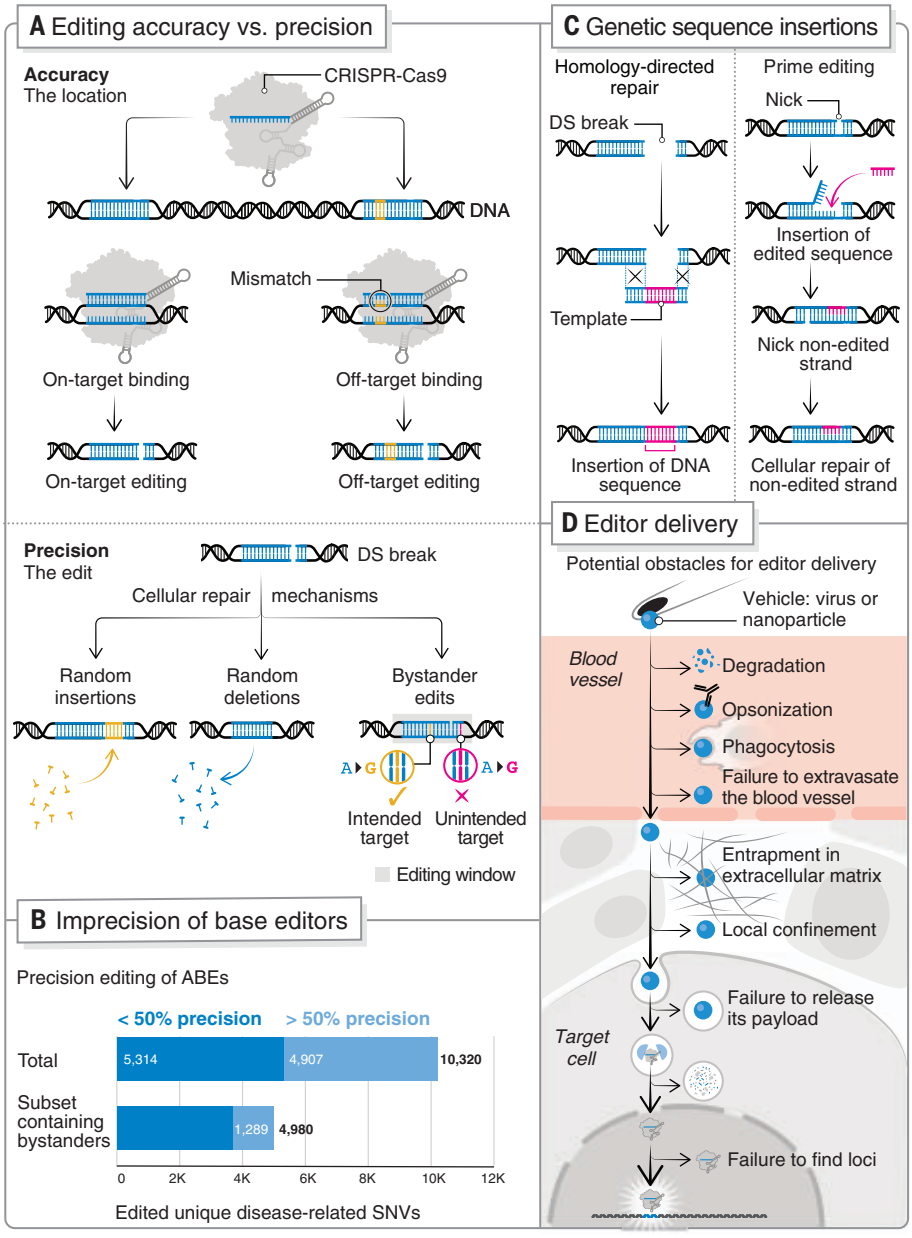


Fig. 3. The genome editing toolbox, part 2: challenges and emerging tools. (**A**) Editing accuracy versus editing precision. Editing accuracy involves specificity for the target location, which is challenged by off-target binding by the Cas9 RNP. Editing precision involves producing the correct intended edit with no unintended edits and is challenged by undesired indels and bystander edits. (**B**) Imprecision of correcting most SNPs using current base editors [data from (*130*), chart contributed by P. H. Yoon]. Correction precision is defined as the fraction of edited reads in their library of disease-associated SNVs correctable by adenine base editors that consists of an exact single-nucleotide correction (*130*). (**C**) Genetic sequence insertions by HDR-mediated CRISPR-Cas9 editing (left) and by prime editing (right) (*195*). (**D**) Challenges of gene editor delivery, including potential biological obstacles that can prevent the gene editor from reaching its target site.

Beyond its success in plants, multiplexed CRISPR-Cas9 editing has been extended to other cell types and organisms. One notable example is the use of multiplexed editing to generate porcine endogenous retrovirus-inactivated pigs, addressing a major safety concern of transplanting pig organs into humans (*82*). Multiplexed editing has the potential to be particularly useful in engineering cell therapy products for cancer and studying the effects of complex polygenic diseases. However, a challenge of CRISPR-Cas9 editing in mammalian cells that is exacerbated by simultaneous DNA cleavages in multiplexed editing is the possibility of triggering DNA damage-response mechanisms governed by transcription factor p53 that can lead to cellular senescence or apoptosis (*82*–*84*). Addressing this challenge and extending the applications of multiplexed editing may become easier with the advancement of newer CRISPR precision editing technologies that do not involve DSBs, including base editing and prime editing (discussed further below), which have already been adapted for multiplexed editing (*85*, *86*).

Site-specific modifications using base editing

Beyond traditional CRISPR-induced knockouts involving double-stranded DNA cutting, base editing (*11*)—which uses Cas effectors fused to enzymes that alter the chemical nature of DNA bases—has provided a successful strategy to generate site-specific and precise point mutations without DSBs, eliminating the need for repair templates and limiting undesired byproducts during editing (Fig. 2E). CRISPR base editors generally consist of fusions between a Cas9 nickase (nCas9), a Cas9 variant that produces a single-stranded rather than a double-stranded break (or a catalytically inactive or "dead" Cas protein, such as dCas9, dCas12a, or dCas13b), and an enzyme that catalyzes a nucleobase deamination reaction [reviewed in (*87*, *88*)]. The sgRNA directs the nCas9-deaminase fusion to the genomic target, where ternary complex formation exposes a region of ssDNA to the deaminase for chemical modification. The resulting base mismatch is then resolved through cellular repair mechanisms. Over the past few years, the toolbox of DNA and RNA base editors has expanded to enable C>T, A>G, C>G, A>I, and C>U conversions (*10*, *11*, *89*–*93*), though there is still need for further improvement, especially for C>G editing. Site-specific modifications broaden the abilities of researchers to study the effects of mutations within genes and can treat genetic disorders by correcting point mutations, which represent the largest class of human pathogenic genetic variants (*94*–*97*). Furthermore, base editing can introduce modifications in dividing and nondividing cells, providing an advantage over HDR, which is restricted to dividing cells (*98*). Base editing has already shown promising results in correcting loss-of-function mutations in a number of mouse models [reviewed in (*87*, *88*)], with a notable example being the recent use of in vivo base editing to correct Hutchinson-Gilford progeria syndrome in mice (*99*). By eliminating the introduction of DSBs, the development of base editing technology represents an important step in precision editing. Refining these tools to the levels of accuracy and precision needed for treatments in humans will be one of the main challenges in the next decade. An early-stage clinical trial using base editing for familial hypercholesterolemia has already begun, and others using base editing for sickle cell disease are set to begin this year (*100*).

Editing accuracy and precision

As we transition into the next decade of CRISPR genome editing, several key challenges demand innovative solutions. Two of these are editing accuracy (i.e., specificity for the target site) and precision (i.e., producing the exact desired editing outcome) (Fig. 3A). To reduce the off-target effects of CRISPR-Cas nucleases resulting from unintended binding and cleavage, researchers have utilized a combination of rational design and selection to develop high-fidelity Cas variants [such as SpCas9-HF1 (*101*), evoCas9 (*102*), HiFiCas9 (*103*)] and the Cas9_R63A/Q768A variant (*104*) and guide optimization methods [such as E-Crisp (*105*, *106*), CasOFFinder (*107*), and sgDesigner (*108*)]. These efforts have been productive: Neither the CRISPR Therapeutics/Vertex nor the Intellia sgRNAs used in clinics today have measurable off-target sites using US Food and Drug Administration (FDA)–grade assays (*109*, *110*). However, off-target editing inaccuracies can also occur as a result of the Cas9-independent behavior of effector domains including deaminases, reverse transcriptases, and transcriptional regulators, as exemplified by analysis of base editing outcomes (*111*, *112*). Progress is currently being made through use of high-fidelity Cas variants and rational engineering of the deaminase domain to reduce nucleic acid binding without Cas assistance (*111*–*115*), and early-stage clinical trials of base editing (*31*, *116*) offer encouragement in this regard. At the same time, innovating new methods to deliver Cas editors to the target site and refining existing ones (described further below) can also minimize off-target effects. Editing precision poses a larger challenge. In traditional CRISPR-Cas9 editing in eukaryotic cells, the scientist still does not fully control the editing outcome following the introduction of the DSB. Recent machine learning tools have been developed to help predict repair outcomes, although these have yet to be demonstrated for in vivo applications (*117*–*119*). Following the DSB, NHEJ—the default repair pathway for human cells—competes with the less-efficient HDR pathway and results in a spectrum of indels at the target site (*120*, *121*). Although this may be acceptable for a number of CRISPR-induced knockout applications (including clinical), many therapeutic applications require much higher levels of precision and cannot afford undesired indels. Increasing editing precision requires a better understanding of DNA repair processes and a combination of innovation and engineering. One approach is improving HDR efficiency and/or suppressing NHEJ. Strategies that have been developed include chemically inhibiting key enzymes in the NHEJ pathway, using single-stranded oligodeoxynucleotide templates (which have been shown to increase HDR efficiency to 60% in human cells for a single-nucleotide substitution) (*122*), utilizing cell cycle stage control to favor HDR repair (*123*–*125*) and using site-specific Cas9-oligonucleotide conjugates to recruit the donor DNA template to the target site (*126*). Even with these strategies, there are still risks of large deletions and chromosomal rearrangements associated with DSB formation that can lead to genome instability (*127*, *128*). Base editing and prime editing (addressed further below) represent another approach for precision editing intended to avoid DSB formation. Base editing and prime editing have reduced indel formation compared with classic Cas9-mediated editing. However, in some cases, unintentional DSBs can still form at the editing site and lead to indels. It has been demonstrated that fusing base editors to Gam—a bacteriophage Mu protein that binds DSBs—can minimize indel formation during base editing (*129*). For base editing, editing precision is also challenged by bystander editing (or undesired conversions of neighboring editable bases within or near the editing window in addition to the target base) (*87*, *88*). The correction precision of base editors decreases by a large margin when there is more than one target base in the editing window, limiting their therapeutic potential (*130*) (Fig. 3B). In a recent study, about half of the pathogenic single-nucleotide variants (SNVs) correctable by adenine base editors revealed ≥50% correction precision (*130*) (Fig. 3B). However, of the subset of SNVs containing more than one target base in the editing window, only 26% revealed ≥50% correction precision (Fig. 3B). Unfortunately, with current base editors, bystanders are a fairly common occurrence. In a recent study of 21 different base editing systems, about half of the targetable pathogenic point mutations had bystanders in the editing window (*131*). Reducing the size of the editing window can increase precision; however, this also limits the genomic sites that can be targeted because of PAM constraints. A variety of strategies involving structure-guided mutagenesis, directed evolution, and computational-aided design are currently being employed to increase the targeting scope of CRISPR-Cas9 and reduce the bystander effects of base editors (*132*–*134*). For base editing to be an effective strategy for a wider breadth of applications, further engineering is required to build on current strategies to develop base editors with narrower editing windows and different PAM compatibilities without compromising on efficiency and targeting specificity.

Genetic sequence insertions

In recent years, emerging technologies are expanding the functional capabilities of the CRISPR toolbox to make precise programmable genetic sequence insertions, and an important challenge in the next decade will be refining and effectively implementing these technologies for genome engineering applications. Traditional Cas9 editing can introduce transgenes by relying on HDR to incorporate genetic material from a co-delivered donor template into the target site (Fig. 3C) (*27*). Currently, this approach is being widely used in many areas of genome engineering. A notable recent example of its application is the use of targeted integration to fluorescently tag more than 1000 human proteins to study their localization and interactions (*135*). HDR-mediated CRISPR-Cas9 editing has also shown promising results in preclinical and clinical testing for therapeutic development, with key examples in correcting alpha1 antitrypsin deficiency (*136*) and in cancer immunotherapy (*137*, *138*). Despite these successes, HDR-mediated CRISPR-Cas9 editing has its limitations, including being restricted to dividing cells (*123*), the difficulty of donor template delivery, and the precision-related challenges introduced by the DSB. Although certain single-nucleotide mutations can be addressed by base editing, many human pathogenic genetic variants require a small sequence insertion to repair an indel, calling for high-precision alternatives to HDR-mediated CRISPR-Cas9.

Prime editing represents one such alternative that can insert and delete DNA sequences without introducing DSBs (*12*), though this technology still needs further refinement (Fig. 3C). Prime editors consist of nCas9 fused to a reverse transcriptase (RT) and a prime editing gRNA (pegRNA) that serves both to direct nCas9 to the target site and act as a template containing the desired edit for the RT (*12*). Unlike HDR, prime editing can introduce modifications in both dividing and nondividing cells, which is useful for correcting mutations in quiescent cells, such as neurons or hematopoietic stem cells (*139*). Prime editing also provides advantages over base editing in situations where there are multiple target bases in the editing window (*132*) and where a PAM sequence is not immediately adjacent to the desired editing site (*140*). Currently, prime editing has shown promise as an accurate and fairly precise editing tool that has been demonstrated to work in multiple cell types, organoids, mouse embryos, and plants but is still limited in its applications as a result of low editing efficiency [reviewed in (*139*, *141*)]. In two separate demonstrations of prime editing in organoids and mice, there were no detectable off-target edits (*142*, *143*). Although low levels of undesired indel formation have been reported, the ratio of correct editing to indel formation was ~30 times higher for prime editing than for HDR (*12*, *142*). Unfortunately, prime editing efficiencies are low for many applications. In one study, prime editing was more than 30 times less efficient than HDR in repairing a mutation in intestinal CF organoids (*144*). Although more efficient prime editors have been developed, these

also result in higher rates of indel formation (*12*). Currently, base editors still have an advantage over prime editors in editing efficiency and precision (*139*). A main goal for prime editing in the next decade is improving efficiency without compromising editing product purity—an outcome that has the potential to turn prime editing into one of the most versatile tools for precision editing. Future studies should also address remaining uncertainties about the mechanism of prime editing. Recent results show that physical untethering of Cas9 and the RT has no effect on prime editing levels in cells, suggesting that the RT could engage the editing site without being fused to Cas9 and raising questions about whether it could induce unintended integration at other RNA-DNA hybrid sites (*13*). Refining prime editing tools will require engineering and optimizing the different constituent components, including the pegRNAs (*145*).

For large gene insertions, an emerging area in CRISPR genome engineering is RNA-guided DNA transposition. CRISPR-associated transposons (CASTs) enable the precise RNA-guided integration of large DNA cargo up to 10 kb (*146*–*149*). So far, this has only been demonstrated in a few prokaryotes and has not yet been reported to work in mammalian cells (*146*, *147*, *150*, *151*). There is potential for new developments as this area is still in the early stages of research with limited mechanistic understanding of how these systems work (*15*), and very few computationally predicted CAST systems have been characterized (*149*). Further discovery, testing, and engineering will be required to harness the potential of CASTs for genome engineering applications.

Recombinases, which perform a wide array of activities, including insertions, deletions, inversions, and replacements (*152*, *153*), are another area of tool development with the potential to combine with Cas proteins and may be able to further diversify the CRISPR toolbox (*154*). This has recently been demonstrated with the development of two new approaches that have enabled programmable integration of large DNA sequences in human cells (*155*, *156*). One approach, programmable addition through site-specific targeting elements, uses engineered fusion proteins of Cas9, a reverse transcriptase, and a serine integrase that have enabled multiplexed insertions of large DNA cargo, including the fluorescent tagging of different endogenous genes (*155*). Another approach uses twin prime editing, which involves a prime editor and two prime editing guide RNAs that have enabled large gene insertions and inversions when combined with a site-specific serine recombinase (*156*). This approach was used to correct a large sequence inversion associated with Hunter syndrome in human cells with up to ~9% efficiency (*156*). Notably, these studies report no detectable off-target insertions. These approaches for programmable gene insertions will require further characterization and refinement to increase editing efficiencies and to serve as potential therapeutic strategies. The potential impact of programmable gene insertion for genome engineering will continue to motivate discovery and innovation in search of new strategies in addition to improving existing technologies.

Delivery of editors ex vivo and in vivo

Despite all of the recent advances in CRISPR editors, delivery of editors remains a major bottleneck for genome editing in organisms; both innovation and engineering are needed to ensure high delivery efficiency, target specificity, and safety. Advances in delivery technologies have played a large role in developing CRISPR-based therapeutics. The liver represents a clear example where efficient delivery of CRISPR editors has been a clinically tractable challenge (*110*). However, for less accessible organs, the feasibility of CRISPR therapeutics is limited by low delivery efficiencies and will largely depend on improved delivery strategies. Current delivery strategies for potential CRISPR-based treatments in humans are divided into two types of approaches: ex vivo, where cells are isolated from and modified outside of the patient before being reintroduced, and in vivo, where cells are edited directly in the patient following delivery of CRISPR components [reviewed in (*157*–*160*)]. Ex vivo approaches, often used for editing hematopoietic stem and progenitor cells and leukocytes, offer higher cell-type specificity and tighter quality control of editing; however, they are limited to cell types that can survive and be expanded in culture (to achieve a minimum number for reengraftment) and retain in vivo function. In vivo approaches expand CRISPR editing to cell types where ex vivo approaches are not possible, allowing CRISPR to treat a wider range of genetic diseases. Two notable examples where in vivo delivery has had some success in humans are the treatment of transthyretin amyloidosis, which represents the first systemic in vivo delivery of CRISPR to the liver using targeted lipid nanoparticle (LNP) delivery (*110*), and treatment of Leber congenital amaurosis type 10, which involves direct injection of an adeno-associated viral vector harboring the RNA-guided enzyme into the eye (*161*, *162*).

These successes show the tremendous potential of in vivo therapeutic genome editing; however, in general, in vivo delivery of CRISPR editors remains a formidable challenge. Many biological obstacles stand in the way of effective in vivo delivery of editors to targets. In the case of systemic delivery, delivery vehicles need to prevent degradation of the cargo, opsonization, and phagocytosis extravasate from the blood vessel; pass efficiently through the interstitial space; and effectively release cargo upon endocytosis (Fig. 3D). Once the CRISPR cargo is released, it also needs to localize to the nucleus and access the target locus in the

Box 1. Cost, regulation, and access.

With the growing therapeutic potential of CRISPR technologies, other important considerations include affordability, regulation, and access. For a more in-depth examination of the ethical and societal challenges of these technologies, we refer readers to refs. (*196*, *197*). One main challenge to the development and accessibility of CRISPR therapeutics is cost. In many cases, manufacturing costs encompass the expenses of producing both the CRISPR editors and the delivery vehicles, which can be difficult to scale up. For instance, viral-based delivery is a popular strategy for developing CRISPR therapeutics, but the manufacturing of viral vectors depends on expensive culture systems and facilities to produce required quantities of virus (*175*). Improving process development and providing the infrastructure to scale up the production of viral vectors will be important to reduce costs. Treatment administration can also be expensive, especially for ex vivo approaches, which consume time and resources to expand cells in culture and, in cases of bone marrow transplantation with autologous hematopoietic stem cells, require preconditioning patients with chemotherapy (*198*).

Manufacturers also face the burden of regulatory costs to provide extensive characterization and strict safety and quality controls, which can be challenging for investigational or academic manufacturing facilities. As increasing numbers of CRISPR-based treatments move to later stages of clinical trials, manufacturers will need to build the infrastructure and bear the costs to support a current good manufacturing practice–compliant operation. The challenges of bearing these costs and getting FDA approval can lead to abandonment of the therapeutic development in a for-profit setting, as was recently the case for a gene therapy for adenosine deaminase severe combined immunodeficiency (ADA-SCID) that had promising long-term results (*199*, *200*). Even if a treatment passes through all clinical trial phases and gets FDA approval, the potential retail price charged to cover manufacturing costs may be unaffordable to most patients without changes to the current health care infrastructure. Although many of the costs associated with developing new therapies are unavoidable, the motivation to make future CRISPR therapeutics widely accessible to those who need them will drive innovation of more efficient and cost-effective strategies for large-scale production that can meet regulatory standards.

chromosome. Each of these intermediary steps or requirements between the initial intravenous injection and the actual editing of the target locus has its own set of challenges that will necessitate both engineering and innovation to overcome. One of these is controlling delivery vehicle size (often restricted by cargo size), which may present a challenge for bypassing the vascular endothelium and interstitial space between blood vessels and target cells (*157*). This challenge has motivated efforts to engineer and discover smaller CRISPR editors that can be delivered more efficiently (*163*–*167*). Another challenge is preventing uptake and editing in off-target cells; this could be addressed by utilizing a targeting molecule, such as a single-chain variable fragment or glycoprotein, through conjugation to the CRISPR RNP (*168*). Other approaches include engineering or evolving the delivery vehicle to target specific cells (*160*). An alternative strategy to systemic delivery is direct injection to a specific tissue, which largely avoids editing other unintended tissues and organs. However, direct injection results in genome editing of relatively few cells within a localized space and is practical only for organs directly accessible to such injection (*161*, *169*).

In these various approaches, the cargo is delivered as one of three forms: a plasmid DNA (>6400 kDa for a 10-kb plasmid DNA) encoding the CRISPR-Cas9 and gRNA (either together or separately), the Cas9 mRNA (1400 kDa) and gRNA (34 kDa), or the Cas9-gRNA RNP (194 kDa) [reviewed in (*158*, *170*)]. DNA cargo is relatively stable compared with RNA or protein, but its delivery results in the slowest initiation of editing and offers lower control over the functional RNP concentration in the system at any given time. In some cases, DNA cargo can permit prolonged expression of Cas9, but this increases the probability of off-target effects and immunogenic reactions (*170*, *171*). Of the three, RNP delivery results in the fastest initiation of editing and generally results in lower off-target effects, but methods for delivering RNPs are still limited.

Currently, a variety of delivery methods exist for CRISPR gene editing in mammalian systems, though each has its own set of challenges and limitations. These are broadly divided into physical delivery, viral-based delivery, and synthetic material–based delivery [reviewed in (*158*, *159*)]. Common physical delivery approaches include microinjection and electroporation, which can deliver CRISPR editors in all three forms of cargo (*170*). These methods allow for controlled dosage and high-efficiency delivery; however, both methods are effectively limited to ex vivo delivery. Viral-based delivery methods include adeno-associated viruses (AAVs), adenoviruses (AdVs), and lentiviruses, which deliver CRISPR cargo in the form of plasmid DNA and offer high delivery efficiencies by harnessing delivery vehicles that viruses have had thousands of years to evolve. Of these, AAVs are the most promising for in vivo clinical use for CRISPR therapeutics, with notable examples being an ongoing clinical trial for Leber congenital amaurosis type 10 (*161*), a soon-to-begin trial for HIV (*172*), and advanced preclinical work for Hutchinson–Gilford progeria syndrome (*99*). However, one main limitation of AAVs is their low packaging capacity (*173*). Compared with AAVs, AdVs and lentiviruses offer higher packaging efficiencies and have also been used for CRISPR-Cas9 delivery but have faced other challenges, including immunogenicity concerns [reviewed in (*157*, *159*, *174*)]. Another important consideration for viral-based delivery is the high cost and labor-intensive production, especially for AAVs and AdVs (*175*), which represents a major challenge for large-scale manufacturing and patient treatment (Box 1). Compared with viral-based delivery, synthetic material–based delivery methods, which include LNPs (*176*), cationic polymers and peptides (*177*–*181*), and gold nanoparticles (*182*), are often safer and offer high levels of control and flexibility because they can be more easily tailored to all three forms of cargo (DNA, mRNA, and RNP) and be optimized for immunocompatibility (*157*). Notably, as mentioned above, LNPs are the delivery strategy used in the first systemic in vivo delivery of CRISPR in humans for the successful treatment of transthyretin amyloidosis (*110*). In general, however, these synthetic material–based delivery methods offer lower delivery efficiencies compared with viral-based methods (*159*), limiting their effectiveness for in vivo delivery to less accessible target organs. Although improvements can be made with further optimization, their maximum efficiency may be limited by the materials' bulky size and cationic nature, which result in poor interstitial dispersion. Recently, extracellular vesicles and virus-like particles

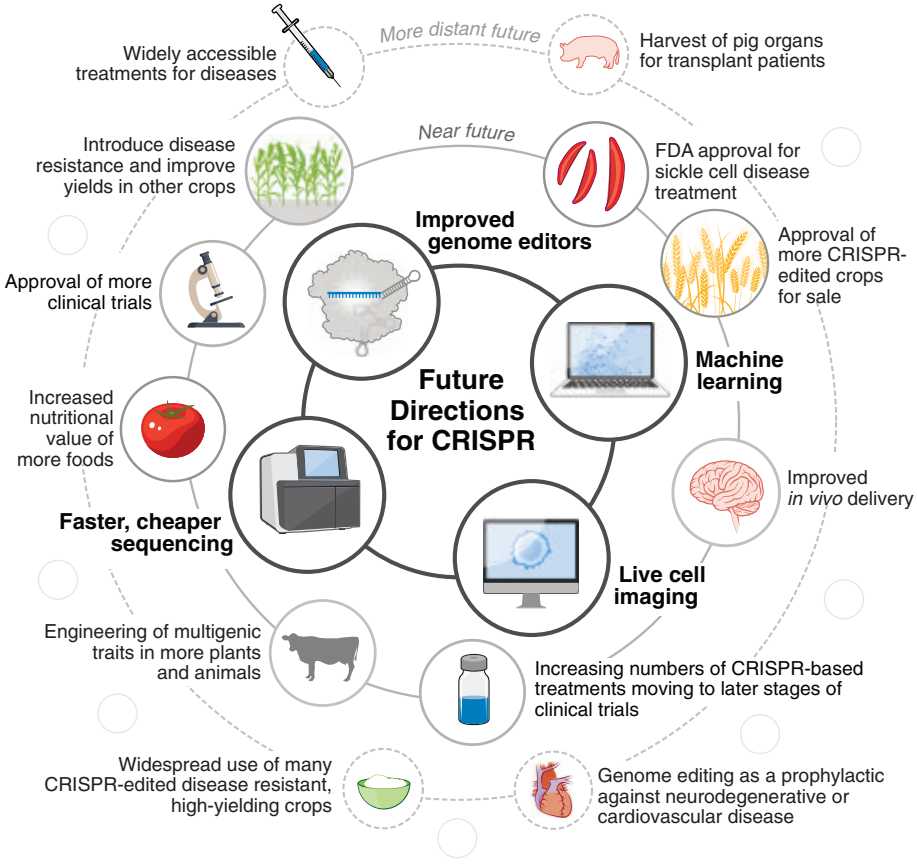


Fig. 4. Future directions: where we could be 10 years from now. The future of CRISPR genome editing will intersect with advances in technologies such as machine learning, live-cell imaging, and sequencing. In the near future, we may witness FDA approval for the first CRISPR-based medicine, as well as increasing numbers of CRISPR treatments moving to later stages of clinical trials and approval of new clinical trials using improved in vivo delivery methods. We expect approval of more CRISPR-edited crops for sale and more demonstrations of CRISPR used to engineer multigenic traits in plants and animals. In the more distant future, we may one day see many widely accessible CRISPR-based treatments and even use genome editing to safely harvest pig organs for transplant patients or as a prophylactic against disease. In agriculture, CRISPR may be routinely used to generate disease-resistant, high-yield crops to increase global food supply and security.

(VLPs) (assemblies of the viral envelope and/or structural proteins that can transduce cells but lack viral genetic material) are emerging as promising delivery platforms that utilize the strengths from both viral-based delivery and synthetic material–based delivery (*183*–*186*). These have the potential to achieve the high-delivery efficiencies of viral-based methods without the safety concerns of random transgene integration or prolonged expression of the editors. One especially exciting area of development is programming the cellular tropism of VLPs by using different envelope glycoproteins to target specific cell types, as recently demonstrated for ex vivo and in vivo delivery and editing (*185*, *186*). The future of CRISPR treatments in humans will largely depend on improving current delivery strategies, innovating new delivery modalities, discovering and engineering more compact CRISPR editors, or a combination of the above.

Beyond mammalian systems, advances are also being made for CRISPR reagent delivery in plants. The plant's cell wall, which has a size exclusion limit of 5 to 20 nm, presents a major challenge for cargo to pass through (*187*). The two predominant methods are delivery of plasmid DNA using agrobacterium, which integrates the transfer DNA into the plant genome, and particle bombardment, which physically breaches the cell wall barrier to introduce the cargo [reviewed in (*187*)]. The main drawback of these methods is random integration of the CRISPR cassette into the plant genome. The goal of enabling transgene-free breeding has motivated alternative approaches, including direct RNP delivery by polyethylene glycol–mediated cell transfection, particle bombardment, electroporation, or lipofection [reviewed in (*188*)]. Though these methods hold promise, the broad application of RNPs in many plant species requires further work in improving delivery efficiency and plant regeneration from edited protoplasts.

Current and future applications

The advent of programmable genome editing technologies has paved the way for applications of cell and gene therapy to treat and even cure disease. Although applications of CRISPR are too numerous to list comprehensively, the treatment of sickle cell disease (SCD) provides an excellent example of the value and risks ahead. There are now at least eight FDA-approved clinical trials of CRISPR-based therapies for SCD and related blood disorders that are ongoing or soon to begin, and FDA approval of the first is expected in 2023 (*189*). However, the challenges to widespread deployment of CRISPR cures for SCD remain formidable. For genome editing to become a standard of care, the field will need to address the difficulty in manufacturing edited cells for each individual patient, the logistical limitations conferred by the need for bone marrow transplantation, and the cost, which can run up to $2,000,000 per patient (*190*). But suppose a new delivery modality for genome editors obviates the need for both ex vivo cell editing and bone marrow transplantation? Such an advance would be game changing, propelling the field into a new era where much broader deployment of genomic therapies would be possible.

Beyond clinical applications, CRISPR is beginning to have impacts in agriculture and animal husbandry. CRISPR-edited foods are already beginning to enter the market. This includes a CRISPR-created tomato with enhanced nutritional qualities and two CRISPR-edited fish (a faster-growing tiger puffer and a red sea bream with greater edible yield), which have been approved for sale in Japan (*191*, *192*). Among these agronomic applications are many examples where CRISPR has enabled the precise "transfer" of small genetic changes that confer desirable traits from one variety of a species to another—a transfer that is either impossible or impractical using any other approach. Beyond small perturbations, CRISPR has also shown potential to generate new genetic variation and complex editing previously not seen in nature. A key example is the recent use of multiplexed editing to simultaneously knock out and activate different genes to introduce disease resistance in wheat and restore growth and yield (*193*). These constitute just the beginning of a wide range of genome editing advances that will increasingly affect our lives in the coming years.

CRISPR-Cas9, part of a bacterial immune system, uses an RNA-guided mechanism to recognize and cut DNA sequences. This fundamental biochemical activity forms the basis for genome editing technology that spans all realms of basic and applied biological research, ranging from developmental biology and plant genetics to sickle cell disease and animal husbandry. The discovery of new CRISPR-based and CRISPR-related enzymes has accelerated rapidly, leading to continued expansion of understanding about the natural biology of these systems in microbes and their utility for genome editing in other cells and organisms. CRISPR-Cas9 and CRISPR-Cas12a, the most widely adopted genome editing enzymes, have become workhorses in research laboratories worldwide. Fundamental research enabled by genome editing speaks to the cross-cutting nature of CRISPR technology and to the timeliness of its arrival as a readily adaptable tool. These wide-ranging applications have in turn enabled expansion of the CRISPR toolbox to enable more precise editing of specific nucleotides or the targeted integration of new genetic information. In the decade ahead, genome editing research and applications will continue to accelerate and will increasingly intersect with technologies including machine learning, live cell imaging, and faster, cheaper DNA sequencing (Fig. 4). Just as the past decade has focused on CRISPR platforms, the decade ahead will increasingly apply those platforms for real-world impacts. In the clinic, we will undoubtedly see increased numbers and types of clinical trials, providing data that will guide next-generation gene and cell therapies. As clinical applications expand, there may be an opening for CRISPR to be used to protect health. For example, as safety and efficacy are established for disease treatment, genome editing might become a prophylactic against neurodegenerative or cardiovascular disease (*31*). Such opportunities would require detailed knowledge of the genetics of multigenic disease and the means to deliver to organs including the brain and heart—neither of which are small tasks. But the potential benefits may drive innovation in these areas well beyond what is possible today. In agriculture, CRISPR screening will provide increasing insights into paths to engineering multigenic traits in both plants and animals. Products generated using CRISPR—whether pig organs for transplant patients (*194*), rice that resists drought with increased yield, or microbiomes fine-tuned for health using CRISPR editing—may all become routine. CRISPR also serves as a notable example of the connection between curiosity-driven research, innovation, and technological breakthroughs. By continuing to explore the natural world, we will discover what cannot be imagined and put it to real-world use for the benefit of the planet.

REFERENCES AND NOTES

1. Y. Ishino, H. Shinagawa, K. Makino, M. Amemura, A. Nakata, Nucleotide sequence of the iap gene, responsible for alkaline phosphatase isozyme conversion in Escherichia coli, and identification of the gene product. *J. Bacteriol.* **169**, 5429–5433 (1987). doi: 10.1128/jb.169.12.5429-5433.1987; pmid: 3316184
2. F. J. M. Mojica, C. Díez-Villaseñor, J. García-Martínez, E. Soria, Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. *J. Mol. Evol.* **60**, 174–182 (2005). doi: 10.1007/s00239-004-0046-3; pmid: 15791728
3. S. J. J. Brouns *et al*., Small CRISPR RNAs guide antiviral defense in prokaryotes. *Science* **321**, 960–964 (2008). doi: 10.1126/science.1159689; pmid: 18703739
4. R. Barrangou *et al*., CRISPR provides acquired resistance against viruses in prokaryotes. *Science* **315**, 1709–1712 (2007). doi: 10.1126/science.1138140; pmid: 17379808
5. M. Jinek *et al*., A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **337**, 816–821 (2012). doi: 10.1126/science.1225829; pmid: 22745249
6. G. Gasiunas, R. Barrangou, P. Horvath, V. Siksnys, Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proc. Natl. Acad. Sci. U.S.A.* **109**, E2579–E2586 (2012). doi: 10.1073/pnas.1208507109; pmid: 22949671
7. G. J. Knott, J. A. Doudna, CRISPR-Cas guides the future of genetic engineering. *Science* **361**, 866–869 (2018). doi: 10.1126/science.aat5011; pmid: 30166482
8. L. S. Qi *et al*., Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell* **152**, 1173–1183 (2013). doi: 10.1016/j.cell.2013.02.022; pmid: 23452860

9. L. A. Gilbert *et al*., CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. *Cell* **154**, 442–451 (2013). doi: 10.1016/j.cell.2013.06.044; pmid: 23849981
10. K. Nishida *et al*., Targeted nucleotide editing using hybrid prokaryotic and vertebrate adaptive immune systems. *Science* **353**, aaf8729 (2016). doi: 10.1126/science.aaf8729; pmid: 27492474
11. A. C. Komor, Y. B. Kim, M. S. Packer, J. A. Zuris, D. R. Liu, Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature* **533**, 420–424 (2016). doi: 10.1038/nature17946; pmid: 27096365
12. A. V. Anzalone *et al*., Search-and-replace genome editing without double-strand breaks or donor DNA. *Nature* **576**, 149–157 (2019). doi: 10.1038/s41586-019-1711-4; pmid: 31634902
13. B. Liu *et al*., A split prime editor with untethered reverse transcriptase and circular RNA template. *Nat. Biotechnol.* **40**, 1388–1393 (2022). doi: 10.1038/s41587-022-01255-9; pmid: 35379962
14. E. V. Koonin, K. S. Makarova, Origins and evolution of CRISPR-Cas systems. *Philos. Trans. R. Soc. B* **374**, 20180087 (2019). doi: 10.1098/rstb.2018.0087; pmid: 30905284
15. J. Y. Wang, P. Pausch, J. A. Doudna, Structural biology of CRISPR-Cas immunity and genome editing enzymes. *Nat. Rev. Microbiol.* **20**, 641–656 (2022). doi: 10.1038/s41579-022-00739-4; pmid: 35562427
16. M. Burmistrz, K. Krakowski, A. Krawczyk-Balska, RNA-Targeting CRISPR-Cas Systems and Their Applications. *Int. J. Mol. Sci.* **21**, 1122 (2020). doi: 10.3390/ijms21031122; pmid: 32046217
17. A. A. Smargon, Y. J. Shi, G. W. Yeo, RNA-targeting CRISPR systems from metagenomic discovery to transcriptomic engineering. *Nat. Cell Biol.* **22**, 143–150 (2020). doi: 10.1038/s41556-019-0454-7; pmid: 32015437
18. B. Chen *et al*., Dynamic imaging of genomic loci in living human cells by an optimized CRISPR/Cas system. *Cell* **155**, 1479–1491 (2013). doi: 10.1016/j.cell.2013.12.001; pmid: 24360272
19. H. Wang *et al*., CRISPR-mediated live imaging of genome editing and transcription. *Science* **365**, 1301–1305 (2019). doi: 10.1126/science.aax7852; pmid: 31488703
20. D. Cologneri, M. Trinidad, J. A. Doudna, Precise transcript targeting by CRISPR-Csm complexes. bioRxiv 2022.06.20.496908 [Preprint] (2022); doi: 10.1101/2022.06.20.496908v1
21. A. East-Seletsky *et al*., Two distinct RNase activities of CRISPR-C2c2 enable guide-RNA processing and RNA detection. *Nature* **538**, 270–273 (2016). doi: 10.1038/nature19802; pmid: 27669025
22. J. S. Chen *et al*., CRISPR-Cas12a target binding unleashes indiscriminate single-stranded DNase activity. *Science* **360**, 436–439 (2018). doi: 10.1126/science.aar6245; pmid: 29449511
23. L. Cong *et al*., Multiplex genome engineering using CRISPR/Cas systems. *Science* **339**, 819–823 (2013). doi: 10.1126/science.1231143; pmid: 23287718
24. M. Jinek *et al*., RNA-programmed genome editing in human cells. *eLife* **2**, e00471 (2013). doi: 10.7554/eLife.00471; pmid: 23386978
25. P. Mali *et al*., RNA-guided human genome engineering via Cas9. *Science* **339**, 823–826 (2013). doi: 10.1126/science.1232033; pmid: 23287722
26. S. W. Cho, S. Kim, J. M. Kim, J.-S. Kim, Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease. *Nat. Biotechnol.* **31**, 230–232 (2013). doi: 10.1038/nbt.2507; pmid: 23360966
27. M. Jasin, R. Rothstein, Repair of strand breaks by homologous recombination. *Cold Spring Harb. Perspect. Biol.* **5**, a012740 (2013). doi: 10.1101/cshperspect.a012740; pmid: 24097900
28. K. S. Pawelczak, N. S. Gavande, P. S. VanderVere-Carozza, J. J. Turchi, Modulating DNA repair pathways to improve precision genome engineering. *ACS Chem. Biol.* **13**, 389–396 (2018). doi: 10.1021/acschembio.7b00777; pmid: 29210569
29. C. D. Yeh, C. D. Richardson, J. E. Corn, Advances in genome editing through control of DNA repair pathways. *Nat. Cell Biol.* **21**, 1468–1478 (2019). doi: 10.1038/s41556-019-0425-z; pmid: 31792376
30. Y. Kang, C. Chu, F. Wang, Y. Niu, CRISPR/Cas9-mediated genome editing in nonhuman primates. *Dis. Model. Mech.* **12**, dmm039982 (2019). doi: 10.1242/dmm.039982; pmid: 31636095
31. K. Musunuru *et al*., In vivo CRISPR base editing of PCSK9 durably lowers cholesterol in primates. *Nature* **593**, 429–434 (2021). doi: 10.1038/s41586-021-03534-y; pmid: 34012082
32. M. R. Capecchi, Gene targeting in mice: Functional analysis of the mammalian genome for the twenty-first century. *Nat. Rev. Genet.* **6**, 507–512 (2005). doi: 10.1038/nrg1619; pmid: 15931173
33. H. Bouabe, K. Okkenhaug, Gene targeting in mice: A review. *Methods Mol. Biol.* **1064**, 315–336 (2013). doi: 10.1007/978-1-62703-601-6_23; pmid: 23996268
34. H. Yang *et al*., One-step generation of mice carrying reporter and conditional alleles by CRISPR/Cas-mediated genome engineering. *Cell* **154**, 1370–1379 (2013). doi: 10.1016/j.cell.2013.08.022; pmid: 23992847
35. H. Wang *et al*., One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. *Cell* **153**, 910–918 (2013). doi: 10.1016/j.cell.2013.04.025; pmid: 23643243
36. H. Yang, H. Wang, R. Jaenisch, Generating genetically modified mice using CRISPR/Cas-mediated genome engineering. *Nat. Protoc.* **9**, 1956–1968 (2014). doi: 10.1038/nprot.2014.134; pmid: 25058643
37. T. Van Dyke, T. Jacks, Cancer modeling in the modern era: Progress and challenges. *Cell* **108**, 135–144 (2002). doi: 10.1016/S0092-8674(02)00621-9; pmid: 11832204
38. H. Lee, D. E. Yoon, K. Kim, Genome editing methods in animal models. *Anim. Cells Syst.* **24**, 8–16 (2020). doi: 10.1080/19768354.2020.1726462; pmid: 32158611
39. A. J. Modzelewski *et al*., Efficient mouse genome engineering by CRISPR-EZ technology. *Nat. Protoc.* **13**, 1253–1274 (2018). doi: 10.1038/nprot.2018.012; pmid: 29748649
40. S. Chen *et al*., CRISPR-READI: Efficient generation of knockin mice by CRISPR RNP electroporation and AAV donor infection. *Cell Rep.* **27**, 3780–3789.e4 (2019). doi: 10.1016/j.celrep.2019.05.103; pmid: 31242412
41. M. Ohtsuka *et al*., i-GONAD: A robust method for in situ germline genome engineering using CRISPR nucleases. *Genome Biol.* **19**, 25 (2018). doi: 10.1186/s13059-018-1400-x; pmid: 29482575
42. H. Mou, Z. Kennedy, D. G. Anderson, H. Yin, W. Xue, Precision cancer mouse models through genome editing with CRISPR-Cas9. *Genome Med.* **7**, 53 (2015). doi: 10.1186/s13073-015-0178-7; pmid: 26060510
43. J. G. Doench, Am I ready for CRISPR? A user's guide to genetic screens. *Nat. Rev. Genet.* **19**, 67–80 (2018). doi: 10.1038/nrg.2017.97; pmid: 29199283
44. C. Bock *et al*., High-content CRISPR screening. *Nat. Rev. Methods Primers* **2**, 8 (2022). doi: 10.1038/s43586-021-00093-4
45. A. Katti, B. J. Diaz, C. M. Caragine, N. E. Sanjana, L. E. Dow, CRISPR in cancer biology and therapy. *Nat. Rev. Cancer* **22**, 259–279 (2022). doi: 10.1038/s41568-022-00441-w; pmid: 35194172
46. M. B. Dong, K. Tang, X. Zhou, J. J. Zhou, S. Chen, Tumor immunology CRISPR screening: Present, past, and future. *Trends Cancer* **8**, 210–225 (2022). doi: 10.1016/j.trecan.2021.11.009; pmid: 34920978
47. N. E. Sanjana, O. Shalem, F. Zhang, Improved vectors and genome-wide libraries for CRISPR screening. *Nat. Methods* **11**, 783–784 (2014). doi: 10.1038/nmeth.3047; pmid: 25075903
48. K. R. Sanson *et al*., Optimized libraries for CRISPR-Cas9 genetic screens with multiple modalities. *Nat. Commun.* **9**, 5416 (2018). doi: 10.1038/s41467-018-07901-8; pmid: 30575746
49. M. Kampmann, CRISPRi and CRISPRa Screens in Mammalian Cells for Precision Biology and Medicine. *ACS Chem. Biol.* **13**, 406–416 (2018). doi: 10.1021/acschembio.7b00657; pmid: 29035510
50. L. Przybyla, L. A. Gilbert, A new era in functional genomics screens. *Nat. Rev. Genet.* **23**, 89–103 (2022). doi: 10.1038/s41576-021-00409-w; pmid: 34545248
51. G. M. Findlay, E. A. Boyle, R. J. Hause, J. C. Klein, J. Shendure, Saturation editing of genomic regions by multiplex homology-directed repair. *Nature* **513**, 120–123 (2014). doi: 10.1038/nature13695; pmid: 25141179
52. G. M. Findlay *et al*., Accurate classification of BRCA1 variants with saturation genome editing. *Nature* **562**, 217–222 (2018). doi: 10.1038/s41586-018-0461-z; pmid: 30209399
53. I. Meitlis *et al*., Multiplexed Functional Assessment of Genetic Variants in CARD11. *Am. J. Hum. Genet.* **107**, 1029–1043 (2020). doi: 10.1016/j.ajhg.2020.10.015; pmid: 33202260
54. R. E. Hanna *et al*., Massively parallel assessment of human variants with base editor screens. *Cell* **184**, 1064–1080.e20 (2021). doi: 10.1016/j.cell.2021.01.012; pmid: 33606977
55. R. Cuella-Martin *et al*., Functional interrogation of DNA damage response variants with base editing screens. *Cell* **184**, 1081–1097.e19 (2021). doi: 10.1016/j.cell.2021.01.041; pmid: 33606978
56. S. Erwood *et al*., Saturation variant interpretation using CRISPR prime editing. *Nat. Biotechnol.* **40**, 885–895 (2022). doi: 10.1038/s41587-021-01201-1; pmid: 35190686
57. S. Chen *et al*., Genome-wide CRISPR screen in a mouse model of tumor growth and metastasis. *Cell* **160**, 1246–1260 (2015). doi: 10.1016/j.cell.2015.02.038; pmid: 25748654
58. R. D. Chow *et al*., AAV-mediated direct in vivo CRISPR screen identifies functional suppressors in glioblastoma. *Nat. Neurosci.* **20**, 1329–1341 (2017). doi: 10.1038/nn.4620; pmid: 28805815
59. B. E. Michels *et al*., Pooled in vitro and in vivo CRISPR-Cas9 screening identifies tumor suppressors in human colon organoids. *Cell Stem Cell* **26**, 782–792.e7 (2020). doi: 10.1016/j.stem.2020.04.003; pmid: 32348727
60. C. Gaillochet, W. Develtere, T. B. Jacobs, CRISPR screens in plants: Approaches, guidelines, and future prospects. *Plant Cell* **33**, 794–813 (2021). doi: 10.1093/plcell/koab099; pmid: 33823021
61. R. T. Manguso *et al*., In vivo CRISPR screening identifies Ptpn2 as a cancer immunotherapy target. *Nature* **547**, 413–418 (2017). doi: 10.1038/nature23270; pmid: 28723893
62. A. Wroblewska *et al*., Protein barcodes enable high-dimensional single-cell CRISPR screens. *Cell* **175**, 1141–1155.e16 (2018). doi: 10.1016/j.cell.2018.09.022; pmid: 30343902
63. A. J. Rubin *et al*., Coupled Single-Cell CRISPR Screening and Epigenomic Profiling Reveals Causal Gene Regulatory Networks. *Cell* **176**, 361–376.e17 (2019). doi: 10.1016/j.cell.2018.11.022; pmid: 30580963
64. A. Dixit *et al*., Perturb-Seq: Dissecting molecular circuits with scalable single-cell RNA profiling of pooled genetic screens. *Cell* **167**, 1853–1866.e17 (2016). doi: 10.1016/j.cell.2016.11.038; pmid: 27984732
65. E. Papalexi *et al*., Characterizing the molecular regulation of inhibitory immune checkpoints with multimodal single-cell screens. *Nat. Genet.* **53**, 322–331 (2021). doi: 10.1038/s41588-021-00778-2; pmid: 33649593
66. P. Datlinger *et al*., Pooled CRISPR screening with single-cell transcriptome readout. *Nat. Methods* **14**, 297–301 (2017). doi: 10.1038/nmeth.4177; pmid: 28099430
67. E. P. Mimitou *et al*., Multiplexed detection of proteins, transcriptomes, clonotypes and CRISPR perturbations in single cells. *Nat. Methods* **16**, 409–412 (2019). doi: 10.1038/s41592-019-0392-0; pmid: 31011186
68. V. Armario Najera, R. M. Twyman, P. Christou, C. Zhu, Applications of multiplex genome editing in higher plants. *Curr. Opin. Biotechnol.* **59**, 93–102 (2019). doi: 10.1016/j.copbio.2019.02.015; pmid: 30978482
69. M. Abdelrahman, Z. Wei, J. S. Rohila, K. Zhao, Multiplex genome-editing technologies for revolutionizing plant biology and crop improvement. *Front. Plant Sci.* **12**, 721203 (2021). doi: 10.3389/fpls.2021.721203; pmid: 34691102
70. Y. Mao, J. R. Botella, Y. Liu, J.-K. Zhu, Gene editing in plants: Progress and challenges. *Natl. Sci. Rev.* **6**, 421–437 (2019). doi: 10.1093/nsr/nwz005; pmid: 34691892
71. Y. Zhang *et al*., Simultaneous modification of three homoeologs of TaEDR1 by genome editing enhances powdery mildew resistance in wheat. *Plant J.* **91**, 714–724 (2017). doi: 10.1111/tpj.13599; pmid: 28502081
72. H. Zhu, C. Li, C. Gao, Applications of CRISPR-Cas in agriculture and plant biotechnology. *Nat. Rev. Mol. Cell Biol.* **21**, 661–677 (2020). doi: 10.1038/s41580-020-00288-9; pmid: 32973356
73. H.-L. Xing *et al*., A CRISPR/Cas9 toolkit for multiplex genome editing in plants. *BMC Plant Biol.* **14**, 327 (2014). doi: 10.1186/s12870-014-0327-y; pmid: 25432517
74. L. G. Lowder *et al*., A CRISPR/Cas9 toolbox for multiplexed plant genome editing and transcriptional regulation. *Plant Physiol.* **169**, 971–985 (2015). doi: 10.1104/pp.15.00636; pmid: 26297141
75. X. Ma *et al*., A robust CRISPR/Cas9 system for convenient, high-efficiency multiplex genome editing in monocot and dicot plants. *Mol. Plant* **8**, 1274–1284 (2015). doi: 10.1016/j.molp.2015.04.007; pmid: 25917172
76. X. Tang *et al*., A single transcript CRISPR-Cas9 system for efficient genome editing in plants. *Mol. Plant* **9**, 1088–1091 (2016). doi: 10.1016/j.molp.2016.05.001 pmid: 27212389
77. K. Xie, B. Minkenberg, Y. Yang, Boosting CRISPR/Cas9 multiplex editing capability with the endogenous tRNA-processing system. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 3570–3575 (2015). doi: 10.1073/pnas.1420294112; pmid: 25733849

78. T. Čermák *et al*., A multipurpose toolkit to enable advanced genome engineering in plants. *Plant Cell* **29**, 1196–1217 (2017). doi: 10.1105/tpc.16.00922; pmid: 28522548
79. T. Li *et al*., Domestication of wild tomato is accelerated by genome editing. *Nat. Biotechnol.* **36**, 1160–1163 (2018). doi: 10.1038/nbt.4273; pmid: 30272676
80. E. Lacchini *et al*., CRISPR-mediated accelerated domestication of African rice landraces. *PLOS ONE* **15**, e0229782 (2020). doi: 10.1371/journal.pone.0229782; pmid: 32126126
81. H. Yu *et al*., A route to de novo domestication of wild allotetraploid rice. *Cell* **184**, 1156–1170.e14 (2021). doi: 10.1016/j.cell.2021.01.013; pmid: 33539781
82. D. Niu *et al*., Inactivation of porcine endogenous retrovirus in pigs using CRISPR-Cas9. *Science* **357**, 1303–1307 (2017). doi: 10.1126/science.aan4187; pmid: 28798043
83. R. J. Ihry *et al*., p53 inhibits CRISPR-Cas9 engineering in human pluripotent stem cells. *Nat. Med.* **24**, 939–946 (2018). doi: 10.1038/s41591-018-0050-6; pmid: 29892062
84. O. M. Enache *et al*., Cas9 activates the p53 pathway and selects for p53-inactivating mutations. *Nat. Genet.* **52**, 662–668 (2020). doi: 10.1038/s41588-020-0623-4; pmid: 32424350
85. Q. Yuan, X. Gao, Multiplex base- and prime-editing with drive-and-process CRISPR arrays. *Nat. Commun.* **13**, 2771 (2022). doi: 10.1038/s41467-022-30514-1; pmid: 35589728
86. R. Mishra, R. K. Joshi, K. Zhao, Base editing in crops: Current advances, limitations and future implications. *Plant Biotechnol. J.* **18**, 20–31 (2020). doi: 10.1111/pbi.13225; pmid: 31365173
87. H. A. Rees, D. R. Liu, Base editing: Precision chemistry on the genome and transcriptome of living cells. *Nat. Rev. Genet.* **19**, 770–788 (2018). doi: 10.1038/s41576-018-0059-1; pmid: 30323312
88. E. M. Porto, A. C. Komor, I. M. Slaymaker, G. W. Yeo, Base editing: Advances and therapeutic opportunities. *Nat. Rev. Drug Discov.* **19**, 839–859 (2020). doi: 10.1038/s41573-020-0084-6; pmid: 33077937
89. X. Li *et al*., Base editing with a Cpf1-cytidine deaminase fusion. *Nat. Biotechnol.* **36**, 324–327 (2018). doi: 10.1038/nbt.4102; pmid: 29553573
90. N. M. Gaudelli *et al*., Programmable base editing of A•T to G•C in genomic DNA without DNA cleavage. *Nature* **551**, 464–471 (2017). doi: 10.1038/nature24644; pmid: 29160308
91. D. B. T. Cox *et al*., RNA editing with CRISPR-Cas13. *Science* **358**, 1019–1027 (2017). doi: 10.1126/science.aaq0180; pmid: 29070703
92. I. C. Kurt *et al*., CRISPR C-to-G base editors for inducing targeted DNA transversions in human cells. *Nat. Biotechnol.* **39**, 41–46 (2021). doi: 10.1038/s41587-020-0609-x; pmid: 32690971
93. C. Xu *et al*., Programmable RNA editing with compact CRISPR-Cas13 systems from uncultivated microbes. *Nat. Methods* **18**, 499–506 (2021). doi: 10.1038/s41592-021-01124-4; pmid: 33941935
94. M. J. Landrum *et al*., ClinVar: Public archive of interpretations of clinically relevant variants. *Nucleic Acids Res.* **44**, D862–D868 (2016). doi: 10.1093/nar/gkv1222; pmid: 26582918
95. A. V. Anzalone, L. W. Koblan, D. R. Liu, Genome editing with CRISPR-Cas nucleases, base editors, transposases and prime editors. *Nat. Biotechnol.* **38**, 824–844 (2020). doi: 10.1038/s41587-020-0561-9; pmid: 32572269
96. J. H. Hu *et al*., Evolved Cas9 variants with broad PAM compatibility and high DNA specificity. *Nature* **556**, 57–63 (2018). doi: 10.1038/nature26155; pmid: 29512652
97. R. T. Walton, K. A. Christie, M. N. Whittaker, B. P. Kleinstiver, Unconstrained genome targeting with near-PAMless engineered CRISPR-Cas9 variants. *Science* **368**, 290–296 (2020). doi: 10.1126/science.aba8853; pmid: 32217751
98. W.-H. Yeh, H. Chiang, H. A. Rees, A. S. B. Edge, D. R. Liu, In vivo base editing of post-mitotic sensory cells. *Nat. Commun.* **9**, 2184 (2018). doi: 10.1038/s41467-018-04580-3; pmid: 29872041
99. L. W. Koblan *et al*., In vivo base editing rescues Hutchinson-Gilford progeria syndrome in mice. *Nature* **589**, 608–614 (2021). doi: 10.1038/s41586-020-03086-7; pmid: 33408413
100. K. Kingwell, Base editors hit the clinic. *Nat. Rev. Drug Discov.* **21**, 545–547 (2022). doi: 10.1038/d41573-022-00124-z; pmid: 35831515
101. B. P. Kleinstiver *et al*., High-fidelity CRISPR-Cas9 nucleases with no detectable genome-wide off-target effects. *Nature* **529**, 490–495 (2016). doi: 10.1038/nature16526; pmid: 26735016
102. A. Casini *et al*., A highly specific SpCas9 variant is identified by in vivo screening in yeast. *Nat. Biotechnol.* **36**, 265–271 (2018). doi: 10.1038/nbt.4066; pmid: 29431739
103. C. A. Vakulskas *et al*., A high-fidelity Cas9 mutant delivered as a ribonucleoprotein complex enables efficient gene editing in human hematopoietic stem and progenitor cells. *Nat. Med.* **24**, 1216–1224 (2018). doi: 10.1038/s41591-018-0137-0; pmid: 30082871
104. M. Bratovič *et al*., Bridge helix arginines play a critical role in Cas9 sensitivity to mismatches. *Nat. Chem. Biol.* **16**, 587–595 (2020). doi: 10.1038/s41589-020-0490-4; pmid: 32123387
105. X.-H. Zhang, L. Y. Tee, X.-G. Wang, Q.-S. Huang, S.-H. Yang, Off-target effects in CRISPR/Cas9-mediated genome engineering. *Mol. Ther. Nucleic Acids* **4**, e264 (2015). doi: 10.1038/mtna.2015.37; pmid: 26575098
106. F. Heigwer, G. Kerr, M. Boutros, E-CRISP: Fast CRISPR target site identification. *Nat. Methods* **11**, 122–123 (2014). doi: 10.1038/nmeth.2812; pmid: 24481216
107. S. Bae, J. Park, J.-S. Kim, Cas-OFFinder: A fast and versatile algorithm that searches for potential off-target sites of Cas9 RNA-guided endonucleases. *Bioinformatics* **30**, 1473–1475 (2014). doi: 10.1093/bioinformatics/btu048; pmid: 24463181
108. K. Hiranniramol, Y. Chen, W. Liu, X. Wang, Generalizable sgRNA design for improved CRISPR/Cas9 editing efficiency. *Bioinformatics* **36**, 2684–2689 (2020). doi: 10.1093/bioinformatics/btaa041; pmid: 31971562
109. H. Frangoul *et al*., CRISPR-Cas9 gene editing for sickle cell disease and β-Thalassemia. *N. Engl. J. Med.* **384**, 252–260 (2021). doi: 10.1056/NEJMoa2031054; pmid: 33283989
110. J. D. Gillmore *et al*., CRISPR-Cas9 In vivo gene editing for transthyretin amyloidosis. *N. Engl. J. Med.* **385**, 493–502 (2021). doi: 10.1056/NEJMoa2107454; pmid: 34215024
111. J. Grünewald *et al*., Transcriptome-wide off-target RNA editing induced by CRISPR-guided DNA base editors. *Nature* **569**, 433–437 (2019). doi: 10.1038/s41586-019-1161-z; pmid: 30995674
112. C. Zhou *et al*., Off-target RNA mutation induced by DNA base editing and its elimination by mutagenesis. *Nature* **571**, 275–278 (2019). doi: 10.1038/s41586-019-1314-0; pmid: 31181567
113. H. A. Rees, C. Wilson, J. L. Doman, D. R. Liu, Analysis and minimization of cellular RNA editing by DNA adenine base editors. *Sci. Adv.* **5**, eaax5717 (2019). doi: 10.1126/sciadv.aax5717; pmid: 31086823
114. J. Grünewald *et al*., CRISPR DNA base editors with reduced RNA off-target and self-editing activities. *Nat. Biotechnol.* **37**, 1041–1048 (2019). doi: 10.1038/s41587-019-0236-6; pmid: 31477922
115. Y. Yu *et al*., Cytosine base editors with minimized unguided DNA and RNA off-target events and high on-target activity. *Nat. Commun.* **11**, 2052 (2020). doi: 10.1038/s41467-020-15887-5; pmid: 32345976
116. S. H. Chu *et al*., Rationally Designed Base Editors for Precise Editing of the Sickle Cell Disease Mutation. *CRISPR J.* **4**, 169–177 (2021). doi: 10.1089/crispr.2020.0144; pmid: 33876959
117. X. Liu, S. Wang, D. Ai, Predicting CRISPR/Cas9 Repair Outcomes by Attention-Based Deep Learning Framework. *Cells* **11**, 1847 (2022). doi: 10.3390/cells11111847; pmid: 35681543
118. V. R. Li, Z. Zhang, O. G. Troyanskaya, CROTON: An automated and variant-aware deep learning framework for predicting CRISPR/Cas9 editing outcomes. *Bioinformatics* **37**, i342–i348 (2021). doi: 10.1093/bioinformatics/btab268; pmid: 34252931
119. W. Chen *et al*., Massively parallel profiling and predictive modeling of the outcomes of CRISPR/Cas9-mediated double-strand break repair. *Nucleic Acids Res.* **47**, 7989–8003 (2019). doi: 10.1093/nar/gkz487; pmid: 31165867
120. S. E. Howden *et al*., Simultaneous reprogramming and gene correction of patient fibroblasts. *Stem Cell Reports* **5**, 1109–1118 (2015). doi: 10.1016/j.stemcr.2015.10.009; pmid: 26584543
121. F. T. Merkle *et al*., Efficient CRISPR-Cas9-mediated generation of knockin human pluripotent stem cells lacking undesired mutations at the targeted locus. *Cell Rep.* **11**, 875–883 (2015). doi: 10.1016/j.celrep.2015.04.007; pmid: 25937281
122. C. D. Richardson, G. J. Ray, M. A. DeWitt, G. L. Curie, J. E. Corn, Enhancing homology-directed genome editing by catalytically active and inactive CRISPR-Cas9 using asymmetric donor DNA. *Nat. Biotechnol.* **34**, 339–344 (2016). doi: 10.1038/nbt.3481; pmid: 26789497
123. S. Lin, B. T. Staahl, R. K. Alla, J. A. Doudna, Enhanced homology-directed human genome engineering by controlled timing of CRISPR/Cas9 delivery. *eLife* **3**, e04766 (2014). doi: 10.7554/eLife.04766; pmid: 25497837
124. T. Gutschner, M. Haemmerle, G. Genovese, G. F. Draetta, L. Chin, Post-translational regulation of Cas9 during G1 enhances homology-directed repair. *Cell Rep.* **14**, 1555–1566 (2016). doi: 10.1016/j.celrep.2016.01.019; pmid: 26854237
125. A. Lomova *et al*., Improving gene editing outcomes in human hematopoietic stem and progenitor cells by temporal control of DNA repair. *Stem Cells* **37**, 284–294 (2019). doi: 10.1002/stem.2935; pmid: 30372555
126. X. Ling *et al*., Improving the efficiency of precise genome editing with site-specific Cas9-oligonucleotide conjugates. *Sci. Adv.* **6**, eaaz0051 (2020). doi: 10.1126/sciadv.aaz0051; pmid: 32494588
127. E. Brunet, M. Jasin, Induction of Chromosomal Translocations with CRISPR-Cas9 and Other Nucleases: Understanding the Repair Mechanisms That Give Rise to Translocations. *Adv. Exp. Med. Biol.* **1044**, 15–25 (2018). doi: 10.1007/978-981-13-0593-1_2; pmid: 29956288
128. G. Cullot *et al*., CRISPR-Cas9 genome editing induces megabase-scale chromosomal truncations. *Nat. Commun.* **10**, 1136 (2019). doi: 10.1038/s41467-019-09006-2; pmid: 30850590
129. A. C. Komor *et al*., Improved base excision repair inhibition and bacteriophage Mu Gam protein yields C:G-to-T:A base editors with higher efficiency and product purity. *Sci. Adv.* **3**, eaao4774 (2017). doi: 10.1126/sciadv.aao4774; pmid: 28875174
130. M. Arbab *et al*., Determinants of base editing outcomes from target library analysis and machine learning. *Cell* **182**, 463–480.e30 (2020). doi: 10.1016/j.cell.2020.05.037; pmid: 32533916
131. A. V. Lavrov, G. G. Varenikov, M. Y. Skoblov, Genome scale analysis of pathogenic variants targetable for single base editing. *BMC Med. Genomics* **13**, 80 (2020). doi: 10.1186/s12920-020-00735-8; pmid: 32948190
132. S.-Y. Yu *et al*., Increasing the targeting scope of CRISPR base editing system beyond NGG. *CRISPR J.* **5**, 187–202 (2022). doi: 10.1089/crispr.2021.0109; pmid: 35238621
133. Q. Wang *et al*., A general theoretical framework to design base editors with reduced bystander effects. *Nat. Commun.* **12**, 6529 (2021). doi: 10.1038/s41467-021-26789-5; pmid: 34764246
134. D. Zhao *et al*., Imperfect guide-RNA (igRNA) enables CRISPR single-base editing with ABE and CBE. *Nucleic Acids Res.* **50**, 4161–4170 (2022). doi: 10.1093/nar/gkac201; pmid: 35349689
135. N. H. Cho *et al*., OpenCell: Endogenous tagging for the cartography of human cellular organization. *Science* **375**, eabi6983 (2022). doi: 10.1126/science.abi6983; pmid: 35271311
136. S. Burns, "Consecutive Genome Editing in Non-Human Primate Achieves Durable Production of Human Alpha-1 Antitrypsin and Reduction of the Native Protein" presented at The European Society Of Gene And Cell Therapy Annual Congress, 20 Oct, 2021.
137. S. P. Iyer *et al*, "The COBALT - LYM study of CTX130: A phase 1 dose escalation study of CD70-targeted allogeneic CRISPR-Cas9–engineered CAR T cells in patients with relapsed/refractory (R/R) T-cell malignancies" presented at the European Hematology Association Annual Meeting, 11 June 2022.
138. L. J. Nastoupil *et al*., First-in-Human Trial of CB-010, a CRISPR-Edited Allogeneic Anti-CD19 CAR-T Cell Therapy with a PD-1 Knock Out, in Patients with Relapsed or Refractory B Cell Non-Hodgkin Lymphoma (ANTLER Study). *Hemasphere* **6**, 1337–1338 (2022). doi: 10.1038/s41434-021-00263-9; pmid: 34031549
139. J. Scholefield, P. T. Harrison, Prime editing - an update on the field. *Gene Ther.* **28**, 396–401 (2021). doi: 10.1038/s41434-021-00263-9; pmid: 34031549
140. A. Kantor, M. E. McClements, R. E. MacLaren, CRISPR-Cas9 DNA base-editing and prime-editing. *Int. J. Mol. Sci.* **21**, 6240 (2020). doi: 10.3390/ijms21176240; pmid: 32872311
141. M. M. Hassan, G. Yuan, J.-G. Chen, G. A. Tuskan, X. Yang, Prime editing technology and its prospects for future applications in plant biology research. *BioDesign Research* **2020**, 1–14 (2020). doi: 10.34133/2020/9350905
142. I. F. Schene *et al*., Prime editing for functional repair in patient-derived disease models. *Nat. Commun.* **11**, 5352 (2020). doi: 10.1038/s41467-020-19136-7; pmid: 33097693
143. P. Gao *et al*., Prime editing in mice reveals the essentiality of a single base in driving tissue-specific gene expression. *Genome Biol.* **22**, 83 (2021). doi: 10.1186/s13059-021-02304-3; pmid: 33722289
144. M. H. Geurts *et al*., Evaluating CRISPR-based prime editing for cancer modeling and CFTR repair in organoids. *Life Sci.*

Alliance **4**, e202000940 (2021). doi: 10.26508/lsa.202000940; pmid: 34373320
145. J. W. Nelson *et al*., Engineered pegRNAs improve prime editing efficiency. *Nat. Biotechnol.* **40**, 402–410 (2022). doi: 10.1038/s41587-021-01039-7; pmid: 34608327
146. S. E. Klompe, P. L. H. Vo, T. S. Halpin-Healy, S. H. Sternberg, Transposon-encoded CRISPR-Cas systems direct RNA-guided DNA integration. *Nature* **571**, 219–225 (2019). doi: 10.1038/s41586-019-1323-z; pmid: 31189177
147. J. Strecker *et al*., RNA-guided DNA insertion with CRISPR-associated transposases. *Science* **365**, 48–53 (2019). doi: 10.1126/science.aax9181; pmid: 31171706
148. J. E. Peters, K. S. Makarova, S. Shmakov, E. V. Koonin, Recruitment of CRISPR-Cas systems by Tn7-like transposons. *Proc. Natl. Acad. Sci. U.S.A.* **114**, E7358–E7366 (2017). doi: 10.1073/pnas.1709035114; pmid: 28811374
149. J. R. Rybarski, K. Hu, A. M. Hill, C. O. Wilke, I. J. Finkelstein, Metagenomic discovery of CRISPR-associated transposons. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2112279118 (2021). doi: 10.1073/pnas.2112279118; pmid: 34845024
150. Y. Zhang *et al*., Multicopy Chromosomal Integration Using CRISPR-Associated Transposases. *ACS Synth. Biol.* **9**, 1998–2008 (2020). doi: 10.1021/acssynbio.0c00073; pmid: 32551502
151. P. L. H. Vo *et al*., CRISPR RNA-guided integrases for high-efficiency, multiplexed bacterial genome engineering. *Nat. Biotechnol.* **39**, 480–489 (2021). doi: 10.1038/s41587-020-00745-y; pmid: 33230293
152. Y. Wang, Y.-Y. Yau, D. Perkins-Balding, J. G. Thomson, Recombinase technology: Applications and possibilities. *Plant Cell Rep.* **30**, 267–285 (2011). doi: 10.1007/s00299-010-0938-1; pmid: 20972794
153. G. Meinke, A. Bohm, J. Hauber, M. T. Pisabarro, F. Buchholz, Cre recombinase and other tyrosine recombinases. *Chem. Rev.* **116**, 12785–12820 (2016). doi: 10.1021/acs.chemrev.6b00077; pmid: 27163859
154. J. L. Bessen *et al*., High-resolution specificity profiling and off-target prediction for site-specific DNA recombinases. *Nat. Commun.* **10**, 1937 (2019). doi: 10.1038/s41467-019-09987-0; pmid: 31028261
155. E. I. Ioannidi *et al*., Drag-and-drop genome insertion without DNA cleavage with CRISPR-directed integrases. bioRxiv 2021.11.01.466786 [Preprint] (2021); doi: 10.1101/2021.11.01.466786
156. A. V. Anzalone *et al*., Programmable deletion, replacement, integration and inversion of large DNA sequences with twin prime editing. *Nat. Biotechnol.* **40**, 731–740 (2022). doi: 10.1038/s41587-021-01133-w; pmid: 34887556
157. S. Tong, B. Moyo, C. M. Lee, K. Leong, G. Bao, Engineered materials for in vivo delivery of genome-editing machinery. *Nat. Rev. Mater.* **4**, 726–737 (2019). doi: 10.1038/s41578-019-0145-9; pmid: 34094589
158. C. A. Lino, J. C. Harper, J. P. Carney, J. A. Timlin, Delivering CRISPR: A review of the challenges and approaches. *Drug Deliv.* **25**, 1234–1257 (2018). doi: 10.1080/10717544.2018.1474964; pmid: 29801422
159. B. H. Yip, Recent Advances in CRISPR/Cas9 Delivery Strategies. *Biomolecules* **10**, 839 (2020). doi: 10.3390/biom10060839; pmid: 32486234
160. J. van Haasteren, J. Li, O. J. Scheideler, N. Murthy, D. V. Schaffer, The delivery challenge: Fulfilling the promise of therapeutic genome editing. *Nat. Biotechnol.* **38**, 845–855 (2020). doi: 10.1038/s41587-020-0565-5; pmid: 32601435
161. M. L. Maeder *et al*., Development of a gene-editing approach to restore vision loss in Leber congenital amaurosis type 10. *Nat. Med.* **25**, 229–233 (2019). doi: 10.1038/s41591-018-0327-9; pmid: 30664785
162. E. A. Pierce, W. F. Chatlos, "BRILLIANCE: A Phase 1/2 Single Ascending Dose Study of EDIT-101, an in vivo CRISPR Gene Editing Therapy in CEP290-Related Retinal Degeneration" presented at The European Society Of Gene And Cell Therapy Annual Congress, 21 Oct, 2021.
163. X. Xu *et al*., Engineered miniature CRISPR-Cas system for mammalian genome regulation and editing. *Mol. Cell* **81**, 4333–4345.e4 (2021). doi: 10.1016/j.molcel.2021.08.008; pmid: 34480847
164. P. Pausch *et al*., CRISPR-CasΦ from huge phages is a hypercompact genome editor. *Science* **369**, 333–337 (2020). doi: 10.1126/science.abb1400; pmid: 32675376
165. A. Shams *et al*., Comprehensive deletion landscape of CRISPR-Cas9 identifies minimal RNA-guided DNA-binding modules. *Nat. Commun.* **12**, 5664 (2021). doi: 10.1038/s41467-021-25992-8; pmid: 34580310
166. M. J. Schmidt *et al*., Improved CRISPR genome editing using small highly active and specific engineered RNA-guided nucleases. *Nat. Commun.* **12**, 4219 (2021). doi: 10.1038/s41467-021-24454-5; pmid: 34244505
167. L. B. Harrington *et al*., Programmed DNA destruction by miniature CRISPR-Cas14 enzymes. *Science* **362**, 839–842 (2018). doi: 10.1126/science.aav4294; pmid: 30337455
168. R. Rouet *et al*., Receptor-Mediated Delivery of CRISPR-Cas9 Endonuclease for Cell-Type-Specific Gene Editing. *J. Am. Chem. Soc.* **140**, 6596–6603 (2018). doi: 10.1021/jacs.8b01551; pmid: 29668265
169. B. T. Staahl *et al*., Efficient genome editing in the mouse brain by local delivery of engineered Cas9 ribonucleoprotein complexes. *Nat. Biotechnol.* **35**, 431–434 (2017). doi: 10.1038/nbt.3806; pmid: 28191903
170. E. A. Taha, J. Lee, A. Hotta, Delivery of CRISPR-Cas tools for in vivo genome editing therapy: Trends and challenges. *J. Control. Release* **342**, 345–361 (2022). doi: 10.1016/j.jconrel.2022.01.013; pmid: 35026352
171. C. T. Charlesworth *et al*., Identification of preexisting adaptive immunity to Cas9 proteins in humans. *Nat. Med.* **25**, 249–254 (2019). doi: 10.1038/s41591-018-0326-x; pmid: 30692695
172. Excision BioTherapeutics, Excision Receives FDA Clearance of IND for Phase 1/2 Trial of EBT-101 CRISPR-Based Therapeutic for Treatment of HIV (2021); https://www.globenewswire.com/news-release/2021/09/15/2297456/0/en/Excision-Receives-FDA-Clearance-of-IND-for-Phase-1-2-Trial-of-EBT-101-CRISPR-Based-Therapeutic-for-Treatment-of-HIV.html.
173. Z. Wu, H. Yang, P. Colosi, Effect of genome size on AAV vector packaging. *Mol. Ther.* **18**, 80–86 (2010). doi: 10.1038/mt.2009.255; pmid: 19904234
174. C. L. Xu, M. Z. C. Ruan, V. B. Mahajan, S. H. Tsang, Viral Delivery Systems for CRISPR. *Viruses* **11**, 28 (2019). doi: 10.3390/v11010028; pmid: 30621179
175. J. C. M. van der Loo, J. F. Wright, Progress and challenges in viral vector manufacturing. *Hum. Mol. Genet.* **25** (R1), R42–R52 (2016). doi: 10.1093/hmg/ddv451; pmid: 26519140
176. A. Pensado, B. Seijo, A. Sanchez, Current strategies for DNA therapy based on lipid nanocarriers. *Expert Opin. Drug Deliv.* **11**, 1721–1731 (2014). doi: 10.1517/17425247.2014.935337; pmid: 25046195
177. S. Ramakrishna *et al*., Gene disruption by cell-penetrating peptide-mediated delivery of Cas9 protein and guide RNA. *Genome Res.* **24**, 1020–1027 (2014). doi: 10.1101/gr.171264.113; pmid: 24696462
178. B. Suresh, S. Ramakrishna, H. Kim, Cell-penetrating peptide-mediated delivery of Cas9 protein and guide RNA for genome editing. *Methods Mol. Biol.* **1507**, 81–94 (2017). doi: 10.1007/978-1-4939-6518-2_7; pmid: 27832534
179. X. Gao *et al*., Treatment of autosomal dominant hearing loss by in vivo delivery of genome editing agents. *Nature* **553**, 217–221 (2018). doi: 10.1038/nature25164; pmid: 29258297
180. Z. Chen *et al*., Targeted delivery of CRISPR/Cas9-mediated cancer gene therapy via liposome-templated hydrogel nanoparticles. *Adv. Funct. Mater.* **27**, 1703036 (2017). doi: 10.1002/adfm.201703036; pmid: 29755309
181. H.-X. Wang *et al*., Nonviral gene editing via CRISPR/Cas9 delivery by membrane-disruptive and endosomolytic helical polypeptide. *Proc. Natl. Acad. Sci. U.S.A.* **115**, 4903–4908 (2018). doi: 10.1073/pnas.1712963115; pmid: 29686087
182. K. Lee *et al*., Nanoparticle delivery of Cas9 ribonucleoprotein and donor DNA *in vivo* induces homology-directed DNA repair. *Nat. Biomed. Eng.* **1**, 889–901 (2017). doi: 10.1038/s41551-017-0137-2; pmid: 29805845
183. P. E. Mangeot *et al*., Genome editing in primary cells and in vivo using viral-derived Nanoblades loaded with Cas9-sgRNA ribonucleoproteins. *Nat. Commun.* **10**, 45 (2019). doi: 10.1038/s41467-018-07845-z; pmid: 30604748
184. P. Gee *et al*., Extracellular nanovesicles for packaging of CRISPR-Cas9 protein and sgRNA to induce therapeutic exon skipping. *Nat. Commun.* **11**, 1334 (2020). doi: 10.1038/s41467-020-14957-y; pmid: 32170079
185. S. Banskota *et al*., Engineered virus-like particles for efficient in vivo delivery of therapeutic proteins. *Cell* **185**, 250–265.e16 (2022). doi: 10.1016/j.cell.2021.12.021; pmid: 35021064
186. J. R. Hamilton *et al*., Targeted delivery of CRISPR-Cas9 and transgenes enables complex immune cell engineering. *Cell Rep.* **35**, 109207 (2021). doi: 10.1016/j.celrep.2021.109207; pmid: 34077734
187. R. Ghogare, Y. Ludwig, G. M. Bueno, I. H. Slamet-Loedin, A. Dhingra, Genome editing reagent delivery in plants. *Transgenic Res.* **30**, 321–335 (2021). doi: 10.1007/s11248-021-00239-w; pmid: 33728594
188. Y. Zhang, B. Iaffaldano, Y. Qi, CRISPR ribonucleoprotein-mediated genetic engineering in plants. *Plant Commun.* **2**, 100168 (2021). doi: 10.1016/j.xplc.2021.100168; pmid: 33898980
189. F. D. Urnov, The Cas9 hammer-and sickle: A challenge for genome editors. *CRISPR J.* **4**, 6–13 (2021). doi: 10.1089/crispr.2021.29120.fur; pmid: 33616446
190. S. Demirci, A. Leonard, K. Essawi, J. F. Tisdale, CRISPR-Cas9 to induce fetal hemoglobin for the treatment of sickle cell disease. *Mol. Ther. Methods Clin. Dev.* **23**, 276–285 (2021). doi: 10.1016/j.omtm.2021.09.010; pmid: 34729375
191. E. Waltz, GABA-enriched tomato is first CRISPR-edited food to enter market. *Nat. Biotechnol.* **40**, 9–11 (2022). doi: 10.1038/d41587-021-00026-2; pmid: 34907351
192. Japan embraces CRISPR-edited fish. *Nat. Biotechnol.* **40**, 10 (2022). doi: 10.1038/s41587-021-01197-8; pmid: 34969964
193. S. Li *et al*., Genome-edited powdery mildew resistance in wheat without growth penalties. *Nature* **602**, 455–460 (2022). doi: 10.1038/s41586-022-04395-9; pmid: 35140403
194. K. Servick, CRISPR slices virus genes out of pigs, but will it make organ transplants to humans safer? *Science* (2017). doi: 10.1126/science.aan7227
195. H. Ledford, Super-precise new CRISPR tool could tackle a plethora of genetic diseases. *Nature* **602**, 464–465 (2019), doi: 10.1038/d41586-019-03164-5
196. C. Brokowski, M. Adli, CRISPR Ethics: Moral Considerations for Applications of a Powerful Tool. *J. Mol. Biol.* **431**, 88–101 (2019).
197. D. Zhang *et al*., Genome editing with the CRISPR-Cas system: an art, ethics and global regulatory perspective. *Plant Biotechnol. J.* **18**, 1651–1669 (2020).
198. A. A. Abraham, J. F. Tisdale, Gene therapy for sickle cell disease: Moving from the bench to the bedside. *Blood.* **138**, 932–941 (2021).
199. D. B. Kohn *et al*., Autologous Ex Vivo Lentiviral Gene Therapy for Adenosine Deaminase Deficiency. *N. Engl. J. Med.* **384**, 2002–2013 (2021).
200. M. Gelman, Orchard discreetly axes license for "bubble boy" gene therapy, despite recent positive long-term data. (Endpoints News, 2021); https://endpts.com/orchard-discretely-axes-license-for-bubble-boy-gene-therapy-despite-recent-positive-long-term-data-update.

ACKNOWLEDGMENTS

Many people have made important contributions to the CRISPR field, and we apologize to those whose work was not cited due to space limitations. We thank T. Tolpa for helpful discussions and artistic renderings of the figures. We thank P. H. Yoon for contributing the analysis of the imprecision of correcting most SNPs using current base editors and M. J. Kan and M. Trinidad for helpful discussions. We thank P. H. Yoon, M. J. Kan, C. A. Tsuchida, F. D. Urnov, and J. H. D. Cate for critical reading of the draft and thoughtful feedback. **Funding:** J.Y.W. was supported by a US National Science Foundation Graduate Fellowship. J.A.D. is a Howard Hughes Medical Institute investigator. **Author contributions:** Both authors contributed to writing and editing the manuscript. **Competing interests:** J.A.D. is a cofounder of Caribou Biosciences, Editas Medicine, Scribe Therapeutics, Intellia Therapeutics and Mammoth Biosciences. J.A.D. is a scientific advisory board member of Vertex, Caribou Biosciences, Intellia Therapeutics, Scribe Therapeutics, Mammoth Biosciences, Algen Biotechnologies, Felix Biosciences, The Column Group and Inari. J.A.D. is the chief science advisor of Sixth Street, is on the board of directors at Altos, Johnson & Johnson, and Tempus, and has research projects sponsored by AppleTree Partners and Roche. The Regents of the University of California have patents issued and pending for CRISPR technologies on which J.A.D. is named as an inventor. J.Y.W. declares no competing interests. **Data and materials availability:** All data are available from the authors. **License information:** Copyright © 2023 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. https://www.sciencemag.org/about/science-licenses-journal-article-reuse

Submitted 9 July 2022; accepted 12 December 2022
10.1126/science.add8643

Your Legacy to Science

AN ESTATE GIFT TO THE
AMERICAN ASSOCIATION FOR THE ADVANCEMENT OF SCIENCE





Since 1848, our founding year, the American Association for the Advancement of Science (AAAS) has been deeply committed to advancing science, engineering and innovation around the world for the benefit of all people.

By making AAAS a beneficiary of your will, trust, retirement plan or life insurance policy, you become a member of our 1848 Society, joining Thomas Edison, Alexander Graham Bell and the many distinguished individuals whose vision led to the creation of AAAS and our world-renowned journal, *Science*, so many years ago.

Unlike many of its peers, *Science* is not for-profit. Your estate gift would provide long-term financial stability and durable annual income that will support operations and competitive innovation for years to come. **This support is vital.**

"As a teacher and instructor, I bear responsibility for the younger generations. If you have extra resources, concentrate them on organizations, like AAAS, that are doing work for all."

—Prof. Elisabeth Ervin-Blankenheim, 1848 Society member

If you intend to include AAAS in your estate plans, provide this information to your lawyer or financial adviser:

Legal Name: American Association for the Advancement of Science
Federal Tax ID Number: 53-0196568
Address: 1200 New York Avenue, NW, Washington, DC 20005

If you would like more information on making an estate gift to AAAS, cut out and return the form below or send an email to philanthropy@aaas.org. Additional details are also available online at **www.aaas.org/1848Society**.

1848 SOCIETY AAAS

cut here

Yes, I would like more information about joining the AAAS 1848 Society.

PLEASE CONTACT ME AT:

Name: _____

Address: _____

City: _____ State: _____ Zip code: _____ Country: _____

Email: _____ Phone: _____

RETURN THIS FORM TO:
AAAS Office of Philanthropy and Strategic Partnerships • 1200 New York Avenue, NW • Washington, DC 20005 USA

RESEARCH ARTICLE SUMMARY

LYMPHOMA

BTG1 mutation yields supercompetitive B cells primed for malignant transformation

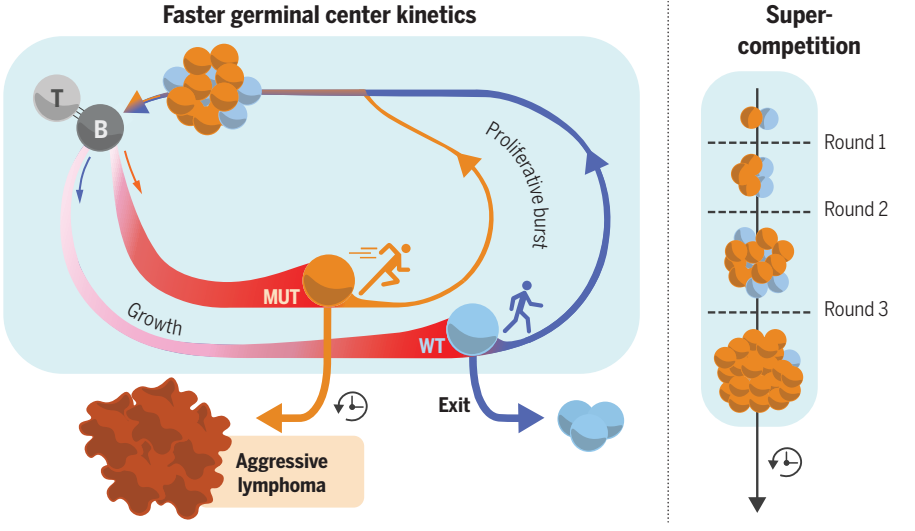
Coraline Mlynarczyk*, Matt Teater, Juhee Pae, Christopher R. Chin, Ling Wang, Theinmozhi Arulraj, Darko Barisic, Antonin Papin, Kenneth B. Hoehn, Ekaterina Kots, Jonatan Ersching, Arnab Bandyopadhyay, Ersilia Barin, Hui Xian Poh, Chiara M. Evans, Amy Chadburn, Zhengming Chen, Hao Shen, Hannah M. Isles, Benedikt Pelzer, Ioanna Tsialta, Ashley S. Doane, Huimin Geng, Muhammad Hassan Rehman, Jonah Melnick, Wyatt Morgan, Diu T. T. Nguyen, Olivier Elemento, Michael G. Kharas, Samie R. Jaffrey, David W. Scott, George Khelashvili, Michael Meyer-Hermann, Gabriel D. Victora, Ari Melnick*

INTRODUCTION: Diffuse large B cell lymphomas (DLBCLs) are aggressive malignancies of which ~40% fail to respond or become refractory to treatment. Mechanisms that mediate these poor outcomes are unknown. Most DLBCLs originate from specialized B cells participating in the germinal center (GC) reaction. GCs are transient structures within which antigen-activated B cells undergo proliferative bursting and somatic hypermutation of their immunoglobulin genes to improve affinity against the encountered antigen. High proliferation activity depletes GC B cells of essential biosynthetic precursors. Hence, GC B cells must endure an intense Darwinian competition process, during which they vigorously compete for access to positive selection signals from a limiting number of T follicular helper (T_{FH}) cells. Only few B cells that win this competition receive T_{FH} cell help, which triggers transient expression of the Myc proto-oncogene and biosynthetic growth programs. Unfortunately, GC B cells acquire abundant off-target mutations, making them prone to malignant transformation. Hence, GC B cell selection by T_{FH} cells requires rigorous, but still largely unknown, control mechanisms to prevent unwanted clonal expansion and malignant transformation of abnormal cells.

RATIONALE: Missense mutations of *BTG1* (B cell translocation gene 1) are specific to GC-derived B cell lymphomas, suggesting that their oncogenic function is cell-context specific. *BTG1* mutations help to genetically define a class of DLBCLs that manifest especially poor clinical outcomes and extensive dissemination. This prompted us to explore how *BTG1* mutations contribute to the pathogenesis of these clinically challenging tumors.

RESULTS: We generated mice for the B cell–conditional expression of the most frequent *BTG1* mutation [Gln^{36}→His (Q36H)]. $Btg1^{Q36H}$ cells almost completely outcompeted their wild-type counterparts, specifically in the GC. This competitive fitness manifested as a stronger induction of MYC-dependent growth programs. Wild-type BTG1 associated with many mRNAs, including those for *MYC* and MYC target genes, a function that was significantly lost by $BTG1^{Q36H}$. Mutant BTG1 expression enhanced *MYC* mRNA loading onto polysomes and induced slightly more-rapid MYC protein induction kinetics and a higher fraction of MYC-positive B cells in murine and human systems. These data suggested a lower threshold for BTG1 mutant cells to respond to T_{FH}-positive selection signals. Accordingly, Btg1 mutant GC B cells experienced faster cell cycle S phase transit and earlier entry into subsequent proliferative bursts. In Bcl2-driven lymphoma mouse models, $Btg1^{Q36H}$ markedly accelerated disease onset, shortened survival, and yielded particularly invasive DLBCL-like lymphomas. $Btg1^{Q36H}$ lymphomas were highly clonal and heavily mutated, reflecting a lymphomagenesis trajectory with increased selection and mutation rounds over time. In patients, *BTG1* mutations scored as strong genetic DLBCL drivers and independently associated with significantly inferior clinical outcomes.

CONCLUSION: Our data suggest that BTG1 serves as a critical gatekeeper controlling a key fitness checkpoint for natural selection of intensely competing B cells during the adaptive immune response. The precise targeting of BTG1 function by missense mutations causes a subtle biochemical impact, whose specific engagement during repetitive rounds of GC B cell positive selection likely explains its substantial oncogenic fitness effect. This reveals the fragility of constraints required to prevent competing B cells from recalling features of uncontrolled natural selection among unicellular organisms and highlights a fine-tuned balance between protection against infection and cancer risk. This effect is reminiscent of the Myc-dependent supercompetition first described during *Drosophila* development. Future therapeutic strategies could potentially take advantage of such evolutionary vulnerabilities to target cancer cell fitness. ■



Supercompetition in germinal centers leads to aggressive B cell lymphomas. GC B cells vigorously compete against each other to receive limiting T cell help, which activates MYC-dependent growth programs in preparation for subsequent proliferative bursting. BTG1 mutant GC B cells respond to T cell help signals with faster MYC induction kinetics and more-rapid entry into the next proliferative phase, which, over many GC rounds, allows them to outcompete wild-type cells, leading to highly fit and disseminated B cell lymphomas.

The list of author affiliations is available in the full article online.
*Corresponding author. Email: chm2065@med.cornell.edu (C.M.); amm2014@med.cornell.edu (A.M.)
Cite this article as C. Mlynarczyk *et al.*, *Science* **379**, eabj7412 (2023). DOI: 10.1126/science.abj7412

READ THE FULL ARTICLE AT
https://doi.org/10.1126/science.abj7412

ILLUSTRATION: CEYDA DURMAZ

RESEARCH ARTICLE

LYMPHOMA

BTG1 mutation yields supercompetitive B cells primed for malignant transformation

Coraline Mlynarczyk[1]*, Matt Teater[1], Juhee Pae[2], Christopher R. Chin[1,3,4,5], Ling Wang[1], Theinmozhi Arulraj[6], Darko Barisic[1], Antonin Papin[7], Kenneth B. Hoehn[8], Ekaterina Kots[9], Jonatan Ersching[2]†, Arnab Bandyopadhyay[6], Ersilia Barin[10], Hui Xian Poh[10], Chiara M. Evans[11,12], Amy Chadburn[7], Zhengming Chen[13], Hao Shen[1], Hannah M. Isles[1], Benedikt Pelzer[1], Ioanna Tsialta[1], Ashley S. Doane[1,5], Huimin Geng[14], Muhammad Hassan Rehman[1,15], Jonah Melnick[1], Wyatt Morgan[1]‡, Diu T. T. Nguyen[11,16], Olivier Elemento[3,17], Michael G. Kharas[11], Samie R. Jaffrey[10], David W. Scott[18], George Khelashvili[5,9], Michael Meyer-Hermann[6,19], Gabriel D. Victora[2], Ari Melnick[1]*

Multicellular life requires altruistic cooperation between cells. The adaptive immune system is a notable exception, wherein germinal center B cells compete vigorously for limiting positive selection signals. Studying primary human lymphomas and developing new mouse models, we found that mutations affecting *BTG1* disrupt a critical immune gatekeeper mechanism that strictly limits B cell fitness during antibody affinity maturation. This mechanism converted germinal center B cells into supercompetitors that rapidly outstrip their normal counterparts. This effect was conferred by a small shift in MYC protein induction kinetics but resulted in aggressive invasive lymphomas, which in humans are linked to dire clinical outcomes. Our findings reveal a delicate evolutionary trade-off between natural selection of B cells to provide immunity and potentially dangerous features that recall the more competitive nature of unicellular organisms.

During the humoral immune response, germinal center (GC) B cells undergo iterative rounds of natural selection to generate high-affinity B cell receptors. This process involves somatic hypermutation and clonal expansion of GC B cells and strictly depends on B cells competing to receive positive selection signals from T follicular helper (T_{FH}) cells (*1*, *2*). T_{FH} cell help allows a few GC B cells to survive and differentiate into memory B cells or plasma cells, or instead undergo additional rounds of proliferative bursting and mutagenesis. Transient induction of Myc- and mammalian target of rapamycin complex 1 (mTORC1)–biosynthetic programs is required to accumulate biomass in preparation for the subsequent proliferative burst (*3*–*5*).

The mutagenic and proliferative nature of GC B cells places them at considerable risk for malignant transformation (*6*, *7*). Hence, most immune system neoplasms arise from B cells that have transited the GC reaction. The most common of these are diffuse large B cell lymphomas (DLBCLs), which have classically been separated into germinal center B-like (GCB) and activated B-like (ABC) DLBCLs according to their transcriptional profiles (*8*). DLBCLs were more recently segregated into genetically defined entities with distinct biological characteristics and mutational profiles (*9*–*11*). Among these, the most clinically challenging, MCD/cluster 5, is an ABC-DLBCL subtype featuring extranodal dissemination and high fraction of proliferating cells. Recurrent somatic mutations in *MYD88* and *CD79B* drive proliferation of these lymphomas through chronic activation of Toll-like receptor and B cell receptor signaling, whereas *TBL1XR1* mutations endow them with a post-GC memory B cell phenotype (*12*, *13*). However, the basis for their marked fitness and aggressive phenotypes remains unknown.

Possible clues to these phenotypes may be gleaned from mechanistic deconvolution of other recurrent mutations. For example, somatic missense mutations of *BTG1* (B cell translocation gene 1) occur in up to 70% of MCD-DLBCL cases (*11*). BTG1 has been reported to regulate gene expression by interacting with transcription factors, mRNA stability regulatory proteins (CNOTs), or the arginine methyltransferase PRMT1 (*14*–*18*). Although BTG1 loss results in minor perturbations of early B cell development (*19*), and a few BTG1 mutants show differential CNOT interaction and tethered-mRNA stability effects in yeast and human embryonic kidney 293T (HEK293T) cells (*20*), the role of BTG1 in mature B cells and how *BTG1* lymphoma mutations contribute to lymphomagenesis remain unknown. Furthermore, *BTG1* is not deleted in DLBCLs, suggesting that mutations are not simple loss-of-function alleles (*11*). In this study, we probed the effect of *BTG1* mutation on GC biology and malignant transformation. *BTG1* mutation conferred a supercompetitive phenotype to GC B cells through a discrete acceleration of their response to T_{FH} cell help, resulting in the formation of highly aggressive lymphomas. Our findings illustrate how subtle biochemical perturbations in GC B cells can confer marked competitiveness and oncogenic transformation potential in concert with microenvironmental cues such as T_{FH} cell help.

Results

BTG1 somatic mutations in DLBCLs are genetic drivers

We surveyed the genomic profiles of 25,670 cancer patients, including publicly available DLBCL datasets (*9*, *21*–*29*). *BTG1* somatic mutations were detected in 11% of DLBCL cases (n = 272/2407), which is frequent given their high genetic heterogeneity, but were rare in other tumor types, suggesting specific BTG1 function in this context (fig. S1A). Within DLBCLs, *BTG1* mutation was enriched in the more clinically unfavorable ABC-DLBCLs (P = 0.0184; fig. S1B), especially those defined as MCD/cluster 5 lymphomas, with 40% of all BTG1 mutant DLBCLs being MCD (fig. S1C). *BTG1* mutations scored among the top DLBCL drivers in a rigorous genetic driver analysis taking into account multiple genomic and epigenomic covariates (Fig. 1A and table S1). Most *BTG1* mutant alleles corresponded to

[1]Division of Hematology and Oncology, Department of Medicine and Meyer Cancer Center, Weill Cornell Medicine, New York, NY, USA. [2]Laboratory of Lymphocyte Dynamics, The Rockefeller University, New York, NY, USA. [3]Department of Physiology and Biophysics, Weill Cornell Medicine, New York, NY, USA. [4]Tri-Institutional PhD Program in Computational Biomedicine, New York, NY, USA. [5]Institute for Computational Biomedicine, Weill Cornell Medicine, New York, NY, USA. [6]Department of Systems Immunology and Braunschweig Integrated Centre of Systems Biology (BRICS), Helmholtz Centre for Infection Research, Braunschweig, Germany. [7]Department of Pathology and Laboratory Medicine, Weill Cornell Medicine, New York, NY, USA. [8]Department of Pathology, Yale School of Medicine, New Haven, CT, USA. [9]Department of Physiology and Biophysics, Weill Cornell Medicine, New York, NY, USA. [10]Department of Pharmacology and Meyer Cancer Center, Weill Cornell Medicine, New York, NY, USA. [11]Molecular Pharmacology Program and Center for Cell Engineering, Center for Stem Cell Biology, Center for Experimental Therapeutics, and Center for Hematologic Malignancies, Memorial Sloan Kettering Cancer Center, New York, NY, USA. [12]Department of Pharmacology, Weill Cornell Medicine, New York, NY, USA. [13]Division of Biostatistics, Department of Population Health Sciences, Weill Cornell Medicine, New York, NY, USA. [14]Department of Laboratory Medicine, University of California, San Francisco, CA, USA. [15]Weill Cornell Medicine–Qatar, Doha, Qatar. [16]Centre for Haemato-Oncology, Barts Cancer Institute, Queen Mary University of London, London, UK. [17]Caryl and Israel Englander Institute for Precision Medicine and Department of Physiology and Biophysics, Weill Cornell Medicine, New York, NY, USA. [18]Centre for Lymphoid Cancer, BC Cancer, Vancouver, BC, Canada. [19]Institute for Biochemistry, Biotechnology and Bioinformatics, Technische Universität Braunschweig, Braunschweig, Germany.
*Corresponding author. Email: chm2065@med.cornell.edu (C.M.); amm2014@med.cornell.edu (A.M.)
†Deceased.
‡Present address: Gordian Biotechnology, San Francisco, CA, USA.

heterozygous missense mutations clustered within the N terminus, which contains several conserved functional motifs (*30*) (Fig. 1B and table S2). Alpha helix 2 (α2) was most heavily mutated, especially at Gln^{36}, which was most frequently (86%) replaced by a histidine [Gln^{36}→His (Q36H); Fig. 1, B and C, and fig. S1D]. Q36H and additional *BTG1* N-terminal mutations can alter the conformational distribution of BTG1 within the α2-α4 region, as determined by in silico atomistic molecular dynamics simulation (*31*).

Btg1^{Q36H} generates supercompetitor GC B cells

DLBCLs originate from B cells having transited the GC reaction. To determine the functional impact of *BTG1* mutation during the humoral immune response, we generated a mouse model for conditional expression of mutant Btg1 from the *Rosa26* locus (*R26*$^{lsl.Btg1Q36H}$), referred to hereafter as Q36H (fig. S2, A to C). We then reverted the Q36H mutation to wild type to generate a wild-type knock-in (WTKI) control line (*R26*$^{lsl.Btg1wt}$; fig. S2, D to G). Q36H and WTKI mice were crossed to the *Cd19*^{Cre} strain to induce Btg1^{Q36H} or Btg1^{WTKI} expression in B cells (*32*). Recombination and expression of Q36H and WTKI alleles was validated in B cells at DNA, RNA, and protein levels (fig. S2, H to M). Notably, the total abundance of *Btg1* transcripts was not increased upon expression of Btg1^{Q36H} or Btg1^{WTKI} from the *Rosa26* locus, indicating an absence of Btg1 dosage effect (fig. S2, J and K). Q36H or WTKI expression in B cells did not perturb early B cell development in the bone marrow nor steady-state mature B cell populations in secondary lymphoid tissues and peritoneal cavity (fig. S3).

To investigate the GC response, we immunized Q36H (*R26*$^{lsl.Btg1Q36H/+}$; *Cd19*$^{Cre/+}$), WTKI (*R26*$^{lsl.Btg1wt/+}$; *Cd19*$^{Cre/+}$), and CRE-control (*R26*$^{+/+}$; *Cd19*$^{Cre/+}$) or CREneg-control (*R26*$^{lsl.Btg1Q36H/+}$; *Cd19*$^{+/+}$) mice with sheep red blood cells (SRBCs), a T cell–dependent antigen, and analyzed spleens 10 days later. The GC response to SRBCs remained unchanged in Q36H and WTKI animals (fig. S4). Using a different T cell–dependent antigen [(4-hydroxy-3-nitrophenyl)acetyl (NP) conjugated to ovalbumin (OVA)] to assess antigen-specific responding cells, we further observed no change in Q36H versus CRE mice over the course of the GC reaction (fig. S5).

Most GC-derived lymphoma oncogenes affect GC size, composition, or output. Hence, the absence of such effects in the Btg1^{Q36H} setting was puzzling. However, these aspects do not necessarily reflect possible selective advantages within B cell clones. To directly test competitive fitness of Q36H GC B cells in vivo, we crossed our mice to the B1-8^{hi} allele. B1-8^{hi} encodes a B cell receptor with high NP affinity in B cells with a lambda immunoglobulin light chain (~10 to 15% of murine B cells) that facilitates their entry into GCs upon NP immunization (*33*). We adoptively transferred a 50:50 ratio of B1-8^{hi}/Q36H (B1-8$^{hi+/-}$; *R26*$^{lsl.Btg1Q36H/+}$; *Cd19*$^{Cre/+}$; CD45.1/2) and B1-8^{hi}/CREneg-control (B1-8$^{hi+/-}$; *R26*$^{lsl.Btg1Q36H/+}$; *Cd19*$^{+/+}$; CD45.1) resting B cells into wild-type CD45.2 (non-B1-8^{hi}) recipients (Fig. 2A). Q36H GC B cells showed a progressive competitive advantage over time, reaching up to ~90% of total GC B cells by day 14 (Fig. 2A and fig. S6A). Q36H did not confer a competitive advantage to non-GC B cells (fig. S6A), and the GC fitness gain was not associated with changes in GC polarity (fig. S6B). A similar advantage of B1-8^{hi}/Q36H GC B cells occurred when competing against B1-8^{hi}/CRE-controls instead (B1-8$^{hi+/-}$; *R26*$^{+/+}$; *Cd19*$^{Cre/+}$; CD45.1; fig. S6C). In contrast, B1-8^{hi}/WTKI cells did not show a competitive advantage in GC B cells over non-GC B cells (fig. S6D). To confirm that the Q36H fitness advantage was not due to the extra *R26*^{Btg1} copy or *Cd19*^{Cre} knock-in, we directly placed B1-8^{hi}/Q36H and B1-8^{hi}/WTKI cells in competition. Q36H again outcompeted WTKI cells in the GC compartment, whether gating or not on antigen-specific (lambda^{+}NP^{+}) cells and even given a lower Q36H abundance in the starting NB or non-GC B compartment (Fig. 2B and fig. S7A). To assess Q36H GC B cell fitness in the context of a polyclonal immune response (without B1-8^{hi}–engineered antigen specificity), we transferred resting B cells from Q36H, WTKI, or CRE-controls into wild-type recipients and measured GC contribution of each genotype versus recipient cells. Q36H cells showed greater expansion within the GC compartment (18% average) as compared with WTKI and CRE at day 21 (~5%; Fig. 2C). Such expansion was similar in NP-specific GC B cells and did not occur in the NB compartment (fig. S7, B and C). In these experiments, *Cd19*^{Cre}-mediated recombination efficiency was as expected (*32*), and the Q36H GC B cell advantage was accompanied by a measurable and significant enrichment of Q36H recombined alleles in GC B versus NB cells (fig. S7D). We next assessed proliferation in competing Q36H and CRE B cells in T cell–independent B cell activation by NP-Ficoll immunization. We observed similar proportion and carboxyfluorescein succinimidyl ester (CFSE) proliferation dye dilution of antigen-specific (lambda^{+}NP^{+}) Q36H and CRE B cells over time (fig. S7E). Thus, Btg1^{Q36H} provided a competitive fitness advantage specifically to GC B cells during T cell–dependent immune responses, independently of antigen affinity. Given the lack of GC-specific phenotype in mice expressing Btg1^{WTKI} in B cells, we focused subsequent functional experiments on mice expressing Btg1^{Q36H}.

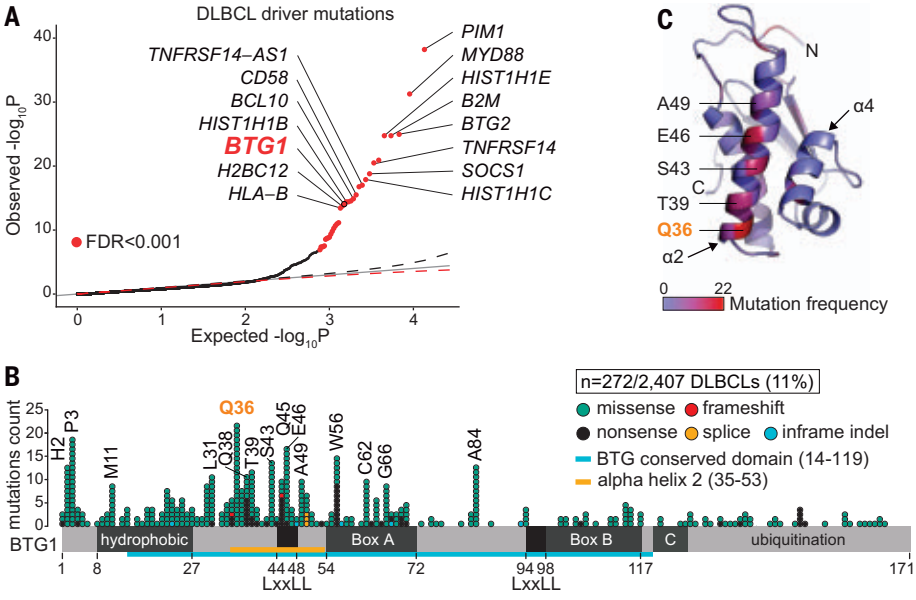


Fig. 1. Characterization of *BTG1* missense mutations in DLBCL. (**A**) Quantile-quantile plot of observed and expected *P* values from a gamma-Poisson distribution for single nucleotide variants across 101 germline-matched whole genome–sequenced DLBCLs. Mutated genes with FDR < 0.001 are depicted with a red dot. (**B**) BTG1 protein and frequency, type, and location of mutations in *n* = 272 unique cases out of *n* = 2407 DLBCLs. LxxLL motifs mediate interaction with nuclear receptors. (**C**) BTG1 homology model (residues 1 to 129) based on a structural reference of BTG2 (Protein Data Bank ID 3DJU) showing mutation frequency. N- and C-terminal ends are denoted. Single-letter abbreviations for the amino acid residues are as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.

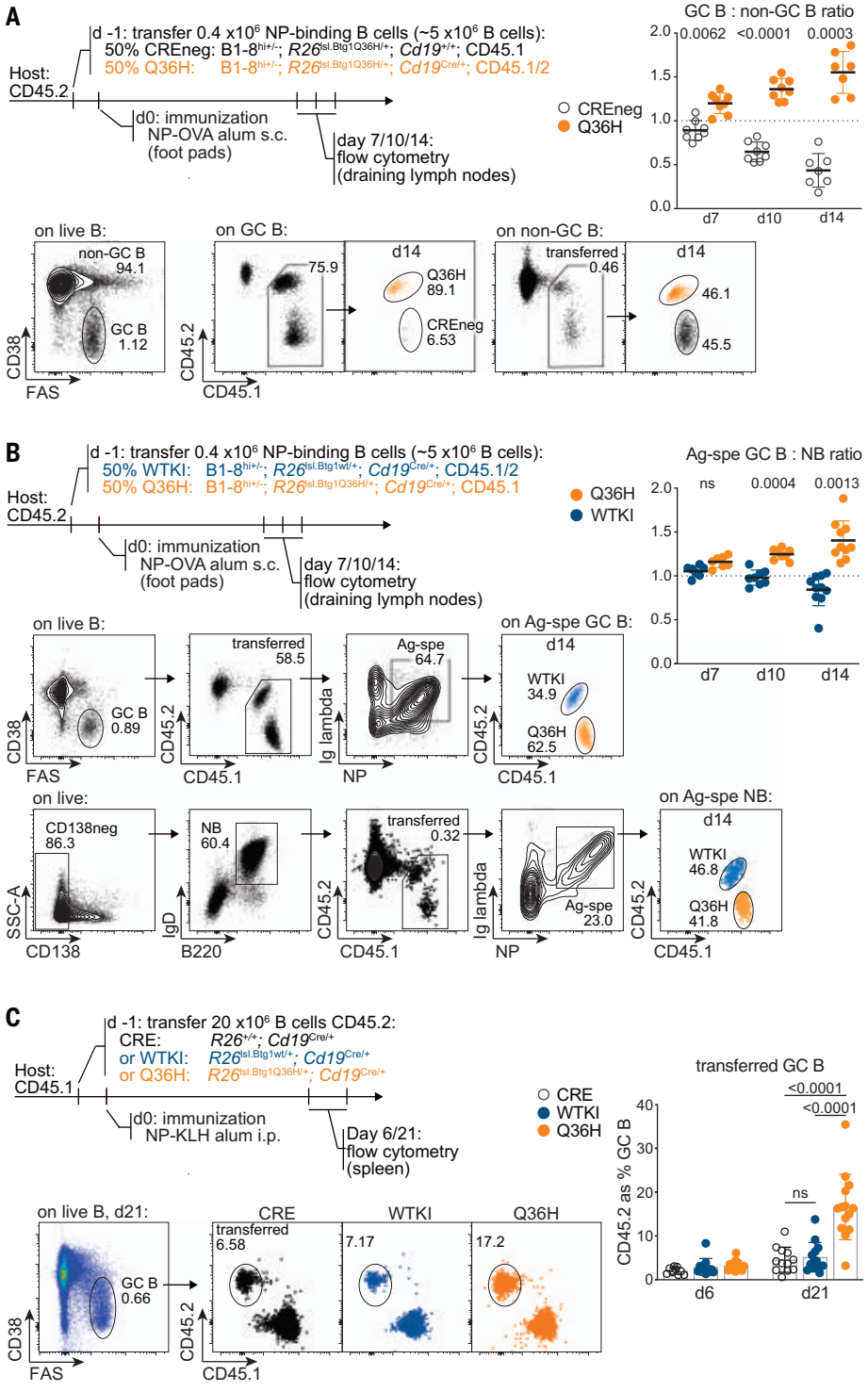


Fig. 2. Btg1^{Q36H} generates supercompetitor GC B cells. (**A**) In vivo competitive assay schematic and representative flow cytometry plots for gating Q36H and CREneg from transferred germinal center (GC) B and non-GC B cells. Graph represents pooled mice from three independent experiments, each with n = 2 or 3 mice per time point. Mean ± SD, paired t test (two-tailed). (**B**) Competitive assay between mature B cells expressing Btg1^{Q36H} versus Btg1^{WT} from *Rosa26* (*R26*) knock-in alleles that differ by a single nucleotide. Representative flow cytometry plots for gating antigen-specific (Ag-spe, lambda$^+$NP$^+$) naïve B (NB, CD138$^-$B220$^+$IgD$^+$) and GC B (FAS$^+$CD38$^-$) WTKI and Q36H cells at day 14 after immunization. Graph represents pooled mice from two independent experiments, representative of three [day 7 (d7) and d14] or two (d10) independent experiments, each with n = 3 to 5 mice per time point. Mean ± SD, paired t test (two-tailed). (**C**) Experimental schematic for the adoptive transfer of B cells with an endogenous B cell receptor repertoire to compete with wild-type recipient cells. Representative flow cytometry plots on live B cells (DAPI-B220$^+$) at day 21 after NP-KLH immunization, for gating transferred cells (CD45.2) of the indicated genotypes within GC B cells. Graph represents mice combined from three independent experiments, with 9 to 14 mice total per genotype and time point. Mean ± SD, unpaired t test (equal SD, two-tailed). s.c., subcutaneously.

Btg1^{Q36H} induces MYC-associated biosynthetic programs in GC B cells

To gain insight into the competitive advantage conferred by mutant Btg1, we next performed RNA sequencing (RNA-seq) in Q36H and CRE-control GC B cells (fig. S8A). As noted earlier, total abundance of *Btg1* transcripts was similar in both groups (fig. S8B). Only ~15% of *Btg1* transcripts carried the Q36H mutation (fig. S8C), which is comparable to the fraction of mutant *BTG1* transcripts (17 to 45%) in BTG1^{Q36H} DLBCL patients (n = 6; fig. S8D). We detected no significant differences in gene expression profiles using both unsupervised and supervised analyses, except for the expected reduction in *Rosa26* due to the knock-in (fig. S8, E and F). Bulk RNA-seq represents a composite of heterogeneous GC B cells in their various transitional states, which could mask effects occurring in subpopulations of cells. We therefore performed gene set enrichment analysis (GSEA) using MSigDB Hallmark and Canonical Pathway gene set collections (*34*–*36*) to determine whether we could detect subtler transcriptional perturbation in groups of coordinately regulated genes. We identified positive enrichment of 94 signatures [false discovery rate (FDR) < 0.05],

most of them related to activation of biosynthetic pathways, mitochondrial function, and MYC target genes (Fig. 3A and table S3). A network connectivity analysis revealed that MYC targets were among the highest connected signatures, suggesting a central role for MYC (Fig. 3A).

Myc is induced transiently together with mTORC1 activation upon receiving strong T_{FH} cell help during selection in the anatomically defined GC light zone (LZ) (*3*–*5*). Myc and mTORC1 biosynthetic programs trigger GC B cell growth, a prerequisite for their subsequent clonal expansion in the GC dark zone (DZ). Accordingly, we observed significant enrichment for signatures induced by T_{FH} cell help in LZ-to-DZ recycling GC B cells, as well as MYC and mTORC1 programs in Q36H versus CRE (Fig. 3B and table S4). The same signatures were significantly enriched among human BTG1 mutant DLBCL patients as compared with BTG wild-type DLBCLs in two independent cohorts (excluding *BTG2* mutant patients; Fig. 3B and fig. S8G). Among ABC-DLBCLs, BTG1 mutant cases also showed higher expression of these signatures (fig. S8H). For functional studies, we generated isogenic human DLBCL cells with ectopic and equivalent expression of $BTG1^{Q36H}$ versus $BTG1^{WT}$ (fig. S8, I and J). In this model, $BTG1^{Q36H}$ expression yielded a similar enrichment of LZ-to-DZ recycling, MYC, and mTORC1 signatures (fig. S8K). Furthermore, genes differentially induced [fold change (FC) > 1.5; FDR < 0.01; table S4] in $BTG1^{Q36H}$ versus $BTG1^{WT}$ DLBCL cells were significantly and exclusively enriched for MYC and MAX DNA binding motifs (fig. S8L). The corresponding $BTG1^{Q36H}$ signature was also significantly enriched in BTG1 mutant DLBCL patients (FDR < 0.001; fig. S8M), indicating consistency of BTG1 mutant effects between murine GC and human DLBCL cases. In line with enhanced biosynthetic rates, $Btg1^{Q36H}$ GC B and $BTG1^{Q36H}$ DLBCL cells showed increased RNA content and cell size (Fig. 3, C and D, and fig. S8, N and O). These data suggested that *BTG1* mutation enhances biosynthetic fitness programs normally associated with positive selection and Myc/mTORC1 activation during the GC reaction, an effect that is maintained in established DLBCLs.

MYC-related signatures could be enriched because of either higher expression of Myc or expansion of Myc-positive (Myc^+) GC B cells. To address this question, we crossed our Q36H mice to the Myc^{GFP} reporter line, in which endogenous Myc protein is fused to green fluorescent protein (GFP) (*37*). To better focus our analysis on GC B cells, we used the $C\gamma1^{Cre}$ line (*38*) to drive $Btg1^{Q36H}$ expression (Fig. 3E). The proportion of Myc^{GFP+} GC B cells was significantly higher in Q36H versus CRE-control mice (Fig. 3F), yet Myc^{GFP+} cells in Q36H and CRE mice manifested similar Myc^{GFP} protein and *Myc* transcript levels (fig. S8, P and Q). As expected, RNA-seq analysis on Myc^{GFP+} versus Myc^{GFP-} GC B cells showed strong enrichment of the LZ-to-DZ recycling, MYC, and mTORC1 signatures in both Q36H and CRE (fig. S8R). However, comparing Myc^{GFP+} versus Myc^{GFP-} cells from Q36H versus CRE mice revealed upregulation of 201 transcripts in Q36H Myc^{GFP+} (Fig. 3G and table S4). These genes were significantly enriched for Myc immediate early direct target genes (*39*), as well as genes activated by T_{FH} cell help through cytokine signaling (Fig. 3H). Thus, the fitness advantage of Btg1 mutant GC B cells was linked to increased proportion of Myc-expressing cells and enhanced activation of early Myc programs, without alteration of steady-state Myc expression levels per cell.

Q36H impairs BTG1 association with MYC and other transcripts involved in LZ-to-DZ recycling

We next explored how $BTG1^{Q36H}$ specifically influences MYC expression and T_{FH} cell help–related transcriptional programs. The BTG1 family proteins TOB1 and TOB2, which share the conserved N-terminal domain with BTG1, were recently shown to interact with RNAs (*40*). We thus performed RNA immunoprecipitation (RIP) assays using V5-tagged $BTG1^{WT}$, $BTG1^{Q36H}$, or enhanced GFP (EGFP) proteins as baits in our DLBCL cells (fig. S9A). As compared with EGFP negative control, $BTG1^{WT}$ enriched ~3000 unique mRNA species, whereas $BTG1^{Q36H}$ enriched only ~700 (fig. S9B). This effect was largely due to loss of $BTG1^{WT}$-associated mRNAs (fig. S9C). Significantly reduced pull-down with $BTG1^{Q36H}$-V5 versus $BTG1^{WT}$-V5 was observed for 732 transcripts, which notably included the *MYC* mRNA (log2FC > 2, q < 0.05; Fig. 4A and table S5). These transcripts were highly enriched for LZ-to-DZ recycling and Myc^{GFP+} signatures and marginally enriched for the centrocyte signature (which includes the small Myc^+ population), but not for other GC or post-GC subpopulations (Fig. 4B). We further validated $BTG1^{WT}$-*MYC* mRNA association and reduced enrichment with $BTG1^{Q36H}$ in independent RIP experiments by quantitative polymerase chain reaction (qPCR; Fig. 4C). We next tested whether BTG1 directly binds RNA by ultraviolet (UV) cross-linking and immunoprecipitation (CLIP) of RNAs interacting with V5-tagged $BTG1^{WT}$, $BTG1^{Q36H}$, and EGFP as compared with canonical RNA-binding proteins Musashi-2 (MSI2) and human antigen R (HuR). No RNA was enriched with EGFP, and very little RNA and no enrichment for RNA-protein complexes right above the expected size of V5-tagged $BTG1^{WT}$ or $BTG1^{Q36H}$ was observed, as opposed to MSI2 and HuR (fig. S9D).

These data suggested that $BTG1^{WT}$ indirectly associates with transcripts involved in T_{FH} cell help and MYC programs to attenuate their expression, a function that is lost by BTG1 lymphoma mutants. In line with this, expressing $BTG1^{WT}$ in our isogenic DLBCL cells resulted in depletion of the LZ-to-DZ recycling, MYC, and mTORC1 signatures (fig. S9E) and reduced cellular RNA abundance, as compared with vector-control cells (fig. S9F). We categorized differentially expressed genes according to their behavior in $BTG1^{WT}$ and $BTG1^{Q36H}$ DLBCL cells. Many transcripts were regulated in a similar fashion (fig. S9G). However, genes repressed by $BTG1^{WT}$ but not $BTG1^{Q36H}$ were largely restricted to the LZ-to-DZ recycling, MYC, and mTORC1 signatures, and complementarily, genes induced by $BTG1^{WT}$ but not $BTG1^{Q36H}$ were mostly enriched for MYC-repressed and anti-mTORC1 signatures (fig. S9G and table S6). Hence, wild-type BTG1 normally associates with *MYC* and transcripts relevant to LZ-to-DZ recycling GC B cells, an effect that is impaired by mutational disruption of its N-terminal domain.

$BTG1^{Q36H}$ lowers the threshold for MYC protein synthesis

We wondered whether loss of *MYC* mRNA association upon BTG1 mutation might alter MYC protein induction kinetics, perhaps enabling cells to reach high MYC plateau levels more rapidly and yielding the observed increased fraction of MYC^+ cells. BTG1 can regulate mRNA stability in T cells (*41*). However, *MYC* mRNA half-life was comparable in $BTG1^{Q36H}$ and $BTG1^{WT}$ DLBCL cells exposed to transcription inhibitor actinomycin D (fig. S10A). To investigate MYC protein expression kinetics, we placed $BTG1^{Q36H}$ and $BTG1^{WT}$ isogenic DLBCL cells under serum deprivation and performed Western blots for MYC and BTG1 (by means of the V5 tag) at serial time points between 15 min and 2 hours after refeeding (fig. S10B). MYC was induced >1.5-fold between 15 and 30 min in $BTG1^{Q36H}$ but did not change significantly in $BTG1^{WT}$-expressing cells (fig. S10B). *MYC* mRNA abundance remained unchanged under similar conditions (fig. S10C). To further investigate the increased MYC protein levels, we treated our DLBCL cells with cycloheximide (CHX) to inhibit protein synthesis, which revealed no effect of $BTG1^{Q36H}$ on MYC protein stability (fig. S10D). We then performed an MG132 treatment time course to inhibit protein degradation. This showed an earlier and higher accumulation of MYC protein levels in $BTG1^{Q36H}$ cells (a 4.5-fold peak at 2 hours MG132 versus dimethyl sulfoxide (DMSO) in $BTG1^{Q36H}$ cells compared with a 3.8-fold peak at 4 hours MG132 versus DMSO in $BTG1^{WT}$ cells; fig. S10E). Both $BTG1^{WT}$ and $BTG1^{Q36H}$ protein levels decreased upon reaching high MYC protein levels, consistent with reports that Myc can down-regulate Btg1 (*42*) (fig. S10E). Accordingly, *Btg1* mRNA levels

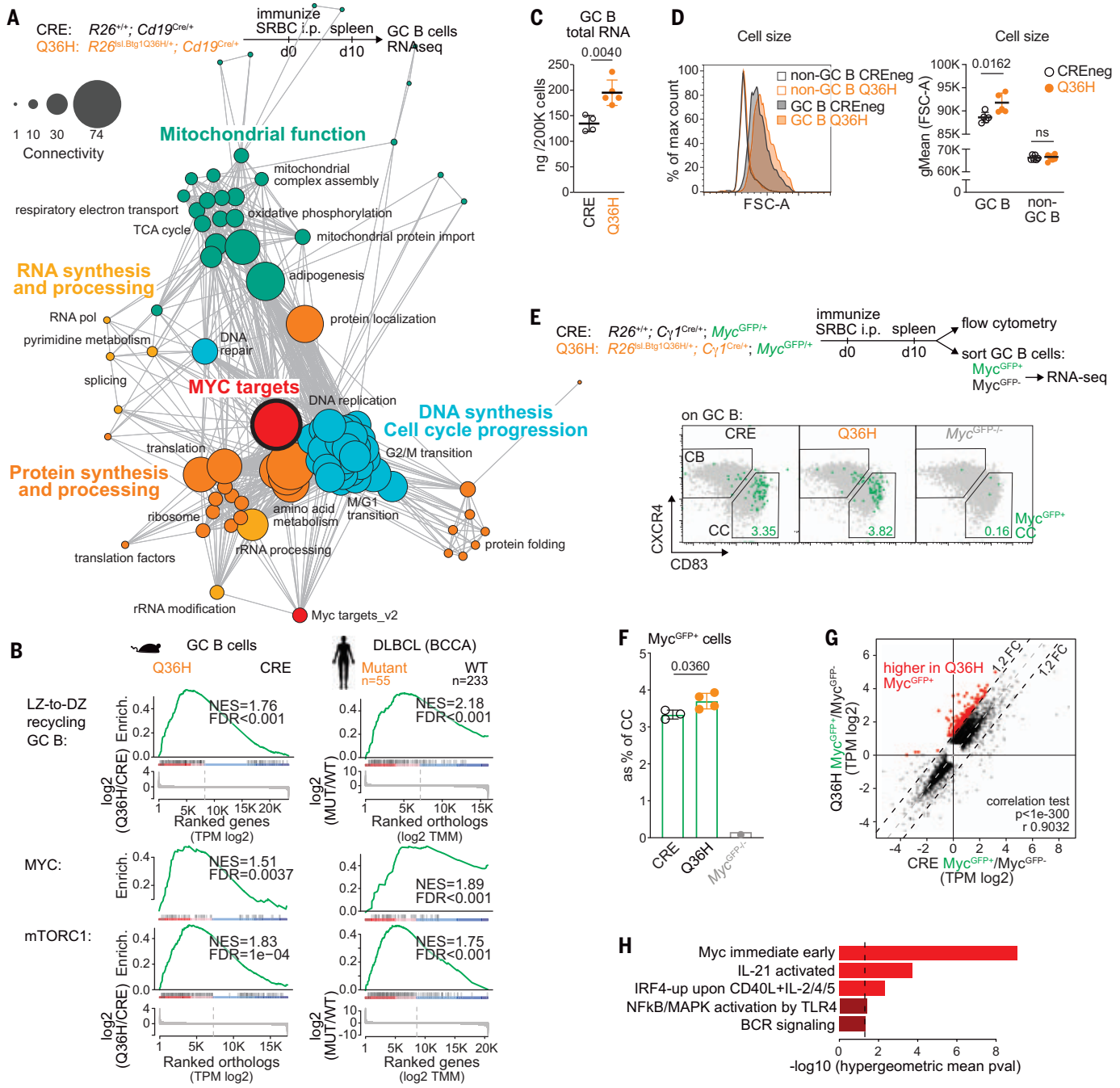


Fig. 3. Mutant BTG1 induces MYC-related biosynthetic programs. (**A**) Gene set enrichment analysis (GSEA) network showing canonical and hallmark gene sets positively enriched in Q36H versus CRE up-regulated genes [normalized enrichment score (NES) ≥ 1.25; FDR < 0.05]. Distance between two enriched gene sets was calculated as the Jaccard index (size intersection: size union) using leading-edge genes. The circle size represents the degree of connectivity to which each gene set is attached to others. (**B**) GSEAs showing LZ-to-DZ recycling GC B, MYC, and mTORC1 activation signatures enrichment in Q36H versus CRE mouse GC B cells (left) and in BTG1 mutant versus WT DLBCL cases in the BC Cancer Agency (BCCA) cohort (right) (*23*). LZ-to-DZ recycling GC B signature is DECP_UPREG [*n* = 221 mouse genes (left); *n* = 201 human orthologs (right)]; MYC signature is SCHUHMACHER_MYC_TARGETS_UP [*n* = 80 mouse orthologs (left); *n* = 75 genes represented in the BCCA dataset (right)]; mTORC1 signature is PENG_RAPAMYCIN_RESPONSE_DN [*n* =242 mouse orthologs (left); *n* = 230 genes represented in the BCCA dataset (right)]. Enrich., enrichment score. (**C**) Graph showing quantification of RNA extracted from 200,000 mouse GC B cells. One experiment, *n* = 4 and 5 mice per genotype. Mean ± SD, unpaired *t* test (equal SD, two-tailed). (**D**) Representative flow cytometry histograms showing cell size measurement as forward scatter area (FSC-A) in GC B (FAS^+CD38^-) and non-GCB (FAS^-CD38^+) cells (left). Graph showing geometric mean (gMean) of FSC-A from one experiment, *n* = 5 and 6 mice per genotype, representative of at least three independent experiments (right). Mean ± SD, unpaired *t* tests (two-tailed). (**E**) Experimental schematic and representative flow cytometry plots showing Myc^{GFP+} cells (green) in centroblasts (CB) and centrocytes (CC). (**F**) Bar plot showing proportion of Myc^{GFP+} cells in CC. One experiment with *n* = 3 and 4 mice per genotype, representative of at least three independent experiments. Mean ± SD, unpaired *t* test (two-tailed). (**G**) Genes (*n* = 201) depicted with red dots present higher expression (FC > 1.2) in GFP^{Myc+} versus GFP^{Myc-} GC B cells comparing Q36H mice (*y* axis) to CRE mice (*x* axis). (**H**) Hypergeometric mean analysis for the 201 genes identified in (G). The Myc immediate early signature consists of Myc-dependent genes induced by 8 hours in Myc WT/flox B cells (*39*). Genes in this and other gene signatures are included in supplementary table S4.

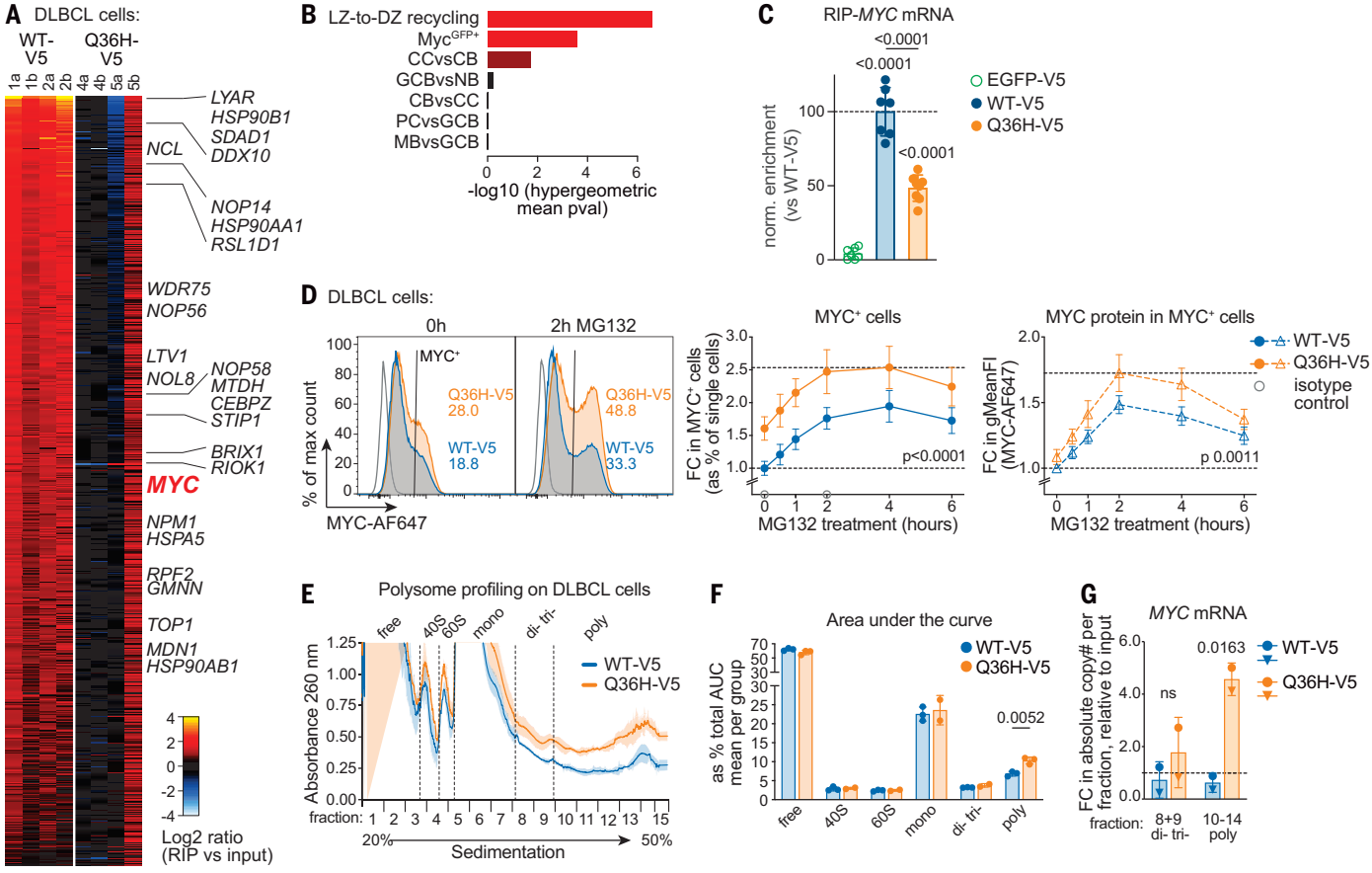


Fig. 4. Q36H mutation disrupts BTG1-*MYC* mRNA association and enhances MYC protein synthesis kinetics. (**A**) Heatmap showing n = 732 $BTG1^{WT}$-associated ($BTG1^{WT}$ RIP/input log2FC > 2, q < 0.05) and $BTG1^{Q36H}$-lost ($BTG1^{Q36H}$/$BTG1^{WT}$ log2FC > 2, q < 0.05) transcripts by RIP-seq, using V5 antibody in SU-DHL4 DLBCL cells overexpressing V5-tagged $BTG1^{WT}$ or $BTG1^{Q36H}$. *MYC* and leading edge genes from positively enriched LZ-to-DZ recycling GC B, MYC, and mTORC1 signatures in $BTG1^{Q36H}$ versus $BTG1^{WT}$ are labeled. (**B**) Signatures enriched for up-regulation of the n = 732 $BTG1^{WT}$-associated and $BTG1^{Q36H}$-lost transcripts as determined by hypergeometric mean analysis. (**C**) RIP-qPCR for *MYC* mRNA in indicated SU-DHL4 V5-tagged DLBCL cells. Mean ± SD of enrichment in RIP over input samples, normalized to $BTG1^{WT}$-V5. Three independent experiments, each with n = 2 or 3 independently generated lines per genotype. Unpaired t test (equal SD, two-tailed, versus EGFP or as indicated). (**D**) Representative MYC flow cytometry histograms of SU-DHL4 DLBCL cells expressing V5-tagged $BTG1^{WT}$ (WT-V5) or $BTG1^{Q36H}$ (Q36H-V5) and treated with 5 μM MG132 for 0 or 2 hours. Proportions of MYC-positive cells (MYC^+) are indicated. Staining with isotype control is included. Graphs show proportion of MYC^+ cells (left) or MYC protein levels within MYC^+ cells (right) as measured by flow cytometry in three independent experiments, each with n = 2 or 3 independently generated lines per genotype, shown as fold change relative to WT-V5 0 hour mean value per experiment. Mean ± SD, two-way analysis of variance (ANOVA) for Q36H-V5 versus WT-V5. gMeanFI, geometric mean fluorescence intensity. (**E**) Polysome profile of $BTG1^{WT}$-V5 and $BTG1^{Q36H}$-V5 expressing SU-DHL4 cells. Mean (line) ± SD (shade) from n = 3 independently generated lines per genotype except fractions 4 to 9: n = 2 Q36H lines owing to technical loss of one line. One experiment representative of two independent experiments. (**F**) Bar plot shows the area under the curve (AUC) for indicated portions of the polysomal traces from (E), as percent of the total AUC average per genotype. Mean ± SD, unpaired t test (equal SD, two-tailed). (**G**) Absolute *MYC* mRNA levels from fractions in (E) and (F), pooled as indicated. n = 2 independently generated lines per genotype. Mean ± SD, unpaired t test (equal SD, two-tailed).

were reduced in Myc^{GFP+} versus Myc^{GFP-} GC B cells (fig. S10F). To measure MYC protein levels per cell, we performed MYC flow cytometry analyses in similar MG132 time courses. The fraction of MYC^+ cells was larger (~1.5-fold) in $BTG1^{Q36H}$ cells than in $BTG1^{WT}$ cells throughout the time course and was accompanied by a mild and faster up-regulation of MYC protein abundance in MYC^+ cells (Fig. 4D and fig. S11A). *MYC* mRNA levels remained comparable between $BTG1^{Q36H}$ and $BTG1^{WT}$ cells (fig. S11B). These observations suggest that $BTG1^{Q36H}$ lowered the threshold for MYC protein synthesis, resulting in a greater fraction of cells expressing MYC. Three additional DLBCL-recurrent BTG1 mutants (S43N, A49T, and T39I) accelerated MYC protein induction in a similar fashion (fig. S11C), and this effect was also observed when we expressed $BTG1^{Q36H}$ versus $BTG1^{WT}$ in freshly immortalized primary human GC B cells (*43*) (fig. S11, D to G).

To directly measure *MYC* mRNA engagement with the translational machinery, we carried out polysome profiling in our isogenic $BTG1^{WT}$ and $BTG1^{Q36H}$ DLBCL cells. We confirmed that polysomal fractions were enriched for ribosomal proteins (e.g., RPS3; fig. S12A). $BTG1^{Q36H}$ cells showed higher RNA abundance in polysomal fractions as compared with $BTG1^{WT}$ cells, indicative of increased overall translation (Fig. 4E and quantification in Fig. 4F). $BTG1^{Q36H}$ cells also contained a significantly greater abundance of *MYC* transcripts within polysomal fractions (Fig. 4G and fig. S12B). Collectively, these data suggested that wild-type BTG1 association with *MYC* mRNA may restrict MYC translation, an effect that is specifically disrupted by *BTG1* mutations. We hypothesized that this might reduce the threshold for T_{FH} cell help–dependent MYC induction, thereby enhancing the response of GC B cells to T_{FH} cell help signals.

More-rapid S phase completion and DZ commitment by Btg1^{Q36H} LZ-to-DZ recycling cells

The degree of T_{FH} cell help, and therefore Myc induction, determines the ability and speed of GC B cells to transit through S phase and subsequently undergo proliferative burst in the DZ (*44*, *45*). Hence, we predicted that enhanced Myc kinetics due to *Btg1* mutation would enable GC B cells to more efficiently progress through S phase and into the DZ proliferative program. To evaluate whether and how such effects might manifest in the heterogeneous GC milieu, we performed targeted single-cell RNA-seq (n = 496 immune-related genes; table S7) on competing B1-8^{hi}/Btg1^{Q36H} and B1-8^{hi}/CREneg control GC B cells sorted at day 10 after NP immunization, when Q36H cells outcompete control cells by ~70 to 30% (Fig. 5A and fig. S6A). A general overview (2982 cells across three biological replicates) revealed a similar general distribution of individual cell gene expression profiles, without bias toward genotype or biological replicate (fig. S13, A and B). Individual cells were assigned to specific GC B cell subpopulations (DZ, LZ, or LZ-to-DZ recycling) according to their distinctive signatures and marker transcript levels (Fig. 5B).

LZ-to-DZ recycling cells enter S phase but generally do not engage G2/M until transiting into the DZ and expressing the corresponding centroblast (CB) transcriptional program (*44*). Accordingly, S and G2/M phase signatures were mostly confined to the DZ compartment (fig. S13C). To compare cell cycle progression of Q36H versus control GC B cells, we plotted their frequency according to cell cycle program expression, across a pseudotime scale starting from G1 phase (Fig. 5C). We found a significant increase in Btg1^{Q36H} cell density from S phase entry and throughout G2/M (P = 0.0071; Fig. 5C). As expected, cells expressing the LZ-to-DZ recycling program localized at the G1-to-S transition and extended into G2/M along this pseudotime axis (pseudotime units 9 to 14; Fig. 5D). We then projected the DZ and G2/M signatures onto the same axis. GC B cells at the LZ-to-DZ recycling pseudotime transition showed earlier induction and greater proportion of Btg1^{Q36H} cells having initiated DZ (P = 0.0117; Fig. 5E) and G2/M (P = 0.0202; Fig. 5F) programs. Along these lines, we identified 28 genes with differential expression in Btg1^{Q36H} versus control cells, which mainly corresponded to changes in the proportion of cells expressing these genes (chi-square $P < 0.05$) and with most genes being related to cell cycle and lymphoma (fig. S13D). Therefore, Btg1^{Q36H} LZ-to-DZ recycling GC B cells engaged the proliferative DZ program earlier and in greater proportions.

These single-cell profiles raised the question of whether Btg1 mutant GC B cells may complete S phase more rapidly or at a higher rate. To explore this, we measured cell cycle kinetics in competition experiments. To distinguish between entry, mid-late, and post S phase cells, we injected 5-ethynyl-2′-deoxyuridine (EdU) and 5-bromo-2′-deoxyuridine (BrdU) at 1.5 hours and 0.5 hours before euthanasia, respectively (*46*) (Fig. 5G). We validated the expected advantage of Btg1^{Q36H} GC B cells under these conditions, with no change in DZ:LZ polarity nor advantage among non-GC B cells (fig. S13, E and F). We observed no difference in S phase entry (BrdU$^+$ only) in either centrocyte (CC) or centroblast (CB) populations (Fig. 5G). In contrast, there was a significant increase in Btg1^{Q36H} CC and CB progressing through mid-late S phase (BrdU$^+$EdU$^+$) and in Btg1^{Q36H} CB post S phase (EdU$^+$ only; Fig. 5G). These observations suggested that Btg1 mutant GC B cells transited and completed S phase more rapidly than control cells. To distinguish whether such effects were specific to the LZ-to-DZ recycling GC B cells, we next performed in vivo Edu incorporation assays in our Q36H/*Cγ1*^{Cre}/*Myc*^{GFP} mice to detect S phase–experienced cells in Myc^{GFP+} versus Myc^{GFP-} GC B populations. To avoid losing the small number of Myc^{GFP+} GC B cells during EdU staining, sorted cells were fixed and stained by immunofluorescence on slides (Fig. 5H and fig. S13G). A significantly higher proportion of EdU$^+$ cells was detected in Q36H versus CRE Myc^{GFP+} GC B cells (P = 0.0304; Fig. 5H), whereas there was no change in Q36H versus CRE Myc^{GFP-} cells. This suggested that Btg1^{Q36H} induced an S phase transit advantage specifically to LZ-to-DZ recycling GC B cells.

Competitive advantage conferred by Btg1^{Q36H} links to faster LZ-DZ kinetics

Our data so far suggest that *Btg1* mutation provides a GC fitness advantage by reducing the threshold to T_{FH} cell help response, as manifested by facilitated Myc protein induction, DZ program commitment, and S phase completion. To determine whether these changes could explain the Btg1^{Q36H} fitness advantage, we used established in silico mathematical models that reflect different aspects of GC temporal and clonal selection dynamics, including B cell receptor and CD40 signaling, antigen uptake, and Myc and mTORC1 induction (*47*). Comparing multiple different combinatorial perturbations of these signals, we found that providing GC B cells with a modest increase in the speed of T_{FH} cell help response, corresponding to a 1.12-fold increase in the rate of Myc induction and mTORC1 activation, recapitulated the Btg1^{Q36H} competitive advantage observed in vivo (fig. S14A). However, under noncompetitive conditions, this model predicted higher GC and CB proportions for Btg1 mutant (fig. S14A), two features that were not observed in vivo (see figs. S4 and S5). Taking into account our observation of a more rapid S phase completion, we determined that shortening the S phase by 21%, in addition to providing 1.12-fold faster T_{FH} cell help response, recapitulated the competitive advantage of Btg1^{Q36H} GC B cells without altering GC volume and polarity under noncompetitive conditions, more accurately reflecting the Btg1^{Q36H} in vivo phenotypes (fig. S14B). These parameters further predicted that Btg1 mutant GC B cells would undergo an increased number of GC rounds over the course of the GC reaction ($P < 2 \times 10^{-16}$ at day 10; Fig. 6A).

To test this scenario functionally, we synchronized competing Q36H and WT GC B cells receiving targeted T_{FH} cell help in vivo and tracked their LZ-to-DZ kinetics. T cell help can be elicited by injecting an OVA-conjugated anti-DEC205 antibody (αDEC205-OVA), which binds to the B cell surface receptor DEC205 and delivers OVA to the major histocompatibility complex class II (MHC-II) presentation pathway, inducing strong interactions with OVA-specific T_{FH} cells (*44*) (Fig. 6B). We adoptively transferred 3% Q36H (B1-8$^{hi+/-}$; *DEC205*$^{+/+}$; *R26*$^{lsl.Btg1Q36H/+}$; *Cd19*$^{Cre/+}$; CD45.1) and 7% WT (fluorescent B1-8$^{hi+/-}$; *DEC205*$^{+/+}$; CD45.1) resting B cells, to reach a 50:50 proportion of Q36H and WT cells at 6 to 7 days into the GC reaction. The remaining 90% were B1-8$^{hi+/-}$; *DEC205*$^{-/-}$; CD45.1/2 cells that lack the DEC205 receptor and do not receive targeted T_{FH} cell help, to better reflect physiological conditions where T_{FH} cell help is only available to a small fraction of GC B cells (Fig. 6C). In this three-way competition system, the different transferred cell populations were distinguished using a combination of congenic and fluorescent markers. Recipient CD45.2 mice were primed with OVA before the B cell transfer to generate OVA-specific T_{FH} cells. The day after transfer, mice were boosted with NP-OVA to recruit OVA-specific recipient T_{FH} cells and NP-specific B1-8^{hi} transferred B cells into the GC reaction.

We delivered T_{FH} cell help to competing DEC205$^{+/+}$ Q36H and WT cells 6 to 7 days into the GC reaction by injecting αDEC-OVA. We then measured the proportion of GC B cells in LZ (CC) and DZ (CB) over time (0, 36, 48, and 72 hours) after T_{FH} cell help delivery (Fig. 6, C and D). In line with previous work (*44*), DEC205$^{-/-}$ cells remained enriched in the LZ over this time course, showing a stable DZ:LZ ratio, and did not expand after anti-DEC205-OVA injection, whereas DEC205$^{+/+}$ cells moved to the DZ (as CB) and expanded between 48 and 72 hours (Fig. 6, D and E, and fig. S14, C and D). Within DEC205$^{+/+}$ cells, T_{FH} cell help induced a more robust increase in DZ:LZ ratio for Q36H GC B cells as compared with WT, peaking at 48 hours (P = 0.0187; Fig. 6E and fig. S14D). These results reflected a greater DZ reentry for Btg1^{Q36H} GC B cells upon receiving T_{FH} cell help signals. Collectively, these data suggested that subtle acceleration of T_{FH} cell help–related Myc induction and S phase completion led to a progressive and marked

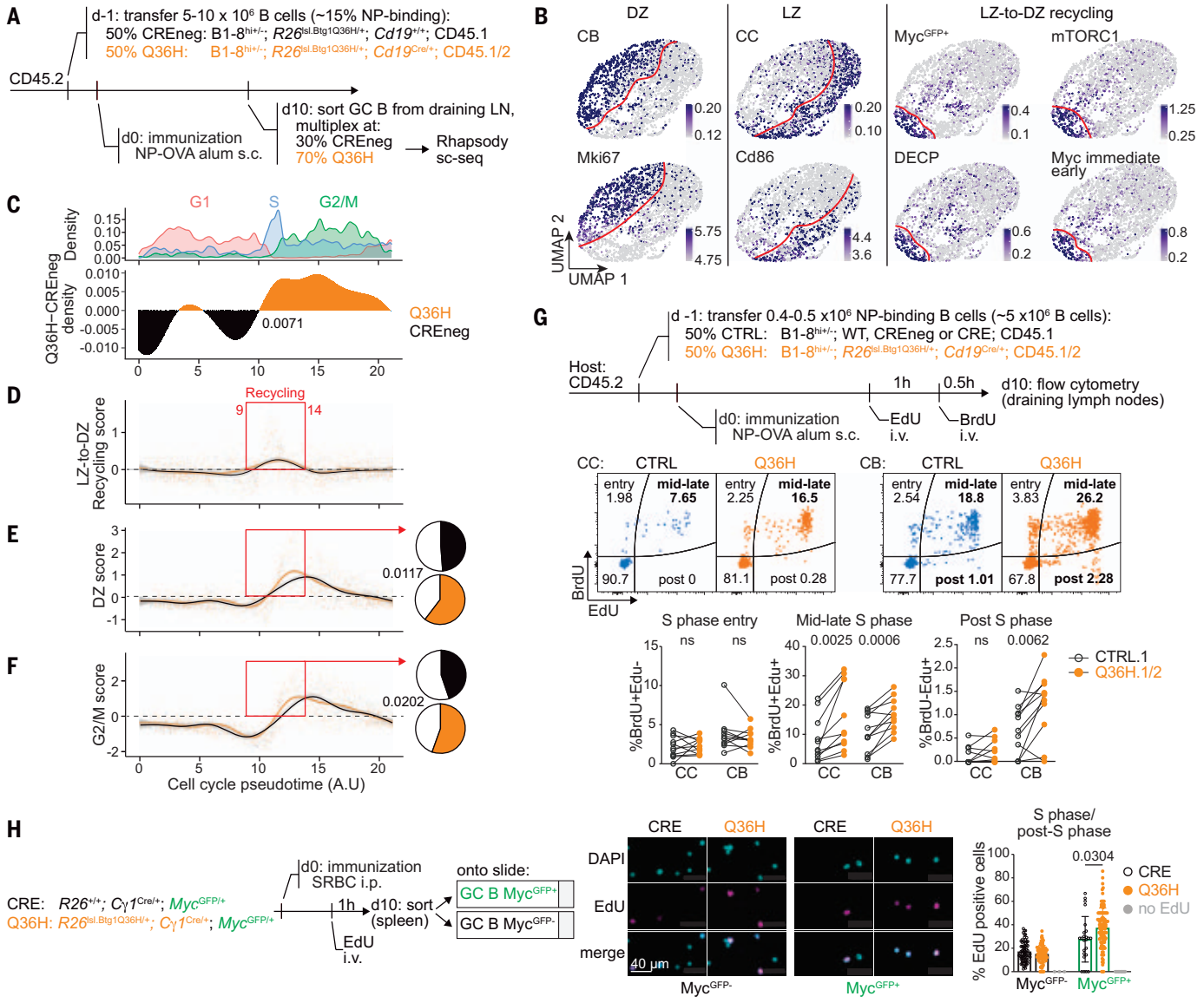


Fig. 5. Faster S phase completion and DZ program commitment in Btg1^{Q36H} LZ-to-DZ recycling cells. (**A**) Experimental design for targeted single-cell RNA sequencing for n = 496 genes in competing Q36H and CREneg GC B cells from n = 4 mice pooled into three biological replicates. (**B**) Dark zone (DZ), light zone (LZ), and LZ-to-DZ recycling GC B cells were defined on the basis of signatures enrichment and of *Mki67* and *Cd86* marker expression levels, projected onto the uniform manifold approximation and projection (UMAP) distribution of cells (n = 2982). Centroblast (CB) and centrocyte (CC) signatures correspond to genes down-regulated or up-regulated in CC versus CB, respectively. Myc^{GFP+} signature represent genes up-regulated in Myc^{GFP+} versus Myc^{GFP-} CRE control GC B cells from Fig. 3E; mTORC1 signature is PENG_RAPAMYCIN_RESPONSE_DN; DECP signature is LZ_DECP_upreg and Myc immediate early signature consists of Myc-dependent genes induced by 8 hours in Myc WT/flox B cells. Genes in these signatures are included in table S4. (**C**) Distribution of cells enriched for G1, S, and G2/M signatures (top) and of Q36H and CREneg GC B cells differential density across a Slingshot cell cycle pseudotime axis. Difference in distribution of Q36H and CREneg cells by Wilcoxon test. (**D** to **F**) Distribution of Q36H and CREneg GC B cells based on their enrichment for (D) LZ-to-DZ recycling GC B signature, (E) DZ signature, or (F) G2/M signature, across the Slingshot cell cycle pseudotime axis. (D) Pseudotime units 9 to 14 correspond to LZ-to-DZ recycling GC cells (score > 0). [(E) and (F)] Pie charts represent Q36H or CREneg GC B cells with (E) a DZ score > 0 or (F) a G2/M score > 0, as a proportion of LZ-to-DZ recycling cells (pseudotime units 9 to 14) from (D). Chi-square P value. (**G**) Experimental design for in vivo cell cycle profiling by EdU/BrdU dual staining of competing Q36H and CTRL GC B cells (top). Representative flow cytometry plots for CC and CB in S phase, mid-late S phase, and post S phase (middle). Graphs show data pooled from three independent experiments, with n = 3 or 4 mice each (bottom). Mean ± SD, paired t test (two-tailed). (**H**) Experimental design to measure S phase-experienced Q36H and CRE Myc^{GFP+} cells in vivo by EdU staining coupled with immunofluorescence (IF). Representative IF images are shown. Bar plot shows quantification from two independent experiments, each with n = 2 or 5 mice per genotype. Each dot represents one individual field of view with a minimum five cells per field of view. Mean ± SD, unpaired t test (two-tailed).

fitness advantage of GC B cells, by conferring greater DZ reentry potential and an opportunity to undergo increased GC repetitive rounds (Fig. 6F).

Btg1^{Q36H} expression induces highly aggressive and rapidly progressive lymphomas

We next explored whether *Btg1*^{Q36H} is a bona fide lymphoma oncogene. Given that MCD/cluster 5 DLBCLs derive from B cells having transited the GC reaction, we induced Btg1^{Q36H} expression using *Cγ1*^{Cre}. MCD-DLBCLs typically express high BCL2 levels (*11*). We crossed our

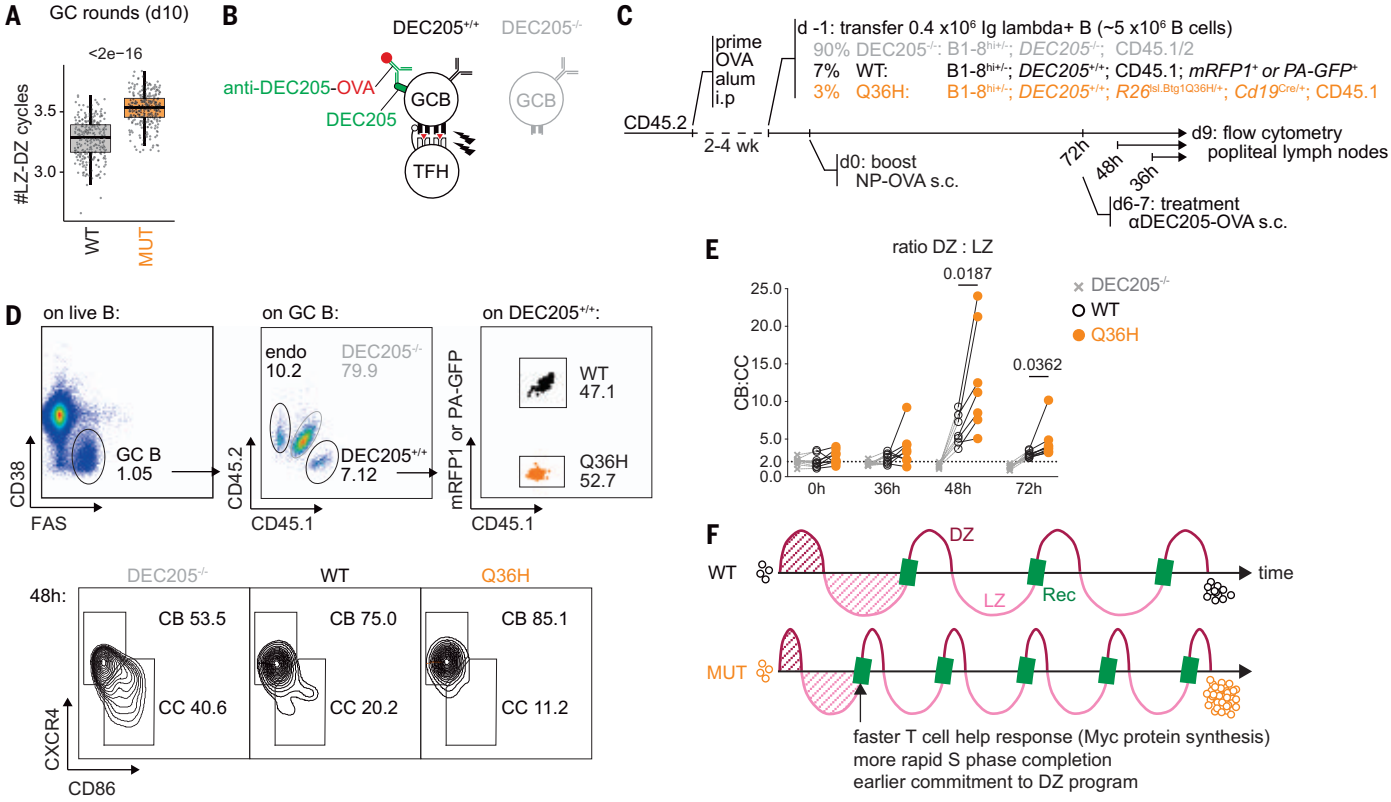
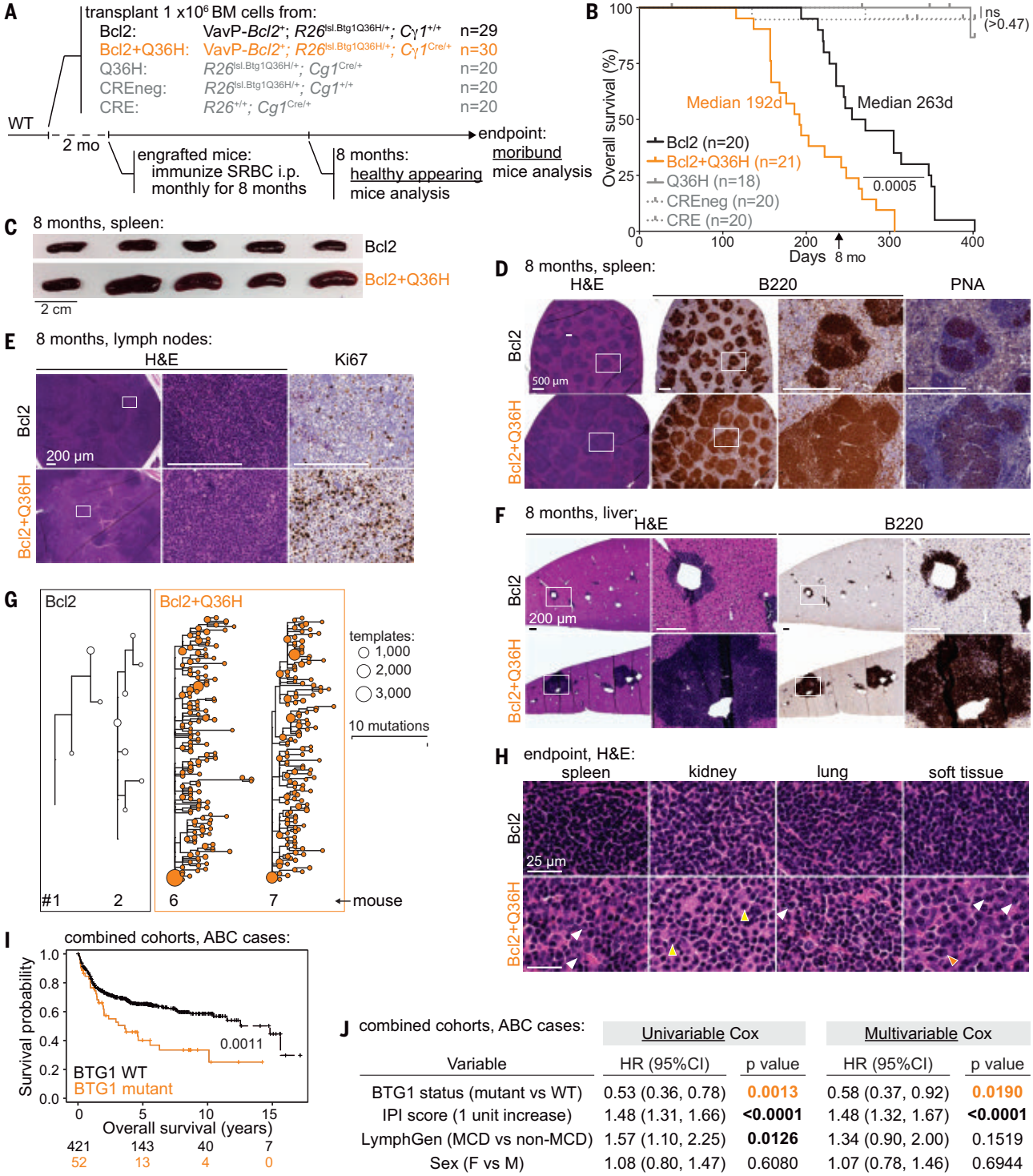


Fig. 6. Accelerated LZ-DZ recycling kinetics in Btg1^{Q36H} GC B cells. (**A**) In silico readout for the number of LZ-DZ cycles undergone by competing Btg1 mutant versus WT GC B cells at day 10 after immunization, using the GC mathematical model in which mutant GC B cells are provided with 1.12 times faster T cell help response (faster up-regulation of Myc and mTORC1) and 21% shorter S phase. Each dot is a readout from a single simulation (n = 300). Difference by Wilcoxon test considering the mutant and WT readout from the same simulation as paired samples. (**B**) Schematic of the system used to deliver targeted T_{FH} cell help to competing GC B cells in vivo. (**C**) Experimental design for the targeted delivery of T_{FH} cell help to competing DEC205$^{+/+}$ WT and Q36H GC B cells in vivo and tracking their centroblast (CB) versus centrocyte (CC) identity over time. (**D**) Representative gating flow cytometry plots for C. (**E**) Graph shows pooled data for (C) and (D), from two experiments, each with n = 3 to 5 mice per time point. Paired t test, two-tailed. (**F**) Schematic model. More-rapid GC LZ-DZ cycles in Btg1 mutant (MUT) GC B cells, because of faster response to T_{FH} cell help via accelerated Myc protein induction, more rapid S phase completion and earlier commitment to DZ transcriptional program, explain the progressive competitive fitness gain of Btg1^{Q36H} GC B cells over the course of the GC reaction. Rec, LZ-to-DZ recycling.

mice to the VavP-*Bcl2* lymphoma mouse model (*48*, *49*) to reflect this biology. We generated cohorts of Bcl2, Bcl2+Q36H, Q36H, CREneg, and CRE mice through bone marrow hematopoietic stem cell transplantation into syngeneic wild-type recipients (Fig. 7A). Bcl2+Q36H animals displayed shorter survival than did Bcl2 animals (P = 0.0005; Fig. 7B). Notably, survival outcomes were indistinguishable between Q36H-only mice and CREneg or CRE control groups (P > 0.47; Fig. 7B). As opposed to Bcl2+Q36H, Bcl2+WTKI animals did not experience more-rapid mortality rates than Bcl2 animals, in a similar transplantation cohort using fetal liver hematopoietic stem cells (fig. S15, A to C).

To determine whether mutant Btg1 accelerated disease onset, we analyzed a subset of seemingly healthy Bcl2+Q36H and Bcl2 animals (n = 5 per group) at 8 months after transplantation. Bcl2+Q36H mice manifested increased splenomegaly (P = 0.0257; Fig. 7C and fig. S15D). Histologically, Bcl2 spleens showed a prominent follicular pattern but otherwise largely intact architecture, whereas Bcl2+Q36H animals manifested distortion and expansion of B cell follicles (B220$^+$) and GC areas (PNA$^+$; Fig. 7D). GC B cells (FAS$^+$CD38$^-$) were increased in proportion to B cells (B220$^+$; fig. S15E), were larger (fig. S15F), contained a higher fraction of mTORC1-activated cells [phosphorylated S6 (pS6$^+$); fig. S15G], and were more clonal (IgVH PCR on cDNA; fig. S15H) in Bcl2+Q36H versus Bcl2 spleens. Unsupervised analysis of RNA-seq performed in these lymphoma cells showed markedly distinct transcriptional profiles (fig. S15I). Mutant BTG1-induced signatures such as positively selected GC B cells and MYC targets were significantly enriched (FDR < 0.001 and 0.005, respectively), suggesting maintenance of BTG1 mutant effects (fig. S15J). Lymph node architecture was severely disrupted in Bcl2+Q36H mice, with expansion of extrafollicular areas containing enlarged and proliferative (Ki67$^+$) lymphoma cells, as compared with Bcl2 (Fig. 7E). Most notably, Bcl2+Q36H mice displayed invasion of malignant B cells into extranodal tissues such as the liver, kidney, and lung, whereas Bcl2 mice only manifested minor perivascular infiltrates in these tissues (Fig. 7F and fig. S15K).

To assess Btg1^{Q36H} effect on the genetic trajectory during lymphomagenesis, we performed B cell receptor (BCR) sequencing (fig. S16, A and B). All lymphomas harbored somatic hypermutation activity, typical of their GC origin (fig. S16C). Bcl2+Q36H lymphoma B cells manifested significantly greater mutational burden (fig. S16D) and clonal expansion when compared with Bcl2, as represented by their lower Simpson's index and expanded phylogenetic trees (Fig. 7G and fig. S16, E and F). Furthermore, Bcl2+Q36H clonal lineages presented lower estimated replacement-to-silent mutation rate ratios (dN/dS) than Bcl2, which is indicative of purifying selection and consistent with more cycles of division and mutation in germinal centers (*50*) (fig. S16G). In line with this, mathematical modeling (fig. S14B) predicted a higher number of divisions in selected mutant versus wild-type GC B cells over the



J combined cohorts, ABC cases:

	Univariable Cox		Multivariable Cox	
Variable	HR (95%CI)	p value	HR (95%CI)	p value
BTG1 status (mutant vs WT)	0.53 (0.36, 0.78)	**0.0013**	0.58 (0.37, 0.92)	**0.0190**
IPI score (1 unit increase)	1.48 (1.31, 1.66)	**<0.0001**	1.48 (1.32, 1.67)	**<0.0001**
LymphGen (MCD vs non-MCD)	1.57 (1.10, 2.25)	**0.0126**	1.34 (0.90, 2.00)	0.1519
Sex (F vs M)	1.08 (0.80, 1.47)	0.6080	1.07 (0.78, 1.46)	0.6944

Fig. 7. Mutant BTG1 drives formation of highly aggressive B cell lymphomas in mice and humans. (**A**) Experimental design for testing contribution of the *R26*$^{lsl.Btg1Q36H/+}$ allele to Bcl2-driven lymphomagenesis. (**B**) Kaplan-Meier curves depicting overall survival of groups described in (A) in days after transplantation, assessed by either time of death or euthanasia upon sickness development. Log-rank (Mantel-Cox) test performed as indicated. (**C**) Image of *n* = 5 spleens per group at 8 months after transplantation. (**D**) Consecutive spleen sections from Bcl2 and Bcl2+Q36H mice at 8 months after transplantation stained with H&E or by immunohistochemistry (IHC) for B cells (B220 antibody) or GC B cells (PNA). Representative of *n* = 5 mice per genotype. (**E**) Consecutive lymph node sections from Bcl2 and Bcl2+Q36H mice at 8 months after transplantation stained with H&E or by IHC for proliferating cells (Ki67). Representative of *n* = 5 mice per genotype. (**F**) Consecutive liver sections from Bcl2 and Bcl2+Q36H mice at 8 months after transplantation stained with H&E or by IHC for B cells (B220). Representative of *n* = 4 mice per genotype. (**G**) B cell lineage trees from BCR sequences analysis (ImmunoSeq) in Bcl2 and Bcl2+Q36H mice at 8 months after transplantation. The B cell clone with the most sequence reads for each mouse (mouse number indicated in the bottom) is shown. Branch lengths represent the estimated number of somatic hypermutations between nodes. Scale bar denotes 10 mutational events. Trees from *n* = 2 mice, representative of *n* = 5 mice per genotype (fig. S16F), are

shown. (**H**) H&E staining on indicated tissue sections from Bcl2 and Bcl2+Q36H moribund mice. Representative of *n* = 6 mice per genotype. Low-grade follicular lymphoma–like centrocytes shown in Bcl2 mice. In Bcl2+Q36H mice, white arrowheads depict large immunoblastic cells; yellow, plasmacytoid cells; and orange, a large-size DLBCL-like cell. (**I**) Kaplan-Meier curves depicting overall survival of ABC-DLBCL patients combined from publicly available cohorts (*9*, *23*, *24*), based on their *BTG1* status. Log-rank test was performed. Censored events are indicated below the graph. (**J**) Univariate and multivariate analyses in ABC-DLBCL cases from the same combined cohorts as in (I). HR, hazard ratio with 95% confidence interval. *P* value from Wald chi-square test for H0:HR = 1.

course of the GC reaction (fig. S16H). Collectively, lymphoid architecture disruption and invasion of extralymphoid tissues by more proliferative, dysplastic, highly mutated, and clonally expanded tumor cells indicated more advanced and phenotypically malignant disease in Bcl2+Q36H mice, with the genetics indicating that $Btg1^{Q36H}$ B cells underwent more rounds of selection and somatic hypermutation, in line with their fitness advantage.

Further along these lines, moribund Bcl2 mice presented with low-grade follicular centrocyte-like lymphoma patterns, whereas moribund Bcl2+Q36H mice featured sheets of larger lymphoid cells, reminiscent of DLBCL with many cells showing immunoblastic or plasmacytoid appearance (Fig. 7H). These differential features were observed in the spleen, as well as extranodal tissues (Fig. 7H). Bcl2+Q36H mice also uniquely manifested cases of extreme disease invasiveness into peripheral organs (fig. S16I), underlining the impact of $Btg1^{Q36H}$ in driving phenotypically high-grade aggressive lymphomas, with features similar to human cases.

Consistent with these aggressive phenotypes, we observed inferior clinical outcomes for *BTG1* mutant patients within ABC-DLBCLs from publicly available cohorts (*9*, *23*, *24*) (P = 0.0011; Fig. 7I). Univariable Cox regression further showed significant association between *BTG1* mutational status and inferior overall survival among ABC-DLBCLs (P = 0.0013; Fig. 7J and fig. S16J). Notably, a multivariable Cox regression analysis, including DLBCL genetic subtype (LymphGen), gender, International Prognostic Index (IPI), and *BTG1* mutation status as covariates, showed that *BTG1* mutation remained significantly associated with overall survival (P = 0.0190; Fig. 7J and fig. S16J). Collectively, these data indicated that *BTG1* mutation presents potent oncogenic properties. Consistent with its role in providing supercompetitive fitness to GC B cells, mutant BTG1 conferred an aggressive and tissue-invasive phenotype to B cell lymphomas, resulting in extranodal spread and inferior clinical outcomes.

Discussion

Multicellular life requires individual cells to sacrifice competitive fitness to ensure homeostasis and survival. Mechanisms that enable "altruistic" cooperation between cells include restraining cell proliferation and preventing acquisition of somatic mutations among other functions and are often viewed as weakened in cancer (*51*). However, certain nonaltruistic, competitive processes take place in normal embryonic development (*52*). The archetypal example of supercompetition was first described during *Drosophila* development, whereby cells with greater Myc expression outgrow and may also cull neighbor cells with slightly lower Myc levels (*53*, *54*). The GC reaction provides an intriguing microcosm of the cancer "return to unicellularity" concept in normal nonembryonic physiology, whereby restraints on proliferation and somatic mutations are attenuated and a Myc-dependent competitive process of clonal diversification and selection takes place. Critical to keeping this process under control, expression of Myc is permitted only transiently in GC B cells, together with mTORC1 activation, to provide time-limited biosynthetic capabilities (*5*).

The absolute arbiter of Myc induction in GC B cells is the intensity of T_{FH} cell help signals received on the basis of affinity for cognate antigen. The resulting range of Myc "doses" can determine initial S phase duration and how many divisions a selected GC B cell undergoes (*45*, *55*). Our findings suggest that in addition to Myc dosage, the rate of Myc protein induction may influence GC B cell fitness, because it is associated with earlier S phase completion, greater DZ reentry, and earlier activation of proliferative programs. Naïve B cell activation models also suggest a critical function for Myc induction kinetics in determining B cell expansion and cell fate (*56*). However, the competitive gain conferred by mutant Btg1 was restricted to the GC B cell compartment. Thus, analogous to embryonic supercompetition, GC B cells with slightly superior Myc induction kinetics have the potential to outcompete and dominate the GC reaction through clonal expansion.

Given these scenarios, it is plausible that B cells would have evolved ways to control immune signaling thresholds for Myc induction, especially because GC B cells are exposed to an environment rich in growth-promoting cytokines and ligands. Our data suggest that wild-type BTG1-mediated restriction of MYC translation may represent one such mechanism, acting as a safeguard or gatekeeper against unrestrained "unicellular-like" fitness. Mutant BTG1 released MYC from this restrictive effect, thus converting GC B cells into supercompetitors that transformed into extremely aggressive lymphoma cells. Moreover, Btg1 mutant cells aberrantly retained MYC signatures while activating DZ and G2/M programs, suggesting that biosynthetic growth and proliferation abnormally coexisted rather than being strictly separated. Future studies will determine whether this supercompetition effect is further associated with eliminating wild-type cells, perhaps depriving them from positive selection signals or resource utilization in their microenvironment. The impact of *BTG1* mutation on the magnitude of T_{FH} cell help response and MYC induction kinetics was quite subtle, implying that competing GC B cells operate at the most extreme limits of multicellular homeostasis, a small perturbation of which is sufficient to tip the balance toward malignancy. The iterative nature of the GC reaction may further amplify this subtle effect over time. These findings suggest a delicate evolutionary trade-off between defending vertebrate organisms from infections versus facilitating certain "atavistic" features in B cells that recall the more intense and dynamic features of unicellular life.

In T cells, BTG1 was reported to maintain quiescence through global mRNA degradation (*41*) and interaction with CCR4-NOT (*15*, *20*, *57*), a deadenylase complex regulating mRNA stability. In HEK293T cells, certain BTG1 lymphoma mutants manifested impaired interaction with CCR4-NOT (*20*). However, mutant BTG1 did not affect *MYC* mRNA levels or mRNA stability in GC or DLBCL cells. The BTG family member TOB1 potentially binds mRNAs through 3′ untranslated region (3′UTR), 5′UTR, and coding sequences, suggesting functions beyond mRNA stability regulation (*40*). However, BTG1 lacks TOB1 C-terminal domain or any known RNA binding domain and instead associated with RNAs indirectly. Given that BTG1 mutants facilitated translation of MYC and potentially other proteins, we hypothesize that wild-type BTG1 association with *MYC* mRNAs restricts MYC translation, until reaching a critical T_{FH} cell help strength threshold, whereas mutant BTG1 lowers this threshold through impaired association with *MYC* mRNAs. These effects are likely mediated through interaction with as of yet unknown RNA binding proteins that might be destabilized by the altered protein conformation induced by mutations such as Q36H (*31*). This would confer a partial dominant-negative effect, whereby mutant BTG1 could occupy wild-type BTG1 space, preventing its ability to form functional RNA regulatory complexes while preserving other BTG1 functions. The dominant effect of these mutations is further underlined by the relatively low fraction of *BTG1* mutant transcripts in primary human DLBCLs

and our animal models. Although *BTG1* is described as a putative tumor suppressor gene in several cancer types (*58*), it behaved instead as a proto-oncogene in DLBCL. This is indicated by its functional gain-of-fitness effects and focal mutational pattern, but only rare truncating mutations or loss of heterozygosity. Other genes can act both as tumor suppressors and proto-oncogenes depending on the cellular context (*59*, *60*). For example, the polycomb protein EZH2 is a tumor suppressor deleted in leukemia, as is *BTG1* (*61*), but is an oncogene with gain-of-function mutations in B cell lymphoma (*6*).

Myc can suppress Btg1 expression through induction of *miR-17-92* (*42*), suggesting a positive feedback mechanism. Indeed, we observed that BTG1 expression was reduced in Myc-positive GC B cells or upon induction of MYC in DLBCL cells, and *miR-17-92* deficiency impairs GC responses (*62*). In line with antagonistic effects, *MYC* and *BTG1* mRNA levels are inversely correlated during cell cycle progression in stimulated peripheral blood lymphocytes and NIH3T3 cells (*63*). However, there may be additional mechanisms through which MYC can repress BTG1 expression, given that in our DLBCL cells, V5-tagged BTG1 lacked the 3′UTR binding sites for *miR-17-92*. *Btg1* transcript levels are evidently tightly controlled in GC B cells because expression of $Btg1^{Q36H}$ or $Btg1^{WT}$ from the *Rosa26* locus did not affect overall *Btg1* transcript abundance.

Finally, the lymphomagenic potential of mutant BTG1–induced fitness was underlined by its powerful cooperative effect with BCL2 overexpression, both of which are associated with the MCD-DLBCL subtype. The competitive advantage conferred by mutant BTG1 might be especially deleterious in cells with constitutive Toll-like receptor and B cell receptor signaling due to *MYD88* and *CD79B* mutations. Alternatively, the aberrant memory and clonal precursor cells induced by *MYD88* and *TBL1XR1* mutations (*12*, *13*, *64*) could gain potential fitness advantage for reentry into subsequent GC reactions. Finally, the supercompetitor phenotype conferred by *BTG1* mutations seems to manifest clinically, because *BTG1* mutation independently associated with inferior outcome among ABC-DLBCLs. It is thus warranted for future studies to explore potential therapeutic vulnerabilities that could be exploited for treatment of these lethal tumors.

Potential limitations

The lymphomagenic potential of $Btg1^{Q36H}$ or $Btg1^{WT}$ was evaluated in separate experiments without affecting our conclusions, because they were evaluated in comparison to their respective positive and negative controls in each experiment (Fig. 7B and fig. S15B) in comparable experimental setups (fig. S15C). In some mouse experiments, we assessed mutant $Btg1^{Q36H}$ effects using animals that did not carry the $Cd19^{Cre}$ allele. This is not expected to be an issue for the following reasons: (i) We observed a gain-of-function phenotype (i.e., gain of competitiveness) rather than a loss-of-function effect, and Cre-associated toxicity would, on the contrary, impair GC B cell competitiveness through cell-killing effects (*65*). (ii) Cre knock-in at the *Cd19* locus results in only one functional *Cd19* allele (*32*). Given its function as a BCR co-receptor, $Cd19^{Cre}$ knock-in would be expected to reduce B cell activation and competitiveness. (iii) Throughout the manuscript, and using multiple orthogonal approaches, Q36H was shown to provide a similar competitive advantage to GC B cells against any tested control, including Cre-positive and Cre-negative cells (e.g., Fig. 2, A to C, and figs. S6, A and C, S7, A and B, and S13E).

Materials and methods

Mouse models

Animal care was in strict compliance with institutional guidelines established by Weill Cornell Medicine, *Guide for the Care and Use of Laboratory Animals* (*66*), and the Association for Assessment and Accreditation of Laboratory Animal Care International. All mouse procedures were approved by the Research Animal Resource Center and Institutional Animal Care and Use Committee of Weill Cornell Medicine (protocol #2011-0031). All mouse experiments were conducted using age- and sex-matched animals 8 to 14 weeks old, unless specified otherwise. All experiments included male and female animals in all groups, except for adoptive transfers, with all recipient mice being males. No sex-based influence or bias was detected in the observations made in this work.

Conditional knock-in $Btg1^{Q36H}$ mice ($R26^{lsl.Btg1Q36H}$) were generated by inGenious Targeting Laboratory Inc. (Ronkonkoma, NY, USA) via insertion of a stop cassette in intron 1 of the *Rosa26* locus followed immediately by a mutated murine $Btg1^{Q36H}$ cDNA-T2A-*Luciferase* sequence (fig. S2A). Murine *Btg1* and human *BTG1* cDNA sequences are slightly different, but their protein sequence is 100% identical. The targeting vector was constructed as follows: The murine $Btg1^{Q36H}$ cDNA-T2A-*Luciferase* sequence was first cloned into the MluI site of a pSa-stop-bGH polyA vector to generate the construct pSa-stop-$Btg1^{Q36H}$ cDNA-T2A-*Luciferase*-bGH polyA. The stop cassette was LoxP-PGK/gb2 promoter-neo/kana resistance gene-PGK polyA-2X SV40 polyA signal-LoxP. The final cassette therefore consisted of a splice acceptor, a LoxP-flanked stop cassette, $Btg1^{Q36H}$ cDNA-T2A-*Luciferase,* and a bGH polyA signal. Then, the entire cassette was inserted into the XbaI site (XbaI was disrupted by this insertion) of the pROSA26-1 vector using the Recombineering technology. The resulting targeting vector contained a short homology arm with a 1.08-kb *Rosa26* genomic sequence upstream of the cassette and a 4.34-kb-long homology arm downstream of the cassette. The targeting vector was confirmed by restriction analysis and sequencing after each modification. iTL IC1 (C57BL/6) embryonic stem cells were electroporated with this targeting vector, stable G418-resistant clones were derived and microinjected into BALB/c blastocysts. Resulting chimeras with a high-percentage black coat color were mated to wild-type C57BL/6N mice to generate F1 heterozygous offspring. Tail DNA was analyzed from pups with black coat color. Short homology arm integration, long arm junction, and targeted allele sequence were confirmed by PCR followed by gel electrophoresis and Sanger sequencing (Genewiz).

Conditional $Btg1^{WTKI}$ mice ($R26^{lsl.Btg1wt}$) were generated at the MSKCC Mouse Genetics Core Facility by introducing a single-nucleotide substitution to revert the Q36H mutation of the cDNA sequence in $R26^{lsl.Btg1Q36H}$ mice to a wild-type allele, by CRISPR-Cas9–mediated genome targeting (fig. S2, D and E). Oocytes from 3- to 5-week-old $R26^{lsl.Btg1Q36H/lsl.Btg1Q36H}$ or $R26^{lsl.Btg1Q36H/+}$ superovulated females and sperm from ≥2-month-old $R26^{lsl.Btg1Q36H/lsl.Btg1Q36H}$ males were used for in vitro fertilization. Resulting zygotes were injected with a mix of Cas9 protein (Integrated DNA Technologies), CRISPR RNA (crRNA, 5′-GGCTCTGGCTGAAAGTCTGC-3′), trans-activating CRISPR RNA (tracrRNA, 5′-AAACAGCAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCU-3′), and a 105–base pair (bp) donor oligo for homologous recombination (HR oligo, 5′-ATCTCCAAGTTCCTCCGCACCAAGGGGCTCACGAGCGAGCGACA**G**CTGCAGACTTTCAGCCAGAGCCTGCAGGAGCTGCTGGCAGAACATTACAAACATCACTGG-3′), which contained a C>G substitution (in bold in the above sequence) to revert H36 (CAC) to Q36 (CAG). The mutated C in $Btg1^{Q36H}$ was also the third base of the protospacer adjacent motif (PAM) sequence on the reverse strand and therefore did not exist in the endogenous *Btg1* sequence or in the *Rosa26* wild-type reverted $R26^{lsl.Btg1wt}$ allele, preventing Cas9 recutting of the reverted allele. The C>G substitution also created a new PvuII site, which we used to genotype the pups. After injection, two-cell embryos were implanted into day 0.5 pseudo-pregnant females. Pups were then checked by PCR on genomic DNA and PvuII digestion (fig. S2, E and F). Potential founders were further checked by Sanger sequencing (Genewiz).

The following strains were obtained from The Jackson Laboratory (Ben Harbor): C57BL/6J (CD45.2, stock 000664), B6.SJL-$Ptprc^{a}Pepc^{b}$/BoyJ (CD45.1, $SJL^{+/+}$, stock 002014), $Cd19^{Cre}$

(*32*) (stock 006785), *Cγ1*^{Cre} (*38*) (stock 010611), *Myc*^{GFP} (*37*) (stock 021935), B1-8^{hi} (*33*) (stock 007594), *DEC205*$^{-/-}$ (*67*) (stock 005528), *PA-GFP* (*44*) (stock 022486), and *mRFP1* (*68*) (stock 005884). The VavP-*Bcl2* (*48*) model was developed by J. M. Adams (Walter and Eliza Hall Institute of Medical Research, Australia).

Genotyping

PCR-based genotyping of the *R26*$^{lsl.Btg1Q36H}$ and *R26*$^{lsl.Btg1wt}$ alleles was performed on mouse tail lysates. Tail tips were lysed in DirectPCR lysis reagent (Viagen 102T) plus proteinase K during overnight incubation at 55°C with shaking, followed by 45 min incubation at 85°C (no shaking) to inactivate proteinase K. After a short spin, supernatant was used for PCR using Taq polymerase (Roche). The forward primer annealed to the *Rosa26* short homology arm upstream of the cassette and was combined with two reverse primers. The first one annealed to the 5′LoxP site of the targeted allele, and the second one annealed to the *Rosa26* long homology arm downstream of the cassette and amplified only the nontargeted allele (because of a >5 kb distance with the forward primer in the targeted allele). Details about primers are included in table S8. The PCR program was 94°C for 2 min, [94°C for 30 s, 65°C for 30 s, and 72°C for 1 min] for 40 cycles, and 72°C for 7 min. Genotyping for *R26*$^{lsl.Btg1Q36H}$ and *R26*$^{lsl.Btg1wt}$ alleles, as well as all other mouse alleles used in this study, was also done by Transnetyx using reverse transcription quantitative polymerase chain reaction (RT-qPCR).

RNA extraction

Cells were collected from culture and resuspended in TRIzol (Invitrogen 15596018), or cells were sorted directly into TRIzol LS (one third more concentrated TRIzol reagent, Invitrogen 10296028), and total RNA was extracted. Purified RNA was resuspended in molecular-grade RNase-free water and analyzed on Nanodrop and Qubit Fluorometric Quantification for quality check and concentration measurement, respectively.

qPCR

cDNA synthesis was performed with equal amounts of total RNA for all samples in each experiment, using the Verso cDNA Synthesis kit (Thermo Fisher AB1453B). RT-qPCR was performed using same volume of cDNA for all samples, on a QuantStudio6 Flex Real-Time PCR System (Thermo Fisher) with Fast SYBR Green Master Mix (Thermo Fisher 4385614). Each biological replicate (either independently generated DLBCL lines or individual mouse samples) was further analyzed in triplicate wells to assess technical variability, of which only the mean was used to calculate the mean ±SD between biological replicates. mRNA expression levels of genes of interests were normalized to that of *TBP* and/or *GAPDH*, as specified, using the delta-delta Ct method. Details about primers used are included in table S8.

Presence of Cre-recombined alleles

Cre-mediated recombination of *R26*$^{lsl.Btg1Q36H}$ and *R26*$^{lsl.Btg1wt}$ alleles was validated on gDNA extracted from sorted naïve B cells and germinal center B cells or from B220^{+}-enriched B cells as indicated, using Puregene Gentra cell kit (QIAGEN 158388). B220^{+} B cells were enriched by positive selection with anti-B220 magnetic beads (Miltenyi Biotec 130-049-501). We performed PCR on gDNA using primers that specifically amplified the recombined targeted allele and Phusion HF polymerase (NEB M0530). The PCR program was 98°C for 30 s, [98°C for 15 s, 70°C for 30 s, and 72°C for 30 s] for 30 cycles, and 72°C for 10 min. PCR products were resolved by agarose gel electrophoresis and visualized on a ChemiDoc Touch imaging system (BioRad) using SYBR Safe DNA stain (Thermo Fisher S33102). The same primers were used to confirm expression of Cre-recombined alleles on total RNA extracted from B220^{+}-enriched B cells, reverse transcribed, and analyzed by qRT-PCR. Forward primer annealed to the *Rosa26* short homology arm upstream of the cassette and reverse primer to the kozak sequence and *Btg1* coding sequence from the *R26*$^{lsl.Btg1Q36H}$ or *R26*$^{lsl.Btg1wt}$ at the *Rosa26* locus (fig. S2, A and D), to ensure only recombined targeted alleles and transcripts were amplified from gDNA and total RNA, respectively. Primer details are included in table S8.

Efficiency of Cre-recombination

The proportion of recombined *R26*$^{lsl.Btg1Q36H}$ allele was measured in the competitive setup using adoptive transfer of B cells with an endogenous BCR repertoire into WT recipient mice. We sorted Q36H naïve B cells (NB) as live B220^{+}IgD^{+}CD45.2 and germinal center B cells (GC B) as live B220^{+}FAS^{+}CD38^{-}CD45.2 from the spleens of CD45.1 recipient mice at day 21 after immunization, when Q36H GC B cells are outcompeting WT recipient cells. gDNA was extracted using Puregene Gentra cell kit (QIAGEN 158388) with volumes adapted for low material. For each sample, a minimum of 240 pg gDNA per well, as measured by Qubit dsDNA HS (Invitrogen Q32854), was used for qPCR amplification using two sets of primers, the first one specifically amplifying the recombined alleles (same as above for PCR on gDNA or for qPCR on cDNA) and the second one specifically amplifying the nonrecombined alleles, sharing the forward primer with the first set, but with the reverse primer annealing to the STOP cassette (fig. S7D). Primer details are included in table S8. A standard curve for each of these primer sets was generated, using their corresponding double-stranded DNA (dsDNA) amplicons to the amount of 10^3 to 10^8 absolute copy number per well. The efficiency of Cre-mediated recombination was calculated as the absolute copy number of recombined alleles over the absolute copy number of recombined plus nonrecombined alleles.

IgH rearrangement

For clonal variance analysis in the lymphomagenesis study cohort, GC B cells from Bcl2-only and Bcl2+Q36H mice were sorted as live B220^{+}FAS^{+}GL7^{+} cells from PNA^{+} enriched splenocytes on BD Influx or BD FACSAria II, 8 months after transplantation. We extracted RNA and synthesized cDNA using equal amounts of RNA for all samples, as described above. In parallel, we processed centroblasts sorted from a wild-type mouse that had been immunized 9 days before with 2% SRBC, to use as a nontumor control. We then performed PCR on equal volume of cDNA for all samples with GoTaq Flexi DNA Polymerase (Promega M8295) to amplify the VHQ52-JH4 and VH7183-JH4 regions of the *IgH* locus, using a set of forward primers that annealed to the framework region of one of the most abundantly used IgVH gene families and a reverse primer annealing to the JH1-4 gene segments, as described before (*69*). Each sample was amplified in duplicates, both of which were resolved by agarose gel electrophoresis on adjacent lanes and visualized on a ChemiDoc Touch imaging system using SYBR Safe DNA stain. The PCR program was 95°C for 2 min, [95°C for 45 s, 60°C for 45 s, and 72°C for 2 min] for 40 cycles, and 72°C for 10 min. Primer details are included in table S8.

Immunizations

To induce GC formation in draining lymph nodes, mice were immunized by subcutaneous injection of the highly substituted hapten NP conjugated to the carrier protein ovalbumin (OVA), NP(16-19)-OVA (Biosearch Technologies N-5051), either into the hind footpad for popliteal lymph nodes collection, or into the base of tail for inguinal lymph nodes collection. We did not observe differences in GC response after NP-OVA immunization in popliteal versus inguinal lymph nodes. In noncompetitive setup, non-B1-8^{hi} mice were injected with 50 μg of NP(16-19)-OVA. In competitive setup, recipients of B1-8^{hi} adoptively transferred B cells were injected 24 hours after transfer with 20 μg of NP(16-19)-OVA subcutaneously for T-dependent stimulation, or with 50 ug of NP(44)-Ficoll (Biosearch Technologies F-1420) intraperitoneally for T cell–independent stimulation. Conjugated NP-OVA or NP-Ficoll was absorbed in alum adjuvant (Thermo Fisher 77161) at a 1:3 ratio of alum:immunogen prior to injection.

To induce germinal center (GC) formation in spleens, mice were immunized by intraperitoneal injection of 500 ml of 2% SRBCs resuspended in sterile 1X Dulbecco's phosphate-buffered saline (DPBS) from a solution of sheep blood in Alsever's (Cocalico Biologicals 20-1334A), or by intraperitoneal injection of 80 μg NP(30-32) conjugated to the carrier protein keyhole limpet hemocyanin (KLH) (Biosearch Technologies N-5060) absorbed in alum adjuvant (Thermo Fisher 77161) at a 1:1 ratio of alum:immunogen prior to injection.

For the experiments with in vivo targeted delivery of T cell help, recipients of a mixture of B1-8^{hi}; *DEC205*$^{+/+}$ and B1-8^{hi}; *DEC205*$^{-/-}$ adoptively transferred B cells were primed by intraperitoneal injection of 50 μg OVA (Biosearch Technologies O-1000) absorbed in alum at a 1:2 ratio of alum:immunogen 2 to 4 weeks prior to the transfer and received a boost immunization by subcutaneous injection of 50 μg NP(19)-OVA without alum into the hind footpad 24 hours after transfer. At the indicated times at days 6 and 7 after boost, mice received 5 μg of anti-DEC-OVA chimeric antibody diluted in phosphate-buffered saline (PBS) by subcutaneous injection into the hind footpad, and popliteal lymph nodes were analyzed at day 9 after boost. Anti-DEC-OVA was produced by transient transfection of HEK293T cells as previously described (*70*).

Competitive adoptive cell transfers

Spleens from mice of the indicated genotypes were mashed and filtered through a 40 μm cell strainer. After treatment with red blood cell lysis solution (QIAGEN 158904), splenocytes were filtered into PBS supplemented with 0.5% bovine serum albumin (BSA) and 2 mM EDTA (PBE). Resting B cells were isolated by negative selection with magnetic cell separation using anti-CD43 beads (Miltenyi Biotec 130-049-801). To determine the proportion of NP-binding B1-8^{hi} cells, we stained a fraction of cells with 5 μg/ml NP(16)-PE (Biosearch Technologies N-5070) and analyzed them by flow cytometry. We then adoptively transferred a total number of isolated B cells corresponding to 0.3×10^6 to 0.5×10^6 NP-binding B cells per recipient, by intravenous injection through the retro-orbital venous sinus, at the proportions indicated in the experimental schematic for each relevant figure. Alternatively, for the transfer of non-B1-8^{hi} mature B cells, we directly adoptively transferred a total number of 20 million isolated B cells per recipient by intravenous injection through the retro-orbital venous sinus.

To measure the proliferation rate of adoptively transferred B cells responding to NP-Ficoll immunization, isolated resting B1-8^{hi} B cells from both genotypes were mixed at equal proportions of NP-binding cells, stained with CellTrace CFSE (Invitrogen C34554) at a final concentration of 5 μM in PBS for 10 min at 37°C, and washed with PBE prior to injection into recipients.

Bone marrow and fetal liver cell transplantation

For lymphomagenesis studies, we transplanted bone marrow cells (Bcl2$^{+/-}$Q36H cohort) or fetal liver cells (Bcl2$^{+/-}$WTKI cohort) from donors of the indicated genotypes into lethally irradiated C57BL/6J females (two doses of 450 rad, on a RS 2000 Biological Research X-ray Irradiator, Rad Source Technologies). Bone marrow cells were harvested from the tibias and femurs of donor mice, filtered through a 40 μm mesh, treated with red blood cell lysis solution (QIAGEN 158904), counted, and frozen down as viable cells at a density of <15 million cells/ml. Genotype of donor mice was double checked by PCR and gel electrophoresis before thawing and transplantation. Fetal livers were isolated from E14.5 embryos and dissociated between the frosted, sandblasted ends of two sterile glass slides (VWR 48312-024). Fetal liver cells were filtered through a 35 μm mesh, counted, and frozen down as viable cells. Genotyping was performed on embryo head lysates prepared the same way as tail tip lysates, by PCR and gel electrophoresis.

Bone marrow or fetal liver cells were then thawed and cultured at 5 million/ml for 2 hours at 37°C in RPMI with 10% fetal bovine serum (FBS), 50 U/ml penicillin, 50 μg/ml streptomycin, freshly added beta-mercaptoethanol at 55 μM, interleukin-3 (IL-3) at 6 ng/ml (R&D Systems 403-ML), SCF at 10 ng/ml (R&D Systems 455-MC), IL-6 at 10 ng/ml (PeproTech 216-16) and IL-7 at 5.4 ng/ml (R&D Systems 407-ML). Cells were then collected, washed three times with 1X DPBS, counted, and resuspended in 1X DPBS for intravenous injection of 1 million bone marrow cells or 0.24 million fetal liver cells in 100 μl per lethally irradiated recipient, through the retro-orbital venous sinus. After 2 months, fully engrafted mice were immunized monthly by intraperitoneal injection of 500 ml of 2% SRBC in sterile 1X DPBS for 3 or 8 months, as specified, to induce GC formation. With the exception of mice euthanized at 8 months where indicated, all mice involved in lymphomagenesis studies were monitored until any one of several criteria for euthanizing were met, including severe lethargy, more than 10% body weight loss, or palpable splenomegaly that extended across the midline.

Live and in situ mouse imaging

Luciferase reporter imaging in live animals carrying the *R26*$^{lsl.Btg1Q36H}$ or *R26*$^{lsl.Btg1wt}$ allele and in situ imaging was performed on an IVIS SpectrumCT In Vivo Imaging System instrument (PerkinElmer) using the Living Image software (PerkinElmer). Mice were injected intraperitoneally with 2.5 mg of luciferin substrate (200 μl at 12.5 mg/ml, Promega P1043) and anesthetized with 2% isoflurane. Bioluminescence signal (radiance, in photons per second per square centimeter per steradian) was acquired 15 min after injection of luciferin. For each experiment, images were set to the same scale and total luminescence signal (flux, in photons per second) was quantified using one region of interest per mouse in the Living Image software. For in situ imaging, mice were euthanized, immediately injected intraperitoneally with 2.5 mg of luciferin substrate, opened in the intraperitoneal area to expose the spleens and surrounding lymph nodes, and imaged within 2 min of euthanasia.

H&E staining and immunohistochemistry

Mouse organs were fixed in 10% neutral buffered formalin (Sigma) for 24 to 48 hours and transferred to 70% ethanol. Tissues were then embedded in paraffin, processed, and stained at the Laboratory of Comparative Pathology [Weill Cornell Medicine (WCM)/Memorial Sloan Kettering Cancer Center (MSKCC)]. Briefly, 5 μm sections were deparaffinized and stained with hematoxylin and eosin (H&E) or processed for immunohistochemistry by heat antigen-retrieval in citrate buffer at pH 6.4, followed by treatment with 3% hydrogen peroxide in methanol to block endogenous peroxidase (HRP) activity. Indirect immunohistochemistry was then performed using biotinylated peanut agglutinin (PNA, Vector Laboratories B1075) or anti-species-specific biotinylated secondary antibodies, followed by avidin-horseradish peroxidase and development with DAB substrate (Vector Laboratories). Sections were counterstained with hematoxylin. The following primary antibodies were used: anti-B220 (BD 550286) and anti-Ki67 (Cell Signaling Technology 12202). Slides were scanned using a Zeiss Mirax Slide Scanner, and photomicrographs were examined using Aperio eSlide Manager (Leica Biosystems). QuPath software (*71*) was used to quantify GC areas.

Flow cytometry analysis and cell sorting

Spleens were mashed and bone marrow cells were flushed from tibias and femurs. Splenocytes and bone marrow cells were filtered through a 40 μm cell strainer and treated with red blood cell lysis solution (QIAGEN 158904). Resulting single-cell suspensions were resuspended in PBE with 0.5 μg/ml rat anti-mouse CD16/CD32 (clone 2.4G2, BD) and incubated for ≥5 min on ice to block Fc receptors. Lymph node cells were directly mashed into the Fc block mix using pellet pestles (Sigma Z359947). Peritoneal B cells were collected by injecting 9 ml of PBS with 0.5% BSA into the peritoneal cavity followed by a subcostal incision and aspiration of the peritoneal cell suspension. After centrifugation, cells were stained for 20 min on ice, in a mix of fluorochrome-conjugated anti-mouse antibodies diluted either

in PBE, or in a 50:50 mixture of Brilliant Stain Buffer (BD 566349) and PBS with 0.5% BSA but without EDTA when the mix contained ≥2 BV- or BUV-conjugated antibodies, to avoid staining artifacts resulting from interaction between these fluorescent dyes. For phosphorylated S6 (pS6) staining, incubation was precisely timed to prevent time-related variations in dephosphorylation by endogenous phosphatases, as previously described (*5*). Cells were first stained for surface markers and then fixed and permeabilized. Fixation and permeabilization using Cytofix/Cytoperm (BD 554722) was timed to occur exactly 30 min after euthanasia for each mouse. We then stained with anti-pS6 diluted in Perm/Wash (BD 554723) for 2 hours. All cells were then washed twice before analysis. DAPI (4′,6-diamidino-2-phenylindole; Thermo Fisher D1306) was added just before acquisition at final 0.5 μg/ml for the exclusion of dead cells. Data were acquired on BD FACS Canto II, BD Fortessa, BD Symphony A5 or BD Symphony A3 and analyzed using FlowJo software package (TreeStar).

For MYC flow cytometry analyses in cell lines, two or three independently generated and validated SU-DHL4 stable lines for each indicated genotype were set at 1 million cells per ml in fresh complete RPMI and incubated at 37°C for 46 hours, for cells to have just doubled. For MYC flow cytometry analyses in human primary GC B cells freshly immortalized, cells were plated at 1 million cells per ml onto fresh feeders and incubated at 37°C for 3 or 4 days. MG132 (Selleck S2619) was then added to cells at 5 μM final MG132. Samples were collected at the indicated times after treatment and immediately fixed for 1 hour at room temperature or overnight at 4°C for all samples, depending on experiments (this did not affect results) using the True-Nuclear Transcription Factor Buffer Set (Biolegend 424401). Cells were then washed twice in 1X perm buffer and kept at 4°C. Fixed and permeabilized cells were blocked with 2% normal goat serum (Cell Signaling Technology 5425) diluted in perm buffer, for 15 min at room temperature. We then stained cells with AF647 anti-MYC (rabbit, clone Y69, Abcam ab190560, diluted to 0.71 μg/ml) for 30 min at room temperature in the dark. Alternatively, we stained cells with unconjugated anti-MYC (rabbit, clone Y69, Abcam ab32072, diluted to 0.45 μg/ml), washed twice with perm buffer, and incubated with AF647 anti-rabbit secondary antibody (Life Technologies A21246, diluted to 2 μg/ml) for another 30 min at room temperature in the dark. We used the following isotype control IgGs at the same concentration as their corresponding anti-MYC antibodies: Abcam ab199093, clone EPR25A, for the AF647 anti-MYC antibody; and Abcam ab172730, clone EPR25A, for the unconjugated anti-MYC antibody. After two washes in perm buffer, we resuspended cells in perm buffer and analyzed them.

For cell sorting from spleen and lymph nodes, cell suspensions were prepared and stained as for flow cytometry analysis, except that splenocytes were pre-enriched either for $B220^+$ B cells by positive selection with anti-B220 magnetic beads (Miltenyi Biotec 130-049-501) or for GC B cells by positive selection with PNA magnetic beads (Miltenyi Biotec 130-110-479) where specified. The indicated populations were then isolated on a BD FACSAria II or a BD Influx sorter at the Flow Cytometry core (WCM) or in the laboratory.

The following antibodies were used: APC-Cy7 anti-B220 (clone RA3-6B2, BioLegend 103224, dilution 1:500), BV786 anti-B220 (clone RA3-6B2, BD 563894, dilution 1:300), FITC anti-human CD2 (clone RPA-2.10, BD 556608, dilution 1:100), BV605 anti-CD5 (clone 53-7.3, BD 563194, dilution 1:500), APC-Cy7 anti-CD19 (clone 6D5, BioLegend 115530, dilution 1:200), FITC anti-CD19 (clone 1D3, BD 553785, dilution 1:500), PE-Cy7 anti-CD21 (clone 7E9, BioLegend 123420, dilution 1:500), FITC anti-CD23 (clone B3B4, BD 553138, dilution 1:500), PE-Cy7 anti-CD25 (clone PC61, BD 561780, dilution 1:500), APC anti-CD38 (clone 90, eBioscience 17-0381-81, dilution 1:500), BUV395 anti-CD38 (clone 90, BD 740245, dilution 1:500), BUV395 anti-CD45.1 (clone A20, BD 565212, dilution 1:200), PerCP-Cy5.5 anti-CD45.1 (clone A20, eBioscience 45-0453-82, dilution 1:200), AF647 anti-CD45.2 (clone 104, BioLegend 109818, dilution 1:500), APC anti-CD45.2 (clone 104, BioLegend 109814, dilution 1:400), PE-Cy7 anti-CD45.2 (clone 104, eBioscience 25-0454-82, dilution 1:300), biotin anti-CD83 (clone Michel-19, BioLegend 121504, dilution 1:200), BV605 anti-CD86 (clone GL-1, BioLegend 105037, dilution 1:200), PE-Cy7 anti-CD86 (clone GL-1, BioLegend 105014 or BD 560582, dilution 1:200), BUV737 anti-CD138 (clone 281-2, BD 564430, dilution 1:500), APC anti-c-Kit/CD117 (clone 2B8, BD 561074, dilution 1:500), biotin anti-CXCR4 (clone 2B11, BD 551968, dilution 1:200), PE anti-CXCR4 (clone 2B11, BD 561734, dilution 1:100), BV421 anti-FAS (clone Jo2, BD 562633, dilution 1:300), PE-Cy7 anti-FAS (clone Jo2, BD 557653, dilution 1:500), AF647 anti-GL7 (clone GL7, BD 561529, dilution 1:500), AF647 anti-MYC (rabbit, clone Y69, Abcam ab190560, diluted to 0.71 μg/ml), unconjugated anti-MYC (rabbit, clone Y69, Abcam ab32072, diluted to 0.45 ug/ml) and AF647 anti-rabbit secondary antibody (Life Technologies A21246, diluted to 2 μg/ml), BV510 anti-IgD (clone 11-26c.2a, BD 563110, dilution 1:500), PE anti-IgD (clone 11-26c.2a, BD 558597, dilution 1:500), PerCP-Cy5.5 anti-IgD (clone 11-26c.2a, BioLegend 405710, dilution 1:500), APC anti-IgG1 (clone A85-1, BD 560089, dilution 1:500), BV421 anti-IgG1 (clone A85-1, BD 562580, dilution 1:500), BV650 anti-IgG1 (clone RMG1-1, BioLegend 406629, dilution 1:500), APC anti-IgM (clone II/41, eBioscience 17-5790-82, dilution 1:500), BV711 anti-IgM (clone II/41, BD 743327, dilution 1:500), BV421 anti-Ig lambda light chain (λ1, λ2, λ3, clone R26-46, BD 744523, dilution 1:400), AF647 anti-pS6 (S6 phosphorylated at S235/236, clone D57.2.2E, Cell Signaling Technology 4851, dilution 1:200), APC streptavidin (BioLegend 405207, dilution 1:300) and PE-Cy7 streptavidin (eBioscience 25-4317-82, dilution 1:400).

To assess progression through S phase of the cell cycle, we performed dual nucleotide pulse and staining as previously described (*46*). Briefly, mice were injected intravenously with 1 mg of EdU (Thermo Fisher E10187) and 1 hour later with 2 mg BrdU (Sigma B5002). Thirty minutes after the second injection, lymph nodes were harvested, and single cell suspensions were prepared. After cell surface staining as described above, cells were fixed and permeabilized using BD Cytofix/Cytoperm and BD Permeabilization Buffer Plus. EdU and BrdU incorporation into DNA was assayed using the Click-iT EdU Pacblue Flow Cytometry Assay Kit (Thermo Fisher C10418) and AF647 mouse anti-BrdU (clone 3D4, BD 560209), respectively.

EdU staining on a slide

Mice were injected intravenously through the retro-orbital venous sinus with 1 mg EdU (200 μl at 5 mg/ml) 1 hour before euthanasia. Spleens were processed and splenocytes stained and sorted as described above. Myc^{GFP+} and Myc^{GFP-} GC B cells were sorted on a BD FACSAria II into PBE and immediately spun onto positively charged glass slides (VWR 48311-703) by cytospin using an Aerospray Pro slide stainer/cytocentrifuge (ELITechGroup) and left to air dry for 2 hours. Cells were fixed with 4% paraformaldehyde for 15 min at room temperature, washed three times with PBS, and left to dry overnight at room temperature. Cells were then washed using PBS plus 3% BSA and permeabilized using 0.5% triton-X for 20 min. EdU was detected using Click-iT Plus EdU Cell Proliferation Kit for Imaging, AF647 dye (Thermo Fisher C10640). This included a staining step with DAPI at 1 μg/ml final. Aqueous mounting media (Fluoromount, Sigma F4680) and coverslips (VWR micro cover glass) were applied to slides right after staining, and imaging was performed immediately.

Images were acquired on a Zeiss LSM 880 laser scanning confocal microscope with a 10× objective lens using 405 nm and 633 nm laser excitation for DAPI and AF647, respectively, using the Zen software (Zeiss). Alternatively, images were acquired on an Evos FL Digital inverted fluorescence microscope (Life Technologies) with a 10× objective lens, a high-sensitivity

monochrome camera (Sony ICX285AL CCD) and DAPI and Cy5 Evos light cubes to detect DAPI and AF647, respectively. Representative images shown were acquired on the Evos instrument. All images were viewed and processed using the Fiji software (*72*). Cyan and magenta lookup tables (LUTs) were applied to DAPI and AF647 channels, respectively, for visualization.

Generation of stable DLBCL cells

All constructs (EF1a-hBTG1, EF1a-hBTG1^{Q36H}, EF1a-EV, EF1a-hBTG1-V5, EF1a-hBTG1^{Q36H}-V5, and EF1a-EGFP-V5) were cloned into the pLEX_307 vector (a gift from D. Root, Addgene 41392), which drives constitutive lentiviral expression of the open reading frames (ORFs) as fusion proteins with a C-terminal V5 epitope tag under the control of the EF1-alpha promoter, as well as puromycin resistance. ORFs with a stop codon were used to express nontagged proteins, and ORFs lacking a stop codon were used to express V5-tagged proteins. Human BTG1 ORF with stop codon was obtained in pENTR221 vector (Transomic) and used to PCR amplify BTG1 ORF without stop codon. Q36H mutation was introduced into BTG1 with or without stop codon by site-directed mutagenesis using Phusion HF polymerase (NEB) followed by DpnI digestion. An empty control vector (EV) was generated by NotI/AscI double digestion of an ORF-containing pENTR/D-TOPO vector, followed by blunting with T4 DNA polymerase and ligation. EGFP ORF without stop codon was obtained in pENTR221 [a gift from D. Root, Addgene 25899 (*73*)]. All of the above entry vectors were recombined into the pLEX_307 destination vector using the Gateway LR Clonase II enzyme mix (Invitrogen 11791-020). T39I, S43N, and A49T mutations were introduced into BTG1-V5 by site-directed mutagenesis using Phusion HF polymerase (NEB) followed by DpnI digestion. Identity of all vectors was verified by Sanger sequencing (Genewiz).

For lentivirus production, 2.5 million HEK293T cells cultured in complete Dulbecco's modified Eagle medium [DMEM (Thermo Fisher 11965118); 10% FBS, 2 mM glutamine, 50 U/ml penicillin, and 50 μg/ml streptomycin] were plated in a 10-cm-diameter plate for 24 hours before cotransfection with the expression construct, pMD2.G, and psPAX2 vectors at a 4:1:3 ratio using Fugene HD (Promega E2311) for a total of 8 μg DNA per plate. After 12 to 16 hours, medium was replaced. Virus-containing medium was collected at 48 and 72 hours after transfection, pooled, filtered, concentrated using PEG-it (System Biosciences LV825A-1), and resuspended at 20X in 1X DPBS. Virus was used right away or stored aliquoted at −80°C.

SU-DHL4 cells cultured in complete RPMI (10% FBS, 2 mM glutamine, 50 U/ml penicillin, 50 μg/ml streptomycin, and 10 mM HEPES) were transduced at a multiplicity of infection of 1 by adding concentrated virus to freshly plated cells. After 72 hours incubation, medium was replaced with fresh medium containing 1 μg/ml puromycin, and cells were selected for 7 days, with fresh addition of puromycin-containing medium every other day. Expression of the constructs was validated by qRT-PCR and Western blot (WB) from an equal number of cells from cell suspension cultures.

Generation of stable human primary GC B cells

Fresh human tonsils were collected from the Department of Pathology at Weill Cornell/New York Presbyterian (IRB protocol 0804009763). To generate human primary GC B cells stably expressing BTG1^{WT} versus BTG1^{Q36H}, we adapted a procedure previously described in detail (*43*). This involved isolating GC B cells from human tonsils, immortalizing them through BCL2/BCL6 retroviral transduction (the CD2$^+$ marker was used to identify immortalized cells), and stably expressing V5-tagged BTG1^{WT}, BTG1^{Q36H}, or EGFP through lentiviral transduction and puromycin selection (fig. S11D).

Tonsillar GC B cells were purified using human B cell negative selection isolation Kit II (Miltenyi Biotec 130-051-151) with addition of IgD-BIOT (SouthernBiotech 9030-08, dilution 1:100) and CD44-BIOT (SouthernBiotech 9400-08, dilution 1:100) antibodies to remove naïve and memory B cells. Purified GC B cells were cultured in RPMI (20%, 2 mM glutamine, 50 U/ml penicillin, 50 μg/ml streptomycin) at a concentration of 1 million cells per ml, on irradiated (30 Gy) YK6-CD40Lg-IL21 feeder cells that carry puromycin resistance (0.2 million cells per ml). GC B cells were replated every 3 to 4 days on freshly irradiated YK6 feeder cells.

For virus production, Lenti-X 293T cells were plated in a 10-cm^2 dish 16 hours before transfection at 0.2 million cells per ml in DMEM (10%, 2 mM glutamine, 50 U/ml penicillin, 50 μg/ml streptomycin). For retrovirus production (BCL2-BCL6), vectors were mixed as followed: 1 μg of pHIT60 (gag-pol), 1 μg of phCMV-GaLV MTR [a gift from D. Hodson, Addgene 163612 (*74*)] and 4 μg of MSCV-BCL6-t2A-Bcl2 [a gift from D. Hodson, Addgene 135305 (*74*)] and incubated for 45 min in 1 ml Opti-MEM with 18 μl of Trans-IT (Mirus Bio MIR 2300) before addition to the Lenti-X 293T dish. For lentivirus production (V5-tagged BTG1^{WT}, BTG1^{Q36H}, EGFP), 8.3 μg of gag-pol vector pCMVDeltaR8.91, 2.8 μg of envelop vector phCMV-GaLV MTR [a gift from D. Hodson, Addgene 163612 (*74*)], and 11 μg of BTG1 constructs in pLEX_307 vector (described above for cell lines) were incubated for 45 min in 1 ml Opti-MEM with 33 μl of Trans-IT (Mirus Bio, MIR 2300) and added to the Lenti-X 293T dish. Viruses were collected and filtered (0.45 μm) 72 hours later and frozen or directly used for infection of GC primary cells.

For infection, 2 million GC B target cells with feeder cells were directly resuspended in virus suspension at a virus:media ratio of 1:4 for retrovirus or 3:4 for lentivirus, with addition of 8 μg/ml of polybrene (Santa Cruz sc-134220) and 25 μM HEPES. Cells and virus mixtures were plated in 12-well plates and centrifuged at 1500*g* for 90 min at 32°C. Virus-containing medium was replaced with fresh medium (RPMI 20%, 2 mM glutamine, 50 U/ml penicillin, 50 μg/ml streptomycin) immediately after centrifugation for retroviral infection and at least 4 hours after centrifugation for lentiviral infection (*43*). Seventy-two hours after lentiviral infections, cells were treated with 1 μg/ml puromycin for 10 days to select for transduced cells. BCL2-BCL6 positivity was continuously monitored by anti-CD2 flow cytometry. Experiments were performed when >95% puromycin-selected cells were obtained within CD2$^+$ GC B cells.

mRNA stability assay

Three independently generated and validated SU-DHL4 stable lines for each indicated genotype were set at 1 million cells per ml in fresh complete RPMI and incubated at 37°C for 24 hours, for cells to be growing exponentially and at similar medium composition at the time of treatment. Cells were then collected (0 hours) or actinomycin D (Cell Signaling Technology 15021) was added at 10 μg/ml final and cells were collected at the indicated time points after treatment. Upon collection, cells were spun and directly resuspended in TRIzol for RNA extraction, followed by cDNA synthesis and qRT-PCR analysis as described above.

RNA immunoprecipitation (RIP)

Two or three independently generated and validated SU-DHL4 stable lines were used for each indicated genotype per experiment. Thirty-five million cells per sample were set at 1 million cells per ml in fresh complete RPMI and incubated at 37°C for 24 hours, for cells to be growing exponentially and at similar medium composition at the time of treatment. Cells were collected, washed twice with cold 1X DPBS, and lysed in 175 μl Magna RIP lysis buffer (from Millipore 17-700) with frequent vortexing for 10 to 30 min on ice. Lysates were directly processed for RIP using magnetic beads, following the Magna RIP kit instructions (Millipore 17-700) or kept at −80°C until processing. Anti-V5 mouse antibody (7.5 μg; Abcam 27671) was used to pull down BTG1^{WT}-V5, BTG1^{Q36H}-V5, or EGFP-V5 (negative control). Ten percent of lysate was saved pre-RIP as input RNA, 175 μl lysate was RIPed, and the remaining lysate volume was saved for protein quantification by Bradford to analyze by Western blot as input protein. Post-RIP, 6%

of the beads suspension was saved during the last wash to analyze by Western blot as RIP protein. RNAs were purified from the beads and input RNA samples according to the Magna RIP kit instructions and resuspended in 25 μl molecular grade RNase-free water. Equal RNA amounts were then used for cDNA synthesis and qRT-PCR as described above. In the first experiment, RIP was done in technical duplicates for each independent line (35 million cells each, two independent lines per genotype). We observed highly similar pull-down efficiency between technical duplicates (compare a and b replicates, fig. S9A), and thus subsequent experiments were done with three independent lines and one RIP per genotype.

Western blot

Three independently generated and validated SU-DHL4 stable lines for each indicated genotype were set at 1 million cells per ml in fresh complete RPMI and incubated at 37°C for 24 hours, for cells to be growing exponentially and at similar medium composition at the time of treatment. For the starvation analyses, after the initial 24 hours incubation, cells were resuspended in fresh complete RPMI without FBS, incubated for 4 hours at 37°C, followed by resuspension in fresh complete RPMI (with 10% FBS) for 15 min, 30 min, 1 hour, and 2 hours, at which times cells were collected, spun, and directly lysed in lysis buffer (50 μl per million cells set up). For the newly protein synthesis analyses, after the initial 24 hours incubation, DMSO or MG132 (Selleck S2619) was added to cells for 5 μM final MG132 (and 0.01% DMSO final for all samples), and samples were collected at the indicated times after treatment, spun, and directly lysed in lysis buffer (50 μl per million cells plated). For the protein stability analyses, after the initial 24 hours incubation, CHX (stock kept in −80°C at 20 mg/ml in H_2O, Santa Cruz Biotechnology 66-81-9) was added to cells at 40 μg/ml final, and samples were collected at the indicated times after treatment, spun, and directly lysed in lysis buffer (50 μl per million cells plated). After development and signal quantification as described below, nonlinear regression curves were generated using exponential one phase decay equation with plateau constraint to zero.

Lysis buffer composition was 20 mM HEPES KOH pH 7.5, 50 mM b-glycerophosphate, 1 mM EDTA pH 8.0, 1 mM EGTA pH 8.0, 0.5 mM Na3VO4, 100 mM KCl, 10% glycerol, and 1% Triton X-100, supplemented with complete protease inhibitor cocktail (Roche). Protein concentration was determined by Bradford (Thermo Fisher 23236) and equal protein amounts were separated by 10 or 12% SDS–polyacrylamide gel electrophoresis. After electrophoretic transfer to BioTrace NT pure nitrocellulose blotting membrane (PALL 66485), membranes were stained with Ponceau, rinsed and blocked with 5% nonfat dry milk (Biorad 1706404) unless indicated otherwise, in PBS containing 0.1% Tween-20. Proteins were probed by overnight incubation at 4°C with the following antibodies: anti-V5 for Western blot after RIP (rabbit, Sigma V8137, dilution 1:2000), anti-V5 for all other Western blots (mouse, abcam 27671, dilution 1:2500), anti-alpha-tubulin (mouse, Sigma T6199, dilution 1:5000), anti-beta-actin (mouse, Sigma 5441, dilution 1:15,000), anti-MYC (rabbit, Cell Signaling Technology 5605, 1:1000), or anti-RPS3 (rabbit, Cell Signaling Technology 9538, 1:3000 in 5% BSA). Membranes were then incubated with HRP-conjugated anti-mouse or anti-rabbit secondary antibody (Cell Signaling Technology 7076 or 7074) or with IRDye 800CW anti-rabbit antibody (LI-COR 926-32211) for RPS3 detection. Development was done with ECL (Thermo Fisher 32106) for HRP-conjugated antibodies or directly for IRDye 800CW, either on a ChemiDoc Touch imaging system, ChemiDoc MP Imaging System (BioRad) or using Amersham Hyperfilm (Thermo Fisher 45001508) and a developer. Protein band intensities were quantified using the chemiluminescence images acquired on the ChemiDoc and the Image Lab software v6.1.0 (Bio-Rad).

Cell volume measurement on Coulter counter

Three independently generated and validated SU-DHL4 stable lines for each indicated genotype were set at 1 million cells per ml (in triplicates or quadruplicates) in fresh complete RPMI and incubated at 37°C for 24 hours. Cells were then resuspended, and an aliquot of cell suspension was used immediately for cell volume (mean fL) analysis on a Multisizer 4e Coulter particle analyzer (Beckman Coulter B43905).

Polysome profiling

Three independently generated and validated SU-DHL4 stable lines for each indicated genotype were set at 1 million cells per ml in fresh complete RPMI and incubated at 37°C for 48 hours. Equal numbers of cells for each line (30 million or 100 million cells, depending on experiment replicate) were treated with CHX at 100 μg/ml for 10 min at 37°C. Cells were then lysed on ice for 10 min in 400 μl lysis buffer containing 20 mM Tris HCl pH 7.4, 100 mM KCl, 5 mM $MgCl_2$, 1% TritonX-100, 100 μg/ml CHX, 10 mM DTT, EDTA-free protease inhibitors (Roche 11873580001,) and 500 U/ml RNasin RNase inhibitor (Promega N2611). Lysates were cleared by centrifugation at 15,000*g*, 4°C for 15 min. Cleared lysate aliquots were saved as inputs, and remaining cleared lysates were loaded onto a linear 20 to 50% sucrose gradient in 40 mM Tris HCl pH 7.5, 300 mM KCl, 20 mM $MgCl_2$, and 100 μg/ml CHX. Samples were then centrifuged at 42,000 rpm for 3 hours at 4°C using a TH641 rotor (Sorvall) in a Sorvall wX+ Ultra Series Centrifuge (Thermo Scientific). After centrifugation, fractions were collected using a BioComp Gradient Station ip. RNA content was measured at A260 concurrent with fraction collection using Triax Flow Cell v1.50A software. AUCs and analyses thereof were calculated using Prism 8 software.

After collection, fractions were prepared for RNA precipitation by adding one-tenth of fraction volume of 3 M NaOAc and 1 volume of fraction volume of isopropanol and incubated at −80°C overnight. Precipitated RNAs were extracted and purified using the Direct-zol RNA Miniprep kit and protocol (Zymo Research, R2050). After adding ERCC spike-ins (Thermo Fisher 4456740) per same volume of RNA for all samples (input and fractions) to be able to normalize for potential differences in efficiency of RT or qPCR amplification, we used equal volumes of samples for RT and qPCR. For MYC, we generated a standard curve using *MYC* dsDNA amplicons at 10^3 to 10^8 absolute copy number per well. Details about primers used are included in table S8.

For WB analysis, fractions were prepared for protein precipitation by adding cold trichloroacetic acid (TCA) at 15% final and incubated at −20°C overnight. Samples were thawed on ice and centrifuged for 30 min at 15,000*g*, 4°C. Pellet was resuspended in 1 ml of cold 15% TCA to fully remove sucrose and salt and centrifuged again for 15 min at 15,000*g*, 4°C. The pellet was washed twice with 500 μl acetone and centrifuged for 10 min at 15,000*g*, 4°C. Samples were then dried using speedvac and resuspended in 1X LDS (Thermo Fisher B0007) diluted in RIPA buffer [50mM HEPES pH 7.4, 150 mM NaCl, 1% sodium deoxycholate, 0.1% SDS, 10 mM sodium pyrophosphate, 10 mM b-glycerophosphate, 1% NP-40, supplemented with 1 mM $MgCl_2$, phosphatase inhibitor cantharidin (Gold Biotechnology C-975-500), benzonase (EMD Millipore 71205-3)] and protease inhibitor cocktail as follows: 115 mM sodium molybdate dihydrate (Gold Biotechnology M-781-100), 100 mM sodium orthovanadate (Gold Biotechnology S-930-25), 100 mM sodium fluoride (Alfa Aesar AAA1301930), 200 mM imidazole (Gold Biotechnology I-902-25), and 400 mM sodium tartrate dihydrate (Alpha Aesar AAA1618730). Samples were then boiled at 75°C for 15 min and analyzed by Western blot.

UV cross-linking and immunoprecipitation (CLIP) and RNA radiolabeling

Immunoprecipitation and radiolabeling of RNAs bound to V5-tagged proteins were performed as previously described (*75*) with the following modifications: 20 million cells were seeded 24 hours prior to cross-linking. Cells were centrifuged at 300*g* for 5 min, washed in 10 ml of 0.22 μm filtered media, then centrifuged

again and resuspended at a final concentration of 4 million cells per ml in 0.22 μm filtered media. Cells were then cross-linked with 150 mJ/cm^2 at 254 nm using a 254 nm UV cross-linker instrument (HL-2000 HybriLinker, UVP) and centrifuged, and pellets snap frozen on dry ice.

Cross-linked cell pellets were resuspended in cold lysis buffer: 50 mM Tris-HCl pH 7.5, 300 mM NaCl, 0.5% Nonidet-P40 (IGEPAL CA-630), supplemented with complete EDTA-free protease inhibitor (1 tablet per 25 ml, Roche 11873580001) and 1X phosphatase inhibitor (100X Halt cocktail, Thermo 8440) prior to use. Cells in lysis buffer were sonicated using a Branson Digital Sonifier Model 450 fitted with 3.125-mm tapered micro tip probe on ice at 10% amplitude (2-s ON, 10-s OFF cycle, total 30 s). A Western blot was performed on lysate aliquots to quantify the levels of V5-tagged proteins. Western blot lysates were first treated with 5 U/μl RNaseI for 10 min at 37°C to completely digest RNA. Cell lysates were normalized according to the levels of V5-tagged proteins. DNase I (0.002 U/μl; Thermo Scientific EN0521) and RNase I (0.02 U/μl; Thermo Scientific AM2294) were used to digest DNA and RNA in the cell lysate for 3 min at 37°C .

Anti-V5 antibody (2.5 μg; mouse, Abcam 27671, dilution 1:2500) was used for immunoprecipitation (IP) of cross-linked RNA-protein complexes per sample. V5 antibodies were first bound to Pierce Protein A/G Magnetic Beads (Thermo Scientific 88802) in lysis buffer for 1 hour at 4°C. IP was performed in lysis buffer at 4°C overnight. As previously described (*75*), bound RNA-protein complexes were washed twice with stringent high-salt buffer (50 mM Tris-HCl pH 7.4, 1 M NaCl, 1 mM EDTA, 1% NP-40, 0.1% SDS, 0.5% sodium deoxycholate) and twice with low-salt buffer (20 mM Tris-HCl pH 7.4, 10 mM $MgCl_2$, 0.2% Tween-20). Half the IP lysate was used for Western blot to verify levels of V5-tagged proteins (anti-V5, Sigma V8137, dilution 1:2000). Western blot lysates were treated with 0.2 U/μl RNaseI for 10 min at 37°C to completely digest RNA. In the remaining half of the IP lysate, RNA bound to pulled down proteins was labeled with 0.25 U/μl T4 PNK (New England Biolabs M0201) and 0.25 μCi/μl [γ-32P]- ATP (Perkin Elmer BLU002Z500UC). RNA-protein complexes were eluted from the beads in 1X NuPage LDS Sample Buffer (Invitrogen NP0008) plus 10 mM DTT and run on an SDS-PAGE gel, before transfer onto a nitrocellulose membrane. The membrane was then exposed using autoradiograph film.

RNA-, RIP-, and BCR-sequencing sample preparation

RNA was extracted using TRIzol as described above, quantified with Qubit, and analyzed for integrity using RNA Pico Chips (Agilent 5067-1513) on an Agilent 2100 Bioanalyzer (Agilent Technologies). Samples with RNA integrity number (RIN) ≥ 8 were further processed. On the basis of the available amount, 2 to 100 ng RNA was used for cDNA synthesis and qRT-PCR to validate samples identity pre-library preparation. Equal amounts of RNA were used for all samples for each experiment for library preparation. Libraries were validated on the Agilent 2100 Bioanalyzer before sequencing and quantified using Quant-iT dsDNA HS Assay (Thermo Fisher Q33120). Each library was generated with a unique index sequence, and libraries were pooled with the same ratio for all samples for sequencing, except for the RIP as described below. Library preparation, sequencing and postprocessing of raw data was performed at the Epigenomics Core and Genomics Resources Core facilities (WCM).

For SU-DHL4 RNA-seq, three independent stable lines were freshly generated as described above, selected with 1 μg/ml puromycin for 4 days, enriched for viable cells by gradient separation (Fico/Lite-LymphoH, R&D Systems I40150), and 200,000 viable cells were used for RNA extraction. For library preparation, 200 ng RNA per sample was used.

For GC B cell bulk RNA-seq, live B220$^+$FAS$^+$ GL7$^+$ cells were sorted from B220$^+$ enriched splenocytes on BD Influx as described above, from four CRE and five Q36H mice. Per sample, 123,000 to 344,000 GC B cells were sorted, of which 120 to 236 ng RNA was extracted.

For both SU-DHL4 and GC B cell bulk RNA-seq, libraries were prepared using Illumina TruSeq stranded-mRNA sample kit and subjected to single read 50 bp sequencing on an Illumina HiSeq 2500.

For Myc^{GFP+} and Myc^{GFP-} GC B cells RNA-seq, live B220$^+$FAS$^+$CD38$^-$ Myc^{GFP+} or Myc^{GFP-} cells were sorted from B220$^+$ enriched splenocytes on BD Influx as described above, from two CRE and three Q36H mice. Per sample, 4000 to 8850 Myc^{GFP+} GC B cells and 155,000 to 250,000 Myc^{GFP-} GC B cells were sorted, of which 1.9 to 3.8 ng (Myc^{GFP+}) and 60 to 300 ng (Myc^{GFP-}) RNA was extracted. Eleven PCR cycles and 0.95 ng RNA were used for all samples for library preparation using the Ultra-low input RNA kit (Takara SMART-Seq v4), and samples were subjected to single read 50 bp sequencing on an Illumina HiSeq.

For Bcl2+Q36H and Bcl2 lymphoma B cells RNA-seq, 300,000 GC B cells from Bcl2 and Bcl2+Q36H mice were sorted as live B220$^+$ FAS$^+$GL7$^+$ cells from PNA$^+$ enriched splenocytes on BD Influx or BD FACSAria II at 8 months after transplantation. Per sample, 150 ng RNA was used for library preparation using Illumina TruSeq stranded-mRNA sample kit followed by single read 50 bp sequencing on an Illumina HiSeq.

For SU-DHL4 RIP-sequencing, input and RIP RNAs were obtained from two independently generated stable lines, each RIPed in duplicates from 35 million cells each as described above. We recovered ~7 to 14 μg RNA from input samples and 500 to 1700 ng RNA from RIP samples. We added ERCC RNA spike-ins (Thermo Fisher 4456740) per same volume of RNA for all input or for all RIP samples to be able to normalize for possible global changes in starting RNA amount and pull down RNA amount between samples. Of these, we used 250 ng RNA in 50 μl for each of the input and RIP samples for library preparation using TruSeq-stranded-mRNA sample kit. We used 22 nmol library for all input and RIP samples, except for the EGFP-V5 RIP samples, for which we added a maximum of 11 nmol each (2:1 ratio, normalized during analysis). Samples were subjected to paired-end 50 bp sequencing on a NovaSeq SP at 700 million reads to achieve 29 million reads per sample.

For BCR sequencing analysis, we used remaining splenocytes from Bcl2 and Bcl2+Q36H mice at 8 months after transplantation that were not used for PNA enrichment and GC B cells sorting, had been treated with red blood cell lysis solution (QIAGEN 158904), and viably frozen in FBS with 10% DMSO. We thawed these splenocytes and extracted gDNA using Puregene Gentra cell kit (QIAGEN 158388). gDNA (1.5 μg) was sent to Adaptive Biotechnologies for ImmunoSeq assay to sequence a 125-bp-long VDJ region around the IgH complementarity-determining region 3 (CDR3; see fig. S16B). Briefly, the somatically rearranged IgH CDR3 was amplified using a two-step amplification, bias-controlled multiplex PCR approach (*76*). The first PCR consisted of forward and reverse amplification primers specific for every V and J gene segment and amplified the hypervariable CDR3 of the immune receptor locus. The second PCR added a proprietary barcode sequence and Illumina adapter sequences.

RNA-sequencing analysis

For all RNA-seq analyses, results were aligned to mm10 or hg19 using STAR (*77*) and annotated to RefSeq using the Rsubread package (*78*). Differential expression was called using edgeR generalized linear model with TMM normalization (*79*, *80*) with thresholds of FC >1.5 and Benjamini-Hochberg corrected $P < 0.01$. Murine Bcl2 and Bcl2+Q236H tumors were additionally modeled using a batch correction for the respective cell sorter and, because of sex imbalance between conditions, only autosomal genes were used for analysis.

Gene set enrichment was assessed using the GSEA algorithm, a computational method based on the Kolmogorov-Smirnov test (*34*). GSEA network analysis was performed by running GSEA using the Hallmark and Canonical Pathways gene sets from MSigDB. Jaccard

similarity index was calculated for all pairs of leading-edge genes from significantly enriched signatures (FDR < 0.05; table S3). Network was then generated using the R igraph package (*81*), adjusting size of node for degree of connectivity.

Unsupervised pathway analysis was performed using information-theoretic pathway analysis approach as described previously (*82*). Briefly, pathways that are informative about nonoverlapping gene groups were identified. Pathways annotations were used from the Biological Process annotations of the Gene Ontology database (http://www.geneontology.org), Human Protein Reference Database (http://www.hprd.org), MSigDB C2 (http://www.gsea-msigdb.org/gsea/msigdb/human/genesets.jsp?collection=C2), and signature categories from the Staudt Lab Signature database (*83*). Only human-curated annotations were used from the Gene Ontology database, and only pathways with five genes or more, and with 300 genes or less were evaluated. This pathway analysis estimates how informative each pathway is about the target gene groups and applies a randomization-based statistical test to assess the significance of the highest information values. We used the default significance threshold of $P < 0.005$. We estimated the FDR by randomizing the input profiles iteratively on shuffled profiles with identical parameters and thresholds, finding that the FDR was always lower than 5%. For each informative pathway, we determined the extent to which the pathway was overrepresented in the target gene group, using the hypergeometric distribution, as described previously (*82*).

Unsupervised hierarchical clustering was performed on RNA-seq using most variable genes (95th percentile for standard deviation across all samples). Euclidean distance was calculated using log-transformed expression values for GC B cell RNA-seq from CRE and Q36H samples [log(TPM+0.1)] and log-transformed FPKM expression values adjusted for sorter batch effect using the ComBat function of the R package sva (*84*, *85*) with parametric empirical Bayes adjustment for the Bcl2 and Bcl2 +Q36H lymphoma RNA-seq. Hierarchical clustering was then performed using Ward's minimum variance method.

Motif analysis

HOCOMOCO (v11) motifs (*86*) were assessed for enrichment in gene promoter (TSS ± 2kb) using CentriMO tool from the MEME framework (*87*, *88*). Comparative enrichment was evaluated within genes found unregulated in $BTG1^{Q36H}$ versus genes that were expressed in SU-DHL4 cells (FPKM > 0.1) and were non-differentially expressed. We performed exact test to evaluate the significance of the number of best matches in the region in the primary set compared with the number of best matches in the same region in the control set of sequences.

RIP sequencing analysis

RIP sequencing results were aligned to a merged hg19-ERCC genome using STAR and annotated to RefSeq using the Rsubread package. Read counts between samples were normalized to ERCC spike-ins using the RUVg method (*89*). Differential RIP abundance was called by a linear additive model using edgeR, adjusting for ERCC spike-ins and subtracting input transcript level. Transcripts found to be lost in $BTG1^{Q36H}$ versus $BTG1^{WT}$ pull-downs were filtered to remove any transcripts that were present less than twofold in $BTG1^{WT}$ RIP versus input.

Targeted single-cell RNA sequencing

We sorted competing Q36H (CD45.1/2) and CREneg control (CD45.1) GC B cells ($DAPI^-B220^+CD4^-FAS^+CD38^-$) as well as T_{FH} cells ($DAPI^-B220^-CD4^+CXCR5^{hi}$ PD-1^{hi}) from n = 4 mice on a BD Influx sorter as described above. Samples were tagged, counted, and multiplexed into three replicates at a Q36H: CREneg:TFH ratio of 60:28:12. BD Rhapsody data were generated according to the manufacturer's protocol. Briefly, a total of 25,000 to 30,000 cells was loaded per BD Rhapsody Cartridge, into two cartridges for single-cell capture, followed by reverse transcription and exonuclease I treatment, library preparation, and sequencing on a NextSeq500, using 75 bp paired-end sequencing. The targeted panel of n = 496 genes (table S7) comprised BD Rhapsody mouse immune response panel (BD 633753, n = 397) plus a custom set of genes implicated in GC or lymphoma biology (n = 99) designed with help from BD. Fastq files were uploaded to Seven Bridges Genomics. Data were demultiplexed and sequences were analyzed with BD's Rhapsody pipeline (BD Rhapsody Analysis Pipeline 1.4 Beta) on Seven Bridges (www.sevenbridges.com). This pipeline included RSEC and DBEC filtration of UMI counts and generated a sparse matrix file of features by barcodes. This sparse matrix data were read into R using the R package Seurat v3.0.2 (*90*). Data were then scaled and log normalized.

We focused on GC B cell clusters for downstream analysis, on the basis of *Cd19* expression and absence of *Cd4* or *Cd8* expression. Linear dimensional reduction was performed by calculation of PCA from the most variable genes. Cells were then clustered using a resolution value of 0.5 and visualized by UMAP determined from dims 1:10. Module scores were calculated using the AddModuleScore function with a control value of 5. Individual genes and module scores were projected and used for appropriate classification of clusters. Single cells were assigned to populations (DZ, LZ, or LZ-to-DZ recycling) on the basis of high score of gene markers and module scores per cluster.

Single-cell trajectory analysis

Cell cycle phase identities were assigned by the Seurat CellCycleScoring function. RNA trajectory analysis was performed using the R package Slingshot v1.2.0 (*91*). This package was used to create a "pseudotime" based on a combination of PCAs 1 and 2 calculated by Seurat, using cells identified as S phase cells as the anchor point. Cells that belonged to a pre-memory B cell cluster on the basis of *Cd38* expression were excluded, and lineage plots were generated by projecting pseudotime onto cells mapped by PCAs 1 and 2. Pseudotime density plots were generated using cell cycle phase identity and the ggplot2 v3.2.1 geom_density function. The difference in density between genotypes was determined by using the density function in R on the pseudotime density plot for each genotype and subtracting the CREneg density Y values from the Q36H density Y values. These differences were then plotted by ggplot2 geom_col function. Difference in density distribution was tested by Wilcoxon rank sum test on the pseudotime values for each genotype. Scatterplots were created by plotting pseudotime on the x axis and module scores on the y axis. Average score and confidence interval per genotype were plotted with the geom_smooth function.

WGS driver analysis

Whole-genome sequencing (WGS) alignments for 101 DLBCL and matched normal sample pairs were retrieved as hg19-aligned BAM files from the European Genome Archive (https://ega-archive.org/, EGAD00001002123) and stripped of alignment information via Picard tools RevertSam, according to GATK best practices (https://software.broadinstitute.org/gatk/best-practices) to generate unmapped BAM files. The unmapped BAM files subsequently underwent alignment to hg38 and somatic variant calling for SNVs, indels, and structural variants (SVs) via the Sarek pipeline v2.5.1 (https://github.com/nf-core/sarek). Briefly, the Sarek pipeline (*92*) applies BWA mem alignment, duplicate removal, base recalibration, somatic SNV/indel calling via Strelka2 (*93*), somatic SV calling via Manta (*94*), and variant annotation with VEP (*95*). In summary, this pipeline yielded VCF files containing 1.36 million SNVs across 101 DLBCL cases, which were used for downstream analyses.

FishHook (https://github.com/mskilab/fishHook) was used to model background mutational processes and nominate DLBCL mutational hotspots across the cohort of 101 DLBCL cases (see above). FishHook takes as input mutation calls (e.g., VCF), a set of hypothesis intervals (.bed files), an eligible territory

(.bed file), and a set of genomic covariates (.bed, .bw files) and outputs a model of mutation density and hypothesis intervals associated with a *P* value and enrichment score. Briefly, fishHook (*96*) models genome-wide somatic mutation density in tumor genomes across the intersection of an eligible territory (e.g., coding sequences, mappable genomic regions) and set of hypothesis intervals (e.g., genes) as a function of genomic covariates, which can represent sequence context composition, chromatin features, or the fractional overlap with reference genomic annotations. The maximum likelihood fit of a fishHook model, implemented as a gamma-Poisson regression, assigns weights to covariates and an expected mutation density to each hypothesis interval. An enrichment value is computed at each hypothesis interval as the ratio of observed to expected mutation density. The model is further used to define a cumulative distribution function (CDF) for mutation density at each interval. Each interval is then assigned a one-sided *P* value as the probability that the mutation density is greater or equal than the observed density.

FishHook was applied to analyze the distribution of 1.36M DLBCL SNVs across 24,498 protein coding genes (GRCh38) using coding sequences as the eligible territory a set of genomic covariates representing B cell–specific transcriptional and chromatin state. To generate the set of covariates, we annotated .bed files of protein coding genes with the number of ATAC-seq peaks from human purified GC B cells (unpublished data) within 100 kb, the number of GC B cell H3K27ac ChIP-seq defined super-enhancers within 100 kb, and the number of TSS within 100 kb for genes expressed >1 TPM in human GC B cells (RNA-seq). Super-enhancers were called from previously reported H3K27ac ChIP-seq data from human purified GC B cells (*97*) using the ROSE method (*98*).The resulting fishHook model with three covariates was fit to annotate genes with *P* values and enrichment scores (table S1). We labeled genes with (Benjamini-Hochberg) FDR < 0.01 as significant. We generated Q-Q plots by pairing observed −log10 transformed quantiles of observed *P* values (*y* axis) with their corresponding −log10 transformed quantiles from the uniform distribution (*x* axis). A genomic inflation factor λ was computed as the slope of a least-squares regression line fitting these data while intercepting the origin. The value of λ (<1.05) was consistent with minimal statistical inflation (*99*).

BCR sequencing analysis

Initial immunoSeq BCR sequencing data processing was performed using the immunoSeq Analyzer 3.0 (Adaptive Biotechnologies). Only sequences associated with at least five templates were retained for further analysis. BCR repertoire sequencing data was then analyzed using the Immcantation (www.immcantation.org) framework. VDJ genes for each sequence were aligned to the IMGT GENE-DB database accessed 22 June 2021 (*100*) using IgBlast v1.17.1 (*101*) and Change-O v1.1.0 (*102*). Only IMGT V-gene sequences obtained from C57BL/6 mice were included in the reference database. Nonproductive BCRs were removed. Sequences were grouped into clonal clusters, which contain B cells that relate to each other by somatic hypermutations from a common VDJ ancestor. Sequences were first grouped by common IGHV gene annotations, IGHJ gene annotations, and junction lengths. Using the hierarchicalClones function of *scoper* v1.2.0 (*103*), sequences within these groups differing by a length normalized Hamming distance of 0.1 within the CDR3 region were defined as clones using single-linkage hierarchical clustering (*104*). This threshold was determined through manual inspection of distance to nearest neighbor plots using shazam v1.1.0 (*105*). Within each clone, germline sequences were reconstructed with D segment and N/P regions masked (replaced with "N" nucleotides) using the createGermlines function within *dowser* (*106*).

Somatic hypermutation was calculated as the Hamming distance between each sequence and its predicted unmutated germline ancestor. Because immunoSEQ data contains only the partial V-gene sequence, only between 69 and 108 nucleotides of region FWR3 were used in this calculation. To quantify B cell clonal diversity, we calculated Simpson's diversity for each sample using the alphaDiversity function of *alakazam* v1.2.0 (*102*). Lower values of Simpson's diversity indicate a greater probability of two random sequences belonging to the same clone, consistent with more large clones. To account for differences in sequence depth, samples within each comparison were down sampled to the same number of sequences, and the mean of 1000 such resampling repetitions was reported. Samples with fewer than 30 sequences were excluded. Clonal lineage trees were built using *dowser* v1.0.0 (*106*) and IgPhyML v1.1.4 (*50*). Because predicted germline junction sequences are masked, only V-gene sites were used to build lineage trees and estimate substitution model parameters. Codons split by insertions relative to the IMGT germline reference sequences were masked using the maskSequences function in *dowser*. Lineage tree topologies, branch lengths, and subject-wide substitution model parameters using maximum likelihood under the GY94 model (*107*). Using fixed tree topologies estimated from the GY94 model, we then estimated branch lengths and subject-wide parameter values under the HLP19 model (*50*). Only clones containing at least two unique sequences and a total >100 sequence reads were included. Trees were visualized using *dowser* and *ggtree* v3.3.1.901 (*108*).

In silico GC modeling

An agent-based GC model with molecular events of GC B cell selection based on DisseD theory (*47*) was used to perform in silico simulations. The model uses elements described before (*109*–*111*) and differential equations describing evolution of the signaling molecules mTOR, Myc, and FoxO. A three-dimensional lattice with chemokine gradients is used to represent the GC area and incorporates randomly distributed T_{FH} cells and follicular dendritic cells. Founder GC B cells enter the GC area with a fixed rate for an initial time period, undergo six divisions, and eventually acquire a LZ phenotype. LZ B cells interact with follicular dendritic cells, resulting in the acquisition of antigen with an affinity-based probability. A four-dimensional shape space (*112*) is used to represent the B cell affinity, where the position of a B cell with respect to a predefined optimal position denoting the antigen corresponds to the affinity of the cell. Successful antigen collection is followed by interactions with T_{FH} cells. When multiple B cells are bound to the same T_{FH} cell, the T_{FH} cell polarizes toward and signals the B cell with the highest amount of antigen collected. Acquisition of antigen and interaction with T_{FH} cell induces changes in the levels of mTOR, Myc, and FoxO molecules. Time period of interaction with follicular dendritic cells and T_{FH} cells ends, and a fate decision is initiated, either at the time when mTOR reaches a threshold value or when the time exceeds a maximum time period of 18 hours. GC B cells with Myc and mTOR levels reaching predefined threshold values are selected. Selected cells return to the DZ phenotype and divide a certain number of times, as determined by Myc levels at the time of selection by T_{FH} cells. Detailed description of the GC model and parameter values are provided in the supplementary text of (*47*).

Two models were considered to recapture the behavior of BTG1 mutant GC B cells. In the first model, we considered a faster up-regulation of mTOR and Myc: The rate of mTOR and Myc up-regulation in Btg1 mutant GC B cells upon contact with TFH cells was increased 1.12 times as compared with WT cells. In the second model, we considered a faster up-regulation of mTOR and Myc, coupled with a shorter S phase duration: In addition to the 1.12 times faster up-regulation of mTOR and Myc, Btg1 mutant GC B cells were assumed to undergo accelerated S phase progression. Average S phase duration of mutant GC B cells was shortened by 21% as compared with WT cells.

Each of the two models were tested with different simulation setups, corresponding to the competitive and noncompetitive in vivo

experiments. In the competitive setup, founder GC B cells at the start of the simulation were assigned as either Btg1 mutant or WT with a probability of 0.5, for equal starting proportions of mutant and WT GC B cells. Positions of GC B cells in the shape space were chosen at a distance of 2 from the optimal position, to mimic the high affinity of B1-8^{hi} cells. For the noncompetitive setup, separate simulations were performed with 100% cells being WT or Btg1 mutant GC B cells, and affinities were randomly chosen. Each simulation was repeated 300 times, and the mean and SD values were calculated. To quantify the relative difference between mutant and WT, mutant readouts were normalized with corresponding WT readouts. In vivo experimental data were normalized in a similar way for comparison to in silico simulation results. Simulations were performed with C++, and statistical tests were performed using R.

Survival analyses on human ABC-DLBCL cases

Kaplan-Meier estimator was used to estimate the overall survival probability followed by log-rank test to compare the difference between *BTG1* mutational status. Univariable and multivariable Cox regression were used to assess the overall survival difference between *BTG1* mutational status, adjusting for IPI, LymphGen call, and sex. All analyses were performed in statistical software SAS v9.4 (SAS Institute).

Statistics

For all statistical analyses, sample size (n) for each experimental group, the nature of replicates, and details for statistical tests used were specified either in the figure legend, main text, or the methods section.

REFERENCES AND NOTES

1. L. Mesin, J. Ersching, G. D. Victora, Germinal center B cell dynamics. *Immunity* **45**, 471–482 (2016). doi: 10.1016/j.immuni.2016.09.001; pmid: 27653600
2. R. Nakagawa *et al*., Permissive selection followed by affinity-based proliferation of GC light zone B cells dictates cell fate and ensures clonal breadth. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2016425118 (2021). doi: 10.1073/pnas.2016425118; pmid: 33419925
3. D. Dominguez-Sola *et al*., The proto-oncogene *MYC* is required for selection in the germinal center and cyclic reentry. *Nat. Immunol.* **13**, 1083–1091 (2012). doi: 10.1038/ni.2428; pmid: 23001145
4. D. P. Calado *et al*., The cell-cycle regulator c-Myc is essential for the formation and maintenance of germinal centers. *Nat. Immunol.* **13**, 1092–1100 (2012). doi: 10.1038/ni.2418; pmid: 23001146
5. J. Ersching *et al*., Germinal center selection and affinity maturation require dynamic regulation of mTORC1 kinase. *Immunity* **46**, 1045–1058.e6 (2017). doi: 10.1016/j.immuni.2017.06.005; pmid: 28636954
6. C. Mlynarczyk, L. Fontán, A. Melnick, Germinal center-derived lymphomas: The darkest side of humoral immunity. *Immunol. Rev.* **288**, 214–239 (2019). doi: 10.1111/imr.12755; pmid: 30874354
7. L. Pasqualucci, Molecular pathogenesis of germinal center-derived B cell lymphomas. *Immunol. Rev.* **288**, 240–261 (2019). doi: 10.1111/imr.12745; pmid: 30874347
8. A. A. Alizadeh *et al*., Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. *Nature* **403**, 503–511 (2000). doi: 10.1038/35000501; pmid: 10676951
9. R. Schmitz *et al*., Genetics and pathogenesis of diffuse large B-cell lymphoma. *N. Engl. J. Med.* **378**, 1396–1407 (2018). doi: 10.1056/NEJMoa1801445; pmid: 29641966
10. B. Chapuy *et al*., Molecular subtypes of diffuse large B cell lymphoma are associated with distinct pathogenic mechanisms and outcomes. *Nat. Med.* **24**, 679–690 (2018). doi: 10.1038/s41591-018-0016-8; pmid: 29713087
11. G. W. Wright *et al*., A probabilistic classification tool for genetic subtypes of diffuse large B cell lymphoma with therapeutic implications. *Cancer Cell* **37**, 551–568.e14 (2020). doi: 10.1016/j.ccell.2020.03.015; pmid: 32289277
12. L. Venturutti *et al*., TBL1XR1 mutations drive extranodal lymphoma by inducing a pro-tumorigenic memory fate. *Cell* **182**, 297–316.e27 (2020). doi: 10.1016/j.cell.2020.05.049; pmid: 32619424
13. L. Venturutti *et al*., An aged/autoimmune B-cell program defines the early transformation of extranodal lymphomas. *Cancer Discov.*OF1–OF28 (2022). doi: 10.1158/2159-8290.CD-22-0561; pmid: 36264161
14. W. J. Lin, J. D. Gary, M. C. Yang, S. Clarke, H. R. Herschman, The mammalian immediate-early TIS21 protein and the leukemia-associated BTG1 protein interact with a protein-arginine *N*-methyltransferase. *J. Biol. Chem.* **271**, 15034–15044 (1996). doi: 10.1074/jbc.271.25.15034; pmid: 8663146
15. J. P. Rouault *et al*., Interaction of BTG1 and p53-regulated BTG2 gene products with mCaf1, the murine homolog of a component of the yeast CCR4 transcriptional regulatory complex. *J. Biol. Chem.* **273**, 22563–22569 (1998). doi: 10.1074/jbc.273.35.22563; pmid: 9712883
16. D. Prévôt *et al*., The leukemia-associated protein Btg1 and the p53-regulated protein Btg2 interact with the homeoprotein Hoxb9 and enhance its transcriptional activation. *J. Biol. Chem.* **275**, 147–153 (2000). doi: 10.1074/jbc.275.1.147; pmid: 10617598
17. M. Busson *et al*., Coactivation of nuclear receptors and myogenic factors induces the major BTG1 influence on muscle differentiation. *Oncogene* **24**, 1698–1710 (2005). doi: 10.1038/sj.onc.1208373; pmid: 15674337
18. K. Hata, K. Nishijima, J. Mizuguchi, Role for Btg1 and Btg2 in growth arrest of WEHI-231 cells through arginine methylation following membrane immunoglobulin engagement. *Exp. Cell Res.* **313**, 2356–2366 (2007). doi: 10.1016/j.yexcr.2007.03.021; pmid: 17466295
19. E. Tijchon *et al*., Tumor suppressors BTG1 and BTG2 regulate early mouse B-cell development. *Haematologica* **101**, e272–e276 (2016). doi: 10.3324/haematol.2015.139675; pmid: 27036158
20. H. A. Almasmoum, B. Airhihen, C. Seedhouse, G. S. Winkler, Frequent loss of BTG1 activity and impaired interactions with the Caf1 subunit of the Ccr4-Not deadenylase in non-Hodgkin lymphoma. *Leuk. Lymphoma* **62**, 281–290 (2021). doi: 10.1080/10428194.2020.1827243; pmid: 33021411
21. N. Kotlov *et al*., Clinical and biological subtypes of B-cell lymphoma revealed by microenvironmental signatures. *Cancer Discov.* **11**, 1468–1489 (2021). doi: 10.1158/2159-8290.CD-20-0839; pmid: 33541860
22. M. C. J. Ma *et al*., Subtype-specific and co-occurring genetic alterations in B-cell non-Hodgkin lymphoma. *Haematologica* **107**, 690–701 (2022). doi: 10.3324/haematol.2020.274258; pmid: 33792219
23. D. Ennishi *et al*., *TMEM30A* loss-of-function mutations drive lymphomagenesis and confer therapeutically exploitable vulnerability in B-cell lymphoma. *Nat. Med.* **26**, 577–588 (2020). doi: 10.1038/s41591-020-0757-z; pmid: 32094924
24. A. Reddy *et al*., Genetic and functional drivers of diffuse large B cell lymphoma. *Cell* **171**, 481–494.e15 (2017). doi: 10.1016/j.cell.2017.09.027; pmid: 28985567
25. J. Zhang *et al*., Genetic heterogeneity of diffuse large B-cell lymphoma. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 1398–1403 (2013). doi: 10.1073/pnas.1205299110; pmid: 23292937
26. J. G. Lohr *et al*., Discovery and prioritization of somatic mutations in diffuse large B-cell lymphoma (DLBCL) by whole-exome sequencing. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 3879–3884 (2012). doi: 10.1073/pnas.1121343109; pmid: 22343534
27. R. D. Morin *et al*., Frequent mutation of histone-modifying genes in non-Hodgkin lymphoma. *Nature* **476**, 298–303 (2011). doi: 10.1038/nature10351; pmid: 21796119
28. E. Cerami *et al*., The cBio cancer genomics portal: An open platform for exploring multidimensional cancer genomics data. *Cancer Discov.* **2**, 401–404 (2012). doi: 10.1158/2159-8290.CD-12-0095; pmid: 22588877
29. J. Gao *et al*., Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci. Signal.* **6**, pl1 (2013). doi: 10.1126/scisignal.2004088; pmid: 23550210
30. A. Rodier *et al*., Identification of functional domains involved in BTG1 cell localization. *Oncogene* **20**, 2691–2703 (2001). doi: 10.1038/sj.onc.1204398; pmid: 11420681
31. E. Kots, C. Mlynarczyk, A. Melnick, G. Khelashvili, Conformational transitions in BTG1 antiproliferative protein and their modulation by disease mutants. *Biophys. J.* **121**, 3753–3764 (2022). doi: 10.1016/j.bpj.2022.04.023; pmid: 35459639
32. R. C. Rickert, J. Roes, K. Rajewsky, B lymphocyte-specific, Cre-mediated mutagenesis in mice. *Nucleic Acids Res.* **25**, 1317–1318 (1997). doi: 10.1093/nar/25.6.1317; pmid: 9092650
33. T. A. Shih, M. Roederer, M. C. Nussenzweig, Role of antigen receptor affinity in T cell–independent antibody responses *in vivo*. *Nat. Immunol.* **3**, 399–406 (2002). doi: 10.1038/ni776; pmid: 11896394
34. A. Subramanian *et al*., Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 15545–15550 (2005). doi: 10.1073/pnas.0506580102; pmid: 16199517
35. A. Liberzon *et al*., Molecular signatures database (MSigDB) 3.0. *Bioinformatics* **27**, 1739–1740 (2011). doi: 10.1093/bioinformatics/btr260; pmid: 21546393
36. A. Liberzon *et al*., The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst.* **1**, 417–425 (2015). doi: 10.1016/j.cels.2015.12.004; pmid: 26771021
37. C. Y. Huang, A. L. Bredemeyer, L. M. Walker, C. H. Bassing, B. P. Sleckman, Dynamic regulation of *c-Myc* proto-oncogene expression during lymphocyte development revealed by a *GFP-c-Myc* knock-in mouse. *Eur. J. Immunol.* **38**, 342–349 (2008). doi: 10.1002/eji.200737972; pmid: 18196519
38. S. Casola *et al*., Tracking germinal center B cells expressing germ-line immunoglobulin γ1 transcripts by conditional gene targeting. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 7396–7401 (2006). doi: 10.1073/pnas.0602353103; pmid: 16651521
39. A. Tesi *et al*., An early Myc-dependent transcriptional program orchestrates cell growth during B-cell activation. *EMBO Rep.* **20**, e47987 (2019). doi: 10.15252/embr.201947987; pmid: 31334602
40. E. C. Luo *et al*., Large-scale tethered function assays identify factors that regulate mRNA stability and translation. *Nat. Struct. Mol. Biol.* **27**, 989–1000 (2020). doi: 10.1038/s41594-020-0477-6; pmid: 32807991
41. S. S. Hwang *et al*., mRNA destabilization by BTG1 and BTG2 maintains T cell quiescence. *Science* **367**, 1255–1260 (2020). doi: 10.1126/science.aax0194; pmid: 32165587
42. Y. Li, P. S. Choi, S. C. Casey, D. L. Dill, D. W. Felsher, MYC through miR-17-92 suppresses specific target genes to maintain survival, autonomous proliferation, and a neoplastic state. *Cancer Cell* **26**, 262–272 (2014). doi: 10.1016/j.ccr.2014.06.014; pmid: 25117713
43. R. Caeser, J. Gao, M. Di Re, C. Gong, D. J. Hodson, Genetic manipulation and immortalized culture of ex vivo primary human germinal center B cells. *Nat. Protoc.* **16**, 2499–2519 (2021). doi: 10.1038/s41596-021-00506-4; pmid: 33837304
44. G. D. Victora *et al*., Germinal center dynamics revealed by multiphoton microscopy with a photoactivatable fluorescent reporter. *Cell* **143**, 592–605 (2010). doi: 10.1016/j.cell.2010.10.032; pmid: 21074050
45. S. Finkin, H. Hartweger, T. Y. Oliveira, E. E. Kara, M. C. Nussenzweig, Protein amounts of the MYC transcription factor determine germinal center B cell division capacity. *Immunity* **51**, 324–336.e5 (2019). doi: 10.1016/j.immuni.2019.06.013; pmid: 31350178
46. A. D. Gitlin, Z. Shulman, M. C. Nussenzweig, Clonal selection in the germinal centre by regulated proliferation and hypermutation. *Nature* **509**, 637–640 (2014). doi: 10.1038/nature13300; pmid: 24805232

47. M. Meyer-Hermann, A molecular theory of germinal center B cell selection and division. *Cell Rep.* **36**, 109552 (2021). doi: 10.1016/j.celrep.2021.109552; pmid: 34433043
48. S. Ogilvy *et al*., Constitutive Bcl-2 expression throughout the hematopoietic compartment affects multiple lineages and enhances progenitor cell survival. *Proc. Natl. Acad. Sci. U.S.A.* **96**, 14943–14948 (1999). doi: 10.1073/pnas.96.26.14943; pmid: 10611317
49. A. Egle, A. W. Harris, M. L. Bath, L. O'Reilly, S. Cory, VavP-*Bcl2* transgenic mice develop follicular lymphoma preceded by germinal center hyperplasia. *Blood* **103**, 2276–2283 (2004). doi: 10.1182/blood-2003-07-2469; pmid: 14630790
50. K. B. Hoehn *et al*., Repertoire-wide phylogenetic models of B cell molecular evolution reveal evolutionary signatures of aging and vaccination. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 22664–22672 (2019). doi: 10.1073/pnas.1906020116; pmid: 31636219
51. A. M. Nedelcu, The evolution of multicellularity and cancer: Views and paradigms. *Biochem. Soc. Trans.* **48**, 1505–1518 (2020). doi: 10.1042/BST20190992; pmid: 32677677
52. N. E. Baker, Emerging mechanisms of cell competition. *Nat. Rev. Genet.* **21**, 683–697 (2020). doi: 10.1038/s41576-020-0262-8; pmid: 32778819
53. C. de la Cova, M. Abril, P. Bellosta, P. Gallant, L. A. Johnston, *Drosophila* myc regulates organ size by inducing cell competition. *Cell* **117**, 107–116 (2004). doi: 10.1016/S0092-8674(04)00214-4; pmid: 15066286
54. E. Moreno, K. Basler, dMyc transforms cells into super-competitors. *Cell* **117**, 117–129 (2004). doi: 10.1016/S0092-8674(04)00262-4; pmid: 15066287
55. A. D. Gitlin *et al*., T cell help controls the speed of the cell cycle in germinal center B cells. *Science* **349**, 643–646 (2015). doi: 10.1126/science.aac4919; pmid: 26184917
56. S. Heinzel *et al*., A Myc-dependent division timer complements a cell-death timer to regulate T cell and B cell responses. *Nat. Immunol.* **18**, 96–103 (2017). doi: 10.1038/ni.3598; pmid: 27820810
57. J. A. Bogdan *et al*., Human carbon catabolite repressor protein (CCR4)-associative factor 1: Cloning, expression and characterization of its interaction with the B-cell translocation protein BTG1. *Biochem. J.* **336**, 471–481 (1998). doi: 10.1042/bj3360471; pmid: 9820826
58. L. Yuniati, B. Scheijen, L. T. van der Meer, F. N. van Leeuwen, Tumor suppressors BTG1 and BTG2: Beyond growth control. *J. Cell. Physiol.* **234**, 5379–5389 (2019). doi: 10.1002/jcp.27407; pmid: 30350856
59. L. Shen, Q. Shi, W. Wang, Double agents: Genes with both oncogenic and tumor-suppressor functions. *Oncogenesis* **7**, 25 (2018). doi: 10.1038/s41389-018-0034-x pmid: 29540752
60. N. Datta, S. Chakraborty, M. Basu, M. K. Ghosh, Tumor suppressors having oncogenic functions: The double agents. *Cells* **10**, 46 (2020). doi: 10.3390/cells10010046; pmid: 33396222
61. E. Waanders *et al*., The origin and nature of tightly clustered *BTG1* deletions in precursor B-cell acute lymphoblastic leukemia support a model of multiclonal evolution. *PLOS Genet.* **8**, e1002533 (2012). doi: 10.1371/journal.pgen.1002533; pmid: 22359517
62. Y. Wu *et al*., MicroRNA-17-92 is required for T-cell and B-cell pathogenicity in chronic graft-versus-host disease in mice. *Blood* **131**, 1974–1986 (2018). doi: 10.1182/blood-2017-06-789321; pmid: 29530952
63. J. P. Rouault *et al*., BTG1, a member of a new family of antiproliferative genes. *EMBO J.* **11**, 1663–1670 (1992). doi: 10.1002/j.1460-2075.1992.tb05213.x; pmid: 1373383
64. L. Venturutti, A. M. Melnick, The dangers of déjà vu: Memory B cells as the cells of origin of ABC-DLBCLs. *Blood* **136**, 2263–2274 (2020). doi: 10.1182/blood.2020005857; pmid: 32932517
65. B. Becher, A. Waisman, L. F. Lu, Conditional gene-targeting in mice: Problems and solutions. *Immunity* **48**, 835–836 (2018). doi: 10.1016/j.immuni.2018.05.002; pmid: 29768166
66. *Guide for the Care and Use of Laboratory Animals* (The National Academies Press, ed. 8, 2011).
67. K. Inaba *et al*., Tissue distribution of the DEC-205 protein that is detected by the monoclonal antibody NLDC-145. I. Expression on dendritic cells and other subsets of mouse leukocytes. *Cell. Immunol.* **163**, 148–156 (1995). doi: 10.1006/cimm.1995.1109; pmid: 7758125
68. J. Z. Long, C. S. Lackan, A. K. Hadjantonakis, Genetic and spectrally distinct *in vivo* imaging: Embryonic stem cells and mice with widespread expression of a monomeric red fluorescent protein. *BMC Biotechnol.* **5**, 20 (2005). doi: 10.1186/1472-6750-5-20; pmid: 15996270
69. J. Hanna *et al*., Direct reprogramming of terminally differentiated mature B lymphocytes to pluripotency. *Cell* **133**, 250–264 (2008). doi: 10.1016/j.cell.2008.03.028; pmid: 18423197
70. G. Pasqual, A. Angelini, G. D. Victora, Triggering positive selection of germinal center B cells by antigen targeting to DEC-205. *Methods Mol. Biol.* **1291**, 125–134 (2015). doi: 10.1007/978-1-4939-2498-1_10; pmid: 25836306
71. P. Bankhead *et al*., QuPath: Open source software for digital pathology image analysis. *Sci. Rep.* **7**, 16878 (2017). doi: 10.1038/s41598-017-17204-5; pmid: 29203879
72. J. Schindelin *et al*., Fiji: An open-source platform for biological-image analysis. *Nat. Methods* **9**, 676–682 (2012). doi: 10.1038/nmeth.2019; pmid: 22743772
73. X. Yang *et al*., A public genome-scale lentiviral expression library of human ORFs. *Nat. Methods* **8**, 659–661 (2011). doi: 10.1038/nmeth.1638; pmid: 21706014
74. R. Caeser *et al*., Genetic modification of primary human B cells to model high-grade lymphoma. *Nat. Commun.* **10**, 4543 (2019). doi: 10.1038/s41467-019-12494-x; pmid: 31586074
75. I. Huppertz *et al*., iCLIP: Protein–RNA interactions at nucleotide resolution. *Methods* **65**, 274–287 (2014). doi: 10.1016/j.ymeth.2013.10.011; pmid: 24184352
76. C. S. Carlson *et al*., Using synthetic templates to design an unbiased multiplex PCR assay. *Nat. Commun.* **4**, 2680 (2013). doi: 10.1038/ncomms3680; pmid: 24157944
77. A. Dobin *et al*., STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013). doi: 10.1093/bioinformatics/bts635; pmid: 23104886
78. Y. Liao, G. K. Smyth, W. Shi, The R package *Rsubread* is easier, faster, cheaper and better for alignment and quantification of RNA sequencing reads. *Nucleic Acids Res.* **47**, e47 (2019). doi: 10.1093/nar/gkz114; pmid: 30783653
79. M. D. Robinson, D. J. McCarthy, G. K. Smyth, edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2010). doi: 10.1093/bioinformatics/btp616; pmid: 19910308
80. D. J. McCarthy, Y. Chen, G. K. Smyth, Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. *Nucleic Acids Res.* **40**, 4288–4297 (2012). doi: 10.1093/nar/gks042; pmid: 22287627
81. G. Csardi, T. Nepusz, The igraph software package for complex network research. *InterJournal* **Complex Systems**, 1695 (2006).
82. H. Goodarzi, O. Elemento, S. Tavazoie, Revealing global regulatory perturbations across human cancers. *Mol. Cell* **36**, 900–911 (2009). doi: 10.1016/j.molcel.2009.11.016; pmid: 20005852
83. A. L. Shaffer *et al*., A library of gene expression signatures to illuminate normal and pathological lymphoid biology. *Immunol. Rev.* **210**, 67–85 (2006). doi: 10.1111/j.0105-2896.2006.00373.x; pmid: 16623765
84. W. E. Johnson, C. Li, A. Rabinovic, Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics* **8**, 118–127 (2007). doi: 10.1093/biostatistics/kxj037; pmid: 16632515
85. J. T. Leek, W. E. Johnson, H. S. Parker, A. E. Jaffe, J. D. Storey, The sva package for removing batch effects and other unwanted variation in high-throughput experiments. *Bioinformatics* **28**, 882–883 (2012). doi: 10.1093/bioinformatics/bts034; pmid: 22257669
86. I. V. Kulakovskiy *et al*., HOCOMOCO: Towards a complete collection of transcription factor binding models for human and mouse via large-scale ChIP-Seq analysis. *Nucleic Acids Res.* **46**, D252–D259 (2018). doi: 10.1093/nar/gkx1106; pmid: 29140464
87. T. L. Bailey *et al*., MEME SUITE: Tools for motif discovery and searching. *Nucleic Acids Res.* **37**, W202–W208 (2009). doi: 10.1093/nar/gkp335; pmid: 19458158
88. T. L. Bailey, P. Machanick, Inferring direct DNA binding from ChIP-seq. *Nucleic Acids Res.* **40**, e128 (2012). doi: 10.1093/nar/gks433; pmid: 22610855
89. D. Risso, J. Ngai, T. P. Speed, S. Dudoit, Normalization of RNA-seq data using factor analysis of control genes or samples. *Nat. Biotechnol.* **32**, 896–902 (2014). doi: 10.1038/nbt.2931; pmid: 25150836
90. T. Stuart *et al*., Comprehensive integration of single-cell data. *Cell* **177**, 1888–1902.e21 (2019). doi: 10.1016/j.cell.2019.05.031; pmid: 31178118
91. K. Street *et al*., Slingshot: Cell lineage and pseudotime inference for single-cell transcriptomics. *BMC Genomics* **19**, 477 (2018). doi: 10.1186/s12864-018-4772-0; pmid: 29914354
92. M. Garcia *et al*., Sarek: A portable workflow for whole-genome sequencing analysis of germline and somatic variants. *F1000Res.* **9**, 63 (2020). doi: 10.12688/f1000research.16665.2; pmid: 32269765
93. S. Kim *et al*., Strelka2: Fast and accurate calling of germline and somatic variants. *Nat. Methods* **15**, 591–594 (2018). doi: 10.1038/s41592-018-0051-x; pmid: 30013048
94. X. Chen *et al*., Manta: Rapid detection of structural variants and indels for germline and cancer sequencing applications. *Bioinformatics* **32**, 1220–1222 (2016). doi: 10.1093/bioinformatics/btv710; pmid: 26647377
95. J. Shamsani *et al*., A plugin for the Ensembl Variant Effect Predictor that uses MaxEntScan to predict variant spliceogenicity. *Bioinformatics* **35**, 2315–2317 (2019). doi: 10.1093/bioinformatics/bty960; pmid: 30475984
96. M. Imielinski, G. Guo, M. Meyerson, Insertions and deletions target lineage-defining genes in human cancers. *Cell* **168**, 460–472.e14 (2017). doi: 10.1016/j.cell.2016.12.025; pmid: 28089356
97. Y. Jiang *et al*., *CREBBP* inactivation promotes the development of HDAC3-dependent lymphomas. *Cancer Discov.* **7**, 38–53 (2017). doi: 10.1158/2159-8290.CD-16-0975; pmid: 27733359
98. W. A. Whyte *et al*., Master transcription factors and mediator establish super-enhancers at key cell identity genes. *Cell* **153**, 307–319 (2013). doi: 10.1016/j.cell.2013.03.035; pmid: 23582322
99. T. A. Pearson, T. A. Manolio, How to interpret a genome-wide association study. *JAMA* **299**, 1335–1344 (2008). doi: 10.1001/jama.299.11.1335; pmid: 18349094
100. V. Giudicelli, D. Chaume, M. P. Lefranc, IMGT/GENE-DB: A comprehensive database for human and mouse immunoglobulin and T cell receptor genes. *Nucleic Acids Res.* **33**, D256–D261 (2005). doi: 10.1093/nar/gki010; pmid: 15608191
101. J. Ye, N. Ma, T. L. Madden, J. M. Ostell, IgBLAST: An immunoglobulin variable domain sequence analysis tool. *Nucleic Acids Res.* **41**, W34–W40 (2013). doi: 10.1093/nar/gkt382; pmid: 23671333
102. N. T. Gupta *et al*., Change-O: A toolkit for analyzing large-scale B cell immunoglobulin repertoire sequencing data. *Bioinformatics* **31**, 3356–3358 (2015). doi: 10.1093/bioinformatics/btv359; pmid: 26069265
103. N. Nouri, S. H. Kleinstein, A spectral clustering-based method for identifying clones from high-throughput B cell repertoire sequencing data. *Bioinformatics* **34**, i341–i349 (2018). doi: 10.1093/bioinformatics/bty235; pmid: 29949968
104. N. T. Gupta *et al*., Hierarchical clustering can identify B cell clones with high confidence in Ig repertoire sequencing data. *J. Immunol.* **198**, 2489–2499 (2017). doi: 10.4049/jimmunol.1601850; pmid: 28179494
105. G. Yaari *et al*., Models of somatic hypermutation targeting and substitution based on synonymous mutations from high-throughput immunoglobulin sequencing data. *Front. Immunol.* **4**, 358 (2013). doi: 10.3389/fimmu.2013.00358; pmid: 24298272
106. K. B. Hoehn, O. G. Pybus, S. H. Kleinstein, Phylogenetic analysis of migration, differentiation, and class switching in B cells. *PLOS Comput. Biol.* **18**, e1009885 (2022). doi: 10.1371/journal.pcbi.1009885; pmid: 35468128
107. R. Nielsen, Z. Yang, Likelihood models for detecting positively selected amino acid sites and applications to the HIV-1 envelope gene. *Genetics* **148**, 929–936 (1998). doi: 10.1093/genetics/148.3.929; pmid: 9539414
108. G. Yu, T. T. Lam, H. Zhu, Y. Guan, Two methods for mapping and visualizing associated data on phylogeny using *Ggtree*. *Mol. Biol. Evol.* **35**, 3041–3043 (2018). doi: 10.1093/molbev/msy194; pmid: 30351396
109. S. C. Binder, M. Meyer-Hermann, Implications of intravital imaging of murine germinal centers on the control of B cell selection and division. *Front. Immunol.* **7**, 593 (2016). doi: 10.3389/fimmu.2016.00593; pmid: 28066409
110. M. Meyer-Hermann, Overcoming the dichotomy of quantity and quality in antibody responses. *J. Immunol.* **193**, 5414–5419 (2014). doi: 10.4049/jimmunol.1401828; pmid: 25355924
111. M. Meyer-Hermann *et al*., A theory of germinal center B cell selection, division, and exit. *Cell Rep.* **2**, 162–174 (2012). doi: 10.1016/j.celrep.2012.05.010; pmid: 22840406

112. A. S. Perelson, G. F. Oster, Theoretical studies of clonal selection: Minimal antibody repertoire size and reliability of self-non-self discrimination. *J. Theor. Biol.* **81**, 645–670 (1979). doi: 10.1016/0022-5193(79)90275-3; pmid: 94141
113. C. Mlynarczyk *et al*., Codes used in "BTG1 mutation yields supercompetitive B cells primed for malignant transformation," version 1.0.0, Zenodo (2022); https://doi.org/10.5281/zenodo.7406382.

ACKNOWLEDGMENTS

The authors thank Y. Jiang for support with conditional mouse line generation strategy; A. Intlekofer for advice and use of the Coulter counter; T. C. Lee for technical support; X. Wang and C. Meydan for data curation; J. C. Hellmuth for support in data analysis; C. Wheeler and G. Fayzikhodjaeva for support with mouse colonies management; S. Monette and M. S. Jiao from the Laboratory of Comparative Pathology (WCM/MSKCC) for support in optimizing IHC staining; C. Durmaz for support with graphical illustration; and L. Scourzic and members of the Melnick and Victora laboratories for discussions and suggestions. The authors acknowledge the Epigenomics, Genomics Resources, Flow Cytometry, and Optical Microscopy core facilities (WCM); the Laboratory of Comparative Pathology (WCM/MSKCC); the Mouse Genetics Core Facility (MSKCC); and the David A. Cofrin Center for Biomedical Information at the HRH Prince Alwaleed Bin Talal Bin Abdulaziz Alsaud Institute for Computational Biomedicine (WCM). **Funding:** Funding was provided by ASH Fellow-to-Faculty scholar award ASHI 204241-01 (C.M.); Leukemia & Lymphoma Society Special Fellow fellowship 3384-19 (C.M.); Lymphoma Research Foundation fellowship LRF-17030307-02 (C.M.); ASH Research Restart award (C.M.); Lymphoma Research Foundation fellowship LSRMP-817214 (C.M.); Damon Runyon Cancer Research Foundation, Berger Foundation fellowship DRG-2353-19 (J.P.); NIH NCI award 1F31CA254302-01 (C.R.C.); COSMIC Marie Skłodowska-Curie grant 765158 (T.A.); a Swiss National Science Foundation fellowship (D.B.); a National Cancer Center postdoctoral fellowship (D.B.); NIH NCI K99CA212276 (D.B.); AACR-Incyte Lymphoma Research Fellowship 21-40-46-PAPI (A.P.); NIH NIAID grant R01AI104739 (K.B.H.); a Cancer Research Institute Irvington postdoctoral fellowship (J.E.); an Agency for Science, Technology and Research (A*STAR) award (H.X.P.); NIH NCI award 1F31CA257204-01A1 (C.M.E.); Lymphoma Research Foundation fellowship 817442 (H.M.I.); NIH NCI award T32CA062948 (H.M.I.); Deutsche Forschungsgemeinschaft PE3140/1-1 (B.P.); Deutsche Krebshilfe, grant Mildred Scheel Nachwuchszentrum 70113307 (B.P.); NIH NCI award 1F31CA220981-01 (A.S.D.); ASH Scholar Award FP00000270 (D.T.T.N.); NIH grant UL1TR002384 (O.E.); NIH grant R01CA194547 (O.E.); Leukemia & Lymphoma Society SCOR grants 180078-02 and 7021-20 (O.E.); a Leukemia & Lymphoma Society Scholarship (M.G.K.); an MSK Lymphoma SPORE grant (M.G.K.); NIH grant R01CA186702 (S.R.J.); a Michael Smith Foundation for Health Research Health Professional Investigator Award (D.W.S.); 1923 Fund (G.K.); NIH grant R01AI139117 (G.D.V.); The Starr Foundation Tri-Institutional Stem Cell Initiative 2021-033 (M.G.K. and A.M.); Starr Cancer Consortium grant I11-0027 (O.E., G.D.V., and A.M.); NIH NCI grant R35CA220499 (A.M.); Leukemia & Lymphoma Society TRP grant 6572 (A.M.); Leukemia & Lymphoma Society SCOR grant 7012 (A.M.); the Follicular Lymphoma Consortium (A.M.); and the Samuel Waxman Cancer Research Foundation (AM). **Author contributions:** Conceptualization: C.M. and A.M. Data curation: M.T., C.R.C., Z.C., and A.S.D. Formal analysis: C.M., M.T., J.P., C.R.C,. T.A., D.B., A.P., K.B.H., J.E., E.K., A.B., E.B., H.X.P., A.C., Z.C., H.M.I., B.P., I.T., D.T.T.N., A.S.D., H.G., M.H.R., J.M., W.M., and G.K. Funding acquisition: C.M., A.M., G.D.V., and O.E. Investigation: C.M., J.P., L.W., D.B., A.P., J.E., E.B., H.X.P., C.M.E., H.M.I., B.P., I.T., H.S., D.T.T.N., M.H.R., and W.M. Methodology: C.M., J.P., A.P., J.E., E.B., H.X.P., C.M.E., H.M.I., B.P., D.T.T.N., M.G.K., S.R.J., G.D.V., and A.M. Resources: O.E., M.G.K., D.W.S., S.R.J., M.M.-H., G.K., G.D.V., and A.M. Software: M.T., C.R.C., T.A., K.B.H., E.K., A.B., Z.C., A.S.D., H.G., M.M.-H., and G.K. Supervision: A.M. and G.D.V. Validation: C.M., M.T., J.P., C.R.C., L.W., T.A., J.E., D.B., A.P., K.B.H., E.K., A.B., E.B., H.X.P., C.M.E., A.C., Z.C., H.M.I., B.P., I.T., GC, and A.M. Visualization: C.M., J.P., M.T., C.R.C., T.A., K.B.H., E.K., A.C., Z.C., H.M.I., A.S.D., and A.M. Writing – original draft: C.M. and A.M. Writing – review & editing: C.M., M.T., J.P., C.R.C., L.W., T.A., D.B., A.P., K.B.H., E.K., J.E., A.B., E.B., H.X.P., C.M.E., A.C., Z.C., H.S., H.M.I., B.P., I.T., A.S.D., H.G., M.H.R., J.M., W.M., D.T.T.N., O.E., M.G.K., S.R.J., D.W.S., G.K., M.M.-H., G.D.V., and A.M. **Competing interests:** The authors declare no direct competing financial or nonfinancial interests. K.B.H. receives consulting fees from Prellis Biologics. A.C. is on the immunohistochemistry advisory board for Leica Biosystems and consults for Boehringer Ingelheim Pharmaceuticals. O.E. is supported by Janssen, Johnson and Johnson, Volastra Therapeutics, AstraZeneca, and Eli Lilly research grants; is a scientific adviser to and equity holder in Freenome, Owkin, Volastra Therapeutics, and OneThree Biotech; and consults for and advises Champions Oncology. M.G.K. discloses provision of services with 28-7 Therapeutics (uncompensated), Accent Therapeutics, AstraZeneca, and Kumquat Biosciences; and is a scientific adviser to and equity holder in 858 Therapeutics. S.R.J. is scientific founder of, adviser to, and owns equity in Gotham Therapeutics and 858 Therapeutics. D.W.S. receives research funding from Janssen and NanoString; has consulted for Abbvie, AstraZeneca, Celgene, and Janssen; and was named inventor on patents describing the use of gene expression in subtyping lymphomas. G.D.V. is a scientific adviser for Vaccine Company, Inc. A.M. receives research funding from Janssen Pharmaceuticals, Sanofi, Epizyme, and Daiichi Sankyo; has consulted for Epizyme and Constellation; and is on the advisory board of KDAC Pharma. **Data and materials availability:** All data are available in the manuscript or the supplementary materials or are deposited as follows: sequencing data generated for this study is available on NCBI Gene Expression Omnibus (GEO) under accession number GSE167786 and codes and mathematical algorithms used in this study are accessible on Zenodo (*113*). **License information:** Copyright © 2023 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. https://www.science.org/about/science-licenses-journal-article-reuse

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abj7412
Figs. S1 to S16
Tables S1 to S8
MDAR Reproducibility Checklist

Submitted 29 May 2021; resubmitted 15 July 2022
Accepted 12 December 2022
10.1126/science.abj7412



Pushing the Boundaries of Knowledge

As AAAS's first multidisciplinary, open access journal, *Science Advances* publishes research that reflects the selectivity of high impact, innovative research you expect from the *Science* family of journals, published in an open access format to serve a vast and growing global audience. Check out the latest findings or learn how to submit your research: **ScienceAdvances.org**

Science Advances
AAAS

GOLD OPEN ACCESS, DIGITAL, AND FREE TO ALL READERS

RESEARCH ARTICLES

CANCER GENOMICS

Heritable defects in telomere and mitotic function selectively predispose to sarcomas

Mandy L. Ballinger[1,2]†, Swetansu Pattnaik[1,2]†, Piyushkumar A. Mundra[1,2], Milita Zaheed[3], Emma Rath[1], Peter Priestley[4,5], Jonathan Baber[4,5], Isabelle Ray-Coquard[6], Nicholas Isambert[7], Sylvain Causeret[7], Winette T. A. van der Graaf[8], Ajay Puri[9], Florence Duffaud[10], Axel Le Cesne[11], Beatrice Seddon[12], Coonoor Chandrasekar[13], Joshua D. Schiffman[14], Andrew S. Brohl[15], Paul A. James[16,17], Jean-Emmanuel Kurtz[18], Nicolas Penel[19], Ola Myklebost[20,21,22], Leonardo A. Meza-Zepeda[22], Hilda Pickett[23], Maya Kansara[1,2], Nicola Waddell[24], Olga Kondrashova[24], John V. Pearson[24], Andrew P. Barbour[25], Shuai Li[26,27,28,29], Tuong L. Nguyen[26], Diane Fatkin[2,30,31], Robert M. Graham[2,30], Eleni Giannoulatou[2,32], Melissa J. Green[33,34], Warren Kaplan[1,2], Shyamsundar Ravishankar[1], Joseph Copty[1], Joseph E. Powell[1,35], Edwin Cuppen[4], Kristel van Eijk[36], Jan Veldink[36], Jin-Hee Ahn[37], Jeong Eun Kim[37], R. Lor Randall[38], Kathy Tucker[3], Ian Judson[8], Rajiv Sarin[39], Thomas Ludwig[40], Emmanuelle Genin[40], Jean-Francois Deleuze[41], the French Exome Project Consortium‡, Michelle Haber[42], Glenn Marshall[42,43], Murray J. Cairns[44,45], Jean-Yves Blay[6], the International Sarcoma Kindred Study‡, David M. Thomas[1,2]*

Cancer genetics has to date focused on epithelial malignancies, identifying multiple histotype-specific pathways underlying cancer susceptibility. Sarcomas are rare malignancies predominantly derived from embryonic mesoderm. To identify pathways specific to mesenchymal cancers, we performed whole-genome germline sequencing on 1644 sporadic cases and 3205 matched healthy elderly controls. Using an extreme phenotype design, a combined rare-variant burden and ontologic analysis identified two sarcoma-specific pathways involved in mitotic and telomere functions. Variants in centrosome genes are linked to malignant peripheral nerve sheath and gastrointestinal stromal tumors, whereas heritable defects in the shelterin complex link susceptibility to sarcoma, melanoma, and thyroid cancers. These studies indicate a specific role for heritable defects in mitotic and telomere biology in risk of sarcomas.

Common and rare genetic variation currently explains less than 50% of the familial relative risk for cancer, leaving the majority of heritability unexplained (*1*, *2*). Nonetheless, genetic studies focused on common epithelial cancers have yielded major insights into the biological mechanisms underpinning specific cancer susceptibilities, exemplified by homologous recombination in breast cancer and mismatch repair in colorectal cancer (*3*, *4*). Studies of different cancer populations may therefore yield further insights into cancer biology. Differing fundamentally from epithelial cancers, sarcomas are rare connective tissue malignancies that arise predominantly from embryonic mesoderm and affect a younger population (*5*). Because of their rarity, they have been relatively understudied to date at the population level. For the most part, genetic studies into sarcomas have used either familial linkage approaches or genome-wide association studies (GWAS). Studies of rare sarcoma-associated syndromes have led to discovery of key cancer genes, such as the canonical tumor suppressor, *TP53* (*6*). Whole-exome sequencing (WES) or whole-genome sequencing (WGS) are now being used to catalog rare variants in known genes (*7*, *8*). In principle, combining both population- and family-based WGS approaches could uncover additional genes and pathways by integrating statistical methods with clinical information. In this study, we undertook a comprehensive, population-based, case-control study using WGS to identify penetrant genes and pathways that may explain sarcoma risk.

Results

Clinical findings

In total, 1644 sarcoma probands were recruited from sarcoma clinics in this international multi-institutional study, regardless of family history (tables S1 and S2). Subjects had a median age at first cancer diagnosis of 47 years, and 49 years at first sarcoma diagnosis. Soft-tissue sarcomas constituted 78.2% of diagnoses (Table 1 and table S3). Multiple primary cancers were common, including breast (n = 77), melanoma (n = 37), second connective tissue tumors (n = 37), nonmelanoma skin (n = 35), prostate (n = 23), colorectal (n = 21), and thyroid cancers (n = 17). Because radiation is a known risk factor for sarcomas (*9*), we determined the relationship with therapy for a prior cancer. In 110 (38%) of the 293 individuals with multiple primary cancers, the sarcoma was the first tumor diagnosed. In 183 individuals where sarcoma was the second tumor diagnosed, less than one-third (56) occurred within a prior radiation

[1]Garvan Institute of Medical Research, Sydney 2010, Australia. [2]St Vincent's Clinical School, University of New South Wales, Sydney 2010, Australia. [3]Hereditary Cancer Centre, Prince of Wales Hospital, Sydney 2031, Australia. [4]Hartwig Medical Foundation, 1098 XH Amsterdam, Netherlands. [5]Hartwig Medical Foundation Australia, Sydney 2000, Australia. [6]Department of Adult Medical Oncology, Centre Leon Berard, University Claude Bernard, 69373 Lyon, France. [7]Centre Georges Francois Leclerc, 21000 Dijon, France. [8]Sarcoma Unit, The Royal Marsden NHS Foundation Trust, London SW3 6JJ, UK. [9]Department of Orthopedic Oncology, Tata Memorial Hospital, Mumbai, Maharashtra 400012, India. [10]la Timone University Hospital, 13005 Marseille, France. [11]Gustave Roussy, 94805 Villejuif, France. [12]Sarcoma Unit, University College Hospital, London NW1 2BU, UK. [13]Liverpool University Hospitals NHS Foundation Trust, Liverpool L7 8XP, UK. [14]Division of Pediatric Hematology/Oncology, Department of Pediatrics, Huntsman Cancer Institute, University of Utah, Salt Lake City, UT 84112, USA. [15]Sarcoma Department, Moffitt Cancer Center, Tampa, FL 33612, USA. [16]The Sir Peter MacCallum Department of Oncology, University of Melbourne, Melbourne 3010, Australia. [17]Parkville Familial Cancer Centre, Peter MacCallum Cancer Centre and Royal Melbourne Hospital, Melbourne 3000, Australia. [18]Centre Hospitalier Universitaire Hautepierre, 67200 Strasbourg, France. [19]Centre Oscar Lambret, 59000 Lille, France. [20]Western Norway Familial Cancer Centre, Haukeland University Hospital, 5021 Bergen, Norway. [21]Department of Clinical Science, University of Bergen, 5007 Bergen, Norway. [22]Institute for Cancer Research, Oslo University Hospital, N-0424 Oslo, Norway. [23]Children's Medical Research Institute, The University of Sydney, Westmead 2145, Australia. [24]QIMR Berghofer Medical Research Institute, Brisbane 4006, Australia. [25]Faculty of Medicine. The University of Queensland, Brisbane 4072, Australia. [26]Centre for Epidemiology and Biostatistics, Melbourne School of Population and Global Health, Faculty of Medicine, Dentistry and Health Sciences, University of Melbourne, Melbourne 3010, Australia. [27]Centre for Cancer Genetic Epidemiology, Department of Public Health and Primary Care, University of Cambridge, Cambridge CB1 8RN, UK. [28]Precision Medicine, School of Clinical Sciences at Monash Health, Monash University, Clayton 3800, Australia. [29]Murdoch Children's Research Institute, Royal Children's Hospital, Parkville 3051, Australia. [30]Molecular Cardiology Division, Victor Chang Cardiac Research Institute, Darlinghurst 2010, Australia. [31]Cardiology Department, St Vincent's Hospital, Sydney 2010, Australia. [32]Computational Genomics Division, Victor Chang Cardiac Research Institute, Sydney 2010, Australia. [33]School of Psychiatry, University of New South Wales, Sydney 2052, Australia. [34]Neuorscience Research Australia, Sydney 2031, Australia. [35]UNSW Cellular Genomics Futures Institute, University of New South Wales, Sydney 2052, Australia. [36]Department of Neurology, University Medical Centre Utrecht Brain Center, Utrecht University, 3584 CX Utrecht, Netherlands. [37]Department of Oncology, Asan Medical Centre, Seoul 05505, South Korea. [38]Department of Orthopaedic Surgery, University of California, Davis Health, Sacramento, CA 95817, USA. [39]Cancer Genetics Unit, ACTREC, Tata Memorial Centre, Mumbai, Maharashtra 410210, India. [40]Université de Brest, Inserm, EFS, UMR 1078, GGB, CHU de Brest, 29200 Brest, France. [41]Centre National de Recherche en Génomique Humaine, Institut de Génomique, 91057 Evry, France. [42]Children's Cancer Institute, Lowy Cancer Research Centre, University of New South Wales, Kensington 2033, Australia. [43]Kids Cancer Centre, Sydney Children's Hospital, Randwick 2031, Australia. [44]School of Biomedical Sciences and Pharmacy, The University of Newcastle, Callaghan 2308, Australia. [45]Centre for Brain and Mental Health Research, The Hunter Medical Research Institute, Newcastle 2305, Australia.

***Corresponding author. Email: d.thomas@garvan.org.au**

†These authors contributed equally to this work.

‡International Sarcoma Kindred Study and French Exome Project Consortium authors and affiliations are listed in the supplementary materials.

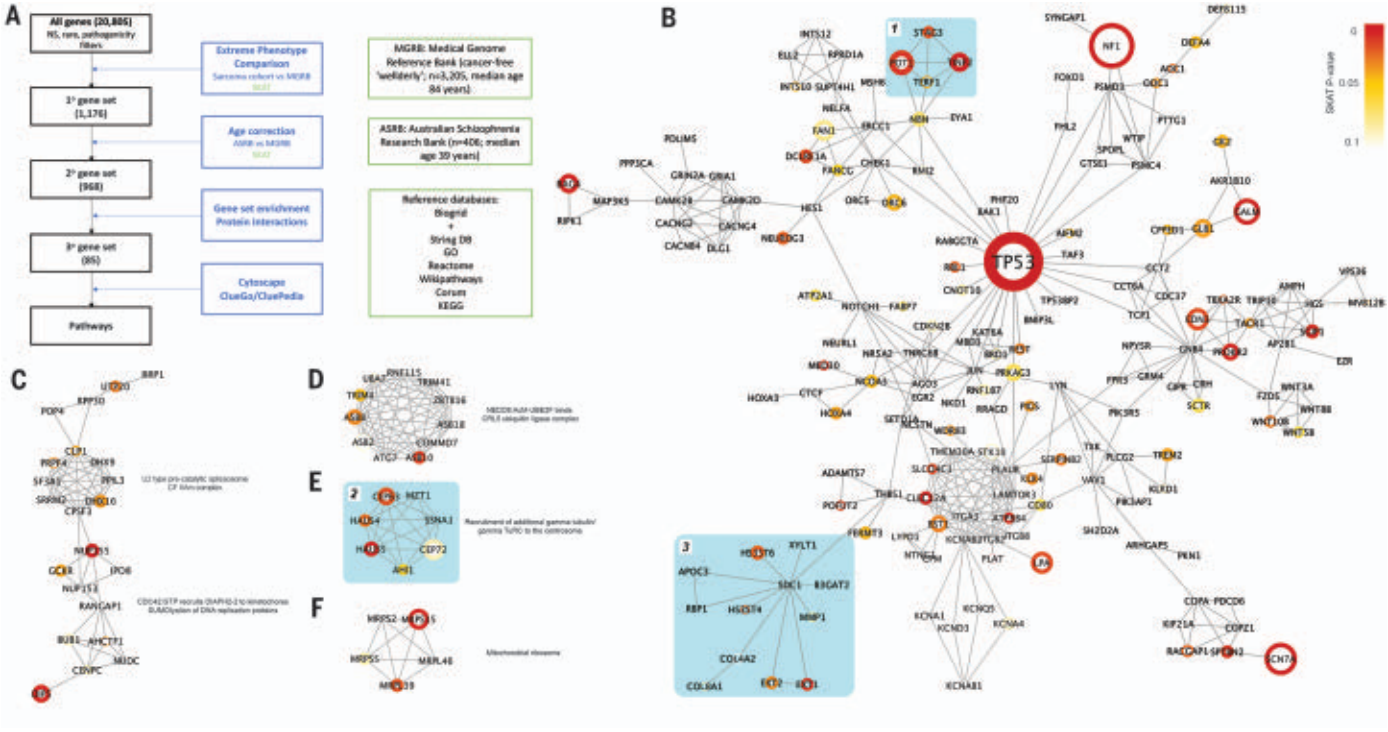


Fig. 1. A systematic analysis of genes and pathways implicated in sarcoma. (**A**) A schematic of the analytic approach of progressively enriching for genes and pathways carrying an excess burden of pathogenic variation in sarcoma probands compared to a universal control cohort (MGRB), including correction for age-depletion of nonsarcoma-related genes. Graph representations of the 968 genes (nodes) in the secondary gene set and their interactions (edges). The analysis was performed in Cytoscape. Node color represents the *P*-value of WRVBT enrichment for each gene in the sarcoma probands relative to the MGRB, the size of the node represents the excess weighted burden of pathogenic variants in the sarcoma probands relative to MGRB, and the width of the node border represents protein interaction and ontologic enrichment by a combined rankscore. Specific groups (1 to 3) relevant to subsequent analyses are highlighted (pale blue boxes). (**B** to **F**) Specific clusters are as follows: (B) Graph representation of the super cluster of 187 genes (nodes) in the secondary geneset centered on *TP53* and their interactions (edges); (C) A 22-gene cluster representing ontologies linked to RNA processing and activity in the M phase of the cell cycle. (D) Eleven genes implicated in antigen processing, ubiquitination, and neddylation. (E) Seven genes defined by centrosomes, spindle organization, and the G_2/M phase of the cell cycle. (F) Five-gene clique with a role in mitochondrial translation and ribosomes.

field. Of 60 women who developed a sarcoma after breast cancer, in more than half (32), the sarcoma arose outside the thorax. These data suggest that the majority of sarcomas in individuals with multiple primary tumors are not therapy related. Upon expert review of pedigrees, more than 20% of families met formal criteria for known hereditary syndromes, including Li Fraumeni Syndrome (LFS), hereditary breast and ovarian cancer (HBOC), and familial melanoma, or were considered to have unusual cancer patterns (Table 1) (*10*). Collectively, these data suggest a strong genetic basis for apparently sporadic sarcomas.

Spectrum of mutated cancer genes in the sarcoma population

After undertaking WGS, we identified 37,820 rare single-nucleotide variants and insertions or deletions (SNV/indels), of which 1033 were known pathogenic or likely pathogenic (C5), 10,702 were new loss-of-function or protein truncating (C4), and 26,085 were possibly pathogenic (C3) (fig. S1). One hundred and nine sarcoma probands (6.6%) carried 112 C4 or C5 variants in clinically relevant cancer genes based on an extended list from the American College of Medical Genetics (ACMG) (table S4) (*11*), including 13 with variants in *ATM* and *TP53*, 10 *NF1*, 9 *BRCA2*, 5 *CHEK2*, 4 *MSH6*, 4 *PALB2*, 4 *LZTR1*, 3 each in *BRIP1* and *EXT2*, 2 each in *APC*, *BRCA1*, *CDKN2B*, *ERCC2* (both compound heterozygotes), *MSH2*, *RHBDF2*, *SDHA*, *SDHB*, and *WT1*, and 1 each in *AIP*, *BAP1*, *BMPR1A*, *BUB1B*, *DICER1*, *EXT1*, *FH*, *FLCN*, *HNF1A*, *MLH1*, *MUTYH* (homozygote), *PMS2*, *POLD1*, *POLE*, *PTEN*, *PTCH1*, *PTPN11*, *RAD51C*, *RAD51D*, *RB1*, *SDHC*, and *SMARCB1* (data file S1). One individual carried C4 or C5 variants in both *SDHB* and *MSH6*. Carriers of C4 or C5 variants were younger at first cancer diagnosis (median age of 41 years) than the remainder of the cohort (median age of 48 years; $P = 0.0003$) and were more likely to develop multiple primary cancers {29 of 101 probands with sufficient clinical information compared with 264 of 1478; odds ratio of 1.85 [95% confidence interval (CI) 1.18–2.91]; $P = 0.011$}. Fewer than one in six of these probands had expected clinical phenotypes that were based on genotypes (*12*). Although 10 out of 13 (77%) carriers of C4 or C5 *TP53* SNV/indels met classic or Chompret criteria for LFS, only 10 out of 164 (6.1%) families meeting either LFS criteria carried a C4 or C5 *TP53* SNV/indel. Other genes in which C4 or C5 variants were detected in families meeting LFS criteria were *MSH6* (2), *ERCC2* (2), *BUB1B*, *LZTR1*, *CHEK2*, *HNF1A*, *PTEN*, *PMS2*, *BRCA2*, *MSH2*, and *SDHB*. Only 1 of 15 (6.7%) probands with C4 or C5 SNV/indels in *BRCA1*, *BRCA2*, or *PALB2* met criteria for HBOC (*13*), whereas only 1 of 27 (4%) families with clinical features of HBOC carried a cognate C4 or C5 variant (*PALB2*). Figure S2 illustrates a sarcoma family carrying a pathogenic *BRCA1* variant, which does not meet HBOC criteria. Twelve sarcoma families met criteria for both HBOC and classic or Chompret criteria for LFS, none of which were explained by C4 or C5 variants in *BRCA1*, *BRCA2*, *PALB2*, or *TP53*. None of 9 probands with pathogenic SNV/indels in mismatch repair genes met Amsterdam criteria for colorectal cancer (*14*), and none of 15 probands meeting familial melanoma criteria (*15*) carried pathogenic

SNV/indels in the gene most strongly associated with familial melanoma, *CDKN2A* (*16*). By contrast with SNV/indels, structural variants appeared to make little contribution to pathogenic burden. Potentially germline pathogenic rare structural variants were identified in *BRCA2* (proband with two breast cancers and an undifferentiated pleomorphic sarcoma), *SDHC* [gastrointestinal stromal tumor (GIST)], and *MSH6* (proband family met Amsterdam criteria for Lynch syndrome). All genes and clinical correlates are shown in data file S1.

Gene and pathway discovery analyses

To maximize power, we applied an extreme phenotype case: control design using 3205 healthy elderly controls [from the Medical Genome Reference Bank (MGRB)], which we previously showed are depleted in cancer-associated genetic variation relative to population-based controls (*17*). A weighted rare-variant burden test (WRVBT) compared the relative burden of C3 to C5 variants in sarcoma probands (discovery set) to the MGRB, after controlling for gender, relatedness, and population structure (Fig. 1A) (*18*). Sensitivity analyses were performed to control for critical parameters including population stratification (*18*). A relaxed P value of < 0.1 was applied to generate an inclusive primary gene set of 1176 genes (Fig. 1A). To control for noncancer, age-related genotypes depleted in MGRB (for example, cardiovascular disease or dementia), we used the Australian Schizophrenia Research Bank (ASRB) (*17*) as a second control set with a median age of 39 years. Genes also enriched in ASRB compared to MGRB were excluded as age- rather than sarcoma-related. A secondary gene set of the top-ranked 968 genes by pathogenic burden was subjected to protein interaction– and pathway-based analysis.

These analyses revealed gene clusters that were composed of highly interrelated subclusters, which we termed cliques. Of 277 (28.6%) genes with at least one first-degree neighbor, 224 were involved in clusters of five or more genes. A supercluster of 187 genes centered on *TP53* (Fig. 1B), the strongest known sarcoma risk gene. *TP53* not only had the greatest burden of C3 to C5 variants but also had the highest number of first-degree neighbors (n = 27). The second largest cluster (Fig. 1C) contained 22 genes, representing ontologies linked to RNA processing and M-phase of the cell cycle, kinetochores, and chromatid separation. A third cluster (Fig. 1D) comprised a clique of 11 genes implicated in antigen processing, ubiquitination, and neddylation. A fourth seven-gene clique (Fig. 1E) was defined by centrosomes, spindle organization, and the G_2/M phase of the cell cycle, and a five-gene clique (Fig. 1F) reflected a role in mitochondrial translation and ribosomes. The degree of interconnectedness within the secondary gene set suggests enrichment in distinct functional pathways.

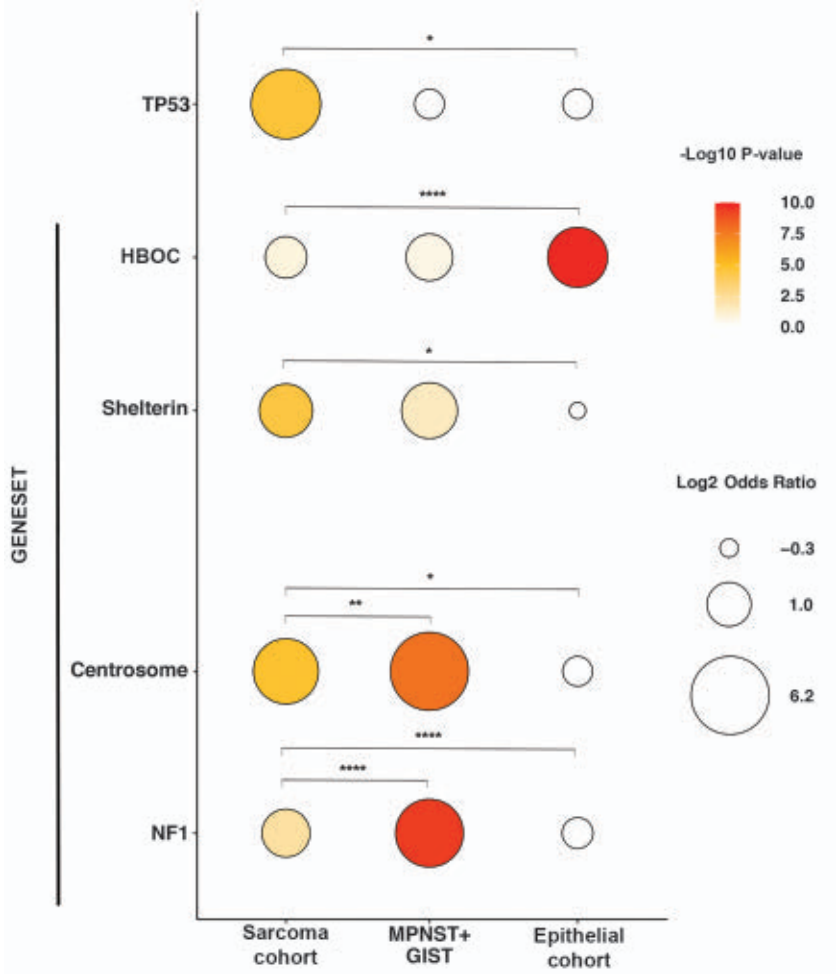


Fig. 2. Sarcoma-specific enrichment in rare pathogenic variants in the shelterin and centrosome pathways. Relative enrichment of pathogenic variants (Log2 Odds Ratio) in sarcoma probands, a subset of 157 cases with GIST or MPNST, and a nonsarcoma cancer population of 632 individuals with epithelial cancers. Gene sets include *TP53* alone (TP53); *BRCA1*, *BRCA2*, and *PALB2* (HBOC); *POT1*, *TINF2*, *TERF1*, *TERF2*, *TERF2IP*, *STAG3*, *SMARCAL1*, and *TIMELESS* (shelterin); *NF1*, *SDHA*, *SDHB*, *SDHC*, *SDHD*, and *LZTR1* (NF1); and *CEP63*, *CEP72*, *HAUS4*, *HAUS5*, *MZT1*, and *SSNA1* (centrosome). Circle size reflects the odds ratio, and the color represents the P value of the enrichment in each group relative to MGRB. Pairwise comparisons between groups are shown (*P < 0.05; **P < 0.01; ****P < 0.0001).

To further enrich for the most important pathways, we focused on the highest-ranked 85 genes with a P value < 0.05 by WRVBT and protein interactions (tertiary gene set) (Fig. 1A). Consistent with known roles in sarcoma susceptibility, the tertiary gene set included *TP53*, *NF1*, *EXT1*, and *EXT2* [UniProt Keywords KW-0043; false discovery rate (FDR) = 2.4×10^{-5}], as well as *PTCH1* (Table 2). Three ontological groups of interest are highlighted in blue (Fig. 1, B and E). Ontologic group 1 comprised the shelterin complex, represented by three genes: *POT1*, *TERF1*, and *TINF2* (GO Component GO.0070187; FDR = 2.8×10^{-5}) (Table 2). Group 2 comprised the mitotic spindle, including the centrosomal genes *CEP63*, *HAUS4*, and *HAUS5* (GO Component term GO.0005819; FDR = 4.5×10^{-5}) (Table 2). Group 3 includes *EXT1* and *EXT2*, linked to hereditary exostoses and bone sarcomas (UniProt Keywords KW-0361; FDR = 2.3×10^{-4}) (Table 2).

We next determined whether these pathways were specific for sarcomas by comparing sarcoma probands to a population with predominantly epithelial cancers (n = 632) (table S5). As a positive control, an enrichment in pathogenic burden was seen for *TP53* in sarcoma probands compared with the epithelial cases. As a negative control, the reverse was observed with a HBOC gene set (*BRCA1*, *BRCA2*, and *PALB2*). Sarcoma-specific enrichment was observed in both the shelterin

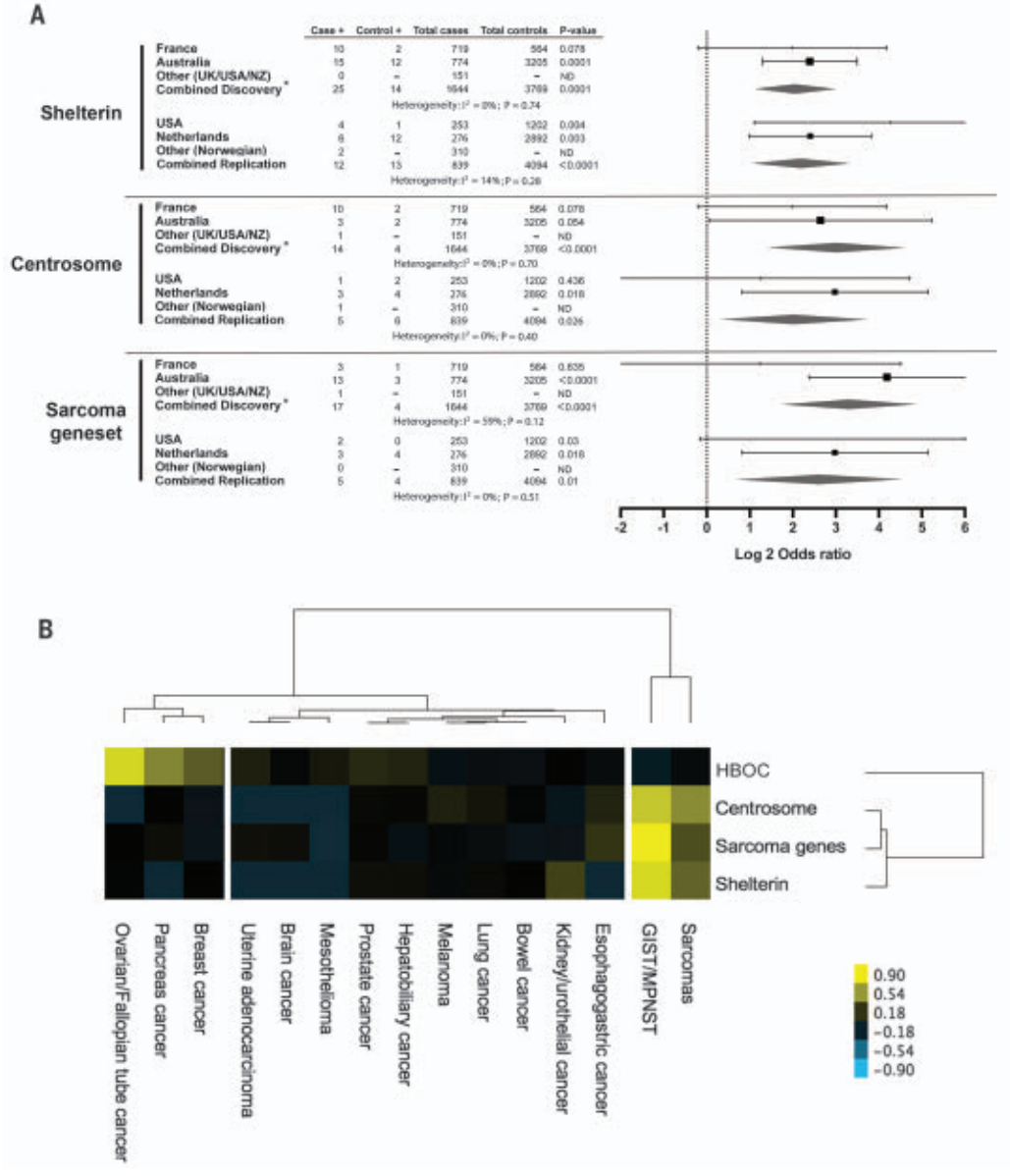


Fig. 3. Replication of enrichment in centrosome and shelterin-complex genes in sarcomas. (**A**) Enrichment of rare C4 or C5 pathogenic variants in sarcoma, shelterin, and centrosome gene sets were estimated for geographically matched discovery (1644 cases and 3769 controls) and replication (839 cases and 4094 controls) datasets. Odds ratios for each geographically matched set are shown with closed squares, and the corresponding horizontal line and whiskers represent 95% confidence intervals (CI). The pooled effect and 95% CI of combined datasets are shown with the diamond. Set heterogeneity was tested using the Cochran-Mantel-Haenszel test without continuity correction and heterogeneity estimates (I^2) and *P* values are presented. Sarcoma cases from the UK, USA, and NZ in the original discovery set and sarcoma cases from Norway in the replication set were excluded from individual analyses due to the lack of geographically matched controls. The combined odds ratio and 95% CI of geographically matched Discovery* and Replication datasets for shelterin, centrosome, and sarcoma gene sets are 5.58 [3.11, 9.64], 4.74 [1.90, 11.83], and 10.59 [4.36, 25.72], respectively. The corresponding I^2 values are 0% for all the gene sets. (**B**) This analysis used the Hartwig Foundation whole-genome germline dataset of 4178 cancers, including sarcomas (276, including 71 GIST or MPNST), breast cancer (801), bowel cancer (650), lung cancer (583), prostate cancer (412), kidney and urothelial cancers (312), melanoma (300), esophagastric cancer (193), ovarian and fallopian tube cancers (174), pancreatic cancer (149), hepatobiliary cancer (138), brain cancer (76), uterine adenocarcinoma (73), and mesothelioma (41). Gene sets were defined as for Fig. 2. Enrichment for each cancer class was determined by normalizing the frequency of variants in that class to the overall frequency in the entire dataset. These data were subject to unsupervised hierarchical clustering for both cancer class and gene sets.

and the centrosome gene sets (Fig. 2). For the centrosome set, further enrichment was specifically observed in malignant peripheral nerve sheath tumors (MPNSTs) and GISTs, which share neural origins and *NF1* as a susceptibility gene (*5*, *19*). The degree of enrichment in centrosome genes in MPNSTs or GISTs was comparable to that seen for *NF1* and related genes such as *LZTR1*, *SDHA*, *SDHB*, *SDHC*, and *SDHD*. A follow-up validation analysis of the discovery set with an iterative (1000-fold) resampling method also recapitulated the enrichment of rare variant burden in the shelterin and centrosome gene sets (fig. S3).

Independent replication of pathway enrichment

We next sought to validate these findings and to exclude population stratification. First, we compiled an independent WGS and WES replication set (n = 839 sarcoma cases from the US, Netherlands, and Norway) and geographically matched independent control set (n = 4094 cancer-free controls) (Fig. 3A and table S2). A WRVBT was performed using C4 and C5 variants in the shelterin, centrosome, and sarcoma gene sets, for each geographically matched case-control discovery and replication set, and after combining these for statistical power (Fig. 3A). Enrichment was confirmed in pathogenic variants in the sarcoma, shelterin, and centrosome gene sets within the combined geographically matched French and Australian discovery sets, as well as in the combined geographically matched replication sets from the Netherlands, US, and Norway, with no significant heterogeneity across these populations (Fig. 3A). Residual population stratification was excluded as a cause for signal enrichment using principal components ancestral matching within the geographically matched populations (figs. S4 and S5) (*18*). We next compared the relative burden of pathogenic variants in the sarcoma, shelterin, centrosome, and HBOC gene sets across a wide range of cancer types using the Hartwig Foundation dataset (table S6)

(*20*). We observed enrichment of HBOC genes in breast, pancreatic, ovarian, and fallopian tube cancers, but not in sarcomas (Fig. 3B). No enrichment was observed in nonsarcoma cancers for the sarcoma, shelterin, and centrosome gene sets. By contrast, a distinct relative enrichment in the shelterin and centrosome gene sets was again observed in the sarcoma cases, and specifically in GISTs or MPNSTs.

Clinical, familial, and molecular genotype–phenotype patterns

Of 19 individuals carrying C4, C5, or essential splice-site variants in the centrosome core genes *CEP63*, *CEP72*, *HAUS4*, and *HAUS5*, as well as in the related genes *CEP89*, *SSNA1*, and *PCM1* (table S7), 8 (42%) developed GIST (7) or MPNST (1). Notably, one individual carried variants in both *HAUS4* and *HAUS5*. An additional 20 individuals carried one C3 variant each in *CEP57*, *CEP63*, *CEP89*, *HAUS4*, *HAUS5*, *PCM1*, *SSNA1*, and *MZT1*. Consistent with the observations in individuals with C4 or C5 variants, 6 of these 20 individuals (30%) with C3 variants developed either GIST (3) or MPNST (3). Altogether, 14 of 39 (36%) carriers of C3 to C5 variants in centrosome genes developed MPNST or GIST, compared with 143 of 1462 (9.8%) in the remaining cohort (RR 4.029 [95% CI 2.372-6.142]; $P < 0.0001$). Of five carriers with GIST, four had somatic c-KIT variants, and one was wild-type. The median age at first cancer diagnosis in C4 or C5 variant carriers (43 years) was comparable to that of the cohort as a whole. Two of the centrosome pedigrees met clinical criteria for LFS, and another two were deemed clinically suspicious. Among individuals with MPNST or GIST, centrosome variants were mutually exclusive with C4 or C5 variants in genes known to cause these tumors: *NF1* (9), *LZTR1*, *SDHA*, and *SDHB*. An analysis of 4179 cancers, including 277 sarcomas, found that 2 of 24 carriers of germline pathogenic variants in centrosome genes showed somatic loss-of-heterozygosity, comparable to 5 of 26 cancers in individuals with pathogenic germline variants in *NF1*, *SDHA*, *SDHB*, and *SDHD*. These data suggest that defects in centrosome pathways appear to confer increased risk of MPNST or GIST.

There were 13 carriers of C4 or C5 variants in four of the six canonical genes in the shelterin complex (Table 3), *POT1* (6), *TERF1* (2), *TINF2* (3), *TERF2IP* (2), and an additional 12 carriers of C4 or C5 variants in genes selected on the basis of their roles in telomere biology, *TIMELESS* (3), *STAG3* (3), and *SMARCAL1* (6) (*21*, *22*). *POT1*, the only established hereditary cancer gene in this group, has been most strongly linked to melanoma, but also has been reported in thyroid cancer and sarcoma (*23*). A loss-of-heterozygosity analysis performed as described above identified 9 somatic loss-of-function events in 55 tumors from carriers of pathogenic germline variants in the shelterin genes, including 2 of 8 in *POT1* and 5 of 19 in *SMARCAL1*.

Amongst carriers of shelterin C4 or C5 variants, formal review of pedigree patterns identified four families meeting GenoMEL criteria for familial melanoma, and four families that met Chompret or classic LFS criteria. Two additional families meeting GenoMEL criteria carried C3 variants in *POT1* (G176R and D224N; fig. S6) (*24*). Notably, three individuals also had thyroid cancer. Several lines of evidence suggest the association between

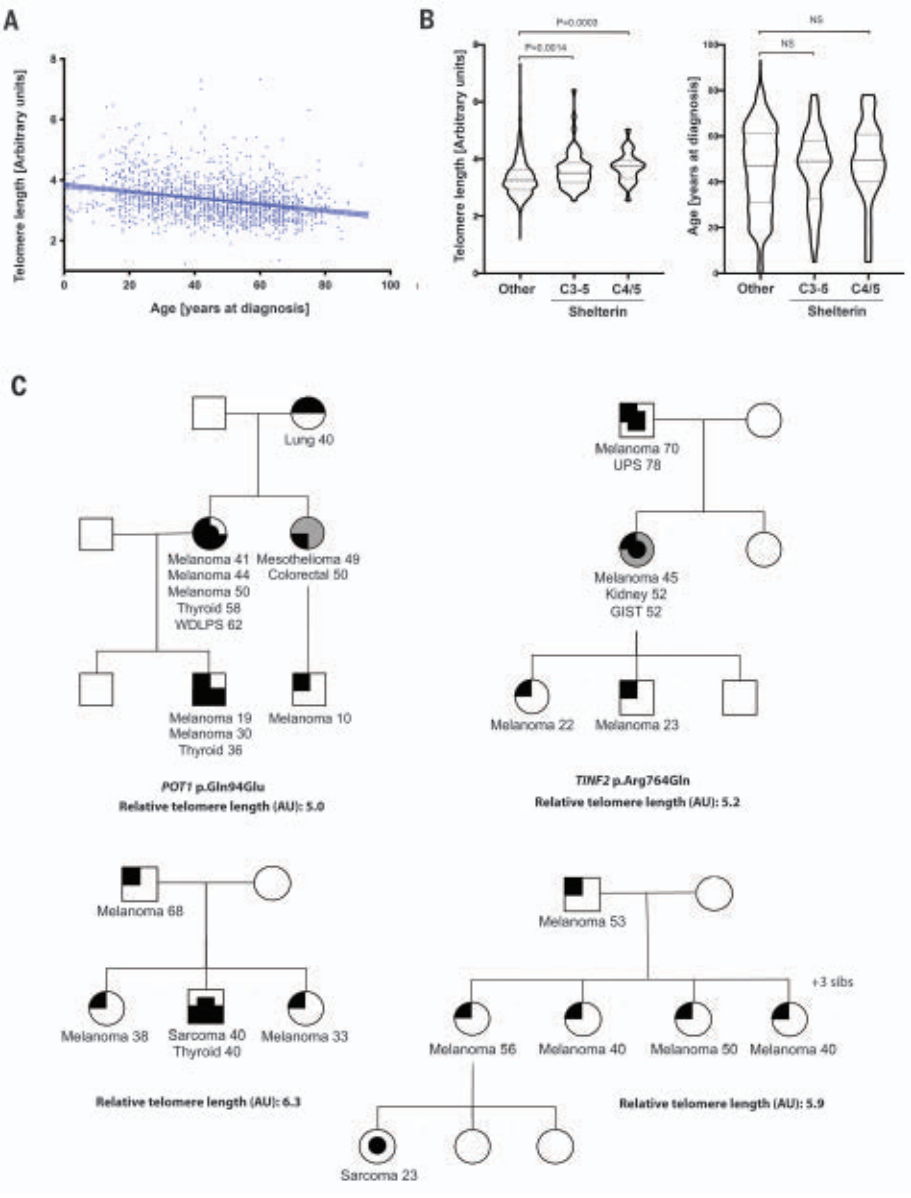


Fig. 4. Leukocyte telomere length in carriers of shelterin-complex variants. (**A**) Relative telomere lengths (RTL) were derived from telomere analysis (TelSeq) of whole-genome sequences on peripheral blood DNA for the sarcoma probands. The *x* axis indicates the age for each individual in years, and the *y* axis indicates RTL in arbitrary units. (**B**) Leukocyte RTL is longer in carriers of variants in the shelterin complex. Left panel: RTL in probands with shelterin complex C3 to C5 ($n = 35$) or C4 or C5 variants ($n = 25$) compared to the remainder of the sarcoma probands. Right panel: Age distribution for these groups. (**C**) Representative pedigrees showing autosomal dominant cancer patterns with an excess of melanoma. Top left panel: family with a *POT1* pathogenic variant; top right panel: family with a *TINF2* pathogenic variant; bottom panels: representative pedigrees of probands with long telomeres and excess of melanoma with no pathogenic variants identified.

Table 1. Clinical and pedigree characteristics of sarcoma probands. HNPCC, Hereditary nonpolyposis colorectal cancer; HBOC, Hereditary breast or ovarian cancer.

Characteristic		N (%)
Gender	Female	855 (52%)
	Male	789 (48%)
	Total	1644
Age at diagnosis, years (mean, SD)	First cancer	45.4, 18.3
	Sarcoma	46.7, 18.7
Number with multiple primary cancers		211
	three primary cancers	63
	≥ four primary cancers	19
	Total	293 (18%)
Pedigree classification	No syndrome	1159
	Classic or Chompret LFS	164
	HNPCC	2
	Familial melanoma	15
	HBOC	27
	Clinically suspicious*	110
	Other**	14
	Uninformative	153
Pathology subtype	Bone	358
	Soft tissue	1286
Genomic class	Complex	836
	Simple	372
	Translocation associated	362
	Unknown	74

*Clinically suspicious: >0.5 cancers or FDR with ≥3 FDR reported and ≥1 FDR <50 years at diagnosis; three or more cancers per proband at any age; 1 or more sarcomas in an FDR; >1 sarcoma or connective tissue tumor at any age per proband; two cancers under 50 years of age per proband. **Neurofibromatosis type 1 (5), retinoblastoma (2), multiple endocrine neoplasia type 1 (2), familial papillary thyroid (1), familial paraganglioma (1), schwannomatosis (1), Gorlin syndrome (1), and McCune Albright syndrome (1).

sarcoma, melanoma, and thyroid cancers is not therapy related. Both melanoma and thyroid cancer are curatively treated by surgery without genotoxic therapy; moreover, relatives of shelterin probands carrying C4 or C5 variants also had an increased incidence of cancer {standardized incidence ratio (SIR) 2.06 [1.50–2.82]; $P = 1.84 \times 10^{-4}$}, comparable to that of *TP53* {SIR 2.59 [1.53–4.37]; $P = 3.69 \times 10^{-4}$}. The risk of melanoma was markedly increased in families carrying C4 or C5 variants in the shelterin complex {SIR 5.60 [3.25–9.65]; $P = 5.61 \times 10^{-10}$}, as was the risk of thyroid cancer {SIR 19.74 [8.22–47.43]; $P = 2.56 \times 10^{-11}$}. An additional individual carried a C4 or C5 variant in *CTC1* and another individual carried a C4 or C5 variant in *NHP2*—, genes that are both implicated in telomere maintenance through the CST and telomerase complexes (*25*, *26*). Taken together, these data suggest that heritable defects in telomere function predispose to a syndrome of sarcoma, melanoma, and thyroid cancers.

We next analyzed relative leukocyte telomere lengths (RLTL) derived from WGS. As expected, RLTL progressively shortened over life span in sarcoma probands (Fig. 4A), unrelated to chemotherapy or radiotherapy exposure (*18*). Despite a similar age at first cancer diagnosis, shelterin variant carriers had longer RLTL than the remainder of the cohort (Fig. 4B). No carriers had clinical or genomic evidence of clonal hemopoiesis, dysplasia, or leukemia (*17*). Independent of genotype, relatives of probands with long telomeres (>2 SD above the mean for age) had an increased risk of melanoma {SIR 3.58 [2.19–5.84]; $P = 3.36 \times 10^{-7}$} and thyroid cancer {SIR 8.43 [3.16–22.45]; $P = 5.96 \times 10^{-9}$}. Representative melanoma-associated pedigree patterns associated with *POT1*, *TINF2*, and *SMARCAL1* variants are shown in Fig. 4C and fig. S6. For two families of probands with long RLTL and cognate tumor patterns in which no causal coding variants were found in shelterin complex genes (Fig. 4C, lower panels), we did not find evidence for causal noncoding variants.

Discussion

Together, these data suggest that heritable defects in telomere and mitotic function increase the risk of sarcoma, in contrast to most epithelial cancers. Telomere maintenance and mitosis are fundamental to chromosome integrity. Sarcomas are predominantly genomically unstable (*27*) and use the alternate mechanism of telomere maintenance (ALT) (*28*). Through replication stress, ALT generates genomic instability that signals to *TP53*, which is strongly linked to sarcoma susceptibility (*29*–*33*). Fifty-two (3.2%) sarcoma probands carried C3 to C5 variants in the shelterin complex and related telomere genes, compared to 13 for *TP53* (0.8%). Pathogenic variation in these genes is rare in the healthy individuals, with the MGRB containing only one C3 to C5 variant in *TP53* (0.03%), and six in shelterin complex genes (0.2%).

Shelterin is a six-subunit protein complex which protects the ends of telomeres (*34*) and is associated with melanoma risk (*35*, *36*). Shelterin variants occurred in 6 out of 15 (40%) sarcoma families with familial melanoma, in contrast to the gene most commonly mutated in familial melanoma (*CDKN2A*), in which no variants were found (*16*). Like *TP53*, germline shelterin pathogenic variants exhibit a dominant pattern of inheritance and cancer risk. Unlike *TP53*, cancer onset occurs at an older age, perhaps the result of an interaction with age-dependent telomere shortening. In addition to shelterin, three genes with roles in telomere maintenance were identified with multiple C4 or C5 variants: *SMARCAL1*, *TIMELESS*, and *STAG3*. *SMARCAL1* and *TIMELESS* regulate telomere stability in ALT cells (*21*, *22*, *37*–*40*), and *STAG3* plays roles in both centromere and telomere–sister chromatid cohesion (*41*, *42*). Shelterin complex pathogenic variant carriers had long leukocyte telomeres, recently associated with increased risk for multiple cancers (*32*). Sarcomas, melanomas, and thyroid cancers may share an association with long-telomere syndrome (*36*). Notably, we did not observe any variants in components of the telomerase complex, the dominant mechanism of telomere maintenance in epithelial cancers.

Overall, 2.3% of sarcoma probands (comprising 10% of MPNSTs and 8.4% of GISTs) carry C3 to C5 variants in centrosome genes. Centrosome variants were mutually exclusive with variants in known MPNST or GIST genes (*NF1*, *LZTR1*, *SDHA*, and *SDHB*), which themselves account for 15% of MPNSTs and 5% of GISTs in our study. Chromosome segregation during mitosis begins with centrosome assembly (*43*). Although somatic centrosome abnormalities are common in cancer (*44*), a germline role for this pathway is limited to *CEP57*. *CEP57* has been associated with mosaic variegated aneuploidy (*45*), a recessive condition associated with chromosomal instability and increased risk of cancer, including sarcomas (*46*). *CEP63* interacts with *CEP72* and *CEP152* to mediate formation of the centrioles, centrosomes, and bipolar spindle

Table 2. Biological pathways enriched in pathogenic variants in the top-ranked 85 sarcoma-specific genes by WRVBT. The top 85 genes by WRVBT and protein interactions with at least one other gene were analyzed in Cytoscape using StringApp. Ontologic groups refer to relevant groups in the secondary gene set (Fig. 1, B and E). Term names are derived from relevant databases (Category). Genes refers to the those found related to the term name, and % term refers to the percentage of total genes in the term represented. The uncorrected *P* values and FDR values are provided.

	Group 1	Group 2	Group 3	Other
Category	GO Component	GO Component	UniProt Keywords	UniProt Keywords
Term name	GO.0070187	GO.0005819	KW-0361	KW-0043
Description	Shelterin complex	Spindle	Hereditary multiple exostoses	Tumor suppressor
Genes	*TINF2 TERF1 POT1*	*CEP63 HAUS4 HAUS5 MAD2L2 TERF1 MYH10 FAM161A HEPACAM RANGAP1 RACGAP1*	*EXT1 EXT2*	*TP53 PTCH1 STARD13 EPB41L3 NF1 RBL1 EXT1 EXT2*
% Term	42.9	3.1	100	4.4
***P* value**	0.0000277	0.0000452	0.00023	0.000024
FDR	0.0016	0.0021	0.0118	0.0029

Table 3. Clinical, molecular, and familial features of probands carrying C4 or C5 variants in the shelterin complex and associated genes. AS, angiosarcoma; CS, chondrosarcoma; FMS, fibromyxosarcoma; GIST, gastrointestinal stromal tumor; LPSWD, Well- or dedifferentiated liposarcoma; LPSM, myxoid liposarcoma; MPNST, malignant peripheral nerve-sheath tumor; OS, osteosarcoma; PNET, primitive neuroectodermal tumor; RMS, rhabdomyosarcoma, not otherwise specified; SS, synovial sarcoma; UPS, undifferentiated pleomorphic sarcoma; LFS, Li Fraumeni Syndrome; NS, no syndrome; U, uninformative.

ID	Gene	Protein change	Gender	Sarcoma	Age at diagnosis	Second malignancies (ages at diagnoses)	Pedigree pattern
017.468.1	*POT1*	p.Gln94Glu	F	LPSWD	62	Thyroid (58); melanoma (41, 44, 50)	Familial melanoma
041.1570.1	*POT1*	p.Tyr501Ter	M	UPS	60	NA	NS
041.1598.1	*POT1*	p.Gln358SerfsTer13	F	AS	50	NA	NS
041.2877.1	*POT1*	p.Tyr89Ter	F	FMS	46	NA	NS
041.2496.1	*POT1*	p.Tyr89Ter	M	MPNST	50	NA	Familial melanoma
017.761.1	*POT1*	p.Ala532HisfsTer13	M	UPS	74	NA	NS
012.3145.1	*SMARCAL1*	p.Arg645ProfsTer46	M	OS	33	NA	NS
041.1055.1	*SMARCAL1*	p.Trp843Ter	M	PNET	22	NA	Chompret
041.3022.1	*SMARCAL1*	p.Lys534Ter	F	UPS	61	NA	NS
012.492.1	*SMARCAL1*	p.Glu848Ter	F	OS	14	NA	NS
045.2520.1	*SMARCAL1*	p.Arg764Gln	M	UPS	52	Thyroid (49)	NS
012.2401.1	*SMARCAL1*	p.Arg764Gln	F	GIST	52	Kidney (52); melanoma (45)	Familial melanoma
013.347.1	*STAG3*	p.Arg1018AspfsTer14	F	OS	52	NHL (50)	NS
L-P001402	*STAG3*	p.Cys141Ter	M	RMS	5	NA	U
013.3129.1	*STAG3*	p.Cys141Ter	F	UPS	49	NA	NS
013.329.1	*TERF1*	p.Arg147Ter	F	UPS	61	Melanoma (64)	Familial melanoma
012.45.1	*TERF1*	p.Arg147Ter	M	CS	74	NA	NS
043.1874.1	*TERF2IP*	p.Arg153GlnfsTer6	F	GIST	43	NA	NS
016.558.1	*TERF2IP*	p.Arg364Ter	F	LPSWD	64	Thyroid (42)	NS
013.297.1	*TIMELESS*	p.Leu867ThrfsTer8	F	CS	78	NA	NS
012.235.1	*TIMELESS*	p.Ala383GlyfsTer19	M	SS	40	NA	Chompret
013.746.1	*TIMELESS*	p.Tyr25Ter	M	CS	58	NA	NS
041.1280.1	*TINF2*	p.Val67TrpfsTer3	F	LPS	66	NA	NS
043.2441.1	*TINF2*	p.Arg256Ter	F	LPSM	37	NA	Chompret
015.1183.1	*TINF2*	p.Arg265Ter	F	UPS	58	NA	Classic LFS

(*47*). *CEP63* and *CEP89* also interact with *HAUS4* and *HAUS5*, augmin-like complex members that regulate centrosome integrity (*48*). Notably, this centrosome group adds to the known genes linked to MPNST and GIST.

The methodologies applied here have broad applications for disease-focused pathway discovery in the emerging era of WGS, allowing mapping of rare SNV/indels and structural variation (SV) in large populations. In our study, SNV/indels (rather than SVs) appear the dominant source of excess pathogenic burden. In contrast to population-based controls, controls of healthy elderly individuals increase statistical power by stripping out pathogenic variation–causing phenotypes (such as cancer) that emerge later in life. Notably, the combination of population-scale and familial cohort designs, historically distinct from each other, allows correlation of genotypes to clinical and familial patterns. The risk-to-relatives analyses

demonstrated an increased cancer incidence in family members of probands with shelterin-complex variants comparable with that seen for carriers of variants in *TP53*. Statistical approaches that combine gene-level pathogenic enrichment with pathway-based approaches based on protein interactions (*49*) and ontologies enhance the power for biological insights into cancer pathogenesis (*50*, *51*). These approaches have identified genes and biologic pathways that appear specific to mesenchymal malignancies, as well as mapping the contributions of known cancer genes. Amongst known and actionable genes, the lack of enrichment in *BRCA1 or BRCA2* is notable, although some sarcoma families meet both HBOC and LFS criteria. We add 14 candidates (including *TERF1*, *TINF2*, *TERF2IP*, *SMARCAL1*, *TIMELESS*, *STAG3*, *CEP63*, *CEP72*, *HAUS4*, and *HAUS5*) to more than 100 known sarcoma-associated genes (*7*). Biologically, these data suggest that the telomeric and mitotic pathways may play specific roles in sarcoma susceptibility, analogous to homologous recombination and mismatch repair in susceptibility to breast and colorectal cancers.

REFERENCES AND NOTES

1. O. G. Bahcall, *Nat. Genet.* **45**, 343 (2013).
2. P. Lichtenstein *et al*., *N. Engl. J. Med.* **343**, 78–85 (2000).
3. A. Paul, S. Paul, *Front. Biosci. (Landmark Ed.)* **19**, 605–618 (2014).
4. C. Wang, Y. Wang, K. S. Hughes, G. Parmigiani, D. Braun, *JNCI Cancer Spectr.* **4**, pkaa027 (2020).
5. World Health Organisation, *WHO Classification of Tumours: Soft Tissue and Bone Tumours*. V. W. Dilani Lokuhetty, Ian Cree, Ed., WHO Classification of Tumours (International Agency for Research on Cancer, Lyon, France., ed. 5, 2020).
6. M. L. Ballinger, M. Pinese, D. M. Thomas, *Genes Chromosomes Cancer* **58**, 130–136 (2019).
7. L. Mirabello *et al*., *JAMA Oncol.* **6**, 724–734 (2020).
8. J. Zhang, K. E. Nichols, J. R. Downing, *N. Engl. J. Med.* **374**, 1391 (2016).
9. C. M. Mery, S. George, M. M. Bertagnolli, C. P. Raut, *Cancer* **115**, 4055–4063 (2009).
10. M. L. Ballinger *et al*., *Lancet Oncol.* **17**, 1261–1271 (2016).
11. S. S. Kalia *et al*., *Genet. Med.* **19**, 249–255 (2017).
12. D. C. Chung, D. A. Haber, Massachusetts General Hospital, *Principles of clinical cancer genetics: a handbook from the Massachusetts General Hospital*. (Springer, New York, 2010), pp. xi, 227 p.
13. D. G. Evans *et al*., *J. Med. Genet.* **46**, 811–817 (2009).
14. H. F. Vasen, P. Watson, J. P. Mecklin, H. T. Lynch, *Gastroenterology* **116**, 1453–1456 (1999).
15. GenoMEL, the Melanoma Genetics Consortium, What is the chance that you have inherited a mutation in CDKN2A? Available from https://genomel.moffitt.org/GenoMELPREDICT. Accessed October 2020.
16. N. J. Taylor *et al*., *J. Am. Acad. Dermatol.* **81**, 386–394 (2019).
17. M. Pinese *et al*., *Nat. Commun.* **11**, 435 (2020).
18. See supplementary materials.
19. J. Andersson *et al*., *Am. J. Surg. Pathol.* **29**, 1170–1176 (2005).
20. P. Priestley *et al*., *Nature* **575**, 210–216 (2019).
21. K. E. Cox, A. Maréchal, R. L. Flynn, *Cell Rep.* **14**, 1032–1040 (2016).
22. M. C. Gadaleta, A. González-Medina, E. Noguchi, *Curr. Genet.* **62**, 725–730 (2016).
23. Y. Wu, R. C. Poulos, R. R. Reddel, *Cancers* **12**, 2739 (2020).
24. C. D. Robles-Espinoza, M del. Castillo Velasco-Herrera, N. K. Hayward, D. J. Adams, *Mol. Cancer Res.* **13**, 211–*222* (2015).
25. C. Huang, P. Jia, M. Chastain, O. Shiva, W. Chai, *Exp. Cell Res.* **355**, 95–104 (2017).
26. T. Vulliamy *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **105**, 8073–8078 (2008).
27. F. Chibon *et al*., *Nat. Med.* **16**, 781–787 (2010).
28. R. L. Dilley, R. A. Greenberg, *Trends Cancer* **1**, 145–156 (2015).
29. E. A. Akbay *et al*., *Oncogene* **32**, 2211–2219 (2013).
30. Y. J. Chen *et al*., *Cancer Res.* **66**, 6473–6476 (2006).
31. M. Lee *et al*., *Nucleic Acids Res.* **46**, 4903–4918 (2018).
32. E. J. McNally, P. J. Luncsford, M. Armanios, *J. Clin. Invest.* **129**, 3474–3481 (2019).
33. A. M. Pinzaru *et al*., *Genes Dev.* **34**, 1619–1636 (2020).
34. M. J. Giraud-Panis, M. T. Teixeira, V. Géli, E. Gilson, *Mol. Cell* **39**, 665–676 (2010).
35. J. Shi *et al*., *Nat. Genet.* **46**, 482–486 (2014).
36. L. G. Aoude *et al*., *J. Natl. Cancer Inst.* **107,** (2014).
37. A. R. Leman *et al*., *Cell Cycle* **11**, 2337–2347 (2012).
38. L. A. Poole *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **112**, 14864–14869 (2015).
39. J. Min, W. E. Wright, J. W. Shay, *Mol. Cell. Biol.* **37**, e00226-17 (2017).
40. E. Feng *et al*., *J. Cell Sci.* **133**, jcs.234914 (2020).
41. B. Liebe, M. Alsheimer, C. Höög, R. Benavente, H. Scherthan, *Mol. Biol. Cell* **15**, 827–837 (2004).
42. T. Winters, F. McNicoll, R. Jessberger, *EMBO J.* **33**, 1256–1270 (2014).
43. P. T. Conduit, A. Wainman, J. W. Raff, *Nat. Rev. Mol. Cell Biol.* **16**, 611–624 (2015).
44. S. A. Godinho, D. Pellman, *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **369**, 20130467 (2014).
45. K. Snape *et al*., *Nat. Genet.* **43**, 527–529 (2011).
46. S. Hanks *et al*., *Nat. Genet.* **36**, 1159–1161 (2004).
47. Z. Wei *et al*., *Mol. Cell. Biol.* **40**, e00535-19 (2020).
48. S. Lawo *et al*., *Curr. Biol.* **19**, 816–826 (2009).
49. R. Oughtred *et al*., *Nucleic Acids Res.* **47** (D1), D529–D541 (2019).
50. D. Szklarczyk *et al*., *Nucleic Acids Res.* **47** (D1), D607–D613 (2019).
51. The Gene Ontology Consortium, *Nucleic Acids Res.* **47**, D330–D338 (2019).

ACKNOWLEDGMENTS

We acknowledge and thank the ISKS families and ISKS investigators. We thank J. Hopper, J. Desai, P. Stalley, T. Corpuz, I. Han, C. Napier, R. Reddel, and the Kinghorn Centre for Clinical Genomics for helpful discussions, advice, and technical assistance. We gratefully acknowledge access to the Medical Genome Reference Bank, the Australian Schizophrenia Research Bank, the Zero Childhood Cancer Program, the Hartwig Foundation dataset, the cardiovascular disease and spontaneous coronary artery dissection cohorts, the French Exome Project (FREX) consortium, and the Amyotrophic Lateral Sclerosis dataset. We thank O. Holmes and S. Wood from Genome Informatics, QIMR Berghofer for their assistance in public data processing. The results shown here are in part based upon data generated by the TCGA Research Network: https://www.cancer.gov/tcga. The NoSarC samples were exome sequenced by the Genomics Facilities at Oslo University Hospital and University of Bergen. **Funding:** Australian National Health and Medical Research Council APP1004017, APP1081477 (D.M.T., M.L.B.), Bioplatforms Australia (D.M.T.), NSW Office of Health and Medical Research Cancer Moonshot Program (D.M.T.), Cancer Australia Priority–driven Collaborative Cancer Research Scheme APP1067094 (D.M.T., M.L.B.), Australian Genomics Health Alliance (D.M.T.), Rainbows for Kate Foundation (D.M.T., M.L.B.), Liddy Shriver Sarcoma Initiative (D.M.T., M.L.B., J.D.S., I.J., A.P.), Johanna Sewell Research Foundation (D.M.T., M.L.B.), Victorian Cancer Agency BFRes11_03 (M.L.B., D.M.T.), CanTeen (D.M.T., M.L.B.), NSW Office of Health and Medical Research (D.M.T.), Sydney Genomics Collaborative Grants Scheme (D.M.T.), Lions Club Kids Cancer Genome Project (D.M.T.), Sarcoma UK (I.J., B.S.), National Health Service funding to the National Institute for Health Research Biomedical Research Centre, UK (I.J.), Norwegian Cancer Society 144385 (O.M.), Norwegian Radium Hospital Legacy, 170502 (O.M.), Trond Mohn Foundation (O.M.), Norwegian Research Council 218241 and 221580 (O.M.), Bergen Research Foundation BFS2017TMT04 (O.M.), Australian National Health and Medical Research Council Principal Research Fellowship APP1104364 (D.M.T.), Cancer Institute NSW Career Development Fellowship CDF171109 (M.L.B.), Cure Cancer Australia Foundation through Cancer Australia Priority-driven Collaborative Cancer Research Scheme (APP1159399) (T.L.N.), Victorian Cancer Council Post-Doctoral Fellowship (CCV1139144) (T.L.N.), Picchi Foundation (T.L.N.), Victorian Comprehensive Cancer Centre (T.L.N.), Victorian Cancer Agency Early Career Research Fellowship ECRF19020 (S.L.), National Health and Medical Research Council of Australia Senior Research Fellowship APP1139071 (N.W.), NetSARC+ INCA and DGOS (J.-Y.B., I.R.-C., N.P., N.I., F.D., J.-E.K.), EURACAN EC 739521 (J.-Y.B. and I.R.-C.) LYRICAN INCA-DGOS-INSERM 12563 (J.-Y.B. and I.R.-C.), LabEx DEvweCAN ANR-10-LABX-0061 (J.-Y.B. and I.R.-C.), Institut Convergence PLASCAN 17-CONV-0002 (J.-Y.B. and I.R.-C.), RHU4 DEPGYN (ANR-18-RHUS-0009 (J.-Y.B. and I.R.-C.), TMRG (INCA and DGOS) (J.-Y.B. and I.R.-C.), InterSARC (INCA and DGOS) (J.-Y.B. and I.R.-C.), Foundation ARC Ligue Contre Le Cancer (J.-Y.B. and I.R.-C.), Cancer Institute NSW Cancer Genetics Training Grant (M.Z.), Hyundai Hope on Wheels (J.D.S.), Soccer for Hope Foundation (J.D.S.), Li-Fraumeni Syndrome Association (J.D.S.), Kneaders Bakery & Café Hope Campaign (J.D.S.), 5 For The Fight (Qualtrics) (J.D.S.), Elephant p53 (EP53) Program funded through Huntsman Cancer Institute by the State of Utah (J.D.S.), Australian National Health and Medical Research Council Senior Research Fellowship APP1121474 (M.J.C.), and University of Newcastle Faculty of Health and Medicine Gladys M Brawn Senior Fellowship (M.J.C.). The FREX consortium was supported by France Génomique National infrastructure, funded as part of "Investissement d'avenir" program managed by Agence Nationale pour la Recherche (contrat ANR-10-INBS-09), https://www.france-genomique.org/. **Author contributions:** Conceptualization: D.M.T. Methodology: D.M.T., M.L.B., S.P., and P.A.M. Software: S.P. and P.A.M. Validation: D.M.T., M.L.B., S.P., P.A.M., P.P., J.B., D.F., R.M.G., E.G., E.C., K.v.E., J.V., T.L., E.G., and J.-F.D. Formal analysis: D.M.T., M.L.B., S.P., P.A.M., E.R., M.Z., K.T., H.P., M.K., S.L., T.L.N., W.K., S.R., J.C., and J.E.P. Resources: I.R.C., N.I., S.C., W.T.A.v.d.G., A.P., F.D., A.L.C., B.S., C.C., J.D.S., A.S.B., P.A.J., J.E.-K., N.P., O.M., L.A.M.-Z., N.W., O.K., J.V.P., A.P.B., M.J.G., J.-H.A., J.E.-K., R.L.R., I.J., R.S., M.H., G.M., M.J.C., J.-Y.B., and J.V. Data Curation: M.L.B., S.P., and P.A.M. Writing – Original Draft: D.M.T., M.L.B., S.P., and P.A.M. Writing – Review and Editing: All authors. Visualization: D.M.T., M.L.B., S.P., and P.A.M. Project administration: D.M.T. and M.L.B. Funding acquisition: D.M.T. **Competing interests:** J.V.P. and N.W. are founders and shareholders of genomiQa Pty Ltd, and members of its board. J.D.S. is employee, stock option holder, and board member of PEEL Therapeutics, Inc. J.V. has sponsored research agreements with Biogen and Astra Zeneca. O.K. has consulted for XING Technologies on development of diagnostic assays for HR deficiency. Other authors declare they have no competing interests. **Data and materials availability:** ISKS, RisC, and ZCCP summary variant frequency data and phenotypic data are available upon application to the respective data access committees at isks@garvan.org.au and agcmp@garvan.org.au and subject to MTA. Summary variant frequency data for the MGRB cohort are available at the web portal: https://sgc.garvan.org.au/explore. Raw genomic data have been deposited at the European Genome-Phenome Archive under study ID EGAS00001003511. Phenotype data are available upon application to the MGRB Data Access Committee at mgrb@garvan.org.au. ASRB data are available upon application at https://www.neura.edu.au/discovery-portal/asrb/. TCGA data are available from the Genomic Data Commons at https://gdc.cancer.gov. NoSarc dataset is available from https://sarcoma.no/en/english-data-access/ upon application and subject to MTA. Hartwig Foundation dataset is available upon application at https://www.hartwigmedicalfoundation.nl/en/. The cardiovascular disease (CVD) and spontaneous coronary artery dissection (SCAD) cohort data will be made available upon reasonable request and in accordance with other publication embargoes. The Wellderly cohort data are available upon application to atorkama@scripps.edu. The FREX consortium data are available upon application at french.exome@gmail.com. The Amyotrophic Lateral Sclerosis dataset is the Dutch cohort of Project MinE. Data access can be requested here: https://www.projectmine.com/research/download-data/. **License information:** Copyright © 2023 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. https://www.science.org/about/science-licenses-journal-article-reuse

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abj4784
Collaborators
Materials and Methods
Text
Tables S1 to S13
Figs. S1 to S11
References (*52*–*113*)
MDAR Reproducibility Checklist
Data Files S1 to S3

Submitted 8 June 2021; resubmitted 25 November 2021
Accepted 16 November 2022
10.1126/science.abj4784

ORGANIC CHEMISTRY

Total synthesis of lissodendoric acid A via stereospecific trapping of a strained cyclic allene

Francesca M. Ippoliti†, Nathan J. Adamson†, Laura G. Wonilowicz, Daniel J. Nasrallah, Evan R. Darzi‡, Joyann S. Donaldson§, Neil K. Garg*

Small rings that contain allenes are unconventional transient compounds that have been known since the 1960s. Despite being discovered around the same time as benzyne and offering a number of synthetically advantageous features, strained cyclic allenes have seen relatively little use in chemical synthesis. We report a concise total synthesis of the manzamine alkaloid lissodendoric acid A, which hinges on the development of a regioselective, diastereoselective, and stereospecific trapping of a fleeting cyclic allene intermediate. This key step swiftly assembles the azadecalin framework of the natural product, allows for a succinct synthetic endgame, and enables a 12-step total synthesis (longest linear sequence; 0.8% overall yield). These studies demonstrate that strained cyclic allenes are versatile building blocks in chemical synthesis.

Geometrically constrained organic molecules have long fascinated chemists. One subclass of such compounds are small cyclic molecules that bear functional groups of typically linear geometry. Embedding typically linear functional groups within small ring systems forces bent geometries, leading to a high degree of strain. Small rings that contain triple bonds were speculated in the early 1900s (*1*), ultimately leading to the validation of benzyne (**1**) (*2*, *3*) and cyclohexyne (**2**) (*4*) in the 1950s (Fig. 1A) (*5*–*9*). Although these in situ–generated intermediates exhibit substantial strain energies of ~40 to 50 kcal/mol (*10*, *11*) and short lifetimes (*12*, *13*), scientists have learned how to leverage these highly reactive intermediates in a range of useful chemical transformations and biological applications. Examples of aromatic and nonaromatic cyclic alkyne trapping experiments can be seen in the synthesis of important ligands, such as XPhos, medicines, agrochemicals, materials, naturally occurring compounds, and other areas (*9*). Moreover, bioorthogonal chemistry (*14*), recently recognized by the Nobel Prize in Chemistry, relies on cyclooctynes (**3**) and other strained intermediates.

Cyclic allenes, such as 1,2-cyclohexadiene (**4**), are a related class of strained molecules that contain a typically linear functional group that is constrained in a small ring (Fig. 1A). Cyclic allene **4** was first validated in 1964 (*15*) and again in 1966 (*16*). This intermediate, along with other cyclic allenes, has received far less attention compared with that of strained cyclic alkynes, despite having a number of advantages that warrant their further investigation. Our laboratory has been interested in heterocyclic allenes, such as azacyclic allene **5**, because of their properties and potential use in chemical synthesis (*9*, *17*, *18*). The geometry-optimized structure of this species is depicted in Fig. 1B. As shown, the allene C=C bonds are each 1.32 Å, and the central allene carbon bears an internal angle of 133°. Additionally, the allene C–H and C–CH_3 bonds are twisted out of the allene C=C=C plane by 39° and 36°, respectively. Several features result from this distinctive structure, including the high reactivity of cyclic allenes due to strain; their capacity to undergo cycloadditions, nucleophilic additions, or metal-catalyzed reactions; and the capability to form two new bonds in a single transformation, with introduction of up to three C(sp^3) stereocenters. As such, strained cyclic allenes have recently been used to prepare highly substituted, sp^3-rich compounds (*9*, *17*–*24*) and have even been used to access DNA-encoded libraries (*25*). Further underscoring their attractiveness is that cyclic allenes are axially chiral and could serve as unconventional yet valuable building blocks for stereoselective synthesis. Moreover, from a practical perspective, strained cyclic allenes can possess heteroatoms within the ring and can be accessed under mild conditions (such as fluoride-based reactions or ambient temperatures) (*19*). Last, because of the relatively scarce body of literature surrounding strained cyclic allenes, there exist numerous opportunities for discovery with regard to reactivity and selectivity patterns in synthetic applications. Cyclic allenes have rarely been used in the synthesis of complex natural products, which is notable given that their close relatives, arynes and strained cyclic alkynes, have been used in more than 100 total syntheses to date (*6*–*8*). One notable example of the use of a cyclic allene in total synthesis was reported in 1997 by Marshall and Sehon and involves a relatively unstrained allene embedded in a macrocyclic 14-membered ring (*26*).

With the aim of developing the chemistry of strained cyclic allenes and evaluating their utility in total synthesis, we considered lissodendoric acid A (**6**) (*27*), a structurally complex member of the manzamine family of alkaloids (*28*) (Fig. 1C). Manzamine alkaloids have been popular targets for total synthesis because of their complex structures and promising biological activities (*29*, *30*). The cumulative synthetic efforts toward manzamine alkaloids, as recently highlighted by Dixon and Jakubec (*29*), have involved more than 25 research groups to date. With regard to lissodendoric acid A (**6**), this natural product was isolated in 2017 from the marine sponge *Lissodendoryx florida* and has been shown to reduce reactive oxygen species (ROS) in a Parkinson's disease model consisting of Neuro 2a cells treated with 6-hydroxydopamine (~50% reduction in ROS levels at concentrations of 0.1 and 10 μM) (*27*). The natural product's complex structure poses a daunting synthetic challenge that in turn would push the limits of strained cyclic allene chemistry. The core of lissodendoric acid A (**6**) is an azadecalin scaffold bearing a conjugated diene, a carboxylic acid substituent, and two stereogenic centers, one of which is quaternary at C8a. In addition, the natural product bears a 14-membered macrocycle tethered beneath the C8a aminodecane substituent. A total synthesis of lissodendoric acid A (**6**) has not been reported.

We questioned whether the azadecalin core and C8a quaternary stereocenter (for example, **7**) could be introduced simultaneously by means of a [4+2] cycloaddition between cyclic allene **8** and 2-pyrone **9** (Fig. 1C) (*31*). The strain associated with cyclic allene **8** would be crucial to the success of the proposed reaction because such intermolecular cycloadditions of analogous cyclic alkene substrates are challenging (*32*, *33*). The use of 2-pyrones in [4+2] cycloadditions with cyclic allenes was previously unknown but, if successful, could enable control of regioselectivity through an inverse electron-demand Diels-Alder process. In addition, the use of a pyrone would furnish a [2.2.2]-bicyclic product (**7**), which in turn would function as a masked 1,3-diene, enabling the strategic unveiling of this potentially sensitive motif at a later stage through thermal expulsion of CO_2. Last, we were intrigued by the possibility of accessing cyclic allene **8** in enantioenriched form and developing enantiospecific trapping reactions.

Our successful total synthesis of lissodendoric acid A (**6**) was enabled by the development of a regioselective, diastereoselective, and stereospecific trapping reaction of a fleeting cyclic allene intermediate. The total synthesis is concise owing to generation of

Department of Chemistry and Biochemistry, University of California, Los Angeles, CA 90095, USA.
*Corresponding author. Email: neilgarg@chem.ucla.edu
†These authors contributed equally to this work.
‡Present address: ElectraTect, Phoenix, AZ 85004, USA.
§Present address: Pfizer Oncology Medicinal Chemistry, San Diego, CA 92121, USA.

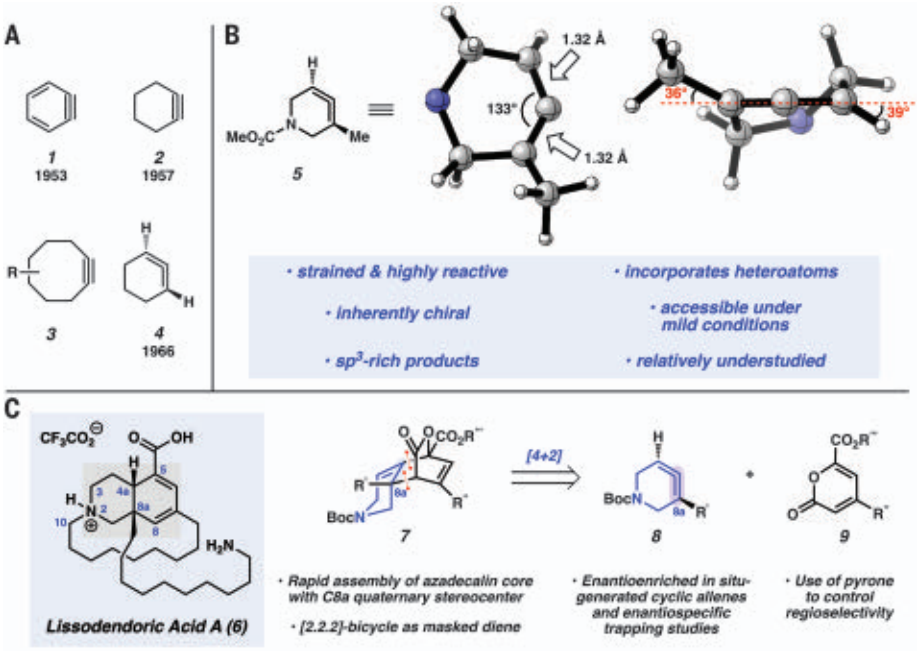


Fig. 1. Strained intermediates and overview of current study. (**A**) Seminal strained cyclic intermediates **1**, **2**, and **4** and bioorthogonal cyclooctyne reagents **3**. (**B**) Geometry-optimized structure of strained azacyclic allene **5** [ωB97XD/6-31G(d)] and key features. CO_2Me is omitted from the three-dimensional representations for clarity. (**C**) Our synthetic target, lissodendoric acid A (**6**), and our synthetic approach by using a [4+2] cycloaddition to access **7** from cyclic allene **8** and pyrone **9** (R′, R″, R‴ = alkyl).

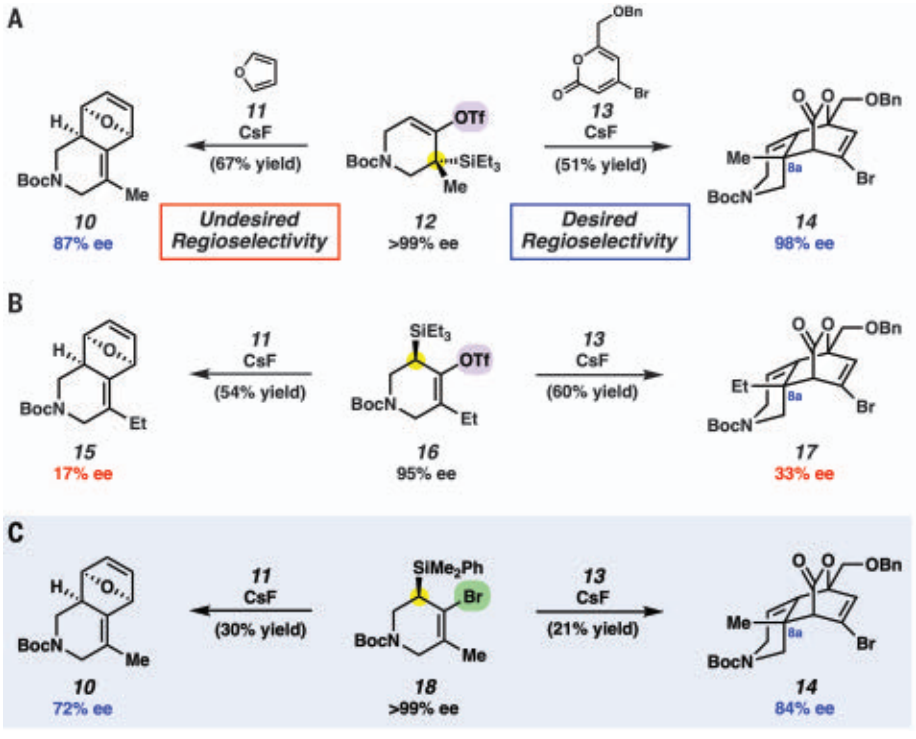


Fig. 2. Regioselectivity and stereospecificity studies using variable cyclic allene precursors and [4+2] cycloaddition partners. (**A**) Use of cyclic allene precursor **12** allows for efficient and stereospecific cycloadditions. (**B**) Stereospecificity is modest when using pseudo-isomeric silyltriflate **16**. (**C**) Switching to silylbromide **18** as the cyclic allene precursor allows for regioselective and stereospecific cycloaddition reactions, while providing a plausible entryway to enantioenriched material. Highlights indicate the primary differences between cyclic allene precursors **12**, **16**, and **18**. Reaction conditions are provided in the supplementary materials.

substantial structural complexity in the key cyclic allene trapping step.

Development of regio- and stereoselective cycloadditions

Because the success of our approach would hinge on the ability to control the regioselectivity and stereoselectivity in the key [4+2] cycloaddition step, we performed preliminary studies with furan **11** and pyrone **13** (Fig. 2). These dienes were selected so that we could compare and contrast the electronic effects of the cyclic allene trapping partners (electron-rich furan versus electron-poor pyrones). Pyrone **13** was chosen because it was easily accessible and possesses potentially useful synthetic handles, whereas furan (**11**) had previously been used in cyclic allene [4+2] cycloadditions (*17*). The syntheses of **12**, **16**, and **18**, which function as precursors to cyclic allene **8** (Fig. 1C), are detailed in the supplementary materials and relied on chiral separation technology to access enantioenriched material. Treatment of silyltriflate **12** with furan (**11**) and CsF in acetonitrile at 23°C delivered the undesired regioisomer **10**, which is consistent with prior experiments that used the N-Cbz derivative of **12** (*17*). In a crucial result, we found that treatment of **12** with readily available pyrone **13** (*34*) gave cycloadduct **14**, which possesses the desired connectivity and the necessary C8a quaternary center, without substantial loss of stereochemical integrity. The reaction is endo-selective, which is consistent with theoretical studies of 1,2-cyclohexadienes (*35*), and occurs with the expected regioselectivity based on this being an inverse electron-demand process. These results demonstrate that regioselectivity in cyclic allene [4+2] cycloadditions can be modulated by judicious selection of the trapping agent.

Although these results were encouraging, we were unable to access compounds such as **12** in high enantiomeric excess (ee) without the use of chiral separation technologies. As such, we evaluated alternative precursors **16** and **18**, in which the silicon substituent would be placed on a less-substituted carbon, in the corresponding cycloaddition reactions (*36*). As shown in Fig. 2B, the use of pseudo-isomeric silyltriflate **16** (albeit with Et in place of Me owing to substrate synthesis logistics), furnished the expected cycloadducts **15** and **17**. However, regardless of the trapping agent, poor stereoretention was observed for reasons that are not well understood presently. Last, we examined silylbromide **18**, which is analogous to an approach to cyclic allene generation developed by West and coworkers (Fig. 2C) (*22*). Although yields were modest in these initial studies, notable stereoretention was observed in the formation of cycloadducts **10** and **14** (*37*). Silylbromide precursors to cyclic allenes had not been synthesized previously in

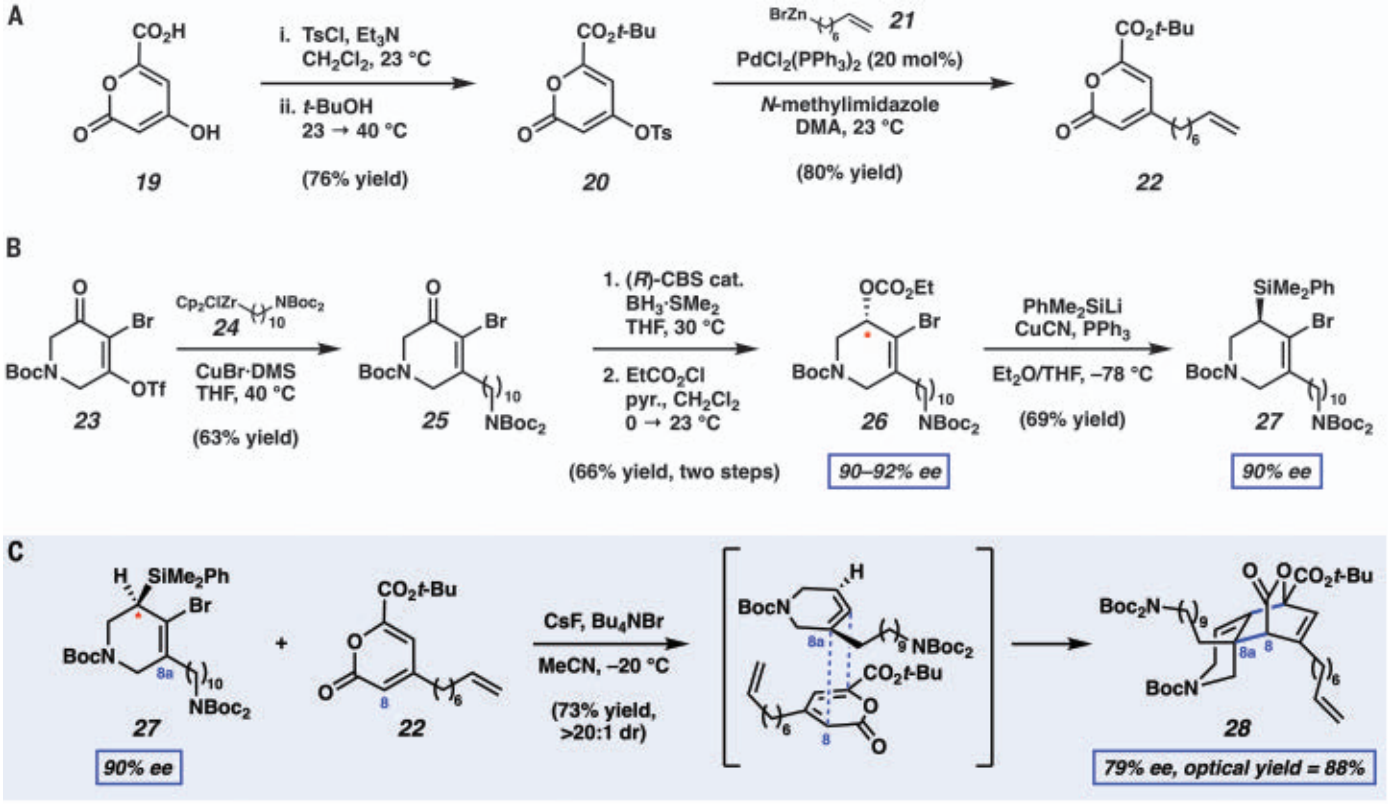


Fig. 3. Assembly of the azadecalin core of lissodendoric acid A (6) by using a strained cyclic allene. (**A**) Synthesis of pyrone **22**. (**B**) Enantioselective route to silyltriflate **27**. (**C**) Cycloaddition of the strained cyclic allene derived from **27** with pyrone **22** proceeds with regioselectivity, diastereoselectivity, and stereospecificity to deliver enantioenriched adduct **28** with the desired quaternary stereocenter at C8a.

enantioenriched fashion, but we were optimistic that derivatives of **18** could be prepared without the need for chiral separations. Before describing the translation of these studies to the total synthesis of lissodendoric acid A (**6**), we also highlight several key features of cyclic allene chemistry that are reflected in the results shown in Fig. 2. Specifically, all reactions proceed under mild conditions, lead to the formation of two C–C bonds, occur with high endo diastereoselectivity (*35*), and provide access to complex heterocyclic products containing three stereocenters, one of which is quaternary in the cases of the pyrone cycloadducts.

Fragment synthesis and assembly of azadecalin core

Having established the feasibility of the key cyclic allene trapping step, we sought to perform analogous studies on a more elaborate substrate. Pursuing this endeavor could allow us to push the limits of strained cyclic allene chemistry while enabling the synthesis of lissodendoric acid A (**6**). Pyrone **22** was prepared in two steps from commercially available carboxylic acid **19** (Fig. 3A). Double tosylation, followed by the addition of *t*-BuOH, furnished tosylate **20**, bearing a *t*-butyl ester substituent. Subsequent Negishi coupling with organozinc reagent **21** provided the alkylated pyrone **22**, which contains a terminal alkene necessary for the eventual installation of the macrocycle by using ring-closing metathesis. The desired cyclic allene precursor, silylbromide **27**, was prepared from bromotriflate **23**, which was obtained commercially (Fig. 3B). Treatment of **23** with alkylzirconium reagent **24** gave ketone **25** through 1,4-addition and ejection of the triflate leaving group (*38*). Corey-Bakshi-Shibata (CBS) reduction (*39*), followed by treatment with ethylchloroformate, gave **26** in 90 to 92% ee. Displacement of the carbonate with a silyl cuprate nucleophile (*22*, *40*, *41*) delivered cyclic allene precursor **27** without loss in enantioenrichment. This step was thought to proceed with inversion of stereochemistry on the basis of prior studies by Oestreich (*41*). The conversion of bromoketone **25** to silylbromide **27** parallels a general racemic strategy pioneered by Lofstrand *et al.* (*22*), but the enantioselective reduction and stereospecific displacement of the carbonate had not been demonstrated previously.

With pyrone **22** and cyclic allene precursor **27** in hand, we evaluated the key [4+2] cycloaddition (Fig. 3C). We found that by simply treating these reactants with CsF, the desired cycloadduct, **28**, could be obtained. Through optimization, we found that the optical yield (*42*, *43*), or percent stereoretention, was greatly improved at lower reaction temperatures, which in turn required the use of Bu_4NBr to improve the solubility of fluoride ion in solution. Ultimately, –20°C was found to be the optimal temperature, enabling the formation of **28** in 73% isolated yield, presumably by means of the depicted transition structure proposed on the basis of prior computational studies of cyclic allene Diels-Alder reactions (*17*, *35*, *44*).

Several aspects of this complexity-generating step should be noted, especially the regioselectivity and stereoselectivity. The regioselectivity is thought to depend on both the cyclic allene and pyrone components. First considering the cyclic allene component, either olefin of the allene could potentially undergo trapping; however, the more electron-rich substituted olefin of the cyclic allene engages with the electron-deficient pyrone, presumably because of the inverse electron-demand nature of this cycloaddition. With regard to the pyrone component, the diene in pyrone **22** aligns with the cyclic allene to promote bond formation between C8a and C8. The pyrone carbonyl, which is conjugated to the diene in **22**, is thought to provide a dominant electronic effect

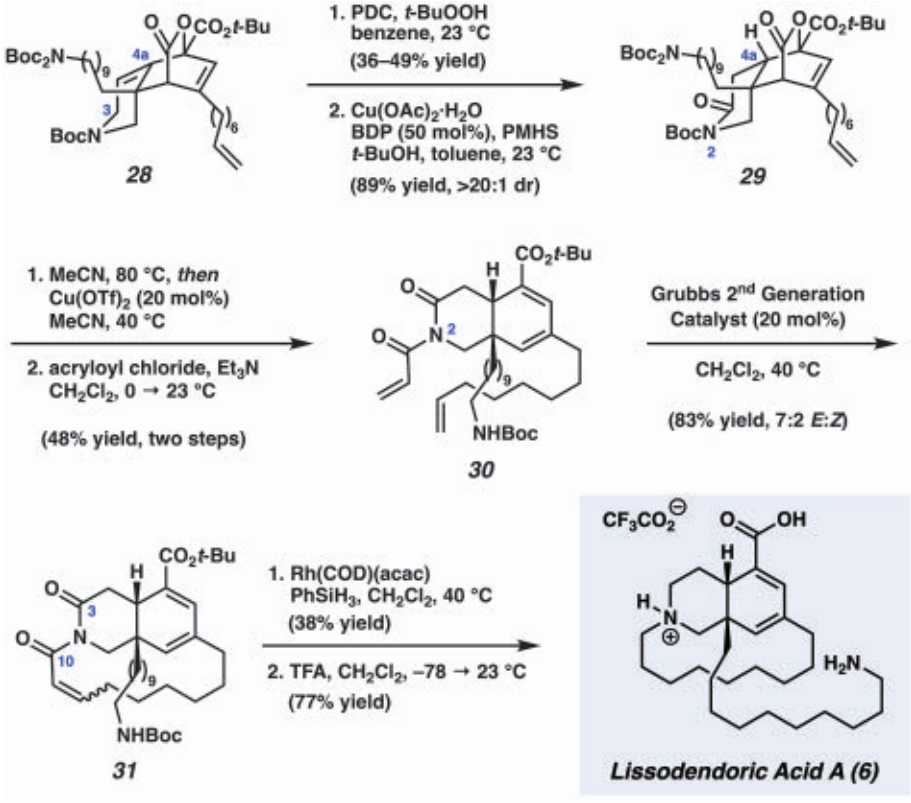


Fig. 4. Completion of the total synthesis of lissodendoric acid A (6).

that leads to this selectivity. The origins of stereoselectivity in this transformation are also notable. The reaction is thought to proceed in an endo fashion, with favorable orbital overlap between the diene and the nonreactive olefin of the cyclic allene, and with favorable approach of the pyrone opposite the C8a alkyl substituent. This gives rise to **28** in >20:1 diastereomeric ratio. The endo approach and facial selectivity are consistent with prior computational studies of cyclic allene Diels-Alder reactions (*17*, *35*). Despite that four cycloadducts could be envisioned (not accounting for enantiomers discussed below), cycloadduct **28**, which bears considerable structural complexity and functionality, is formed selectively as the sole observable product of the reaction.

The absolute stereochemistry and optical yield also warrant particular attention. Cyclic allene precursor **27** bears a single stereocenter (Fig. 3B, asterisk). Upon generation of the cyclic allene intermediate, point chirality in **27** is transmitted to axial chirality in the cyclic allene intermediate. Then, cycloaddition reintroduces point chirality, with the generation of two tertiary stereocenters and the C8a quaternary stereocenter. The sole stereocenter present in reactant **27** is ablated in this process. The optical yield (*42*, *43*), or stereoretention, in forming **28** at –20°C is 88%, indicating that strain-driven cycloaddition between two highly substituted reactants is favorable over facile enantiomerization of the strained cyclic allene at low temperature. The barrier for enantiomerization for related cyclic allenes has been estimated to be ~14 kcal/mol (*45*).

Completion of total synthesis

To complete the total synthesis of lissodendoric acid A (**6**), we developed the concise late-stage sequence shown in Fig. 4 using (±)-**28**. The first challenge was to introduce the C4a hydrogen substituent chemoselectively, without reduction of the other olefins. The C4/C4a olefin proved recalcitrant toward a range of reduction conditions; however, we ultimately achieved a selective reduction by first oxidizing C3 with pyridinium dichromate (PDC)/*t*-BuOOH (*46*, *47*) and then treating the resulting α,β-unsaturated lactam intermediate with a copper hydride source (*48*) to effect diastereoselective 1,4-reduction and furnish cis-azadecalin **29**. The CO_2 bridge embedded in the [2.2.2]-bicycle of **28** (installed in the aforementioned cyclic allene trapping) was important for the success of this sequence because it masked the 1,3-diene motif that would otherwise be in conjugation with the C4/C4a olefin. Next, we arrived at tetraene **30** through a sequence involving thermal extrusion of CO_2 (*49*), copper triflate–promoted removal of two of the three Boc groups (*50*), and amidation of the piperidine nitrogen with acryloyl chloride. Access to **30** set the stage for macrocyclization, which we achieved using ring-closing metathesis promoted by the robust Grubbs' 2nd Generation catalyst (Sigma-Aldrich) (*51*, *52*). The desired macrocycle **31** was obtained in 83% yield and as an inconsequential mixture of *E*:*Z* olefin isomers (ratio ranging from 2:1 to 3.5:1). All that remained to complete the total synthesis was to selectively reduce the α,β-unsaturated imide and perform a global deprotection. We achieved the former using a reduction protocol reported by Das *et al*. (*53*), which led to saturation of the C3 amide and the C10 α,β-unsaturated amide selectively without reduction of the ester or diene functional groups. Although the overall yield of this transformation is modest (38%), given that three functional groups are being reduced (~72% yield per reduction), we deemed the yield acceptable. Last, treatment of the reduced intermediate with trifluoroacetic acid furnished lissodendoric acid A (**6**). The total synthesis is 12 steps from **23** (longest linear sequence; 0.8% overall yield).

Another matter of interest is the absolute stereochemical configuration of lissodendoric acid A. On the basis of experimental and computational circular dichroism studies, the isolation chemists had postulated the absolute configuration of (–)-**6** to be as we have depicted throughout this study (*27*). To validate this hypothesis, our enantioenriched material was elaborated to the natural product. The circular dichroism spectrum for our synthetic sample matched that of a natural sample. We conclude that the natural configuration of (–)-**6** proposed by the isolation chemists was indeed correct.

These studies demonstrate that cyclic allenes are powerful tools for complex molecule synthesis and should prompt the further investigation and application of these unusual compounds.

REFERENCES AND NOTES

1. R. Stoermer, B. Kahlert, *Ber. Dtsch. Chem. Ges.* **35**, 1633–1640 (1902).
2. J. D. Roberts, H. E. Simmons Jr., L. A. Carlsmith, C. W. Vaughan, *J. Am. Chem. Soc.* **75**, 3290–3291 (1953).
3. G. Wittig, L. Pohmer, *Angew. Chem.* **67**, 348 (1955).
4. F. Scardiglia, J. D. Roberts, *Tetrahedron* **1**, 343–344 (1957).
5. H. H. Wenk, M. Winkler, W. Sander, *Angew. Chem. Int. Ed.* **42**, 502–528 (2003).
6. P. M. Tadross, B. M. Stoltz, *Chem. Rev.* **112**, 3550–3577 (2012).
7. C. M. Gampe, E. M. Carreira, *Angew. Chem. Int. Ed.* **51**, 3766–3778 (2012).
8. H. Takikawa, A. Nishii, T. Sakai, K. Suzuki, *Chem. Soc. Rev.* **47**, 8030–8056 (2018).
9. S. M. Anthony, L. G. Wonilowicz, M. S. McVeigh, N. K. Garg, *JACS Au* **1**, 897–912 (2021).
10. J. F. Liebman, A. Greenberg, *Chem. Rev.* **76**, 311–365 (1976).
11. R. O. Angus Jr., M. W. Schmidt, R. P. Johnson, *J. Am. Chem. Soc.* **107**, 532–537 (1985).
12. R. S. Berry, J. Clardy, M. E. Schafer, *J. Am. Chem. Soc.* **86**, 2738–2739 (1964).
13. E. W.-G. Diau, J. Casanova, J. D. Roberts, A. H. Zewail, *Proc. Natl. Acad. Sci. U.S.A.* **97**, 1376–1379 (2000).

14. E. M. Sletten, C. R. Bertozzi, *Angew. Chem. Int. Ed.* **48**, 6974–6998 (2009).
15. M. W. Moser, The reactions of gem-dihalocyclopropanes with organometallic reagents. thesis, Massachusetts Institute of Technology (1964).
16. G. Wittig, P. Fritze, *Angew. Chem. Int. Ed.* **5**, 846 (1966).
17. J. S. Barber *et al*., *Nat. Chem.* **10**, 953–960 (2018).
18. M. M. Yamano *et al*., *Angew. Chem. Int. Ed.* **58**, 5653–5657 (2019).
19. I. Quintana, D. Peña, D. Pérez, E. Guitián, *Eur. J. Org. Chem.* **2009**, 5519–5524 (2009).
20. V. A. Lofstrand, F. G. West, *Chem. Eur. J.* **22**, 10763–10767 (2016).
21. J. S. Barber *et al*., *J. Am. Chem. Soc.* **138**, 2512–2515 (2016).
22. V. A. Lofstrand, K. C. McIntosh, Y. A. Almehmadi, F. G. West, *Org. Lett.* **21**, 6231–6234 (2019).
23. M. M. Yamano *et al*., *Nature* **586**, 242–247 (2020).
24. A. V. Kelleghan, D. C. Witkowski, M. S. McVeigh, N. K. Garg, *J. Am. Chem. Soc.* **143**, 9338–9342 (2021).
25. M. V. Westphal *et al*., *J. Am. Chem. Soc.* **142**, 7776–7782 (2020).
26. J. A. Marshall, C. A. Sehon, *J. Org. Chem.* **62**, 4313–4320 (1997).
27. E. G. Lyakhova *et al*., *Org. Lett.* **19**, 5320–5323 (2017).
28. M. Radwan, A. Hanora, S. Khalifa, S. H. Abou-El-Ela, *Cell Cycle* **11**, 1765–1772 (2012).
29. D. J. Dixon, P. Jakubec, Total synthesis of manzamine alkaloids, in *Marine Natural Products*, H. Kiyota, Ed. (Springer, 2021), pp. 23–59.
30. P. Ashok, S. Ganguly, S. Murugesan, *Drug Discov. Today* **19**, 1781–1791 (2014).
31. H. Hopf, M. S. Sherburn, *Synthesis* **54**, 864–886 (2002).
32. Control experiments are available in the supplementary materials.
33. K. Afarinkia, V. Vinader, T. D. Nelson, G. H. Posner, *Tetrahedron* **48**, 9111–9171 (1992).
34. T. Luo, M. Dai, S.-L. Zheng, S. L. Schreiber, *Org. Lett.* **13**, 2834–2836 (2011).
35. M. Ramirez, D. Svatunek, F. Liu, N. K. Garg, K. N. Houk, *Angew. Chem. Int. Ed.* **60**, 14989–14997 (2021).
36. Strategically, an enantioenriched cyclic allene precursor bearing a tri-substituted stereocenter was expected to be more easily accessible as compared with a precursor with a fully substituted stereocenter.
37. The relationship between stereoretention and the structure of the cyclic allene precursor (identity and location of the silyl group and the nature of the leaving group) is not well understood presently and will be the subject of further investigation.
38. P. Wipf, J. H. Smitrovich, *J. Org. Chem.* **56**, 6494–6496 (1991).
39. E. J. Corey, C. J. Helal, *Angew. Chem. Int. Ed.* **37**, 1986–2012 (1998).
40. H. Ito, Y. Horita, M. Sawamura, *Adv. Synth. Catal.* **354**, 813–817 (2012).
41. A. Weickgenannt, M. Oestreich, *Chem. Eur. J.* **16**, 402–412 (2010).
42. Although the term "stereoretention" is commonly used by synthetic chemists, "optical yield" is the preferred International Union of Pure and Applied Chemistry (IUPAC) term when comparing the ratio of the optical purity of the product with that of the reactant.
43. IUPAC, *Compendium of Chemical Terminology*, A. D. McNaught, A. Wilkinson, Eds. (Blackwell Scientific Publications, ed. 2, 1997); https://doi.org/10.1351/goldbook.
44. The use of five or more equivalents of pyrone **22** led to optimal results. However, using fewer equivalents of pyrone **22** was also synthetically useful. For example, when using two equivalents of pyrone **22**, cycloadduct **28** was obtained in slightly lower isolated yield and optical yield of 62 and 85%, respectively.
45. K. J. Daoust *et al*., *J. Org. Chem.* **71**, 5708–5714 (2006).
46. This oxidation step proceeds by means of a peroxide intermediate, which can also be isolated and recycled (supplementary materials).
47. N. Chidambaram, S. Chandrasekaran, *J. Org. Chem.* **52**, 5048–5051 (1987).
48. B. A. Baker, Z. V. Bosković, B. H. Lipshutz, *Org. Lett.* **10**, 289–292 (2008).
49. M. H. Abdullahi, L. M. Thompson, M. J. Bearpark, V. Vinader, K. Afarinkia, *Tetrahedron* **72**, 6021–6024 (2016).
50. V. Evans, M. F. Mahon, R. L. Webster, *Tetrahedron* **70**, 7593–7597 (2014).
51. A. K. Chatterjee, J. P. Morgan, M. Scholl, R. H. Grubbs, *J. Am. Chem. Soc.* **122**, 3783–3784 (2000).
52. A. K. Chatterjee, T.-L. Choi, D. P. Sanders, R. H. Grubbs, *J. Am. Chem. Soc.* **125**, 11360–11370 (2003).
53. S. Das, Y. Li, L.-Q. Lu, K. Junge, M. Beller, *Chem. Eur. J.* **22**, 7050–7053 (2016).
54. F. M. Ippoliti *et al*., Data from: Total synthesis of lissodendoric acid A via stereospecific trapping of a strained cyclic allene. Dryad (2023); doi:10.5068/D1Q38N.

ACKNOWLEDGMENTS

We thank E. Lyakhova (G. B. Elyakov Pacific Institute of Bioorganic Chemistry, Russia) for providing an authentic sample of (–)-**6**. We thank A. Doyle (UCLA) for use of her laboratory's chiral HPLC. **Funding:** The authors are grateful to the NIH-NIGMS (R01 GM123299 and R35 GM139593 to N.K.G., F32 GM139289 to N.J.A., and F32-GM122245 to E.R.D.), the NSF (DGE-2034835 to F.M.I. and L.G.W.), Bristol-Myers Squibb (to J.S.D.), the UCLA Graduate Division (to J.S.D.), and the Trueblood family (to N.K.G.). These studies were supported by shared instrumentation grants from the NSF (CHE-1048804) and the NIH NCRR (S10RR025631). **Author contributions:** F.M.I., N.J.A., L.G.W., D.J.N., E.R.D., and J.S.D. designed and performed experiments and analyzed experimental data. N.K.G. directed the investigations and prepared the manuscript with contributions from all authors; all authors contributed to discussions. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** Experimental procedures and characterization data are provided in the supplementary materials and (*54*). Correspondence and requests for materials should be addressed to N.K.G. (neilgarg@chem.ucla.edu).

License information: Copyright © 2023 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. https://www.science.org/about/science-licenses-journal-article-reuse

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.ade0032
Materials and Methods
Figures S1 to S72
References (*55*–*59*)

Submitted 18 July 2022; accepted 18 November 2022
10.1126/science.ade0032

LIGHT POLLUTION

Citizen scientists report global rapid reductions in the visibility of stars from 2011 to 2022

Christopher C. M. Kyba,[1,2]* Yiğit Öner Altıntaş,[1]† Constance E. Walker,[3] Mark Newhouse[4]

The artificial glow of the night sky is a form of light pollution; its global change over time is not well known. Developments in lighting technology complicate any measurement because of changes in lighting practice and emission spectra. We investigated the change in global sky brightness from 2011 to 2022 using 51,351 citizen scientist observations of naked-eye stellar visibility. The number of visible stars decreased by an amount that can be explained by an increase in sky brightness of 7 to 10% per year in the human visible band. This increase is faster than emissions changes indicated by satellite observations. We ascribe this difference to spectral changes in light emission and to the average angle of light emissions.

Over much of Earth's land surface, the night sky no longer fully transitions to starlight and moonlight after sunset (*1*). Instead, the sky also glows with an artificial twilight caused by the scatter of anthropogenic light in the atmosphere (*2*). The radiance of skyglow grew exponentially for much of the 20th century (*3*) as a result of population growth, expansion of settlements, and deployment of new lighting technologies (*4*, *5*). The character of the night sky is now different from what it was when life evolved and civilization developed.

Many of the behavioral and physiological processes of life on Earth are connected to daily and seasonal cycles. For example, visual predation requires sufficient light to see, and predator–prey interactions are therefore expected to be affected by skyglow (*6*). There are few controlled field studies of the ecological impacts of skyglow, but it has been have shown to affect plants, animals, and their interactions (*7*), and laboratory studies have demonstrated changes in the physiology of fish at skyglow-like nighttime illuminance of 0.01 lux (*8*). In addition to its environmental consequences, skyglow limits human observation of starry skies and the Milky Way. The increase in skyglow has affected human culture (*9*), not only by restricting stargazing and astronomy but also by changing the overall appearance of the night sky.

Effective methods for reducing light pollution are well understood (*10*, *11*), and many of them also reduce electricity consumption. These measures have been implemented on local scales (*12*, *13*) but have not seen widespread adoption. Nevertheless, awareness of light pollution has led some policy-makers to

[1]Remote Sensing and Geoinformatics, Deutsches GeoForschungsZentrum Potsdam, 14473 Potsdam, Germany. [2]Geographisches Institut, Ruhr-Universität Bochum, 44780 Bochum, Germany. [3]Office of Observatory Site Protection, National Optical-Infrared Astronomy Research Laboratory, Tucson, AZ 85719, USA. [4]Communications, Education and Engagement, National Optical-Infrared Astronomy Research Laboratory, Tucson, AZ 85719, USA.
***Corresponding author. Email: kyba@gfz-potsdam.de**
†Present address: European Organization for the Exploitation of Meteorological Satellites, 64295 Darmstadt, Germany.

introduce measures that attempt to control light pollution (*14*).

During the 2010s, many outdoor lights were replaced by light-emitting diodes (LEDs): Global LED market share for new general lighting grew from under 1% in 2011 to 47% in 2019, and LED market share for new outdoor lighting in the United States was 66% in 2020 (*15*). The impact on skyglow from this transition to LEDs is unclear. Some researchers have predicted that it will be beneficial (*16*); others, that it could be harmful because of spectral changes (*17*) or a rebound effect (*18*), in which the high luminous efficacy (more light emitted for a given power) of LEDs leads to more or brighter lights being installed or longer hours of operation.

The generation of skyglow and changes in its character are related to social, economic, and technological processes, and we therefore expect skyglow trends to differ within and across countries. Time-series measurements of skyglow from individual sites, although useful for some purposes, might not be representative of how skyglow is changing on larger scales. It would therefore be beneficial to measure changes in skyglow on continental and global scales.

In principle, it is possible to directly measure skyglow through satellite observations of Earth at night (*19*). Unfortunately, the only satellite instruments that currently monitor the whole Earth have limited resolution and sensitivity and cannot detect light with wavelengths below 500 nm (*20*). This is a problem for three reasons: (i) Shorter wavelengths scatter more effectively in the atmosphere, increasing the chance that a photon emitted upward returns to Earth as skyglow (*2*); (ii) LEDs marketed as white usually have an emission peak between 400 and 500 nm, where the satellite sensor is insensitive; and (iii) human visual sensitivity shifts toward shorter wavelengths at night (*11*). The first two effects could mean that changes in ground radiance observed by satellite (*21*, *22*) differ from changes in skyglow. The third effect means that ground-based radiometers (and photometers) face a similar problem: If skyglow darkens at longer wavelengths but brightens at shorter wavelengths, it would be unclear whether the number of stars visible to humans would increase or decrease (*23*, *24*).

We analyzed a citizen science dataset in which the human visual system was directly used as a sensor (*24*, *25*). In the Globe at Night project [operated by the National Optical-Infrared Astronomy Research Laboratory (NOIRLab)], participants are presented with a set of star maps (example shown in fig. S1) and asked which one best matches the night sky at their location (*26*). This provides an estimate of the naked eye limiting magnitude (NELM), the visual apparent magnitude of the faintest star that can be seen. Astronomical magnitudes are an inverted logarithmic scale, so the NELM is smaller for brighter skies. The NELM is related to skyglow because as background radiance increases, faint point sources of light become invisible (*27*). The limiting magnitude estimated by citizen scientists using this method correlates with the locations of skyglow determined using satellite datasets (*24*, *25*). We grouped different regions of the globe according to their sky brightness in 2014, as determined by the (satellite-based) World Atlas (*1*), to examine how the NELMs in similarly bright areas change over time (Fig. 1 and fig. S2).

Our method (*26*) accounts for differences in the set of citizen scientists participating each year (Fig. 2) and allows us to measure changes in stellar visibility on global or continental spatial scales. The overall number of observations limits the spatial and temporal scales over which trends can be determined (particularly for developing countries, where rapid skyglow change is suspected but little observational evidence is available). The Globe at Night data have a spatial bias toward Europe, North America (especially the United States), and a small number of other countries; there is also a bias toward inhabited areas. For example, 50.6% of the Asian contributions are from Japan, and contributions from Australia are overwhelmingly from coastal areas (Fig. 2B and fig. S3) during 2020 (Fig. 2A). Because Europe and North America have sufficient data in both time and space, we report trends for those continents and combine all the others (referred to as Rest of World hereafter).

Although this dataset does not represent an average of either the land area or human population distribution, participants are concentrated in regions where skyglow is most prevalent (*1*). The global trend in skyglow that we measure likely underestimates the trend in countries with the most rapid increases in economic development, because the rate of change in light emission is highest there (*21*) and we expect the addition of new lights to have a greater impact on skyglow than would result from the replacement of existing lights.

To convert the NELM measurements to rates of change in effective skyglow radiance for human vision, we fitted a model to a subset of the dataset [observations from January 2011 to August 2022, without twilight, moonlight, or reported snow on the ground (*26*)], using a maximum likelihood method. There are five free parameters in the model: (i) the average limiting magnitude reported in regions with no light pollution (*y* intercept in Fig. 1); (ii) the slope of the relationship between NELM and World Atlas skyglow radiance (similar to the slope in Fig. 1, but with a time-adjusted radiance); (iii) the annual rate of change in artificial skyglow; (iv) the standard deviation in the residuals between measured and predicted NELM; and (v) an estimate of the error rate in the Globe at Night dataset (e.g., due to participants reporting their location or NELM incorrectly).

We find that the change in the number of visible stars reported by Globe at Night participants is equivalent to a 9.6% per year annual increase in sky brightness, averaged over the locations of participants (Table 1). For an 18-year period (such as the duration of a human childhood), this rate of change would increase sky brightness by more than a factor of 4. A location with 250 visible stars would see

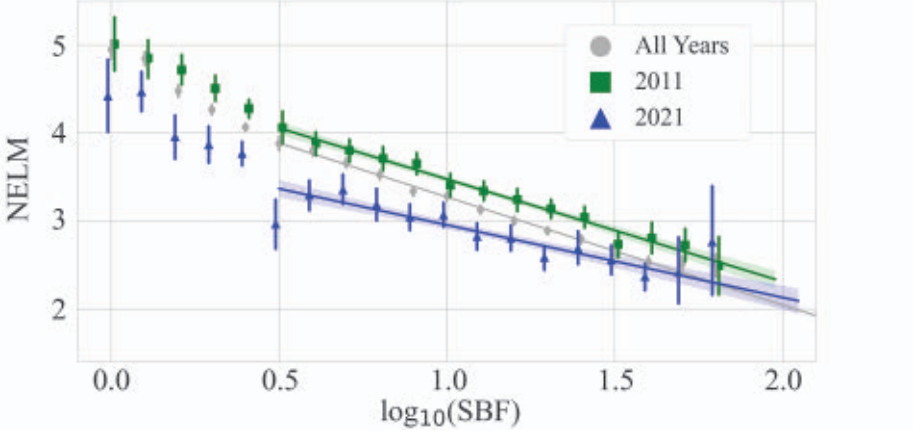


Fig. 1. Naked eye limiting magnitude estimated by Globe at Night participants as a function of the night sky brightness in 2014. The sky brightness factor (SBF) is the ratio of total radiance to natural sky radiance, so SBF = 1 indicates starlight and SBF = 10 (plotted as $\log_{10}$SBF = 1) indicates that the sky is 10 times as bright as starlight (plotted as $\log_{10}$SBF = 0 and 1, respectively) (*26*). The relationship is shown for 2011 (green squares), 2021 (blue triangles), and the average of all years from 2011 to 2022 (gray circles). Smaller NELM values mean that fewer stars are visible. Lines indicate linear models fitted to the data for $\log_{10}$SBF >0.5, which corresponds to ~3 times as bright as starlight. Shaded regions show the 95% confidence interval.

Table 1. Summary of results for different geographic areas. Best-fitting values of our model are listed for three continental groupings and the global average. $\hat{r}$ is the rate of increase, $\hat{N}_n$ is the NELM for natural skies, $\hat{s}$ relates NELM to the World Atlas, σ is the standard deviation of residuals, and E is the error rate (*26*). Uncertainties are ±1 standard deviation and are statistical only. Rest of World includes four continents because of insufficient coverage (see text). More detailed results, before combining continents, are listed in table S1.

Region	$\hat{r}$ (%)	$\hat{N}_n$ (NELM)	$\hat{s}$	σ (NELM)	E (%)
Europe	6.5 ± 1.0	4.77 ± 0.03	−1.33 ± 0.03	0.953 ± 0.012	2.4 ± 0.5
North America	10.4 ± 0.5	4.95 ± 0.02	−1.52 ± 0.02	1.009 ± 0.008	1.7 ± 0.3
Rest of World	7.7 ± 0.7	4.66 ± 0.03	−1.37 ± 0.02	1.077 ± 0.011	2.5 ± 0.5
Global average	9.6 ± 0.4	4.825 ± 0.014	−1.429 ± 0.014	1.022 ± 0.006	2.1 ± 0.2

that number reduce to 100 visible stars over the same period. Because our method uses measurements made with human vision, it accounts for changes in both the radiance and spectrum of the night sky.

We confirmed this finding by performing an alternative analysis of the data. Instead of using a maximum likelihood method to fit a model of the skyglow change, we performed least-squares fitting of a linear model to the monthly (observed minus expected) NELM residuals under the assumption that skyglow remains constant (Fig. 3A) or increases at a rate of 9.6% per year (Fig. 3B). The best-fitting rate of change in NELM residuals is −0.044 ± 0.007 magnitude per year for the uncorrected model and 0.001 ± 0.007 magnitude per year for the model corrected for skyglow increase. The even distribution of points above and below zero in the corrected model shows that the trend is not being driven by outliers at the start or end of the analysis period (Fig. 3B).

The rate of skyglow increase that we find is much larger than the rates of growth in light emissions observed by satellite in the 500- to 900-nm band, which were 2.2% per year globally during 2012–2016 (*21*) and >1.6% per year during 1992–2017 (with the possibility of faster increases in the visual band from 2012 onward) (*22*). For a more direct comparison, we analyzed the satellite radiance trends at the Globe at Night observation locations during 2014–2021 (*26*). Even after controlling for the locations, rates of change in upward radiance measured by satellite were also much smaller than our calculated rate of skyglow increase (table S2). For example, for North America during 2011–2021, we find the rate of increase in skyglow radiance is 10.4 ± 0.5%, compared to −0.80 ± 0.04% in the surface radiance measured by satellite at these locations during 2014–2021 (table S2). We ascribe the smaller satellite radiance change at Globe at Night locations compared to that found in other studies (*21*, *22*) to the tendency of Globe at Night participants to observe from residential areas (see materials and methods).

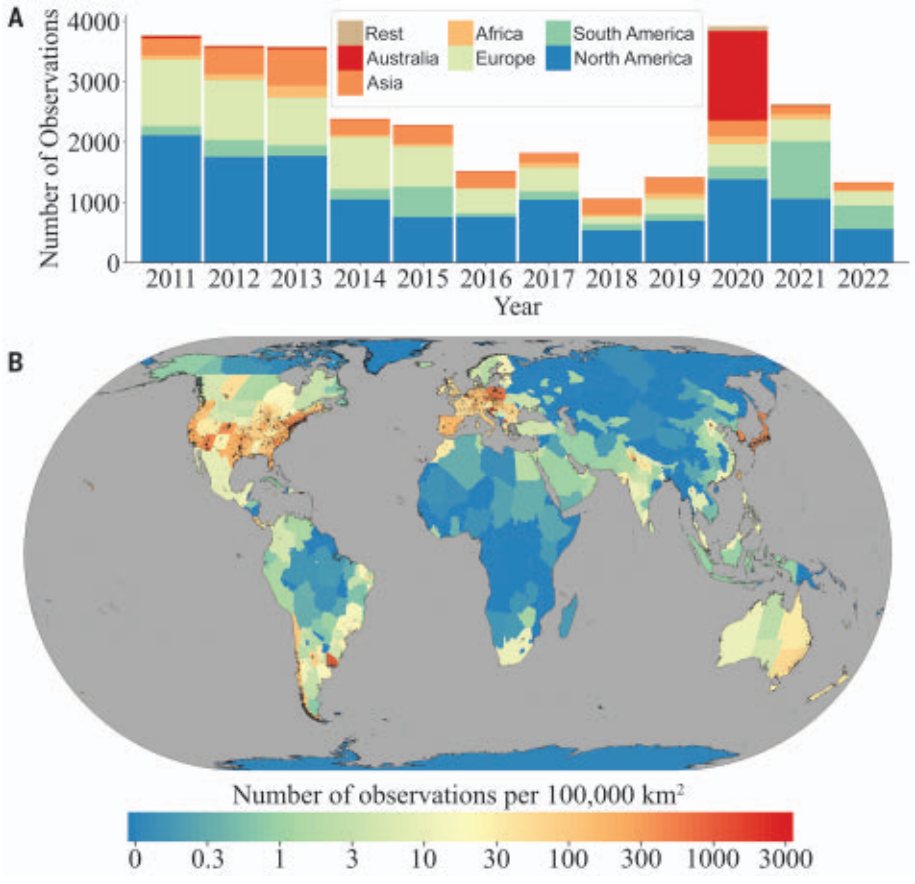


Fig. 2. Participation in Globe at Night from January 2011 to August 2022. (**A**) Bar chart showing the temporal distribution. Participation for each year is subdivided by continent, as indicated in the legend. (**B**) Spatial distribution of all years combined, on an equal-area Eckert IV map projection. Colors indicate the participation normalized by land area, on a logarithmic scale. Black points show the locations of individual observations (more visible in fig. S3). Some of the largest countries have been divided into smaller jurisdictions.

These different results are not incompatible, because there are several differences between observing surface radiance with satellites and sky radiance as seen by humans on the ground. For example, the widespread conversion of streetlights from gas discharge lamps to LEDs (*16*, *28*) could result in spectral changes that affect the two datasets differently, as discussed above. If luminance is maintained after converting street lighting to LEDs, the spectral shift toward shorter-wavelength (bluer) light causes the radiance observed by the satellite to

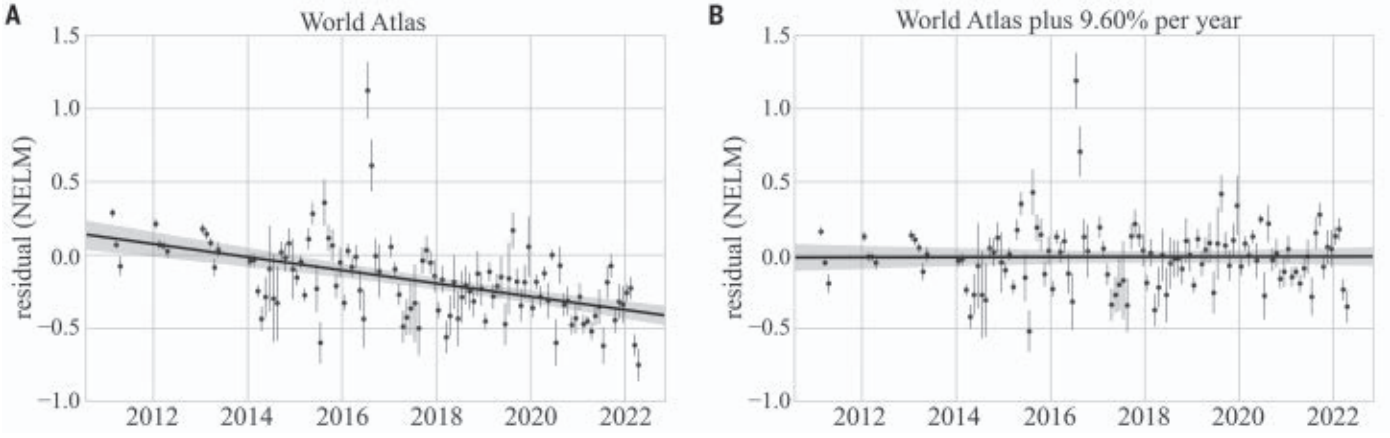


Fig. 3. Difference from expected limiting magnitude under different skyglow growth models. The monthly residuals (observed minus expected) between reported NELM and the expectation based on the World Atlas (Fig. 1) under two models: (**A**) the standard 2014 World Atlas; (**B**) the World Atlas radiance multiplied by an exponential increase of 9.6% per year relative to January 2014 (*26*). Positive residuals indicate that observers reported more stars than expected. Error bars show the standard error, and data points before 2014 include higher numbers of observations than those made after 2014. Black lines show a linear model fitted to the residuals, and the shaded region shows the 95% confidence interval.

decrease (*17*). By contrast, skyglow mesopic luminance after installation of LEDs could potentially either increase as a result of increased atmospheric scattering of blue light (*2*, *29*) or decrease as a result of improved lighting fixtures that reduce horizontal emission (*16*).

The contributions to ground-observed skyglow and satellite-observed surface radiance depend on the lighting type. Most satellite radiometers have little to no sensitivity to light emitted toward the horizon (e.g., from a window or self-luminous sign) (*30*). However, light propagating toward the horizon is the largest contributor to skyglow because of its longer path length (by an order of magnitude) from ground to space at such angles. In the early evening, a large fraction of the light that escapes cities is emitted by sources other than streetlights (*31*). Some of these lighting applications, such as decorative and advertising lighting, produce a larger fraction of horizontally propagating light than modern street lighting does. It is therefore likely that some of the differences between the rates of change for skyglow that we calculate and those estimated from satellite data arise from changes in lighting practices or deployment.

We draw two conclusions from these results. First, the visibility of stars is deteriorating rapidly, despite (or perhaps because of) the introduction of LEDs in outdoor lighting applications. Existing lighting policies are not preventing increases in skyglow, at least on continental and global scales. Second, the use of naked-eye observations by citizen scientists provides complementary information to the satellite datasets.

REFERENCES AND NOTES

1. F. Falchi *et al.*, *Sci. Adv.* **2**, e1600377 (2016).
2. J. C. Barentine, *Nat. Astron.* **6**, 1120–1132 (2022).
3. F. Hölker *et al.*, *Ecol. Soc.* **15**, 13 (2010).
4. K. W. Riegel, *Science* **179**, 1285–1291 (1973).
5. J. Y. Tsao, P. Waide, *Leukos* **6**, 259–281 (2010).
6. C. C. M. Kyba, F. Hölker, *Landsc. Ecol.* **28**, 1637–1640 (2013).
7. A. A. Wilson *et al.*, *Integr. Comp. Biol.* **61**, 1122–1133 (2021).
8. F. Kupprat, F. Hölker, W. Kloas, *Environ. Pollut.* **262**, 114324 (2020).
9. T. Gallaway, *J. Econ. Issues* **44**, 71–88 (2010).
10. O. P. Cleaver, Control of coastal lighting in anti-submarine warfare. *Tech. Rep.*, Engineer Board, Fort Belvoir, VA (1943). Report No. 746, GN 373 (accessed 16 May 2019).
11. F. Falchi, P. Cinzano, C. D. Elvidge, D. M. Keith, A. Haim, *J. Environ. Manage.* **92**, 2714–2722 (2011).
12. M. Aubé, J. Roby, *J. Quant. Spectrosc. Radiat. Transf.* **139**, 52–63 (2014).
13. R. F. Green, C. B. Luginbuhl, R. J. Wainscoat, D. Duriscoe, *Astron. Astrophys. Rev.* **30**, 1 (2022).
14. M. Morgan-Taylor, in *Urban Lighting, Light Pollution and Society*, J. Meier *et al.*, Eds. (Routledge, 2015), chap. 9, pp. 159–176.
15. G. Zissis, P. Bertoldi, T. Serrenho, Update on the status of LED-lighting world market since 2018. *Tech. Rep.* (2021). https://data.europa.eu/doi/10.2760/759859.
16. B. Kinzey *et al.*, An investigation of LED street lighting's impact on sky glow. *Tech. Rep.*, US Department of Energy (2017). PNNL-26411.
17. L.-W. Hung, S. J. Anderson, A. Pipkin, K. Fristrup, *J. Environ. Manage.* **292**, 112776 (2021).
18. J. Y. Tsao, H. D. Saunders, J. R. Creighton, M. E. Coltrin, J. A. Simmons, *J. Phys. D Appl. Phys.* **43**, 354001 (2010).
19. A. Sánchez de Miguel, C. C. M. Kyba, J. Zamorano, J. Gallego, K. J. Gaston, *Sci. Rep.* **10**, 7829 (2020).
20. S. D. Miller *et al.*, *Remote Sens.* **5**, 6717–6766 (2013).
21. C. C. M. Kyba *et al.*, *Sci. Adv.* **3**, e1701528 (2017).
22. A. Sánchez de Miguel, J. Bennie, E. Rosenfeld, S. Dzurjak, K. J. Gaston, *Remote Sens.* **13**, 3311 (2021).
23. A. Sánchez de Miguel *et al.*, *Mon. Not. R. Astron. Soc.* **467**, 2966–2979 (2017).
24. C. C. M. Kyba, *Nat. Astron.* **2**, 267–269 (2018).
25. C. C. M. Kyba *et al.*, *Sci. Rep.* **3**, 1835 (2013).
26. Materials and methods are available as supplementary materials.
27. A. Crumey, *Mon. Not. R. Astron. Soc.* **442**, 2600–2619 (2014).
28. N. Schulte-Römer, J. Meier, M. Söding, E. Dannemann, *Sustainability* **11**, 6160 (2019).
29. C. B. Luginbuhl, P. A. Boley, D. R. Davis, *J. Quant. Spectrosc. Radiat. Transf.* **139**, 21–26 (2014).
30. C. C. M. Kyba *et al.*, *J. Geophys. Res. Atmos.* **127**, e2021JD036382 (2022).
31. J. C. Barentine *et al.*, *J. Quant. Spectrosc. Radiat. Transf.* **253**, 107120 (2020).
32. C. C. M. Kyba, Y. O. Altıntaş, C. Walker, M. Newhouse, Software for analyzing Globe at Night data, GFZ Data Services (2022); https://doi.org/10.5880/GFZ.1.4.2022.008.
33. F. Falchi *et al.*, Supplement to: The new world atlas of artificial night sky brightness. GFZ Data Services, (2016); https://doi.org/10.5880/GFZ.1.4.2016.001.

ACKNOWLEDGMENTS

We thank the thousands of participants in Globe at Night for their observations, which made this analysis possible. We thank K. Tockner for providing feedback on a draft and J. Hollan for producing the Globe at Night star charts. C.C.M.K. thanks A. Wisniewska for her 2012 visit to Berlin under the Erasmus student exchange program, which sparked his interest in Globe at Night and citizen science. We thank the referees for the suggestion to analyze DNB data at the observation locations. **Funding:** Globe at Night and C.E.W. are supported by an NSF Award (AST-1421197) to M. Mountain. C.C.M.K. acknowledges funding from the Federal Ministry of Education and Research, Germany (PT-Jülich under grant BMBF-033L038A), through the European Union's Horizon 2020 research and innovation program ERA-PLANET, grant agreement no. 689443 via the GEOEssential project, and from the Helmholtz Association Initiative and Networking Fund under grants ERC-RA-0031 and CS-0003. **Author contributions:** C.C.M.K. developed the analysis methodology and drafted the manuscript. C.C.M.K. and Y.O.A. wrote analysis software. Y.O.A. produced the figures. C.E.W. initiated and manages the Globe at Night project. M.N. manages the Globe at Night database. All authors contributed to writing and editing the manuscript. **Competing interests:** C.E.W. is a director of the International Dark-Sky Association, and C.C.M.K. is a former director of the same body. We declare that there are no other competing interests. **Data and materials availability:** The Globe at Night web application is at https://www.globeatnight.org/webapp/, and the data are available at https://www.globeatnight.org/maps.php. The subset of the dataset we used, and our analysis code, are archived at Deutsches GeoForschungsZentrum Potsdam (GFZ) Data Services (*32*). We also used the World Atlas data (*33*). **License information:** Copyright © 2023 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. https://www.science.org/about/science-licenses-journal-article-reuse

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abq7781
Materials and Methods
Figs. S1 to S3
Tables S1 and S2
References (*34*–*37*)

Submitted 11 May 2022; accepted 17 November 2022
10.1126/science.abq7781

BEHAVIORAL ADAPTATION

Phenotype–environment mismatch errors enhance lifetime fitness in wild red squirrels

Lauren Petrullo[1]*, Stan Boutin[2], Jeffrey E. Lane[3], Andrew G. McAdam[4], Ben Dantzer[1,5]*

Mismatches between an organism's phenotype and its environment can result in short-term fitness costs. Here, we show that some phenotype–environment mismatch errors can be explained by asymmetrical costs of different types of errors in wild red squirrels. Mothers that mistakenly increased reproductive effort when signals of an upcoming food pulse were absent were more likely to correctly increase effort when a food pulse did occur. However, mothers that failed to increase effort when cues of an upcoming food pulse were present suffered lifetime fitness costs that could only be offset through food supplementation. In fluctuating environments, such phenotype–environment mismatches may therefore reflect a bias to overestimate environmental cues and avoid making the costliest error, ultimately enhancing lifetime fitness.

An organism's ability to match its phenotype to environmental conditions is central to hypotheses related to the evolution of adaptive phenotypic plasticity (*1*). Organisms producing phenotypes that are adaptive in a particular environment (i.e., a "match") can enhance their reproductive success and survival. Alternatively, organisms that mismatch their phenotypes to environmental conditions often suffer fitness costs. For example, birds that mistime breeding to a fluctuating food supply experience poorer offspring growth (*2*) and recruitment (*3*). Such mismatches have historically been viewed as a cost of phenotypic plasticity (*4*). Yet, emerging research demonstrates that mismatched phenotypes may also promote long-term population persistence (*5*, *6*). Organisms may therefore aim to manage the costs and potential benefits of phenotype–environment mismatches over their life course to maximize fitness.

Existing theory about how organisms balance the costs of making different types of errors can help explain the occurrence of mismatches (*7*–*9*). For example, in the presence or absence of information about the future environment, an organism may respond or not respond, creating a minimum of four conditions (two matches, two mismatch errors). Each type of mismatch error is expected to have a distinct fitness cost, and organisms should err toward making the least costly of the two (*7*). For a mismatched phenotype to be adaptive, the costs of not responding to the cue (false negative) should be greater than the cost of responding erroneously in the absence of the cue (false positive) (*7*). Empirical tests of this framework

[1]Department of Psychology, University of Michigan, Ann Arbor, MI, USA. [2]Department of Biological Sciences, University of Alberta, Edmonton, Canada. [3]Department of Biology, University of Saskatchewan, Saskatoon, Canada. [4]Department of Ecology and Evolutionary Biology, University of Colorado, Boulder, CO, USA. [5]Department of Ecology and Evolutionary Biology, University of Michigan, Ann Arbor, MI, USA.
*Corresponding author. Email: petrullo@umich.edu (L.P.); dantzer@umich.edu (B.D.)

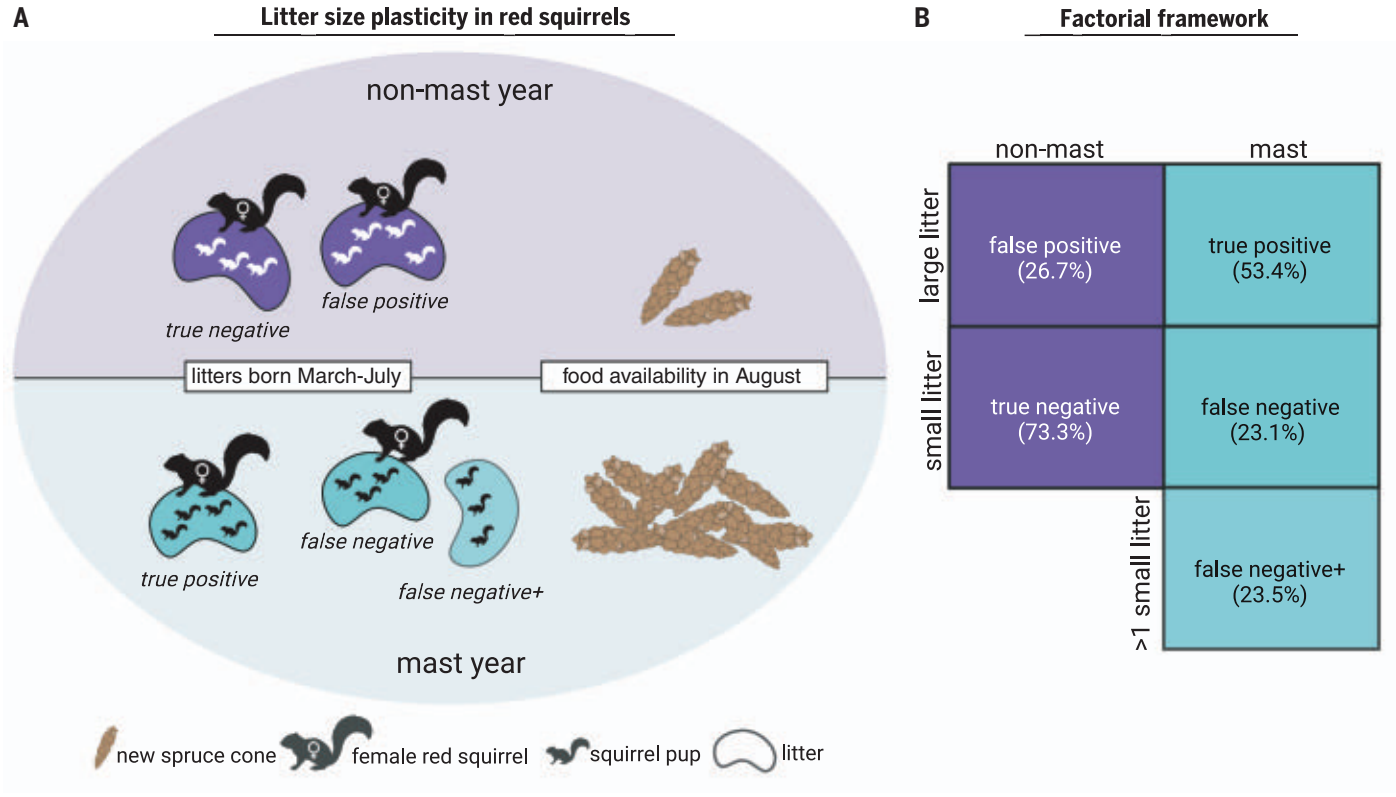


Fig. 1. Litter size plasticity generates phenotype–environment mismatches. (**A**) Yukon red squirrels inhabit a resource pulse system where fluctuations in new food production (no pulse = nonmast year, pulse = mast year) combine with variation in litter sizes (large = three or more pups, small = less than three pups) to create matched (true positive, true negative) and mismatched (false positive, false negative) phenotypic responses. Some females that produced small litters in mast years went on to produce a second litter later that year (false negative+), potentially mitigating the costs of an initial reproductive mismatch. (**B**) A factorial framework showing frequency rate of each response category (percentage of litters falling into each category in a given year) in parentheses (N = 2729 litters).

are rare because they require a system where individuals experience within-generation environmental fluctuations [fine-grained environments (*10*)] and where lifetime fitness can be quantified to determine the net effects of matches and mismatches. Individuals must also receive reliable cues about the future environment (*11*), and phenotypic responses must vary enough to permit a factorial assessment of relative fitness costs across all conditions.

Assessing the costs and benefits of mismatching in a resource pulse system

Here, we used a long-term study of wild North American red squirrels (*Tamiasciurus hudsonicus*) living in the Yukon, Canada, to probe whether phenotype–environment mismatches can be adaptive. Red squirrels in this region have been monitored continuously since 1987 (*12*). Using 31 years of life-history and fitness data from over 1000 female squirrels, we tested whether reproductive phenotype–environment mismatch errors allow females to avoid making the costliest type of error and thus enhance lifetime fitness. Yukon red squirrels inhabit a resource pulse ecosystem where they experience profound fluctuations in food availability as a result of a swamp-and-starve strategy by their primary food source, seeds from the cones of masting white spruce (*Picea glauca*) trees (*13*). Squirrels store spruce cones containing seeds underground in middens and rely on these resources in lieu of bodily fat reserves (*14*). In nonmast years, few to no new seeds are produced, intraspecific competition for territories is high, and offspring overwinter survival is at its lowest (*15*). In these nonmast years, litter sizes experience stabilizing selection around a population mean of 3.0 pups (*16*) (fig. S1), where the fecundity benefit of producing more pups is balanced by the benefit of producing small litters of fast-growing offspring that can acquire territories and survive winter (*16*, *17*). By contrast, in mast years [approximately every 4 to 7 years (*13*)], a superabundance of cones is produced in the autumn, leading to higher offspring survival, and directional selection favoring increased litter sizes [optimum litter size = 5.2 pups (*16*); population mean = 3.5 pups (fig. S1)]. Females use ecological cues to adaptively increase litter sizes in the months before the mast occurs, capitalizing on the future food boom by recruiting more pups into the breeding population (tables S1 and S2) (*16*). These increases in litter size may be caused by physiological responses to phytohormonal cues of an upcoming mast (*18*) but not by energetic differences (*19*) or differences in the number of hoarded spruce cones in an individual's midden (table S3).

Most females responded to mast cues by increasing litter sizes before the food pulse (true positive, 53.4%), but others mismatched by failing to increase litter size (false negative, 23.1%) or by increasing litter size in the absence of mast cues [false positive, 26.7% (Fig. 1)].

Divergent costs of phenotype–environment mismatching

When an upcoming mast was imminent, females that failed to respond (false negative) suffered a 97% decrease in their annual reproductive success (recruiting only 1.07 ± 0.09 pups) compared to females that responded by increasing litter sizes [true positive, recruiting 2.11 ± 0.10 pups (Fig. 2A and tables S4 and S5)]. In nonmast years, females producing large litters (false positive) had annual reproductive success similar to that of females producing small litters [true negative (Fig. 2A and tables S4 and S5)], suggesting a marginal or otherwise undetectable cost of false positive errors (*20*). A small number of females (6.6%) produced a second successful litter in a mast year. When litter sizes were small but still a part of this multilitter strategy (false negative+), mothers minimized, but did not eliminate, the costs of a false negative error by having an additional litter in that year (Fig. 2A and tables S4 and S5).

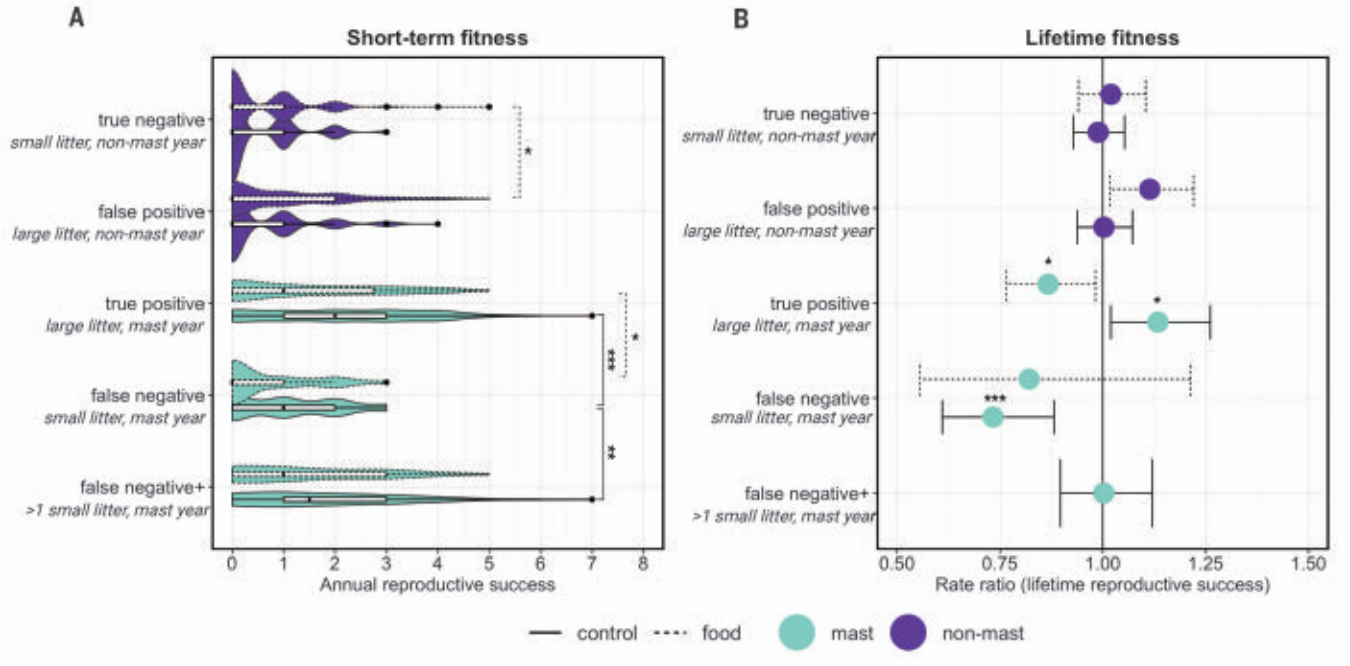


Fig. 2. The fitness costs of mismatch errors are asymmetrical, cumulative, and ameliorated by food supplementation. (**A**) In nonmast years, the difference in annual reproductive success (number of pups recruited into the breeding population that year) between control females that produced large (false positive) and small (true negative) litters was negligible. In mast years, control females that failed to respond to mast cues and produced small litters (false negative) exhibited significantly lower annual reproductive success than those that produced large litters [true positive (tables S4 and S5), *N* = 2729 litters]. Violin plots show raw data; inset boxplots depict median and outliers. (**B**) There was no cumulative cost to maternal lifetime fitness (number of total pups recruited into the breeding populaton over the life course) of repeated overestimation of mast cues (false positive), but each missed mast cue (false negative) significantly decreased lifetime fitness among control but not food-supplemented females (*N* = 1055 females). Filled circles show point estimates of rate ratios (relative difference). Bars denote 95% confidence intervals (CIs). Estimates >1.0 indicate a positive effect on lifetime reproductive success; estimates <1.0 indicate a negative effect. Asterisks denote significance (*$P > 0.05$, **$P > 0.01$, ***$P < 0.001$).

Because squirrels breed annually and fitness accrues with each breeding attempt, we assessed whether mismatched responses had cumulative effects on female lifetime fitness. Although each true positive (responding in presence of cue) produced across the life span increased maternal lifetime reproductive success by 12.5%, each false negative (failure to respond in presence of cue) mismatch decreased lifetime reproductive success by 30% (Fig. 2B and table S6). Females would therefore have to produce three true positive matches to recuperate the lifetime fitness costs of missing a single mast cue, an impossibility given the frequency of mast events [once every ~4 to 6 years (fig. S5)] and the maximum red squirrel life span [8 years, median = 3.5 years (*21*)].

Avoiding the costliest error benefits mothers but not offspring

The rare nature of spruce mast events provides additional ecological context for the maintenance of false positive mismatches. Females that did not experience a mast event before dying (54.2%) exhibited considerably lower lifetime reproductive success than females that did (table S2). The uncertainty of encountering a mast during a squirrel's lifetime, coupled with the importance of mast years for maximizing lifetime fitness, should favor individuals that are highly sensitive to mast cues even if they mismatch in years when cues are absent (*8*, *22*). Indeed, each false positive error produced over the lifetime increased a female's probability of avoiding the costliest error (false negative) by 67% (Fig. 3A and table S7). Females therefore benefit from an error-avoidance bias (*7*) or adaptive gullibility (*8*), such that overestimating cues and responding when cues are absent reduces the likelihood of failing to respond when cues are present.

Using data on 8455 squirrel pups born to these females, we found that this error-avoidance bias, despite enhancing maternal lifetime fitness, came at a cost to offspring. Although females that produced large litters in nonmast years (false positive) reduced the likelihood that they made a mismatch error in a mast year (Fig. 3A), their pups grew slower [$\beta = -0.15 \pm 0.07$, $P = 0.03$ (Fig. 3B and tables S8 and S9)] and were less likely to survive their first winter, thus failing to contribute to the breeding population [$\beta = -1.26 \pm 0.44$, $P = 0.005$ (Fig. 3C and tables S10 and S11)]. Explaining why traits that increase maternal fitness but decrease offspring fitness (i.e., "selfish maternal effects") exist has been contentious (*23*). Here, we show that traits promoting maternal fitness at the expense of offspring fitness, such as overestimating environmental cues, may be more likely to evolve in systems where mothers must manage asymmetrical costs of reproductive mismatch errors.

It is currently unclear why heterogeneity in female reproductive responses to cues of spruce masting persists in this population, but it may reflect variation in the physiological integration of mast cues. Being born in a mast year did not affect a female's ability to respond to mast cues in the future (table S12). However, first-time (primiparous) mothers were ~1.5 times less likely to make a litter-size mismatch error than were multiparous mothers [$\beta = -0.42 \pm 0.17$, $P = 0.01$ (fig. S3 and table S12)]. This may reflect parity-dependent differences in neuroendocrine sensitivity to phytohormonal (estrogenic) mast cues (*18*); it may also indicate that, because most females breed only once in their lifetime (fig. S2A), first-time mothers have more to lose by mismatching what may be their only attempt at contributing to the gene pool.

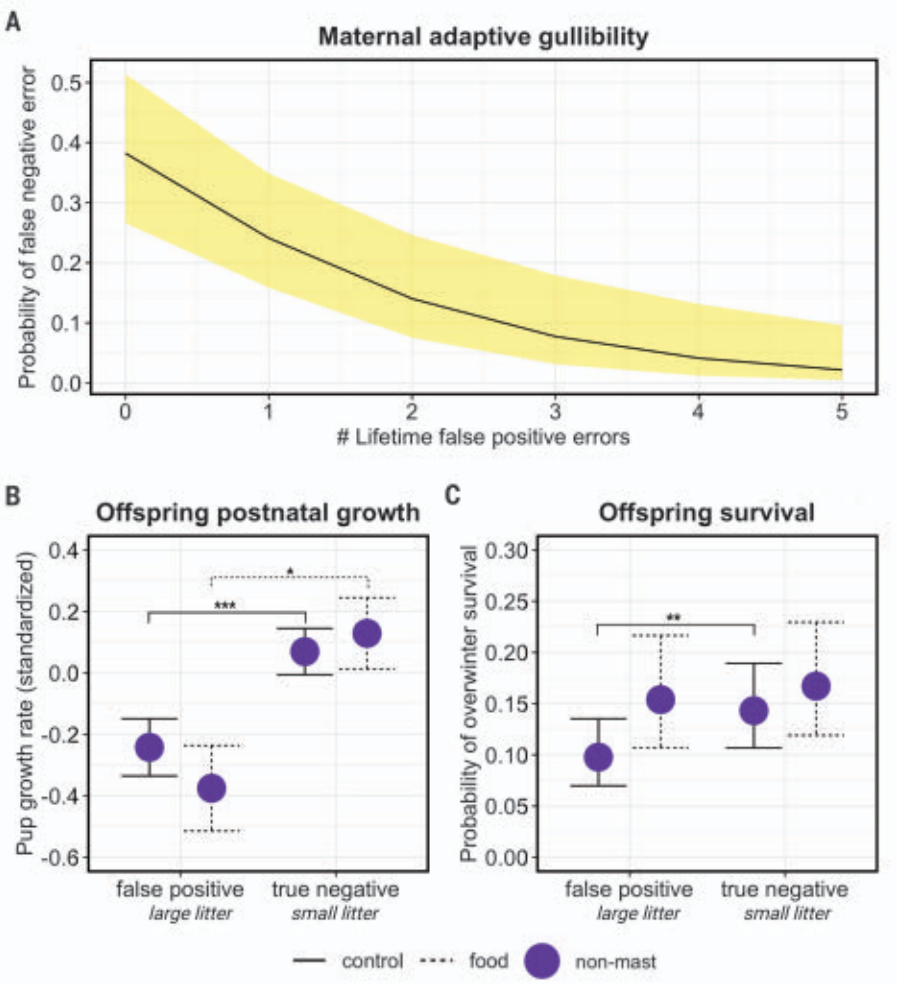


Fig. 3. Overestimating mast cues enhances maternal lifetime fitness despite costs to offspring. (**A**) Each false positive error (erroneously producing a large litter in a nonmast year) made across a control female's lifetime significantly decreased her probability of making the costliest error [false negative (Fig. 2B), $N = 361$ squirrels]. Plot shows predicted probabilities from a binomial model (table S7); ribbon indicates 95% CIs. (**B**) In nonmast years, when new food is low and territory competition is high, pups born into large litters (false positive) had poorer postnatal growth [true negative (tables S8 and S9), $N = 5048$ pups] and (**C**) were less likely to survive their first winter than pups born into small litters [true negative (tables S10 and S11), $N = 8455$ pups]. Pups of supplemented females that made false positive errors did not suffer a survival cost. Asterisks denote significance (*$P > 0.05$, **$P > 0.01$, ***$P < 0.001$).

Food supplementation modifies patterns of error avoidance

In addition, environmental fluctuations that are episodic but predictable can create well-adapted systems in which the manipulation of resources is necessary to uncover mismatch costs. We tested the hypothesis that providing squirrels with additional food would disrupt selection for an error-avoidance bias by reducing the fitness costs of mismatching. To do so, we provided ad libitum peanut butter to females living on a nearby but distinct experimental study area for 14 consecutive years. In mast years, food supplementation relaxed selection for larger litters, and directional selection on litter size was negative rather than positive [$\beta = -0.26 \pm 0.12$, $P = 0.03$ (table S13)].

Providing an additional food source diminished and in some cases reversed the costs of mismatch errors. Producing a large litter in the absence of a cue (false positive) was associated with an increase in maternal annual reproductive success (Fig. 2A and tables S4 and S5) and cumulatively positive effects on lifetime reproductive success (Fig. 2B and table S6). Supplemented females who failed to respond to mast cues (false negative) incurred short-term fitness costs similar to those incurred by control females, but without any cumulative cost over the life span (Fig. 2B and table S6). An amelioration of the trade-off between offspring quality and quantity appears to be the mechanism that mitigated these costs. In nonmast years, pups born in large litters (false positive) on the food-supplemented study area did not exhibit poorer survival as compared to pups born into small litters [true negative (Fig. 3C and tables S10 and S11)].

In environments where selective pressures are unpredictable, producing a mix of matched and mismatched phenotypes (i.e., bet-hedging) can maximize fitness (*24*). By contrast, in environments where selective pressures are predictable, adaptive plasticity is expected to evolve in a manner dependent on the spatiotemporal grain of the environment (*25*). Organisms inhabiting predictable coarse-grained environments (across-generation fluctuations) should exhibit polymorphisms that are environmentally induced (*26*) or genetically determined (*27*). Organisms inhabiting predictable fine-grained environments (within-generation fluctuations), such as Yukon red squirrels, are instead expected to exhibit adaptive phenotypic plasticity (*10*).

What has remained unclear is why phenotypes that are mismatched to environmental conditions persist. Mismatch errors can occur because the cues that organisms receive about their future environment are imprecise or unreliable (*11*), a growing phenomenon in the wake of global anthropogenic climate change (*28*). Or as our results indicate, mismatches can persist because organisms must balance asymmetrical costs of different types of errors to avoid making the costliest error over their lifetimes.

REFERENCES AND NOTES

1. R. Levins, *Evolution in Changing Environments* (Princeton Univ. Press, 1968).
2. L. McKinnon, M. Picotin, E. Bolduc, C. Juillet, J. Bêty, *Can. J. Zool.* **90**, 961–971 (2012).
3. T. E. Reed, S. Jenouvrier, M. E. Visser, *J. Anim. Ecol.* **82**, 131–144 (2013).
4. T. J. DeWitt, *J. Evol. Biol.* **11**, 465–480 (1998).
5. C. K. Ghalambor *et al*., *Nature* **525**, 372–375 (2015).
6. A. M. Leonard, L. T. Lancaster, *BMC Evol. Biol.* **20**, 47 (2020).
7. M. G. Haselton, D. M. Buss, *J. Pers. Soc. Psychol.* **78**, 81–91 (2000).
8. R. H. Wiley, in *Behavioral Mechanisms in Evolutionary Ecology*, L. A. Real, Ed. (Univ. of Chicago Press, 1994), pp. 157–189.
9. P. Bateson, P. Gluckman, M. Hanson, *J. Physiol.* **592**, 2357–2368 (2014).
10. B. S. Baythavong, *Am. Nat.* **178**, 75–87 (2011).
11. T. E. Reed, R. S. Waples, D. E. Schindler, J. J. Hard, M. T. Kinnison, *Proc. Biol. Sci.* **277**, 3391–3400 (2010).
12. B. Dantzer, A. G. McAdam, M. M. Humphries, J. E. Lane, S. Boutin, . *J. Anim. Ecol.* **89**, 2397–2414 (2020).
13. J. M. Lamontagne, S. Boutin, *J. Ecol.* **95**, 991–1000 (2007).
14. M. M. Humphries, D. W. Thomas, D. L. Kramer, *Physiol. Biochem. Zool.* **74**, 283–292 (2001).
15. A. G. McAdam, S. Boutin, *Evolution* **57**, 1689–1697 (2003).
16. A. G. McAdam, S. Boutin, B. Dantzer, J. E. Lane, *Am. Nat.* **194**, 574–589 (2019).
17. J. G. Hendrix *et al*., *J. Anim. Ecol.* **89**, 1408–1418 (2020).
18. S. Boutin *et al*., *Science* **314**, 1928–1930 (2006).
19. M. M. Humphries, S. Boutin, *Ecology* **81**, 2867–2877 (2000).
20. R. M. Nesse, *Ann. N. Y. Acad. Sci.* **935**, 75–85 (2001).
21. A. G. McAdam, S. Boutin, A. K. Sykes, M. M. Humphries, *Écoscience* **14**, 362 (2007).
22. D. D. P. Johnson, J. H. Fowler, *Nature* **477**, 317–320 (2011).
23. D. J. Marshall, T. Uller, *Oikos* **116**, 1957–1963 (2007).
24. B.-E. Sæther, S. Engen, *Trends Ecol. Evol.* **30**, 273–281 (2015).
25. P. H. Van Tienderen, *Evolution* **45**, 1317–1331 (1991).
26. D. W. Pfennig, *Evolution* **46**, 1408–1420 (1992).
27. O. Leimar, P. Hammerstein, T. J. M. Van Dooren, *Am. Nat.* **167**, 367–376 (2006).
28. M. E. Visser, C. Both, M. M. Lambrechts, *Adv. Ecol. Res.* **35**, 89–110 (2004).
29. L. Petrullo, S. Boutin, A. G. McAdam, J. E. Lane, B. Dantzer, *Figshare* (2022); https://doi.org/10.6084/m9.figshare.16915357.v1.

ACKNOWLEDGMENTS

We thank A. MacDonald and her family for long-term access to her trapline and the Champagne and Aishihik First Nations for allowing us to conduct our work within their traditional territory. Thanks to all of the field technicians who collected data and to D. Delaney, Q. Webber, J. Feder, and four anonymous reviewers and the editor for their helpful feedback. This is KRSP paper 123. **Funding:** NSF PRFB DEB-2010726 (L.P.), Natural Sciences and Engineering Research Council of Canada (S.B., J.E.L., A.G.M.), NSF IOS-1749627 (B.D.), NSF DEB-0515849 (A.G.M.). **Author contributions:** L.P. and B.D. conceived of the project. All authors performed the research related to this manuscript, and L.P. and B.D. analyzed the data. The manuscript was written by L.P. and B.D. with review and edits provided by all authors. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** All data and code are available at the figshare digital repository (*29*). **License information:** Copyright © 2023 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. https://www.science.org/about/science-licenses-journal-article-reuse

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abn0665
Materials and Methods
Supplementary Text
Figs. S1 to S5
Tables S1 to S17
References (*30*–*35*)
MDAR Reproducibility Checklist

Submitted 2 November 2021; accepted 15 December 2022
10.1126/science.abn0665

PLANT SCIENCE

Nanobody-driven signaling reveals the core receptor complex in root nodule symbiosis

Henriette Rübsam[1]†, Christina Krönauer[1]†, Nikolaj B. Abel[1], Hongtao Ji[1,2], Damiano Lironi[1], Simon B. Hansen[1], Marcin Nadzieja[1], Marie V. Kolte[1], Dörte Abel[1], Noor de Jong[1], Lene H. Madsen[1], Huijun Liu[1], Jens Stougaard[1], Simona Radutoiu[1]*, Kasper R. Andersen[1]*

Understanding the composition and activation of multicomponent receptor complexes is a challenge in biology. To address this, we developed a synthetic approach based on nanobodies to drive assembly and activation of cell surface receptors and apply the concept by manipulating receptors that govern plant symbiosis with nitrogen-fixing bacteria. We show that the *Lotus japonicus* Nod factor receptors NFR1 and NFR5 constitute the core receptor complex initiating the cortical root nodule organogenesis program as well as the epidermal program controlling infection. We find that organogenesis signaling is mediated by the intracellular kinase domains whereas infection requires functional ectodomains. Finally, we identify evolutionarily distant barley receptors that activate root nodule organogenesis, which could enable engineering of biological nitrogen-fixation into cereals.

Living cells must constantly monitor their environment to integrate changes and to communicate with neighboring cells and microorganisms. Cell surface receptors perceive external signals and convert stimuli into intracellular responses. There is, however, a limited understanding and a lack of methodology for exploring how individual receptors assemble and function in signaling-competent complexes that coordinate the downstream responses. We tackle this problem in vivo using single-pass transmembrane cell surface receptors that govern plant-microbe interactions. Plants interact with a wealth of microbes and need to distinguish between those that pose a risk and those that offer a potential benefit. To this end, plants use pattern-recognition receptors, including lysine motif (LysM) receptors, that perceive microbial-derived carbohydrate signals and mount an intracellular response (*1*–*4*). Plant LysM receptors recognize conserved cell wall components, such as chitin from pathogenic fungi (*5*, *6*), Myc factors in arbuscular mycorrhizal symbiosis (*7*, *8*), and Nod factors in nitrogen-fixing symbiosis with rhizobia (*2*, *9*–*12*). Genetic studies in the model legume *Lotus japonicus* (*Lotus*) have identified two LysM Nod factor

[1]Department of Molecular Biology and Genetics, Aarhus University, 8000 Aarhus C, Denmark. [2]National Key Laboratory of Crop Genetic Improvement, College of Plant Science and Technology, Huazhong Agricultural University, Wuhan 430070, China.
***Corresponding author. Email: radutoiu@mbg.au.dk (S.R.) and kra@mbg.au.dk (K.R.A.)**
†These authors contributed equally to this work.

receptors, NFR1 and NFR5 (*10*), along with the symbiosis receptor-like kinase SYMRK (*13*), which initiate and control the two developmental programs leading to nodule organogenesis and infection (intracellular accommodation), respectively. Loss-of-function mutations in these three receptors render plants incapable of establishing root nodule symbiosis. To study how the receptors collaborate during signaling and to define the active receptor complexes, we developed a synthetic approach based on single-domain antibody fragments from camelids (called nanobodies) to actively promote receptor assembly. We used the nanobody intervention to identify the core receptor complex that induces signaling in root nodule symbiosis in legumes.

Nanobody-driven assembly defines the receptor complex initiating organogenesis

To investigate receptor complex activation, we developed technology that enables the controlled assembly of receptors in living cells and used symbiotic receptors as proof-of-principle targets. To drive assembly, we used llama-derived nanobodies as they are small, specific, high-affinity binders that are active in the reducing environment of the cytosol (*14*). We generated a nanobody with high affinity to the *Lotus* NFR5 kinase (Nb_{NFR5}) and used it as a C-terminal tag on the *Lotus* NFR1 receptor (Fig. 1A and fig. S1). We reasoned that Nb_{NFR5} fused to the C-terminal tail of NFR1 would specifically recognize and bind the kinase domain of the native NFR5 receptor and drive the formation of a heteromeric complex. This was tested by expressing nanobody-tagged NFR1 receptor constructs from the native *Nfr1* promoter and terminator in *Agrobacterium rhizogenes*–induced roots of *nfr1* mutant plants (Fig. 1B). We observe nodule formation independent of rhizobia or Nod factor, illustrating that receptor assembly is the key step in symbiotic signaling (Fig. 1C). In a control experiment, we omit the nanobody and observe that no spontaneous nodules are formed (Fig. 1C). To establish a general approach that does not rely on nanobody generation for each protein of interest, we next used the well-characterized LaG16 nanobody that binds superfolder green fluorescent protein (GFP) (*15*). We fused GFP to NFR5 and the LaG16 nanobody to NFR1 and expressed the constructs from their native promoters and terminators in *nfr1 nfr5* double mutant plants (Fig. 1E). When NFR1 is associated to NFR5 through nanobody binding to GFP, nodules are formed independently of rhizobia (Fig. 1, D and F). Similarly, organogenesis is activated in the reverse situation with GFP fused to NFR1 and the nanobody fused to NFR5 (fig. S2). To test whether this phenotype is dependent on the direct binding between the nanobody and GFP, we engineered a mutant version of LaG16

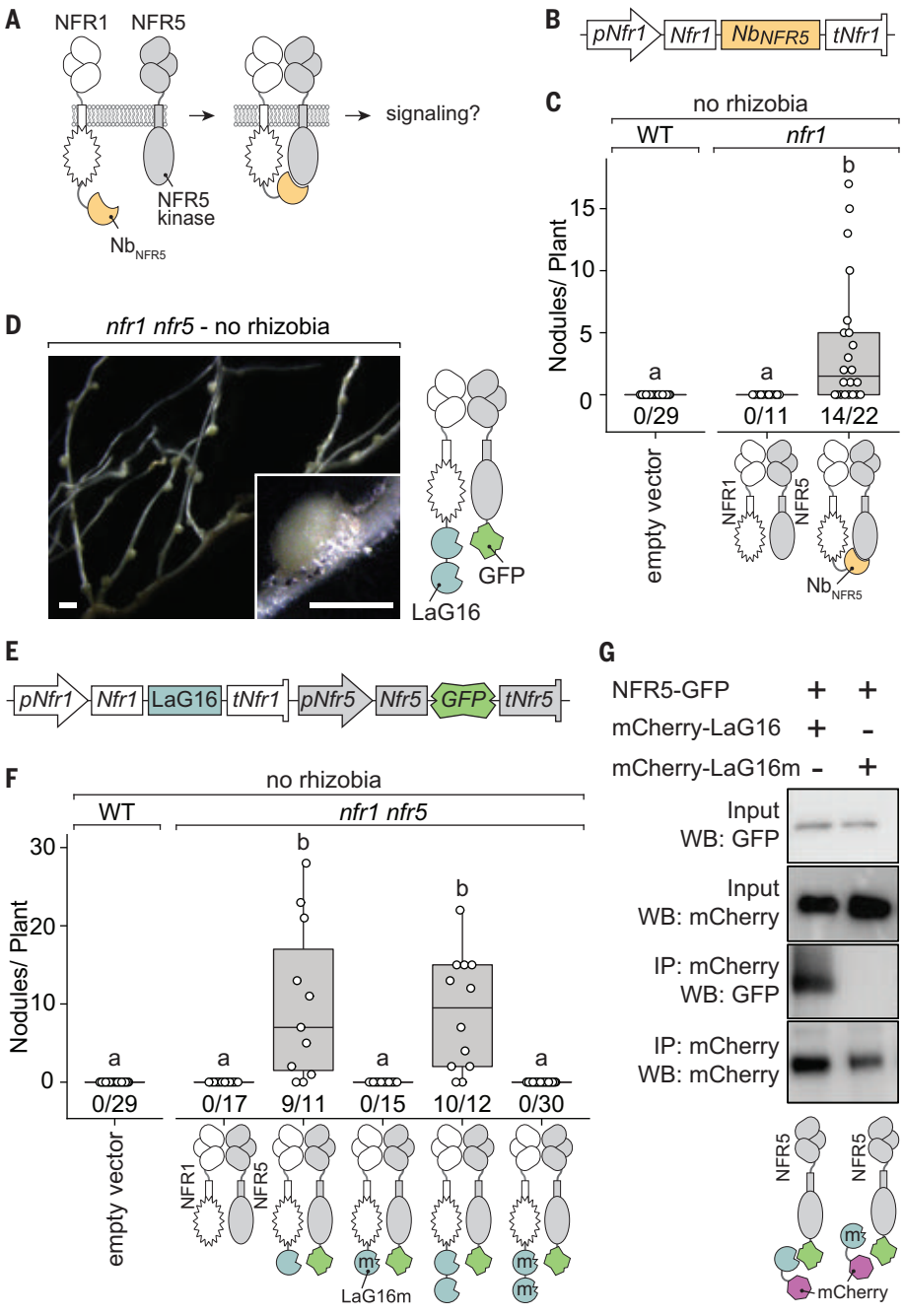


Fig. 1. Nanobody-driven assembly defines the receptor complex activating organogenesis. (**A**) Experimental design for nanobody-induced complex formation to activate symbiotic signaling. (**B**) Schematic of nanobody-tagged NFR1 receptor construct for expression in *Lotus* roots. (**C**) Number of nodules formed on WT and *nfr1* mutant plants expressing the indicated constructs in the absence of rhizobia. Nodules were counted 9 weeks after hairy root transformation. Circles represent individual plants. Numbers below boxplots specify number of nodulating plants out of the total number of plants. (**D**) Representative brightfield images of *Lotus* roots that express the depicted constructs, showing the nodulation phenotype in the absence of rhizobia. Scale bar: 1 mm. (**E**) Schematic of NFR1 and NFR5 receptor constructs for expression in *Lotus* roots. (**F**) Number of nodules formed on WT and *nfr1 nfr5* double mutant plants expressing the indicated constructs in the absence of rhizobia. The letter "m" represents the mutated nanobody with no binding to GFP. (**G**) Coimmunoprecipitation of LaG16- or LaG16m-tagged mCherry and GFP-tagged NFR5 from *Nicotiana benthamiana* leaf extracts. Anti-GFP and anti-mCherry antibodies were used for immunoblotting. LaG16m is the mutated nanobody with no binding to GFP. In (C) and (F), lowercase letters indicate significant differences between samples [analysis of variance (Kruskal-Wallis) and post-hoc analysis (Dunn's test), $P < 0.05$].

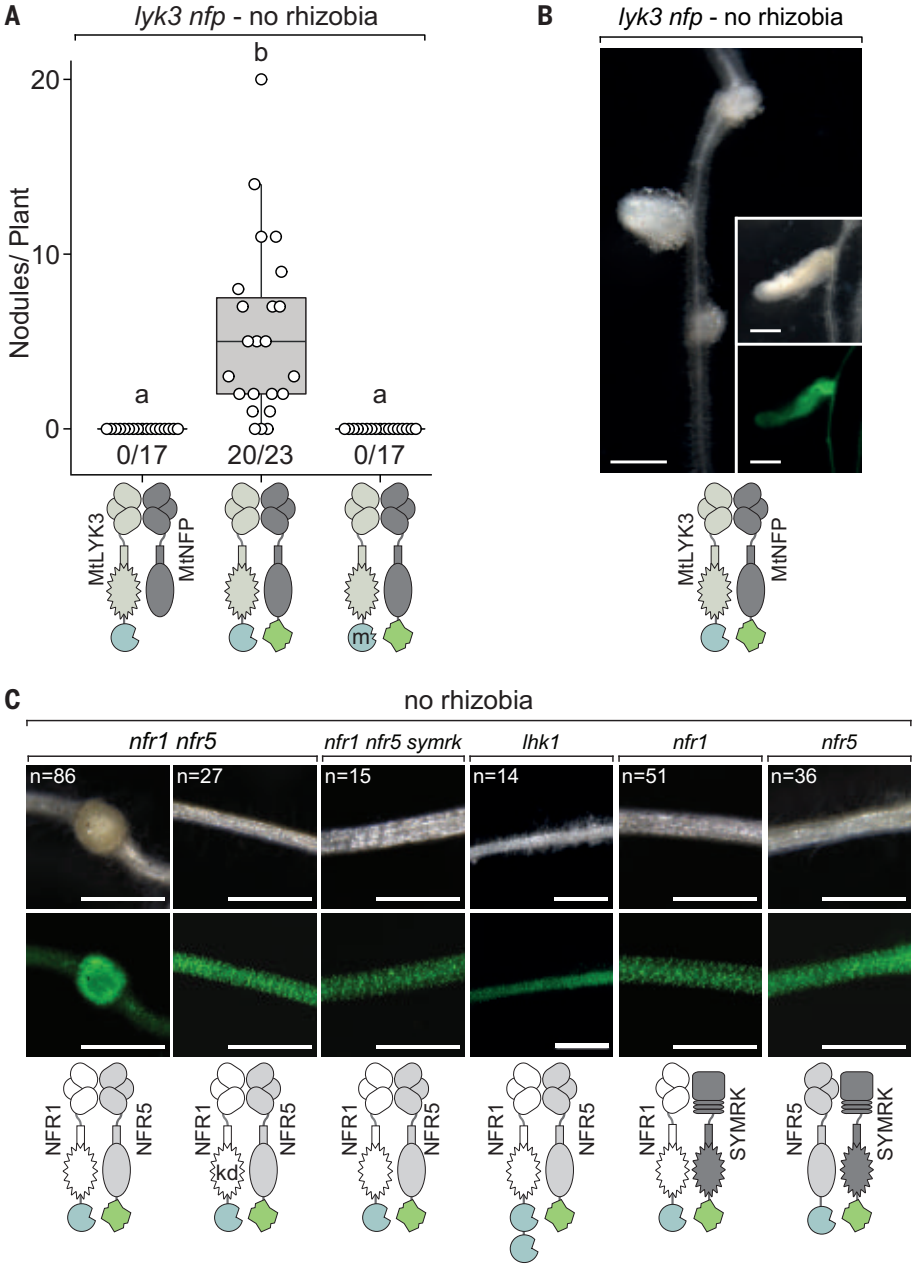


Fig. 2. The core symbiotic signaling complex activating organogenesis is conserved. (**A**) Number of nodules formed on *Medicago lyk3 nfp* double mutant plants expressing the indicated constructs in the absence of rhizobia. Nodules were counted 9 weeks after hairy root transformation. Circles represent individual plants. Numbers below the boxplots specify number of nodulated plants out of the total number of plants. The letter "m" represents the mutated nanobody with no binding to GFP. Lowercase letters indicate significant differences between samples [analysis of variance (Kruskal-Wallis) and post-hoc analysis (Dunn's test), $P < 0.05$]. (**B**) Representative brightfield and YFP fluorescence (transformation control) images of *Medicago* roots, showing the nodulation phenotype in the absence of rhizobia. Scale bar: 1 mm. (**C**) Representative brightfield and YFP fluorescence (transformation control) images of *Lotus* roots expressing NFR1-LaG16 with NFR5-GFP, kinase-dead (kd) NFR1-LaG16 with NFR5-GFP, NFR1-LaG16 with SYMRK-GFP or NFR5-LaG16 with SYMRK-GFP. *n* = total number of plants. Scale bars: 1 mm.

with three bulky tryptophan residues in the nanobody paratope. In contrast to functional LaG16, mutated LaG16 (LaG16m) lost the ability to bind GFP-tagged NFR5 in tobacco pull-down assays (Fig. 1G). Furthermore, when LaG16m is fused to NFR1, it fails to initiate signaling leading to organogenesis (Fig. 1F), demonstrating that NFR1-NFR5 receptor association is essential for signal transduction. To explore whether signaling is enhanced by assembling more receptors, we next coupled two LaG16 nanobodies in series to either NFR1 or NFR5, separated by a 40-residue-long (~130 Å) flexible linker which should sterically allow one receptor to associate with two co-receptors (Fig. 1F and fig. S2). We see no difference in nodule formation when one or two nanobodies are used, indicating that a single association event between NFR1 and NFR5 receptors is sufficient to initiate nodule organogenesis (Fig. 1F and fig. S2).

The core signaling complex is conserved in legumes

To understand whether the heteromeric receptor complex corresponds to the activated receptor state in other legume-rhizobia symbioses, we performed analogous experiments in *Medicago truncatula* (*Medicago*), in which we associated the two Nod factor receptors LYK3 and NFP (*9*, *16*) using nanobodies. We observe that a heteromeric LYK3-NFP receptor complex is sufficient to initiate organogenesis in *Medicago* in the absence of external signals (Fig. 2, A and B), demonstrating that the triggering event for signal activation leading to organogenesis is the association of a conserved core receptor complex comprising one NFR1/LYK3 receptor and one NFR5/NFP receptor. To understand whether symbiotic signaling is exclusively dependent on the physical assembly of the receptors or if it also requires a phosphorylation event, we introduced a point mutation in NFR1 impairing its kinase function (*17*). We find that plants expressing the kinase-dead NFR1 fail to form nodules (Fig. 2C), showing that the protein kinase activity mediates downstream signaling. The SYMRK receptor containing a malectin-like domain and leucine-rich repeats in the ectodomain was shown to interact with NFR5 and to induce spontaneous nodules when overexpressed (*18*). To decipher the composition and functional requirements of the core complex, we tested nanobody-assembled NFR1-NFR5 signaling in *nfr1 nfr5 symrk* triple mutant plants. The transformed roots show no nodulation (Fig. 2C), verifying the dependence of signaling on the common symbiotic pathway and confirming the vital downstream function of SYMRK. To investigate whether it is possible to drive symbiotic signaling from SYMRK independently of the formed NFR1-NFR5 complex, we next used nanobodies to associate SYMRK to either NFR1 or NFR5 (Fig. 2C). We observe that neither the SYMRK-NFR1 nor the SYMRK-NFR5 complex can drive organogenesis (Fig. 2C), showing that the essential role of SYMRK in nodulation requires either a ligand- or nanobody-induced NFR1-NFR5 complex. To verify that the activated receptor complex operates through the known organogenesis pathways, we next explored its dependency on cytokinin signaling (*19*, *20*). We find that

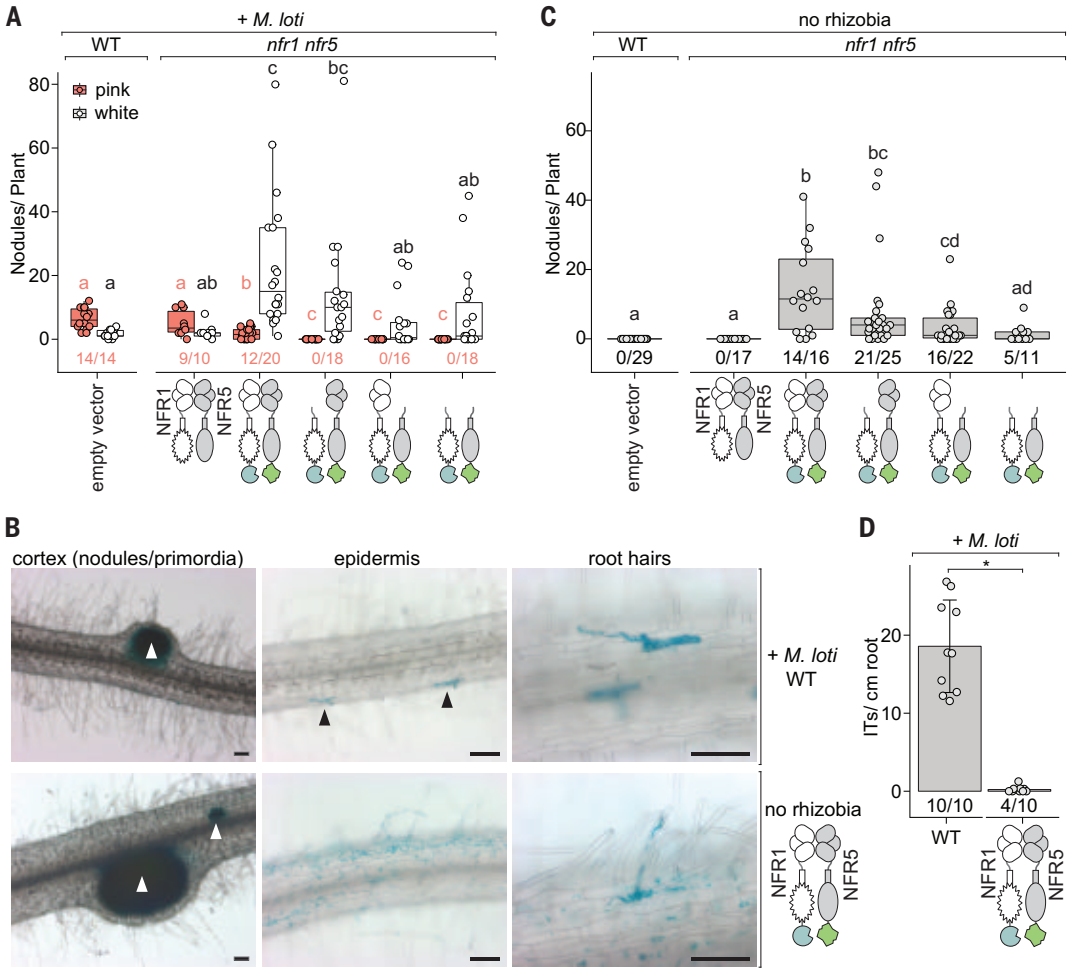


Fig. 3. Signaling activates both cortical and epidermal programs. (**A**) Number of nodules formed on WT and *nfr1 nfr5* double mutant plants expressing the indicated constructs. Nodules were counted 6 weeks after inoculation with *M. loti*. Red numbers below boxplots specify the number of plants with pink (infected) nodules out of the total number of plants. (**B**) Representative images of *pNIN*:GUS expression in WT plants inoculated with *M. loti* and non-inoculated roots of stable *Lotus* lines expressing NFR1-LaG16 and NFR5-GFP. Triangles pinpoint cells expressing GUS. For the stable line root hair image, three images were overlaid. Scale bars: 0.1 mm. (**C**) Number of nodules formed on WT and *nfr1 nfr5* double mutant plants expressing the indicated constructs. Nodules were counted 9 weeks after hairy root transformation. Numbers below the boxplots specify number of nodulated plants out of the total number of plants. In (A) and (C), lowercase letters indicate significant differences between samples [analysis of variance (Kruskal-Wallis) and post-hoc analysis (Dunn's test), $P < 0.05$]. Note that the empty vector and complementation controls are the same as in Fig. 1F. (**D**) Number of infection threads (ITs) formed per cm of root at 10 days post infection with *M. loti* in WT and stable *Lotus* lines expressing NFR1-LaG16 and NFR5-GFP. Error bars represent standard deviation. For statistical analysis a Kruskal-Wallis test was performed, $*P < 0.001$. Numbers below the bars specify the number of plants with ITs out of the total number of plants.

the activated NFR1-NFR5 complex is unable to initiate organogenesis in the *lhk1* loss-of-function cytokinin receptor mutant (Fig. 2C). Together, the data show that the NFR1-NFR5 complex is essential to initiate nodule organogenesis and that the downstream signaling cascade requires the symbiotic pathway.

Activated complex initiates both cortical and epidermal programs

We next investigated the effect of the nanobody-induced NFR1-NFR5 complex on the infection program by inoculating plants with *Mesorhizobium loti* (*M. loti*) bacteria, the symbiont of *Lotus*. We monitored infection by *M. loti* through the appearance of pink, leghemoglobin-producing nodules on the roots. Although an increased number of white nodules form on plants expressing the nanobody-associated NFR1-NFR5 complex, we see that formation of infected pink nodules is reduced compared with plants expressing the native receptors (Fig. 3A and fig. S3). This suggests that the activated NFR1-NFR5 complex initiates parallel developmental programs for both infection and organogenesis. But compared with organogenesis, infection is hampered by the presence of the nanobody-driven heteromeric NFR1-NFR5 state. We find that plants expressing the nanobody-induced complexes of either NFR1-SYMRK or NFR5-SYMRK are efficiently infected by rhizobia (fig. S4), suggesting that only the activated NFR1-NFR5 complex affects the infection program. Next, we investigated whether rhizobia that use an intercellular infection mechanism and circumvent the epidermal root hair infection program can colonize nodules induced by the nanobody-mediated receptor complexes. For this, we analyzed infection with IRBG74, which uses an intercellular, crack-entry mechanism (*21*, *22*). The transgenic *nfr1 nfr5* mutant roots form only white, uninfected nodules (fig. S5), suggesting that both infection modes are impaired when NFR1 and NFR5 are part of a nanobody-bound complex. To understand how downstream processes are affected by the presence of the nanobody-activated receptor complex, we created stable *Lotus* lines expressing GFP-tagged NFR5 and LaG16-tagged NFR1. The stable lines were generated in the *nfr1 nfr5* double mutant background with a β-glucuronidase (*GUS*) reporter gene driven by the *NIN* promoter. Looking at expression of the transcription factor *NIN*, which orchestrates both the epidermal infection program and the cortical organogenesis program, we can evaluate the effect of the activated receptor state on both programs (*23*). We observe nodule organogenesis independent of rhizobia 7 to 10 days after germination in three independent stable lines. GUS staining reveals *NIN* activation in nodules as well as in dividing cortical cells, which mark sites of emerging nodule primordia (Fig. 3B). This cortical *NIN* activation is very similar to the native condition observed for wild-type (WT) plants inoculated with *M. loti* (Fig. 3B). In the wild type, we see that the epidermal infection program is tightly

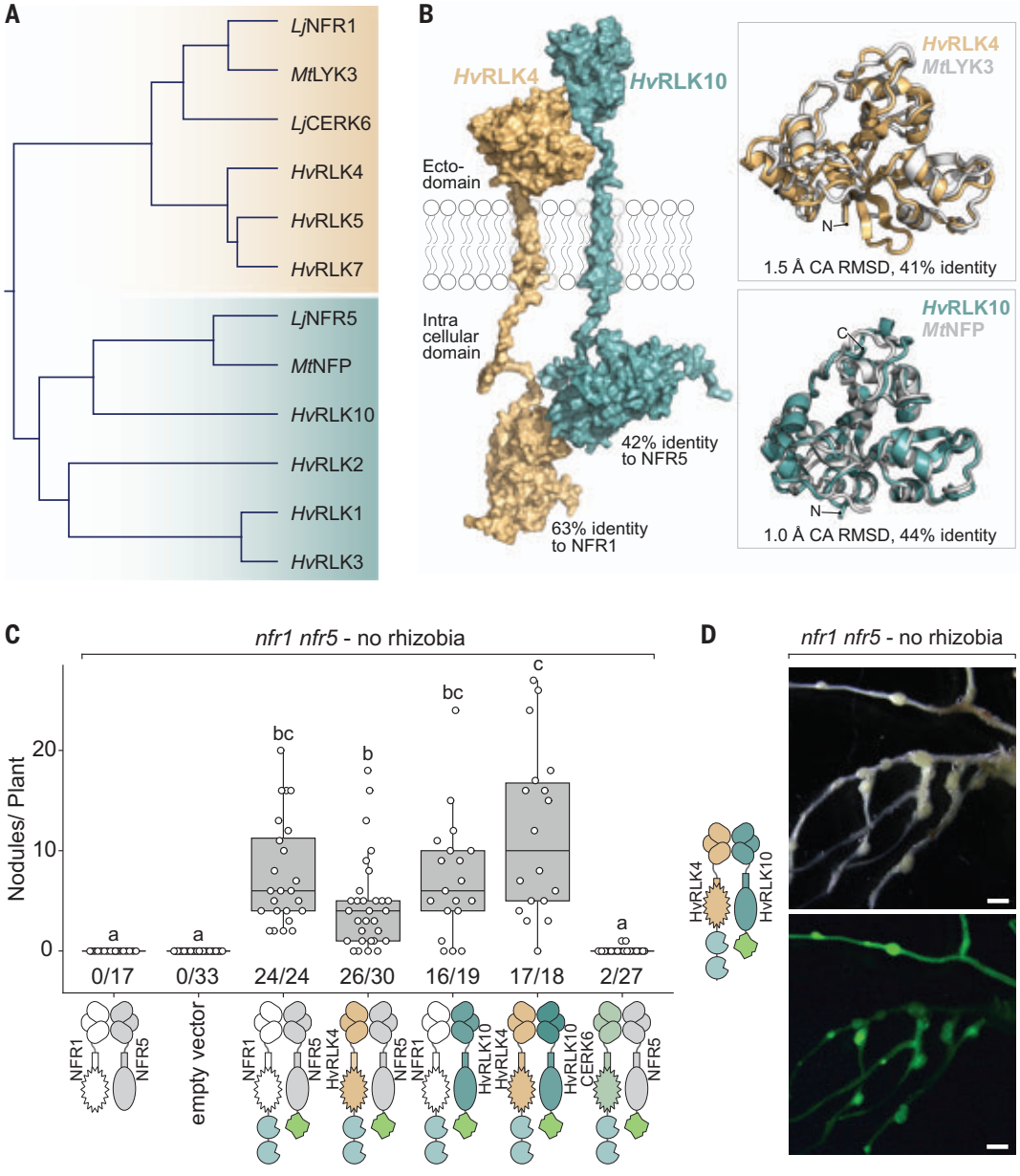


Fig. 4. Barley receptors can activate the symbiotic signaling pathway. (**A**) Phylogeny of selected barley (*Hv*), *Lotus* (*Lj*) and *Medicago* (*Mt*) full-length LysM receptors. Yellow and teal shaded backgrounds highlight NFR1- and NFR5-type receptors, respectively. (**B**) Structural modelling of barley RLK4 and RLK10. The amino acid sequence identities compared with the intracellular parts of *Lotus* NFR1 and NFR5 are indicated. Superposition of the modelled ectodomains of barley RLK4 and RLK10 compared with the crystal structures of *Medicago* LYK3 (PDB-ID: 6XWE) and NFP (PDB-ID: 7AU7), respectively. C-α root-mean-square deviations (RMSD) are indicated in angstroms (Å). (**C**) Number of nodules formed on *nfr1 nfr5* double mutant plants expressing the indicated constructs in the absence of rhizobia. Nodules were counted 9 weeks after hairy root transformation. Numbers below the boxplots specify number of nodulated plants out of the total number of plants. Lowercase letters indicate significant differences between samples [analysis of variance (Kruskal-Wallis) and post-hoc analysis (Dunn's test), $P < 0.05$]. Note that the complementation control is the same as in Fig. 1F. (**D**) Phenotype of nodules formed in *nfr1 nfr5* roots expressing barley RLK4-LaG16 with RLK10-GFP are depicted to the left. YFP fluorescence serves as a transformation control. Scale bars: 1 mm.

regulated, as *NIN* is expressed only in a few epidermal root hair cells. By contrast, plants expressing the nanobody-driven NFR1-NFR5 complex show broad *NIN* expression in numerous epidermal cells along the root (Fig. 3B). This wider, less-regulated epidermal *NIN* activation might interfere with the coordinated timing of organogenesis and infection. We quantified infection efficiencies by comparing WT plants with the stable *Lotus* lines expressing the NFR1-NFR5 complex 10 days after inoculation with *M. loti*. We observe a strong reduction in the number of infection threads formed when nanobodies are used to constitutively activate NFR1-NFR5 (Fig. 3D). Because both organogenesis and infection are initiated in transgenic roots expressing the nanobody-driven receptor complex, we took the opportunity to ask whether certain receptor domains contribute differently to the two developmental processes. For this, we engineered nanobody-mediated receptor complexes lacking the ectodomains of NFR1, NFR5, or both and expressed them in transgenic *Lotus* roots. We find that all ectodomain-lacking versions of the complexes are still able to activate organogenesis, albeit with a reduced efficiency compared with full-length receptors (Fig. 3C). By contrast, removal of any ectodomains results in complete loss of infection by *M. loti* as no infected pink nodules appear (Fig. 3A). These data show that the infection program is dependent on functional Nod factor-recognizing ectodomains, whereas organogenesis signaling can be mediated solely from the intracellular NFR1-NFR5 domains. Furthermore, our findings demonstrate that it is possible to bypass the ligand-induced signaling function conferred by the ectodomains to induce the cortical organogenesis program.

Cereal receptors can activate root nodule organogenesis

The ability to develop nitrogen-fixing root nodules is restricted to a single clade of flowering plants (*24*). However, most plants, including cereals, can form arbuscular mycorrhizal symbiosis by activating the common symbiosis pathway that is shared with root nodule symbiosis in legumes (*25*, *26*). This dichotomy indicates that the LysM receptors from cereals are unable to activate downstream signaling leading to nodulation. To test this, we analyzed the barley genome and found seven LysM

receptor-like kinases (RLKs) belonging to the NFR1 and NFR5 families (Fig. 4A). Phylogenomic analysis and structural modeling of these barley RLKs revealed that RLK4 and RLK10 show the greatest similarity to NFR1/LYK3 and NFR5/NFP, respectively (Fig. 4, A and B, and fig. S8) (*11*, *12*, *27*). We first tested the ability of these individual barley RLKs to functionally complement the *nfr1* and *nfr5* single or *nfr1 nfr5* double mutants when expressed in *Lotus* from the native *Nfr1* and *Nfr5* promoters and terminators. Despite the similarity, RLK4 and RLK10 are unable to functionally complement NFR1 and NFR5 (fig. S6 and table S1). We then asked whether nanobody-driven association of *Lotus* NFR5 with barley RLK4 expressed from the native *Nfr1* promoter and terminator induces nodule formation in *nfr1 nfr5* double mutant plants. We observe that the RLK4-NFR5 complex leads to nodule organogenesis independent of rhizobia (Fig. 4C). Likewise, nodules form when barley RLK10 is in a nanobody-mediated complex with *Lotus* NFR1 (Fig. 4C). Moreover, we find that the barley RLK4-RLK10 receptor complex associated through nanobodies leads to nodule organogenesis independent of the Nod factor receptors in both *Lotus* and *Medicago* (Fig. 4, C and D, and fig. S7). Barley RLK4 and RLK10 fail to support infection as no pink nodules appear when plants are inoculated with *M. loti* (fig. S6). Together, these results demonstrate that LysM receptors from barley—which diverged from *Lotus* 200 million years ago (*28*)—have retained the ability to activate the symbiotic pathway leading to root nodule organogenesis in legumes. We next investigated whether the *Lotus* chitin receptor CERK6 (*29*)—the closest homolog to *Lotus* NFR1 and in the same receptor family as barley RLK4—can support nodulation. *Lotus* CERK6 shares 62% sequence identity with *Lotus* NFR1 (table S1), but only 2 out of 27 plants form a single nodule when CERK6 and NFR5 are associated through nanobodies in the *nfr1 nfr5* mutants (Fig. 4C). Because RLK4 shares only 51% sequence identity with *Lotus* NFR1 (table S1), signal activation cannot be solely explained by sequence similarity. Although the molecular details that enable certain receptors to activate organogenesis remain elusive, these results show that signaling specificity is ensured by receptor complex composition.

In this study, we developed a synthetic approach to investigate the composition and activation of multicomponent and branched receptor systems. Our nanobody technique for driving complex formation, together with other approaches such as chemical control dimerization (*30*), will be applicable when studying other cell surface receptor systems. We used the nanobody-based approach to address a longstanding question in the field of plant-microbe interactions: namely, what constitutes an activated LysM receptor signaling complex. We demonstrate that NFR1 and NFR5 form the core receptor complex initiating root nodule symbiosis in *Lotus*. Our data match earlier findings showing that root hair deformation and calcium influx as the initial responses to rhizobia are maintained in the *symrk* mutant but abolished in the *nfr1* and *nfr5* mutants (*13*, *31*, *32*), suggesting that the NFR1-NFR5 complex has a signaling function independent of SYMRK. The data support that the first critical event for signal activation involves assembly of a core NFR1-NFR5 complex, which then interacts with SYMRK. We show that although the heteromeric NFR1-NFR5 complex activates the epidermal program essential for infection by rhizobia, nanobody-driven activation leads to inefficient colonization by rhizobia. A tight spatial and temporal synchronization between the infection program and the nodule organogenesis program is essential for efficient root nodule symbiosis. *NIN* is critical for both infection and organogenesis and continuous signaling from NFR1-NFR5 through nanobody activation and misregulation of *NIN* likely compromises the synchronization of bacterial attachment, root hair curling, and infection thread progression with the formation of nodule primordia leading to inefficient infection. Our data show that the intracellular kinase domains are necessary and sufficient for organogenesis signaling when artificially assembled using nanobodies whereas functional receptor ectodomains are important for infection. The main function of the NFR1 and NFR5 ectodomains is to perceive Nod factors, which naturally allow synchronization between receptor complex activation and infection and organogenesis. The ectodomains may serve additional functions like interaction scaffolds to mark the site of infection. We demonstrate that barley RLK4 and RLK10 receptors have the capacity to activate symbiotic signaling in legumes, suggesting that the ability to activate organogenesis is deeply rooted in the LysM receptor family, and was co-opted by plants in the nitrogen-fixing clade for root nodule symbiosis. Evidence for this conserved signaling function in barley suggests routes for engineering biological nitrogen-fixation into cereals for more sustainable food production in the future.

REFERENCES AND NOTES

1. E. B. Madsen *et al.*, *Nature* **425**, 637–640 (2003).
2. A. Broghammer *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **109**, 13859–13864 (2012).
3. C. Zipfel, G. E. D. Oldroyd, *Nature* **543**, 328–336 (2017).
4. L. Buendia, A. Girardin, T. Wang, L. Cottret, B. Lefebvre, *Front. Plant Sci.* **9**, 1531 (2018).
5. H. Kaku *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **103**, 11086–11091 (2006).
6. R. Willmann *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **108**, 19824–19829 (2011).
7. J. He *et al.*, *Mol. Plant* **12**, 1561–1576 (2019).
8. F. Feng *et al.*, *Nat. Commun.* **10**, 5047 (2019).
9. E. Limpens *et al.*, *Science* **302**, 630–633 (2003).
10. S. Radutoiu *et al.*, *Nature* **425**, 585–592 (2003).
11. Z. Bozsoki *et al.*, *Science* **369**, 663–670 (2020).
12. K. Gysel *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2111031118 (2021).
13. S. Stracke *et al.*, *Nature* **417**, 959–962 (2002).
14. J. R. Ingram, F. I. Schmidt, H. L. Ploegh, *Annu. Rev. Immunol.* **36**, 695–715 (2018).
15. P. C. Fridy *et al.*, *Nat. Methods* **11**, 1253–1260 (2014).
16. B. B. Amor *et al.*, *Plant J.* **34**, 495–506 (2003).
17. E. B. Madsen *et al.*, *Plant J.* **65**, 404–417 (2011).
18. M. K. Ried, M. Antolín-Llovera, M. Parniske, *eLife* **3**, e03891 (2014).
19. L. Tirichine *et al.*, *Science* **315**, 104–107 (2007).
20. J. D. Murray *et al.*, *Science* **315**, 101–104 (2007).
21. S. P. Cummings *et al.*, *Environ. Microbiol.* **11**, 2510–2525 (2009).
22. J. Montiel *et al.*, *Plant Physiol.* **185**, 1131–1147 (2021).
23. L. Schauser, A. Roussis, J. Stiller, J. Stougaard, *Nature* **402**, 191–195 (1999).
24. D. E. Soltis *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **92**, 2647–2651 (1995).
25. C. Kistner *et al.*, *Plant Cell* **17**, 2217–2229 (2005).
26. G. V. Radhakrishnan *et al.*, *Nat. Plants* **6**, 280–289 (2020).
27. X. R. Li *et al.*, *Nat. Commun.* **13**, 6421 (2022).
28. K. H. Wolfe, M. Gouy, Y. W. Yang, P. M. Sharp, W. H. Li, *Proc. Natl. Acad. Sci. U.S.A.* **86**, 6201–6205 (1989).
29. Z. Bozsoki *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **114**, E8118–E8127 (2017).
30. S. Kim *et al.*, *Mol. Plant* **14**, 1379–1390 (2021).
31. H. Miwa, J. Sun, G. E. Oldroyd, J. A. Downie, *Mol. Plant Microbe Interact.* **19**, 914–923 (2006).
32. L. H. Madsen *et al.*, *Nat. Commun.* **1**, 10 (2010).

ACKNOWLEDGMENTS

The authors thank F. Pedersen and M. K. Sørensen for plant care at the greenhouse facility; I. Jensen for help with the statistical analysis, and A. F. Nielsen and T. G. Fitzgerald for proofreading the manuscript. **Funding:** This work was supported by the Bill and Melinda Gates Foundation and the UK government's Foreign, Commonwealth and Development Office (FCDO) through the Engineering Nitrogen Symbiosis for Africa project (ENSA; OPP11772165), the Danish Council for Independent Research (9040-00175B), the Molecular Mechanisms and Dynamics of Plant-microbe interactions at the Root-Soil Interface project (InRoot), supported by the Novo Nordisk Foundation grant NNF19SA0059362, and the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (834221). **Author contributions:** H.R. and C.K. performed molecular cloning, in planta functional analyses and analyzed the data with help from H.J., D.L., M.N., M.K., D.A., N.d.J., L.H.M., and H.L. N.B.A. performed the co-immunoprecipitation experiment. S.B.H. and K.R.A. developed the NFR5 nanobody. K.R.A. and S.R. conceived the experiments, analyzed data, and coordinated studies. K.R.A., S.R., H.R, C.K., and J.S. wrote the manuscript with input from all authors. **Competing interests:** Some findings in this manuscript are considered for patent application. **Data and materials availability:** All data are available in the main text or the supplementary materials. Gene constructs used in this study are available from S.R. and K.R.A. under a material agreement with Aarhus University, Denmark. **License information:** Copyright © 2023 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. https://www.sciencemag.org/about/science-licenses-journal-article-reuse

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.ade9204
Materials and Methods
Figs. S1 to S8
Table S1
References (*33*–*54*)
MDAR Reproducibility Checklist
Data S1 and S2

Submitted 19 September 2022; accepted 15 December 2022
10.1126/science.ade9204

QUANTUM SIMULATION

A superconducting quantum simulator based on a photonic-bandgap metamaterial

Xueyue Zhang[1,2]†, Eunjong Kim[1,2]†, Daniel K. Mark[3], Soonwon Choi[3], Oskar Painter[1,2,4]*

Synthesizing many-body quantum systems with various ranges of interactions facilitates the study of quantum chaotic dynamics. Such extended interaction range can be enabled by using nonlocal degrees of freedom such as photonic modes in an otherwise locally connected structure. Here, we present a superconducting quantum simulator in which qubits are connected through an extensible photonic-bandgap metamaterial, thus realizing a one-dimensional Bose-Hubbard model with tunable hopping range and on-site interaction. Using individual site control and readout, we characterize the statistics of measurement outcomes from many-body quench dynamics, which enables in situ Hamiltonian learning. Further, the outcome statistics reveal the effect of increased hopping range, showing the predicted crossover from integrability to ergodicity. Our work enables the study of emergent randomness from chaotic many-body evolution and, more broadly, expands the accessible Hamiltonians for quantum simulation using superconducting circuits.

Realizing a scalable architecture for quantum computation and simulation is a central goal in the field of quantum information science. Although architectures with nearest-neighbor (NN) coupling between quantum particles on a lattice are prevalent, quantum systems with long-range interactions can realize a richer set of computational tasks and physical phenomena (*1*–*4*). For instance, in the case of gate-based quantum computation, coupling beyond the NN level enables nonlocal gate operations between qubits, which can reduce the overhead of quantum algorithms and lift the restrictions on code rate and distance of local-interaction–based quantum error-correcting codes (*5*, *6*). In the case of analog quantum simulation, the inclusion of long-range interactions can alter the behavior of otherwise integrable many-body systems (*7*, *8*), resulting in quantum chaotic dynamics, which is at the root of such topics as quantum thermalization (*9*) and quantum information scrambling (*10*). Furthermore, control over the range of lattice connectivity grants access to different physical regimes and the crossover between them, such as in many-body quantum phase transitions (*11*–*13*) and the hydrodynamics of nonequilibrium quantum states (*14*).

For engineered quantum systems consisting of interacting quantum particles on a lattice, it is often challenging to scale to larger lattice sizes while maintaining a high degree of lattice connectivity and single-site control. One common approach, developed for trapped-ion and neutral-atom systems, is to use resonant modes of either vibrational (*1*) or optical (*15*) cavities as a quantum bus for mediating interactions between the internal states of atoms across the lattice. Similar schemes have been adopted in superconducting quantum circuits, realizing systems as large as 20 qubits with all-to-all coupling via a common microwave cavity (*16*). Increasing the number of lattice sites in this case, however, leads to either parasitic coupling arising from dense placement of sites in a fixed-volume cavity or frequency-crowding effects stemming from the increased spectral density of cavity modes when increasing the cavity size (*17*).

An alternative approach for connecting quantum particles on a lattice is to construct a quantum bus from an intrinsically extensible structure, such as a waveguide. Along this direction, engineered photonic-bandgap waveguides have been proposed as a quantum bus that simultaneously protects quantum particles from radiative damping through the waveguide while allowing for extended-range lattice connectivity (*18*). The waveguide-bus concept has been investigated in the context of many-body simulation with cold atoms coupled to engineered nanophotonic waveguides (*18*, *19*), and recent experiments have explored qubit-photon bound states in superconducting quantum circuits with microwave photonic-bandgap waveguides (*20*–*24*). However, the realization of a scalable many-body quantum simulator, with single-site quantum-particle control and a high level of lattice connectivity, has remained an open challenge.

We demonstrate a scalable many-body quantum simulator consisting of a one-dimensional (1D) lattice of superconducting transmon qubits coupled to a common metamaterial waveguide. This system provides both tunable-range connectivity between qubits and full single-site control and state measurement of individual qubits. The waveguide acts both as a bus for mediating exponentially decaying long-range interactions between qubits and as a Purcell filter enabling multiplexed, rapid readout of the qubit states with high fidelity. This system realizes an extended version of the Bose-Hubbard model with tunable hopping range and on-site interaction. Using our ability to efficiently collect measurement outcomes from many-body quench dynamics—enabled by the fast experimental repetition rate of our system—we perform direct analysis of outcome statistics to learn Hamiltonian parameters in situ and study the effect of hopping range on the evolution of randomness across the system. Specifically, we observe a distribution of outcome bit-string probabilities reflecting the ergodic nature of the Hamiltonian with long-range hopping. This result experimentally confirms the expectation from quantum chaos for interacting many-particle systems, highlighting the connection between ergodic unitary dynamics and its effective statistical description in terms of random matrix theory (*25*).

Metamaterial-based quantum simulator

The backbone of the many-body quantum simulator in this work is a metamaterial waveguide formed from a chain of lumped-element inductor-capacitor (LC) microwave resonators. The waveguide can be described by a generic model (Fig. 1A) of a 1D cavity array with NN coupling t (*26*, *27*). The corresponding dispersion relation (Fig. 1B) is given by $\omega_k = \omega_c + 2t\cos(kd)$, exhibiting a passband centered around the cavity frequency ω_c with a bandwidth of $4t$, where k is the wave vector and d is the lattice constant of the array. The bandgap at frequencies below $\omega_{e,-} = \omega_c - 2t$ (above $\omega_{e,+} = \omega_c + 2t$) is denoted as the lower (upper) bandgap, abbreviated as LBG (UBG). Inside the bandgaps, the off-resonant coupling between a bare quantum emitter and the waveguide modes gives rise to an emitter-photon bound state (*28*) whose photonic tail is localized around the emitter. Localization follows a spatial profile $(\mp 1)^{\Delta x} e^{-|\Delta x|/\xi}$ in the LBG or UBG (*27*), where Δx is the displacement in the number of unit cells from the emitter and ξ is the localization length controlled by the detuning Δ between the band-edge frequency and the transition frequency of the bound state. The overlap of two bound states results in photon-mediated coupling with a range covering multiple unit cells, i.e., long-range coupling, exhibiting a greater strength and a more extended range ξ at a smaller detuning $|\Delta|$ (Fig. 1C).

The metamaterial waveguide consists of a 42–unit-cell array of capacitively coupled lumped-element microwave resonators (Fig. 1, D and E) and is equipped at both ends with engineered tapering sections, designed to reduce the impedance mismatch to external 50-ohm

[1]Thomas J. Watson, Sr., Laboratory of Applied Physics and Kavli Nanoscience Institute, California Institute of Technology, Pasadena, CA 91125, USA. [2]Institute for Quantum Information and Matter, California Institute of Technology, Pasadena, CA 91125, USA. [3]Center for Theoretical Physics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA. [4]AWS Center for Quantum Computing, Pasadena, CA 91125, USA.
*Corresponding author. Email: opainter@caltech.edu
†These authors contributed equally to this work.

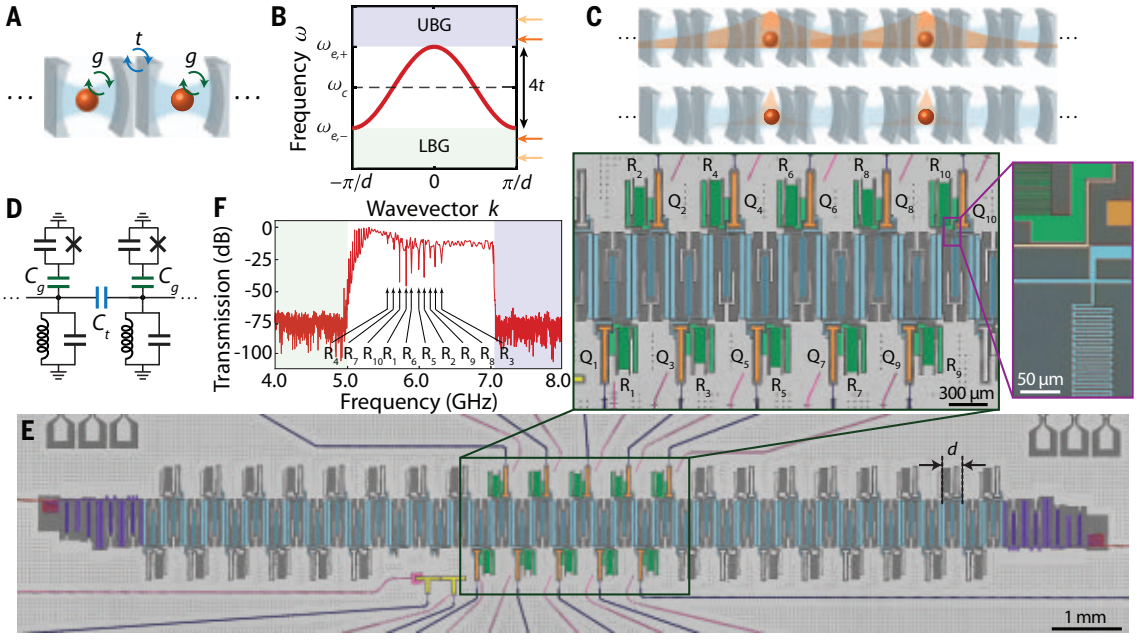


Fig. 1. Metamaterial-based quantum simulator. (**A**) Schematic showing a 1D array of coupled cavities with nearest-neighbor coupling t. Each cavity is occupied by a quantum emitter (orange ball) with coupling g to the cavity. (**B**) Dispersion relation of the coupled cavity array in (A) with a passband between $\omega_{e,\pm}$ centered at ω_c (bandwidth of $4t$). The LBG (UBG) below (above) the passband is shaded in green (purple). (**C**) Top (Bottom): Cartoon of two emitter-photon bound states at small (large) detuning $|\Delta|$, indicated by dark (light) orange arrows in (B), exhibiting an extended (restricted) spatial range and large (small) photonic component in the bound states. (**D**) Electrical circuit realization of (A) with capacitively coupled LC resonators and transmon qubits corresponding to the cavity array and the quantum emitters, respectively. The coupling capacitors are color coded in accordance with (A). (**E**) Optical micrograph (false colored) of the fabricated quantum simulator with 42 metamaterial resonators (lattice constant $d = 292$ μm) colored blue connected to input-output ports (red) via tapering sections (purple). Ten qubits (Q_i, colored orange), controlled by individual charge drive lines (pink) and flux bias lines (dark blue), and their readout resonators (R_i, colored green) couple to the 10 inner unit cells of the metamaterial waveguide with a zoomed-in view in the left inset. Detailed view of the coupling region is shown in the right inset. Two auxiliary qubits (yellow) are not used in this experiment. (**F**) Transmission spectrum through the metamaterial waveguide (red curve) with black arrows indicating the 10 resonances of the readout resonators R_i.

input-output ports at frequencies lying within the passband of the waveguide (*24*, *29*). Each of the middle 10 metamaterial resonators (unit cells labeled by i = 1 to 10) couples to a transmon qubit (*30*), which serves as the quantum emitter. Individual addressing of each qubit is achieved by excitation (XY control) from a charge drive line and frequency tuning (Z control) from a flux bias line. Dispersive qubit readout is enabled by capacitively coupling each qubit Q_i to a compact readout resonator R_i, which itself is then coupled to the metamaterial resonator of the same unit cell. The entire metamaterial and transmon qubit system (the device) is fabricated by using evaporated thin-film aluminum on a high-resistivity silicon substrate, with fabrication procedures detailed in (*21*, *29*). Further details of the device modeling and the experimental setup used to measure and test the device are discussed in supplementary text I and II, respectively.

This realization of the device enables qubit readout using the passband of the metamaterial waveguide with built-in protection against Purcell decay channels (supplementary text III). The transmission spectrum through the waveguide (Fig. 1F) shows a passband ranging from $\omega_{e,-}/2\pi \approx 5.01\,\mathrm{GHz}$ to $\omega_{e,+}/2\pi \approx 7.08\,\mathrm{GHz}$ with ripples smaller than 8 dB near the center. The extinction ratio of the transmission between that measured in the passband and that measured in the bandgaps is greater than 65 dB, with a sharp transition in the transmission occurring within 100 MHz of the band edges. In the middle of the passband, resonances associated with the readout resonators are observed between 5.574 and 6.328 GHz. The average decay rate of 10 readout resonators is $\overline{\kappa_{R_i}}/2\pi = 11.8$ MHz, enabling fast, high-fidelity multiplexed readout while maintaining a low level of readout cross-talk. For details of readout methods and characterization, refer to supplementary text IV.

Bose-Hubbard model with long-range hopping

The spatially extended bound-state excitations, formed between transmon-qubit excitations and waveguide photons of the metamaterial-waveguide bus, creates a lattice of interacting microwave photons (*31*). This quantum system is described by an extended version of the 1D Bose-Hubbard model with tunable long-range hopping and on-site interaction. Specifically, each bound state formed from qubit Q_i, inheriting the level structure of an anharmonic oscillator from a transmon qubit (*30*), serves as a bosonic site with local site energy $\epsilon_i = \omega_{01,i}$ and the on-site interaction $U_i = \omega_{12,i} - \omega_{01,i}$. Here, $\omega_{01,i}$ and $\omega_{12,i}$ are the transition frequencies of the bound state on site Q_i from its ground state $|0\rangle$ to the first excited state $|1\rangle$ and that from the first to the second excited state $|2\rangle$, respectively. In addition, the long-range hopping $J_{i,j}$ is enabled by the overlap between a pair of qubit-photon bound states on sites Q_i and Q_j. The Hamiltonian of this model that captures the basic processes mentioned above can be written as

$$\hat{H}/\hbar = \sum_{i,j} J_{ij}\hat{b}_i^\dagger \hat{b}_j + \sum_i \frac{U_i}{2}\hat{n}_i(\hat{n}_i - 1) + \sum_i \epsilon_i \hat{n}_i \quad (1)$$

where $\hat{b}_i^\dagger$ ($\hat{b}_i$) is the creation (annihilation) operator and $\hat{n}_i \equiv \hat{b}_i^\dagger \hat{b}_i$ is the number operator on site Q_i. The parameters of the Hamiltonian realized in this simulator can be learned through experiments enabled by the precise, single-site–level control over qubits.

We measure the on-site interaction U_i (Fig. 2A) by performing spectroscopy of ω_{12} after initializing Q_i in its first excited state $|1\rangle$. From within either bandgap, $|U_i|$ decreases as ω_{01} approaches the closest band edge owing to dressing from the passband modes of the metamaterial (*22*)—i.e., the Lamb shift. In the UBG, a wide tuning range of $|U_i|$ is achievable from the strong hybridization between the $|1\rangle$-$|2\rangle$ transition and the band-edge modes at $(\omega_{01} - \omega_{e,+})/2\pi < 300$ MHz. The magnitude of hopping $|J_{i,j}|$ is measured from vacuum Rabi oscillations between sites Q_i and Q_j by

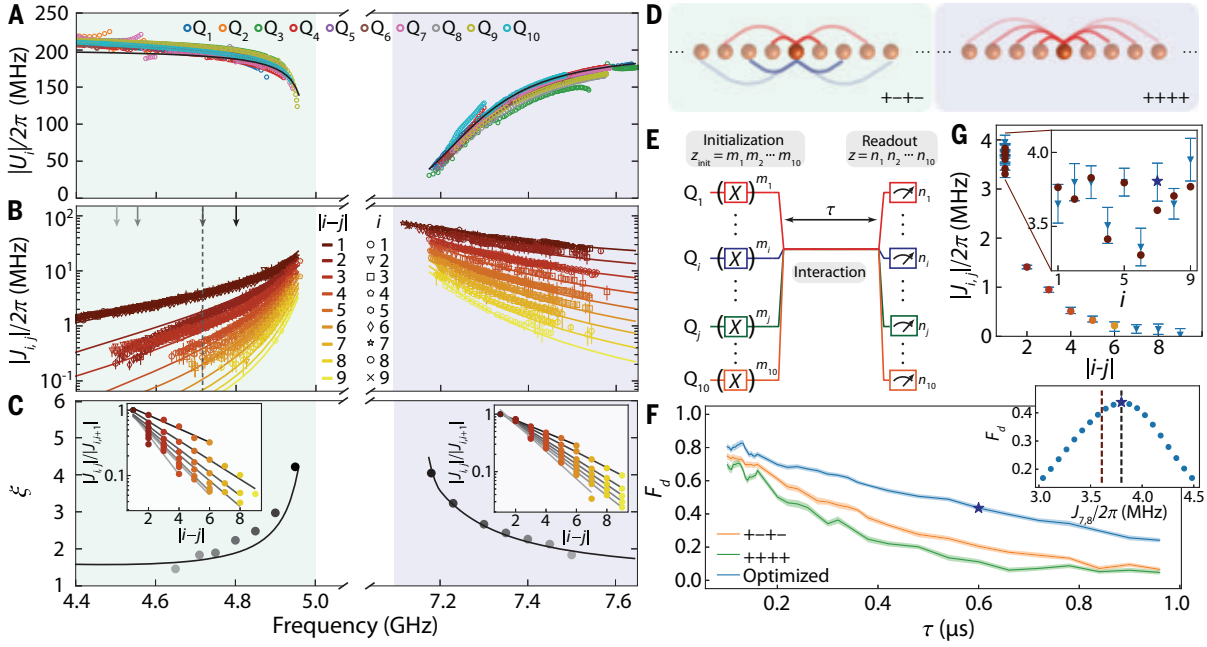


Fig. 2. Hamiltonian learning. (**A**) On-site interaction $|U_i|$ versus frequency ω_{01} with measured values indicated by colored circles. (**B**) Hopping amplitude $|J_{ij}|$ versus frequency with experimental data shown as markers (error bars indicate a standard deviation). Colors represent the distance $|i-j|$ between sites. Four gray-scale arrows specify frequencies in Fig. 3. (**C**) Localization length ξ extracted by fitting the exponential decay of measured hopping rate [results of polynomial fitting of datapoints in (B)] as a function of distance $|i-j|$ (insets) at a few different frequencies. The darkness of a marker matches a fitting curve in the inset at the same frequency. In (A) to (C), green (purple) shading on the left (right) corresponds to LBG (UBG), and theory curves (solid) are obtained from numerical calculations using an identical circuit model. (**D**) Left (Right): Cartoon illustrating hopping from a site inside the LBG (UBG) with positive or negative sign represented in red or blue and the amplitude represented by opacity, with alternating (all positive) signs denoted as $+-+-$ ($++++$). (**E**) Pulse sequence for many-body evolution. Ten sites, Q_1 to Q_{10}, are initialized (X gate) in a bit-string z_{init} at their idle frequencies, then tuned to resonance for time τ during the interaction stage, followed by a site-resolved single-shot readout at their idle frequencies to obtain a final bit-string z. (**F**) Many-body fidelity estimator F_d at $\omega_{01}/2\pi = 4.72$ GHz versus evolution time τ. The F_d curves assume $J_{i,j}$'s from two-qubit measurement indicated by the dashed line in (B) with alternating signs (orange) and all positive signs (green), and from the numerical optimization (blue). The shading corresponds to a standard deviation for 40 randomly chosen z_{init}'s in the five-excitation sector. Inset: F_d at $\tau = 0.6$ μs versus $J_{7,8}$ with optimized parameters on the remaining $J_{i,j}$'s, where the brown and the black dashed lines indicate $J_{7,8}$ extracted from (B) and the numerical optimization, respectively. (**G**) Comparison of $|J_{i,j}|$ interpolated from two-qubit measurements (colored circles) and from numerical optimization (blue triangles with error bars for 68% confidence interval), corresponding to the orange and the blue curves in (F), respectively. All nearest-neighbor hopping $|J_{i,i+1}|$'s are shown (details in the inset), whereas for larger hopping distances $|i-j| > 1$ only the average values are indicated. The navy stars represent the maximum F_d at $\tau = 0.6$ μs in (F) and the optimized $J_{7,8}$ in the insets of (F) and (G).

initializing one site with a π-pulse and tuning ω_{01} of both sites on resonance for a duration τ with fast flux pulses. For a fixed distance $|i-j|$, $|J_{i,j}|$ increases with a decreasing $|\Delta|$, resulting from larger photonic components of the bound states (Fig. 2B). Compared to the LBG, the UBG exhibits larger $|J_{i,j}|$ at the same $|\Delta|$, owing to a stronger coupling g of the bare qubits to the metamaterial at higher frequencies and the breakdown of the tight-binding cavity array model in the circuit realization (Fig. 1D; also see supplementary text I). At a specific ω_{01}, $|J_{i,j}|$ decreases exponentially as a function of distance $|i-j|$, resembling the profile of the photonic tail in a qubit-photon bound state (Fig. 2C). From fitting the exponential decay curve we extract the localization length ξ, which ranges from $\xi = 1.4$ to 4.2, with the largest localization length occurring at the smallest achievable band-edge detunings. For even smaller detunings, the eigenstate of the two interacting bound states merges into the passband and becomes radiative to the waveguide.

Many-body Hamiltonian learning

Beyond the above single- and two-qubit measurements, we perform in situ many-body characterization of Hamiltonian parameters (*32*, *33*), which are otherwise hard to access. For example, the sign of the hopping term $J_{i,j}$ inherits the spatial profile of the photonic component of the bound states. In the case of the bound states in the UBG, the sign of the hopping terms are all uniform (positive), whereas for bound states in the LBG, the hopping terms alternate sign as the distance between lattice sites increases by one (Fig. 2D). This is due to the photonic component of the bound state behaving as a defect mode inside the bandgap, exhibiting a spatial profile resembling the wave vector at the nearest band edge ($k = 0$ at the upper band-edge and $k = \pi/d$ at the lower band edge). Although insignificant in measurements involving only two lattice sites, the sign of the hopping terms does alter the many-body dynamics of the system. Here, we use a many-body fidelity estimator F_d proposed in (*33*) to reveal this information. This fidelity estimator, which closely tracks the true many-body fidelity, is obtained for ergodic quench evolution of simple initial states (supplementary text VI).

We follow the sequence described in Fig. 2E to perform the many-body quench evolution. The sequence consists of preparing a set of five randomly chosen sites in their first excited state, followed by using flux pulses to align ω_{01} of all 10 sites for time τ, and then finally performing site-resolved single-shot measurement on all lattice sites to obtain a 10-bit string $z = n_1 n_2 \cdots n_{10}$. The many-body fidelity estimator F_d is calculated by comparing bit-string

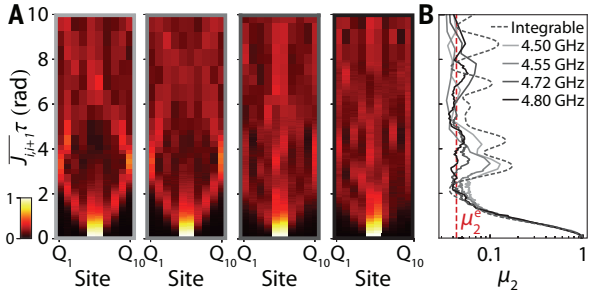


Fig. 3. Two-particle quantum walk with increasing hopping range. (**A**) Evolution of the population $\langle \hat{n}_i \rangle$ on sites Q_1 to Q_{10} as a function of normalized evolution time $\overline{J_{i,i+1}}\tau$. The system is initialized in $z_{init} = 0000110000$ and the evolution occurs at $\omega_{01}/2\pi = 4.50$, 4.55, 4.72, and 4.80 GHz with the longest evolution times of 904, 781, 430, and 200 ns from left to right. (**B**) The second moment μ_2 as a function of normalized evolution time $\overline{J_{i,i+1}}\tau$. Results calculated from the data in (A) are shown as solid curves with gray scales corresponding to frames in (A) and arrows in Fig. 2B. Result from numerical simulation of the integrable Hamiltonian is shown as the dotted curve, and μ_2^e for a generic ergodic system is indicated by the red dashed line.

statistics of repeated measurements with numerical simulation of the evolution assuming a set of Hamiltonian parameters in Eq. 1. The maximum F_d is achieved at the parameter values closest to the Hamiltonian realized in the experiment. The fast repetition rate of this experiment enables us to perform a large number of measurements (1.6×10^5 in total), reducing statistical error and increasing sensitivity to small Hamiltonian parameter variations (see supplementary text V for details of qubit control, pulse sequence, and repetition rate).

We compare F_d at $\omega_{01}/2\pi = 4.72$ GHz using three different parameter sets for $J_{i,j}$ in Fig. 2F: a first set with amplitudes derived from the two-qubit experiments in Fig. 2B assuming alternating signs (Fig. 2D, left); a second set with the same amplitudes as the first but all positive signs (Fig. 2D, right); and a third set of optimized parameter values that maximize F_d. The optimized hopping terms are restricted to be real-valued, with independent $J_{i,i+1}$ for each i = 1 to 9 and $J_{i,j}$ for each distance $|i-j| > 1$ (all qubit pairs of the same distance having the same $J_{i,j}$). An alternating sign of $J_{i,j}$ with distance is favored, yielding a higher many-body fidelity compared to hopping terms with all positive signs. This is further evidenced by the alternating signs of the resulting optimized parameter set. Although we find small differences between the set of optimized hopping amplitudes and those from the two-qubit experiments with alternating signs (Fig. 2G), F_d of the optimized parameter set is markedly better. The sensitivity of F_d to the hopping terms is highlighted in the inset of Fig. 2F, where the variation of the fidelity versus $J_{7,8}$ is shown. For details of the F_d calculation and parameter optimization, refer to supplementary text VI.

Ergodic many-body dynamics with long-range hopping

We now use the platform to study the effect of long-range hopping on the many-body dynamics. Specifically, the ergodicity of the 1D Bose-Hubbard model in the hardcore limit ($|U/J| \gg 1$) depends on the range of hopping, which exhibits integrable behavior with NN hopping, and chaotic behavior with long-range hopping. We study this crossover with various hopping ranges and investigate the resulting dynamics using both conventional one- and two-site correlators, and the statistics of the global bit-strings resulting from qubit-state measurement outcomes across the lattice. This latter technique is particularly useful in identifying universal signatures of ergodicity and the effect of decoherence at long evolution times. The crossover between integrable and ergodic dynamics can be qualitatively visualized by a two-particle quantum walk (*34*–*36*) with initial excitations on sites Q_5 and Q_6 using the sequence shown in Fig. 2E. The measured quantum walk at a few different ω_{01}'s indicated by arrows in Fig. 2B is shown (Fig. 3A) as a function of normalized evolution time $\overline{J_{i,i+1}}\tau$, where $\overline{J_{i,i+1}}$ is the average NN hopping rate (the corresponding numerical simulations are provided in supplementary text VII, showing that the quantum walk patterns are not visibly affected by decoherence). The excitation wave packets smear over the system when ω_{01} is close to the band-edge frequency. More quantitatively, this trend can be probed by computing the probability p_z of measuring a certain bit-string z in the two-excitation sector at evolution time τ. For a generic ergodic Hamiltonian, the second moment $\mu_2 \equiv \sum_z p_z^2$ (*25*), which reflects the probability fluctuations, converges to $\mu_2^e = 2/(D+1)$ after initial evolution (*33*) owing to the chaotic nature of its quantum dynamics ($D = 45$ is the dimension of the two-excitation Hilbert space). No such convergence is expected in an integrable Hamiltonian owing to revivals associated with ballistic propagation of wave packets. As an example, we show in Fig. 3B the results from the spin-1/2 XY model obtained from modifying the Hamiltonian in Eq. 1 by keeping only NN hopping terms in the hardcore limit. When ω_{01} is closer to the band edge, the measured second moment deviates from the simulated integrable result and converges to μ_2^e at an earlier normalized evolution time $\overline{J_{i,i+1}}\tau$ consistent with the breaking of integrability due to the extended hopping range. With $|U/J| > 36$ for all the measurements illustrated in Fig. 3, finite on-site interactions of the Bose-Hubbard model play a negligible role in the breaking of integrability (supplementary text VIII).

To further probe this ergodic nature of Hamiltonian with long-range hopping, we use the experimental evolution at $\omega_{01}/2\pi = 4.72$ GHz as an example. At a short time (τ = 16 ns), the excitations remain in their initial sites. This is visualized for a quantum walk with initial excitations on sites Q_5 and Q_6 in the left panel of Fig. 4C (evolution of population $\langle \hat{n}_i \rangle$) and in the bottom left panel of Fig. 4D (two-site correlator $\langle \hat{n}_i \hat{n}_j \rangle$). The histogram $P(p_z)$ of experimentally measured bit-string probabilities $\{p_z\}$ at this early evolution stage (Fig. 4B, left) shows a distribution with a tail of large p_z values, giving a large μ_2 (Experiment curve in Fig. 4A). This is associated with an insufficient scrambling of the initially localized quantum information. At an intermediate time ($\tau = 360$ ns), the excitations are more spread out over the entire 1D lattice (Fig. 4D, middle left panel), forming a "speckle" pattern with site-to-site fluctuation that is associated with quantum interference. The quantitative signatures of this speckle pattern manifest in the histogram $P(p_z)$ following the Porter-Thomas (PT) distribution (*37*) (Fig. 4B, middle) and in the second moment μ_2 settling to the ergodic value μ_2^e. The PT distribution results from the randomness in the distribution of wavefunction magnitudes, which is predicted by Berry's conjecture (*38*) stating that the single-particle eigenstates of a chaotic system behave like random superpositions of plane waves. Similarly, in the many-body settings, the distribution of wave function magnitudes across basis states also follows the PT distribution. Our observation is the first experimental verification of this many-body version of Berry's conjecture in a Bose-Hubbard system, whose extension in the thermodynamic limit provides the modern theory of quantum thermalization such as eigenstate thermalization hypothesis (*39*, *40*). This draws a connection between quantum many-body chaos and random matrix theory, leading to a deeper understanding of the randomness in many-body dynamics (*32*). Distinct from the randomness inherent in random circuits (*25*, *41*), the randomness in our case originates from the ergodicity of the time-independent Hamiltonian. In contrast to the experimental results, theoretical calculations at the same evolution time using the integrable Hamiltonian show aggregated excitations on a few sites (Fig. 4E, middle) and the resulting larger value of μ_2 (Integrable theory curve in Fig. 4A). This comparison highlights the effect of long-range hopping in probing the ergodic regime.

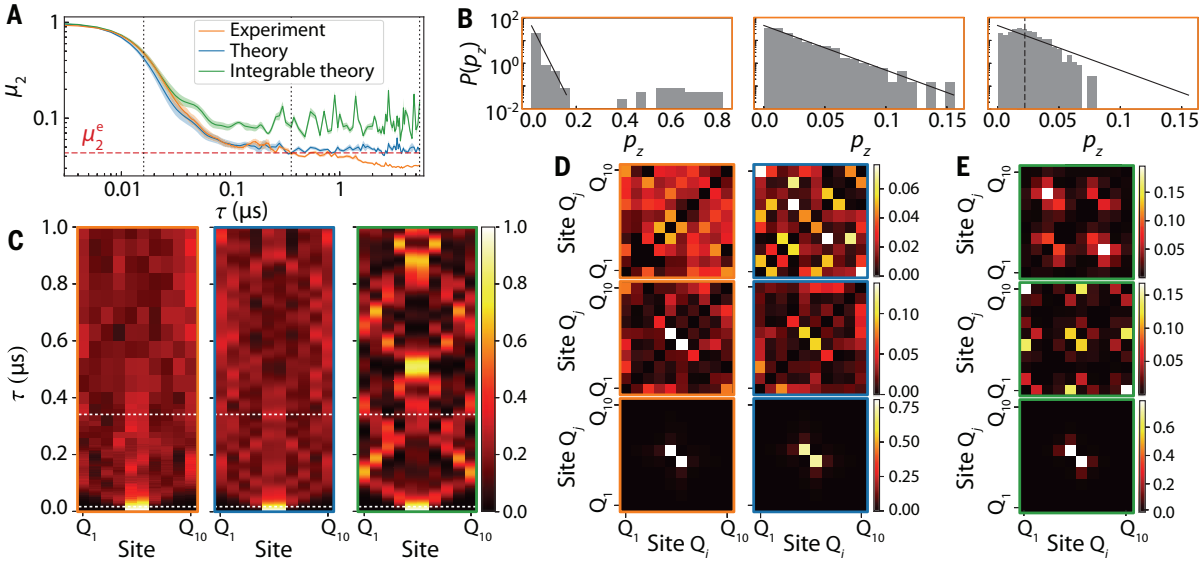


Fig. 4. Ergodic many-body dynamics with long-range hopping at 4.72 GHz. (**A**) Second moment μ_2 as a function of evolution time τ in our system from the experiment (orange) and the theory with the optimized parameter set in Fig. 2G (blue), compared to theoretical predictions of the integrable model (green). The shading on each curve corresponds to a standard deviation of the mean second moment for 20 randomly chosen initial bit-strings z_{init} in the two-particle sector, and the red dashed line represents the ergodic value μ_2^e. (**B**) Density histogram $P(p_z)$ of the distribution of experimental bit-string probabilities $\{p_z\}$ with the 20 initializations z_{init}'s at evolution times $\tau = 16$ ns, 360 ns, and 5.4 μs from left to right [indicated by the dotted lines in (A)]. The solid lines show the PT distribution, and the dashed line in the right plot shows the value $p_z = 1/D$ of a classical uniform distribution. (**C**) Evolution of the population $\langle \hat{n}_i \rangle$ on sites Q_1 to Q_{10} as a function of time τ with $z_{\text{init}} = 0000110000$ in the cases of experiment, theory, and integrable theory from left to right. The white dashed lines at the bottom (in the middle) indicate $\tau = 16$ ns (360 ns). (**D** and **E**) Two-site correlator $\langle \hat{n}_i \hat{n}_j \rangle$ with $z_{\text{init}} = 0000110000$ at evolution times $\tau = 16$ ns, 360 ns, and 5.4 μs from bottom to top in the cases of experiment [left column of (D)], theory [right column of (D)], and integrable theory (E).

In addition, we study the impact of decoherence by juxtaposing the measurement results and the decoherence-free theoretical calculation using the optimal learned Hamiltonian with long-range hopping. Before the evolution time of $\tau \approx 1$ μs, the two cases agree in the second moment μ_2 (Experiment and Theory curves in Fig. 4A), the quantum walk population (Fig. 4C, left and middle), and the two-site correlator (Fig. 4D, middle panels), suggesting that these results are not affected by decoherence. After a long evolution time ($\tau = 5.4$ μs, larger than the averaged Ramsey coherence time $\overline{T_{2,i}^*} = 1.16$ μs), the second moment of the two cases deviates from one another, and the experimental speckle pattern begins to wash out compared to the theoretical modeling (Fig. 4D, top panels). Another probe of the decoherence is the histogram $P(p_z)$ of the measured bit-string probabilities (Fig. 4B, right). Here, the histogram deviates from the PT distribution, narrows substantially, and approaches a uniform distribution corresponding to a completely decohered, maximally mixed state. Additional numerical simulations of μ_2 and $P(p_z)$ for ergodic and integrable systems can be found in supplementary text VIII.

Conclusion and outlook

Our many-body quantum simulator is based on a 1D lattice of transmon qubits connected together using a superconducting metamaterial, which exhibits photonic bandgaps that protect qubit-photon bound states from decay and a transmission passband used for high-fidelity multiplexed qubit-state readout. Furthermore, the metamaterial plays the role of a scalable photonic bus to mediate tunable long-range coupling between qubit-photon bound states. This system of interacting bound states realizes a Bose-Hubbard model with long-range hopping. We characterize the system using conventional single- and two-qubit measurements along with a sample-efficient many-body Hamiltonian learning protocol. Lastly, we study the many-body quench dynamics of the system versus the range of the lattice hopping, revealing the ergodic nature of the extended Bose-Hubbard model, as distinct from its NN-coupling counterpart. The major challenge in probing long-time quantum evolution in our experiment is the short Ramsey coherence time $\overline{T_{2,i}^*} = 1.16$ μs, limited by flux-noise–induced dephasing. Incorporating a single refocusing pulse has been shown to increase the coherence time to $\overline{T_{2E,i}} = 5.64$ μs at the single-qubit level (supplementary text II). The extended quantum evolution times enabled by further dynamical decoupling, combined with the tunable-range coupling investigated in this work, provide valuable opportunities to explore nonequilibrium dynamics with or without coupling to an environment and quantum phases of matter in the presence of frustration.

REFERENCES AND NOTES

1. K. Mølmer, A. Sørensen, *Phys. Rev. Lett.* **82**, 1835–1838 (1999).
2. V. D. Vaidya *et al.*, *Phys. Rev. X* **8**, 011002 (2018).
3. A. Periwal *et al.*, *Nature* **600**, 630–635 (2021).
4. D. Bluvstein *et al.*, *Nature* **604**, 451–456 (2022).
5. S. Bravyi, D. Poulin, B. Terhal, *Phys. Rev. Lett.* **104**, 050503 (2010).
6. N. Delfosse, M. E. Beverland, M. A. Tremblay, arXiv:2109.14599 [quant-ph] (2021).
7. P. Jurcevic *et al.*, *Nature* **511**, 202–205 (2014).
8. P. Richerme *et al.*, *Nature* **511**, 198–201 (2014).
9. R. Nandkishore, D. A. Huse, *Annu. Rev. Condens. Matter Phys.* **6**, 15–38 (2015).
10. M. K. Joshi *et al.*, *Phys. Rev. Lett.* **124**, 240505 (2020).
11. R. Islam *et al.*, *Science* **340**, 583–587 (2013).
12. R. Landig *et al.*, *Nature* **532**, 476–479 (2016).
13. S. Ebadi *et al.*, *Nature* **595**, 227–232 (2021).
14. M. K. Joshi *et al.*, *Science* **376**, 720–724 (2022).
15. S.-B. Zheng, G.-C. Guo, *Phys. Rev. Lett.* **85**, 2392–2395 (2000).
16. Q. Guo *et al.*, *Nat. Phys.* **17**, 234–239 (2021).
17. Y. P. Zhong *et al.*, *Nat. Phys.* **15**, 741–744 (2019).
18. J. S. Douglas *et al.*, *Nat. Photonics* **9**, 326–331 (2015).
19. D. E. Chang, J. S. Douglas, A. González-Tudela, C.-L. Hung, H. J. Kimble, *Rev. Mod. Phys.* **90**, 031002 (2018).
20. Y. Liu, A. A. Houck, *Nat. Phys.* **13**, 48–52 (2017).
21. M. Mirhosseini *et al.*, *Nat. Commun.* **9**, 3706 (2018).
22. N. M. Sundaresan, R. Lundgren, G. Zhu, A. V. Gorshkov, A. A. Houck, *Phys. Rev. X* **9**, 011021 (2019).
23. M. Scigliuzzo *et al.*, *Phys. Rev. X* **12**, 031036 (2022).
24. E. Kim *et al.*, *Phys. Rev. X* **11**, 011015 (2021).
25. S. Boixo *et al.*, *Nat. Phys.* **14**, 595–600 (2018).
26. M. J. Hartmann, F. G. S. L. Brandão, M. B. Plenio, *Nat. Phys.* **2**, 849–855 (2006).
27. G. Calajó, F. Ciccarello, D. Chang, P. Rabl, *Phys. Rev. A* **93**, 033833 (2016).
28. S. John, J. Wang, *Phys. Rev. Lett.* **64**, 2418–2421 (1990).
29. V. S. Ferreira *et al.*, *Phys. Rev. X* **11**, 041043 (2021).
30. J. Koch *et al.*, *Phys. Rev. A* **76**, 042319 (2007).
31. A. A. Houck, H. E. Türeci, J. Koch, *Nat. Phys.* **8**, 292–299 (2012).
32. J. Choi *et al.*, arXiv:2103.03535 [quant-ph] (2021).

33. D. K. Mark, J. Choi, A. L. Shaw, M. Endres, S. Choi, arXiv:2205.12211 [quant-ph] (2022).
34. A. Peruzzo *et al*., *Science* **329**, 1500–1503 (2010).
35. P. M. Preiss *et al*., *Science* **347**, 1229–1233 (2015).
36. Z. Yan *et al*., *Science* **364**, 753–756 (2019).
37. C. E. Porter, R. G. Thomas, *Phys. Rev.* **104**, 483–491 (1956).
38. M. V. Berry, *J. Phys. Math. Gen.* **10**, 2083–2091 (1977).
39. J. M. Deutsch, *Phys. Rev. A* **43**, 2046–2049 (1991).
40. M. Srednicki, *Phys. Rev. E Stat. Phys. Plasmas Fluids Relat. Interdiscip. Topics* **50**, 888–901 (1994).
41. F. Arute *et al*., *Nature* **574**, 505–510 (2019).
42. Data and Simulation Code for, A superconducting quantum simulator based on a photonic-bandgap metamaterial, version 1, Zenodo (2022); https://doi.org/10.5281/zenodo.7502375.

ACKNOWLEDGMENTS

The authors thank A. Gorshkov, A. González-Tudela, D. Chang, O. Motrunich, R. Ma, F. Brandão, G. Refael, S. Meesala, V. Ferreira, G. Kim, A. Butler, and Z. Zheng for helpful discussions. We appreciate MIT Lincoln Laboratories for the provision of traveling-wave parametric amplifiers used for both spectroscopic and time-domain measurements in this work, and the AWS Center for Quantum Computing for the Eccosorb filters installed in the cryogenic setup for infrared filtering. We also thank the Quantum Machines team for technical support and discussions on the Quantum Orchestration Platform. **Funding:** This work was supported by the AFOSR Quantum Photonic Matter MURI (grant FA9550-16-1-0323), the DOE-BES Quantum Information Science Program (grant DE-SC0020152), the Institute for Quantum Information and Matter, an NSF Physics Frontiers Center (grant PHY-1125565) with support of the Gordon and Betty Moore Foundation, the Kavli Nanoscience Institute at Caltech, and the AWS Center for Quantum Computing. D.K.M. acknowledges support from the NSF QLCI program (2016245) and the DOE Quantum Systems Accelerator Center (contract no. 7568717). **Author contributions:** X.Z., E.K., D.K.M., S.C., and O.P. came up with the concept. X.Z. and E.K. planned the experiment, performed the device design and fabrication, and performed the measurements. X.Z., E.K., D.K.M., and S.C. analyzed the data. O.P. supervised the project. All authors contributed to the writing of the manuscript. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** Experimental data shown in the main text and supplementary materials, as well as the simulation code, are available in Zenodo (*42*). **License information:** Copyright © 2023 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. https://www.sciencemag.org/about/science-licenses-journal-article-reuse

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.ade7651
Materials and Methods
Supplementary Text
Figs. S1 to S14
Tables S1 and S2
References (*43*–*88*)

Submitted 6 September 2022; accepted 16 December 2022
10.1126/science.ade7651

ZEOLITES

A 3D extra-large-pore zeolite enabled by 1D-to-3D topotactic condensation of a chain silicate

Jian Li[1,2,3]*†‡, Zihao Rei Gao[2,4]†, Qing-Fang Lin[5,6]†, Chenxu Liu[6], Fangxin Gao[5], Cong Lin[2,3,7], Siyao Zhang[5], Hua Deng[8], Alvaro Mayoral[9,10], Wei Fan[11], Song Luo[11], Xiaobo Chen[12], Hong He[8,13], Miguel A. Camblor[4]*, Fei-Jian Chen[5,6]*, Jihong Yu[6]*

Zeolites are microporous silicates with a large variety of applications as catalysts, adsorbents, and cation exchangers. Stable silica-based zeolites with increased porosity are in demand to allow adsorption and processing of large molecules but challenge our synthetic ability. We report a new, highly stable pure silica zeolite called ZEO-3, which has a multidimensional, interconnected system of extra-large pores open through windows made by 16 and 14 silicate tetrahedra, the least dense polymorph of silica known so far. This zeolite was formed by an unprecedented one-dimensional to three-dimensional (1D-to-3D) topotactic condensation of a chain silicate. With a specific surface area of more than 1000 square meters per gram, ZEO-3 showed a high performance for volatile organic compound abatement and recovery compared with other zeolites and metal-organic frameworks.

The size of the molecules that can enter, diffuse, and react into zeolites is limited by the size of their pores (*1*–*4*), which are typically described as being "of molecular dimensions," meaning that they are the size of small molecules (<7 Å). For many applications, small pores enhance reaction and sorption selectivity (*5*), but for other applications such as processing large molecules from petroleum or sorption and reaction of organic pollutants, stable zeolites with larger pores are in demand (*6*). Natural and synthetic zeolites have a fully connected, three-dimensional (3D) network of corner-sharing SiO_4 tetrahedra that are tectosilicates or framework silicates (*7*), with Si occasionally substituted by other atoms. However, some zeolites are obtained in the form of 2D precursors (phyllosilicates or layered silicates) (*7*) that only become fully connected tectosilicate zeolites by condensation of their layers through a calcination procedure that is "topotactic" because it does not alter the layer topology (*8*–*10*). The condensing layers can be obtained by direct synthesis or by the disassembly of certain zeolites, as in the so-called ADOR (assembly-disassembly-organization-reassembly) process (*11*).

After several decades of extensive and systematic zeolite synthesis studies (*12*), there have still been no reported examples or predictions of a 3D zeolite obtained by condensation from a 1D precursor, either directly synthesized or obtained by disassembly of another zeolite. We report such a 1D-to-3D topotactic condensation from ZEO-2, a directly synthesized complex "zeolitic" chain silicate, into ZEO-3, a fully connected 3D extra-large-pore zeolite ("ZEO-*n*" refers to materials discovered and patented by the Anhui ZEO New Material Technology Co., China). This condensation does not alter the topology of the chain silicate, so it is topotactic. The resulting stable zeolite ZEO-3 exhibits very low density, a multidimensional system of interconnected extra-large pores (Fig. 1), and the presence in its structure of double four-membered ring units (D4Rs), which are small cubes of silica. For pure silica zeolites, this kind of unit is strained and up to now was believed to need a F^- anion near its center to be accessible for crystallization (*13*) because it has never been seen before in a silica zeolite synthesized without the use of F^- anions.

Synthesis and structure of the precursor chain silicate

We synthesized the complex chain silica zeolite precursor ZEO-2 using tricyclohexylmethylphosphonium ($C_{19}H_{36}P^+$, tCyMP) as an organic structure–directing agent (OSDA) from a gel of molar composition 1 SiO_2: 0.5 tCyMPOH: 10 H_2O heated at 175°C (see the supplementary

[1]Berzelii Center EXSELENT on Porous Materials, Department of Materials and Environmental Chemistry, Stockholm University, 10691 Stockholm, Sweden. [2]Anhui ZEO New Material Technology Co., Hefei 230071, China. [3]College of Chemistry and Molecular Engineering, Peking University, Beijing 100871, China. [4]Instituto de Ciencia de Materiales de Madrid, Consejo Superior de Investigaciones Científicas (ICMM-CSIC), 28049 Madrid, Spain. [5]Department of Chemistry, Bengbu Medical College, Bengbu 233030, China. [6]State Key Laboratory of Inorganic Synthesis and Preparative Chemistry, College of Chemistry, International Center of Future Science, Jilin University, Changchun 130012, China. [7]Department of Mechanical Engineering, The Hong Kong Polytechnic University, Kowloon, Hong Kong, China. [8]Center for Excellence in Regional Atmospheric Environment, Key Laboratory of Urban Pollutant Conversion, Institute of Urban Environment, Chinese Academy of Sciences, Xiamen 361021, China. [9]Instituto de Nanociencia y Materiales de Aragón (INMA), CSIC-Universidad de Zaragoza, 50009 Zaragoza, Spain. [10]Laboratorio de Microscopias Avanzadas (LMA-Universidad de Zaragoza), 50018 Zaragoza, Spain. [11]Department of Chemical Engineering, University of Massachusetts, Amherst, MA 01003, USA. [12]State Key Laboratory of Heavy Oil Processing, China University of Petroleum, Qingdao 266580, China. [13]State Key Joint Laboratory of Environment Simulation and Pollution Control, Research Center for Eco-Environmental Sciences, Chinese Academy of Sciences, Beijing 100085, China.
***Corresponding author. Email: jxpxlijian@pku.edu.cn (J.L.); macamblor@icmm.csic.es (M.A.C.); feijian@jlu.edu.cn (F.-J.C.); jihong@jlu.edu.cn (J.Y.)**
†These authors contributed equally to this work.
‡Present address: State Key Laboratory of Coordination Chemistry, School of Chemistry and Chemical Engineering, Nanjing University, Nanjing 210023, China.

materials). The structure of ZEO-2 was successfully solved ab initio by using eight continuous rotation electron diffraction (cRED) (*14*) datasets. The pure silicate ZEO-2 crystals display a needle-like morphology (fig. S1A) and have a *C*-centered monoclinic cell with a = 23.5465(7) Å, b = 24.7446(7) Å, c = 14.4024(4) Å, and β = 115.1974(9)° (tables S1 and S2 and fig. S2). ZEO-2 is a complex 1D chain silicate decorated with silanol and silanolate groups (Fig. 2A) that hold the structure together through numerous H bonds between adjacent chains (Fig. 2, B and C), with the tCyMP cations located in the interchain space (fig. S5). The cations were occluded intact, as demonstrated by ^{13}C and ^{31}P nuclear magnetic resonance (NMR) (fig. S6), and amount to 8.85 OSDA per unit cell according to C analysis (25.0 wt %). H bonds were observed in the ^{1}H magic-angle spinning (MAS) NMR spectrum as a broad resonance around 15.1 parts per million (ppm) (fig. S7), indicating a moderate to strong H bond (*15*) corresponding to O···O distances (*16*) of ~2.51 Å, in good agreement with the crystallographic distances of 2.47 to 2.52 Å.

The 1D pure silica chains in ZEO-2 are aligned along the [001] direction (Fig. 2A) and surrounded by four identical chains in the *ab* plane (Fig. 2C). At the edge of the ZEO-2 chain, four silanols or silanolates form a single four-membered ring (S4R) that face, slightly displaced, an identical S4R from the next chain, with H bonding along the [110] and [1-10] directions connecting adjacent chains (Fig. 2B). The high resolution of the ^{29}Si MAS NMR spectrum of ZEO-2 (Fig. 2D) reveals four Q^3 Si sites (–94.2, –95.8, –98.6, and –100.4 ppm) spanning a chemical shift range unprecedented for Q^3 in zeolites, which are more typically centered at around –102 ± 1 ppm. However, the values are well within the general Q^3 range in silicates (*17*). The spectrum also shows seven Q^4 Si sites (from –106.8 to –116.8 ppm) and is thus in good agreement with the crystallographic results (4 Q^3 and 5 Q^4 all with the same multiplicity plus 2 Q^4 with half multiplicity; table S5). $^{29}Si\{^{1}H\}$ cross-polarization (CP) MAS NMR spectroscopy proved the existence of those four Q^3 Si sites in ZEO-2 (Fig. 2D, top).

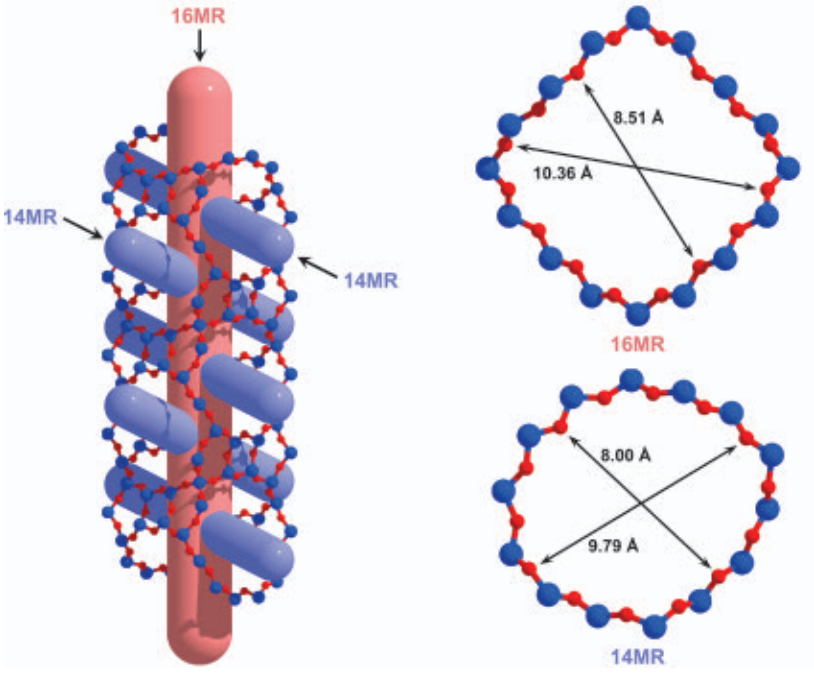


Fig. 1. The extra-large-pore system in ZEO-3. Shown is the 3D system of the interconnected extra-large-pore system (left) and the crystallographic pore size of ZEO-3 (right). The van der Waals radius of O (2 × 1.35 Å) has been subtracted.

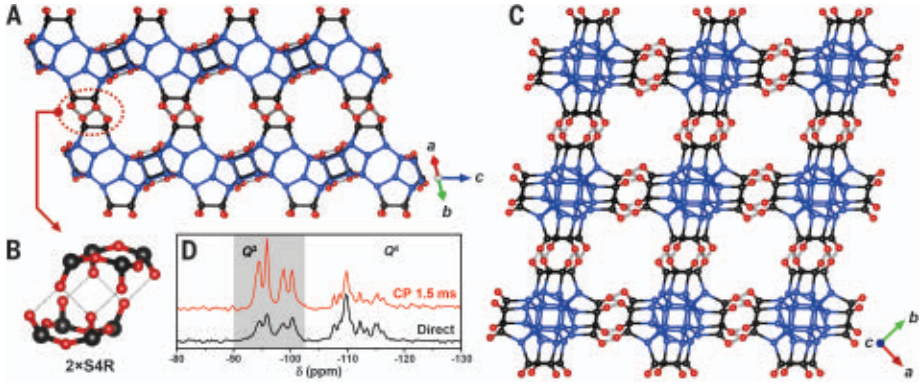


Fig. 2. Structure of the chain silicate ZEO-2. Only O atoms related to the subsequent condensation reaction are shown (small red spheres). Silicon atoms are shown as blue (always Q^4) or black (Q^3 in ZEO-2 converting into Q^4 in ZEO-3). (**A** to **C**) A chain of ZEO-2 (A) is H bonded (B) to four adjacent chains (C). (**D**) ^{29}Si MAS NMR spectrum (bottom) showing the resolution of the Q^3 and Q^4 silicon sites (n = 4 and 7 sites, respectively). The close proximity of Q^3 sites to H atoms is revealed in the $^{29}Si\{^{1}H\}$ CP MAS NMR spectrum by their relative intensity enhanced by polarization transfer from close protons at short contact time (top, 1.5 ms).

Condensation of 1D chains to a 3D extra-large-pore zeolite

Upon calcination in air to remove the OSDA (600°C, 3-hour ramp, 6-hour plateau), silanol groups in adjacent chains condensed into Si-O-Si bridges with H_2O elimination, resulting in the pure silica zeolite ZEO-3 (Fig. 3, A to C), which maintained the needle-like morphology (fig. S1B). The condensation occurred between 370 and 390°C (fig. S8), coincident with the removal of organics (fig. S15). Phosphorus residues were eliminated by washing with water within an autoclave at 100°C for 1 day or by reduction with H_2 (a mixture of H_2/N_2 with 10/90 volume ratio) from the as-made ZEO-2 sample at 600°C with a 2-hour ramp and a 6-hour plateau. The structure of ZEO-3 was also solved ab initio by cRED with five datasets (tables S2 and S3 and fig. S3). The unit cell of ZEO-3 shrank to a = 21.5046(8) Å, b = 21.2757(8) Å, c = 14.4638(4) Å, and β = 108.7196(1)° but maintained the same symmetry as ZEO-2, as well as the topology of the chain. Whereas a 17% contraction of the structure occurred along a and b axes, the c axis underwent only a marginal expansion of 0.4%. To obtain more accurate atomic positions, the structures of ZEO-2 and ZEO-3, including the position of the disordered tCyMP in ZEO-2, were subsequently Rietveld refined against synchrotron powder x-ray diffraction data (SPXRD) (figs. S4 and S5 and tables S4 to S10). All of the unit-cell data given above correspond to the refined structures. The final refined unit-cell compositions of ZEO-2 and ZEO-3 were $[Si_{80}O_{176}H_{24}]|(C_{19}H_{36}P)_8$ and $Si_{80}O_{160}$,

respectively (see the supplementary materials for details).

During thermal treatment, neighboring S4Rs in ZEO-2 connect to each other to form a D4R (Fig. 3A) by condensation of the terminal Si-OH groups, yielding the fully connected framework of ZEO-3 (Fig. 3, B and C). The condensed solid is a true, noninterrupted, 3D extra-large-pore zeolite. The channel system of ZEO-3 is 3D with 16- × 14- × 14-membered ring (MR) channels (Fig. 3, B and C) and full connectivity between channels (Fig. 1). The fulfillment of the condensation process is demonstrated by the ^{29}Si MAS NMR spectrum of ZEO-3, which reveals total condensation: All Si atoms are Q^4 sites (Fig. 3D, bottom) with a negligible amount of Q^3 that could be assigned to connectivity defects, as proven by the very-low-intensity enhancement by cross-polarization (Fig. 3D, middle and top) and Fourier transform infrared spectra from self-supported pellets (fig. S16).

The structural models obtained were fully corroborated by spherical aberration (Cs)–corrected scanning transmission electron microscopy (STEM) (Fig. 4 and fig. S10), with which a faint signal corresponding to the tCyMP (C and P) was also identified between the chains of ZEO-2 (Fig. 4A) in the place that after condensation will become 14-MR pores in ZEO-3. The visualization of ZEO-2 along [001] could not be obtained because it was hampered by the larger thickness (because this is the long needle direction) and the existence in the structure of atoms at different z levels that are displaced from one another along x and y directions. The 14- and 16-MR pores of ZEO-3 are clearly visible (Fig. 4, B and C) together with the smaller 4-, 5-, and 6-MR units in both materials (Fig. 4 and fig. S10).

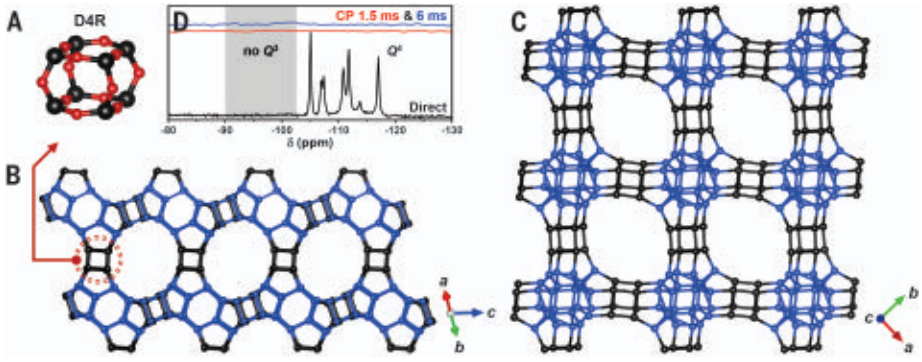


Fig. 3. 1D-to-3D topotactic condensation into the extra-large-pore framework silicate ZEO-3. (**A** to **C**) During calcination of ZEO-2, condensation of Q^3 sites through dehydroxylation connects two S4Rs to make a D4R (A), through which each chain is bonded to four adjacent chains, resulting in the extra-large-pore ZEO-3 with 14-MR (B) and 16-MR (C) channels. (**D**) The corresponding ^{29}Si MAS NMR spectra (bottom) are dominated by Q^4 sites with almost no Q^3 defects and thus little intensity enhancement in the ^{29}Si{^{1}H} CP MAS NMR under short (middle, 1.5 ms) or long (top, 6 ms) contact times.

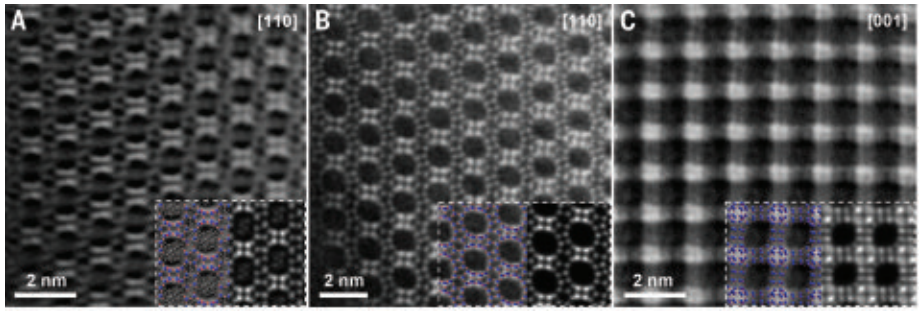


Fig. 4. Cs-corrected STEM visualization of the 1D silicate chain ZEO-2 and the 3D zeolite ZEO-3. (**A** to **C**) ZEO-2 along [110] projection (A) and ZEO-3 showing the extra-large 14-MR along the [110] projection (B) and the 16-MR along the [001] projection (C). All other rings (4-, 5-, and 6-MRs) are clearly visible. To facilitate image interpretation, the schematic models (Si atoms are shown in blue and O atoms in red) have been superimposed, and the simulated STEM images are also included in the right bottom inset.

The σ expansion of beta zeolite polymorphs

The details of the topology are shown in Fig. 5 and tables S11 and S12. The ZEO-2 chain is topologically identical to the one found in polymorph B of the beta zeolite, although in that zeolite, it is not an isolated chain but rather is embedded in the 3D framework. We use here the acronym BEB to refer to that polymorph, although this is not an accepted zeolite topology code. The chain in BEB and ZEO-3 is built by a large composite building unit (Fig. 5A) that is alternately rotated by +90° and −90° (Fig. 5B, dotted arrows). If the rotation were performed always in the same sense (that is, always by either +90° or −90°; Fig. 5C, dotted arrows), then a chiral chain would result, which is in fact embedded in the chiral BEA polymorph of the beta zeolite (space group $P4_122$ or $P4_322$). Thus, condensation of such a chain in a way similar to the condensation of ZEO-2 would result in a new chiral 16- × 14- × 14-MR hypothetical zeolite. ZEO-3 (Fig. 5D) and this hypothetical chiral zeolite (Fig. 5E) correspond to the σ expansion of polymorphs BEB (Fig. 5F) and BEA (Fig. 5G), respectively, of the beta zeolite (*18*, *19*). A minimization of both structures using the general utility lattice program (GULP) suggests that the energies of both ZEO-3 and σ-BEA are close to the energy-density relationship normally encountered in zeolites (fig. S17 and table S14). Because ZEO-3 has been realized, we believe that the hypothetical extra-large-pore σ-BEA zeolite might be a reasonable target for future studies.

Properties of ZEO-3

ZEO-3 is a stable, fully connected silicate zeolite containing 3D-interconnected pores opened only through extra-large windows. The crystallographic pore sizes of ZEO-3 are 10.36 × 8.51 Å and 9.79 × 8.00 Å for the 16- and 14-MR pores, respectively (Fig. 1). The 3D extra-large-pore nature of ZEO-3 resulted in a very low framework density (FD) value (12.76 tetrahedral atoms per 1000 Å^3). Compared with the other known stable, low-density (alumino) silicate zeolites, including **FAU**, **EMT**, ***BEA**, **BEC**, **ISV**, and **IWV** and the recently reported PST-2, PST-32 (*20*), and ZEO-1 (*6*), this value is the lowest and puts ZEO-3 as the crystalline silica polymorph with the most open framework (table S13). The calculated density of ZEO-3 is just 1.27 g/cm^3, less than half the density of quartz (2.65 g/cm^3) and near that of water. In fact, ZEO-3 breaks the observed tendency between the framework density and the size of the smallest rings in the zeolite structure (*21*). For an average smallest ring of 4.25, the predicted minimum FD (*21*) is 13 T atoms per 1000 Å^3, which is greater than the value for ZEO-3. Compared with the real values of noninterrupted zeolites containing 4- and 5-MRs, ZEO-3 is well below the lowest calculated FD of **ISV** and **IWV** (15.0 T-atoms per 1000 Å^3).

The observed N_2 and Ar adsorption/desorption isotherms (type Ia) of ZEO-3 revealed high specific surface areas of 989 and 1032 m^2/g (figs. S11 and S12), respectively. The nonlocal density functional theory method applied to the

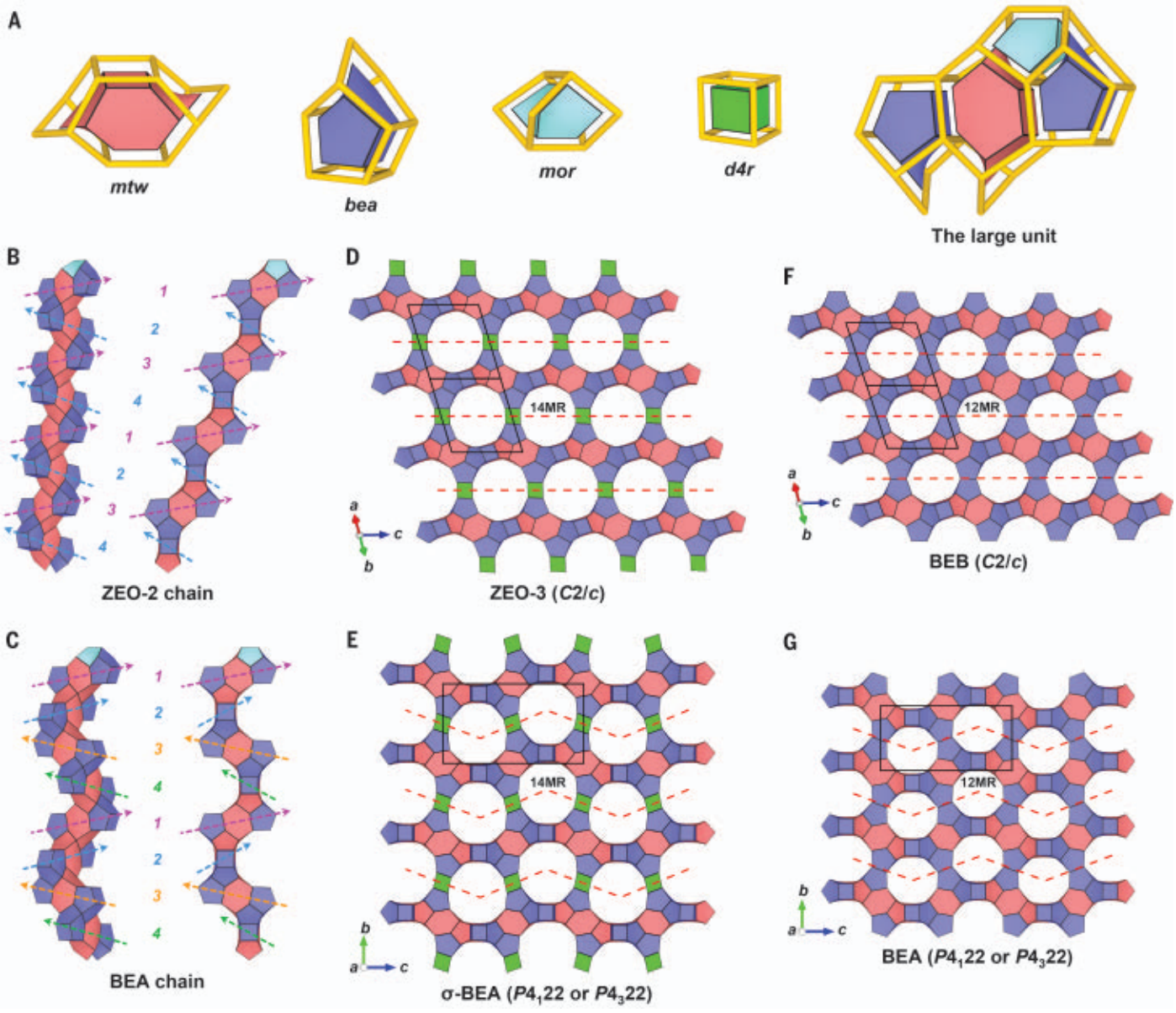


Fig. 5. Topology of ZEO-3 and a hypothetical chiral extra-large-pore zeolite. (**A**) Composite building units: the small units and the large unit made from them. (**B** and **C**) ZEO-2 chain (B) and BEA chain (C) built by attaching successive large units, viewed along two different directions. (**D** to **G**) The tiling structures of ZEO-3 (D), the σ-expanded **BEA** (E), and the polymorphs BEB (F) and **BEA** (G) of the beta zeolite. Dotted arrows indicate topologically identical large units (units 1, 2, 3, and 4) rotated by alternating +90°, −90°, +90°, and −90° in (B), or nonalternating +90°, +90°, +90°, and +90° or −90°, −90°, −90°, and −90° in (C). In (D) to (G), the dotted lines separate neighboring ZEO-2 or BEA chains; in (D) and (E), the newly formed D4Rs between neighboring ZEO-2 or BEA chains are highlighted in green; in (F) and (G), no D4Rs exist because S4Rs are shared between neighboring ZEO-2 or BEA chains.

Ar adsorption data calculated mean pore sizes of 10.8 and 8.8 Å (fig. S13), which match well with the crystallographic results. The extra-large pores of ZEO-3 allowed the diffusion and adsorption of large molecules such as Nile blue (fig. S14), suggesting the potential for the removal of large organic pollutants from waste liquid streams.

Adsorption has been considered as an energy-saving candidate for volatile organic compound (VOC) abatement and recovery (*22*). The development of sorbents with high adsorption capacity, water vapor resistance, and easy regeneration is critical for a successful adsorption technology (*23*). Zeolites are among the best adsorbents for VOCs given their distinctive microporosity, high adsorption capacity, and nonflammability (*23*, *24*). The adsorption equilibrium capacities of toluene and water vapor on ZEO-3 were larger than those on beta zeolites, which have smaller pores (Fig. 6A). ZEO-3 exhibited much longer breakthrough time (better dynamic capacity) than beta zeolites, with little interference from water (Fig. 6B). The main desorption peak of toluene occurred at a lower temperature for ZEO-3 than for beta (Fig. 6C), making its thermal regeneration easier. Thus, ZEO-3 outperforms beta, a reference zeolite for this application (*25*, *26*), in terms of adsorption capacity and regeneration potential. Comparison of ZEO-3 with commercial zirconium 1,4-dicarboxybenzene UiO-66, one of the most stable metal-organic frameworks (MOFs) (*27*), suggested a similar performance for the fresh materials (fig. S18). However, after recycling, the performance of ZEO-3 was maintained, whereas that of UiO-66 was compromised. At the end of five adsorption cycles, the PXRD pattern of UiO-66 revealed

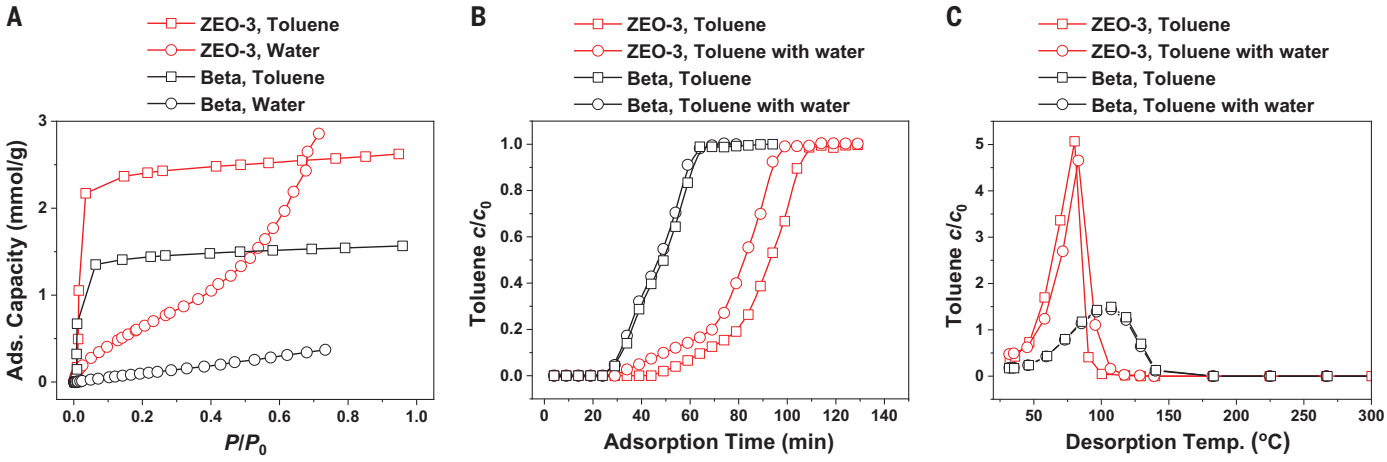


Fig. 6. Application of 3D extra-large-pore zeolite ZEO-3 in VOC removal. Shown are VOC adsorption isotherms (**A**), breakthrough adsorption (**B**), and desorption curves (**C**) on a ZEO-3 (red) and a beta (black) zeolite.

that it had been destroyed, whereas ZEO-3 remained stable (fig. S19). Additionally, the desorption temperature of UiO-66 was significantly higher than that of ZEO-3 (fig. S20), indicating a poorer regeneration ability. Phosphorous-free ZEO-3 had a high thermal stability (1100°C 10°C/min, 1-hour plateau) and hydrothermal stability (760°C, 10% H_2O, 3 hours), which may ease its application in real conditions (fig. S9). It is also possible to introduce active sites (such as Ti) into ZEO-3 through a one-pot synthesis method (figs. S21 and S22). Ultraviolet-visible spectra reveal that Ti-ZEO-3 exhibits both tetra-coordinated (~210 nm) and hexa-coordinated (~270 nm) Ti species (fig. S23). The catalytic properties of Ti-ZEO-3 will be investigated in a future study.

Concluding remarks

The fact that ZEO-3 presents structural features that sharply depart from previous observations (silica D4R units without F^- anions and FD lower than predicted) deserves some consideration. Those previous observations refer to materials directly synthesized by hydrothermal crystallization. The successful synthesis of ZEO-3 demonstrates that materials that may be considered not accessible by direct synthesis can be obtained by postsynthesis transformations. This has also been observed for zeolites obtained by the ADOR process, which afforded "unfeasible" zeolites (*28*) or the hybrid guest-host pure silica **STW**, which was also predicted to be unfeasible by direct synthesis (*29*). This observation allows us to foresee new materials developed by 1D-to-3D topotactic condensation, as was done using the ADOR process. It is worth mentioning that the final step in the ADOR process is a 2D-to-3D condensation that is similar to the 1D-to-3D process reported here but produces a systematic reduction in pore size with regard to the parent material.

REFERENCES AND NOTES

1. J. Zhong, J. Han, Y. Wei, Z. Liu, *J. Catal.* **396**, 23–31 (2021).
2. T.-Y. Cui, A. Rajendran, H.-X. Fan, J. Feng, W.-Y. Li, *Ind. Eng. Chem. Res.* **60**, 3295–3323 (2021).
3. Y. Wu, B. M. Weckhuysen, *Angew. Chem. Int. Ed.* **60**, 18930–18949 (2021).
4. S. Montalvo, C. Huiliñir, R. Borja, E. Sánchez, C. Herrmann, *Bioresour. Technol.* **301**, 122808 (2020).
5. M. Dusselier, M. E. Davis, *Chem. Rev.* **118**, 5265–5329 (2018).
6. Q.-F. Lin *et al.*, *Science* **374**, 1605–1608 (2021).
7. F. Liebau, "Nomenclature and structural formulae of silicate anions and silicates," in *Structural Chemistry of Silicates* (Springer, 1985); pp. 69–75.
8. L. Schreyeck, P. Caullet, J.-C. Mougenel, J.-L. Guth, B. Marler, *J. Chem. Soc. Chem. Commun.* **2187-2188**, 2187 (1995).
9. R. Millini, G. Perego, W. O. Parker Jr., G. Bellussi, L. Carluccio, *Microporous Mater.* **4**, 221–230 (1995).
10. B. Marler, H. Gies, *Eur. J. Mineral.* **24**, 405–428 (2012).
11. P. Eliášová *et al.*, *Chem. Soc. Rev.* **44**, 7177–7206 (2015).
12. C. S. Cundy, P. A. Cox, *Chem. Rev.* **103**, 663–702 (2003).
13. C. M. Zicovich-Wilson, M. L. San-Román, M. A. Camblor, F. Pascale, J. S. Durand-Niconoff, *J. Am. Chem. Soc.* **129**, 11512–11523 (2007).
14. M. Gemmi *et al.*, *ACS Cent. Sci.* **5**, 1315–1329 (2019).
15. T. Steiner, *Angew. Chem. Int. Ed.* **41**, 49–76 (2002).
16. H. Eckert, J. P. Yesinowski, L. A. Silver, E. M. Stolper, *J. Phys. Chem.* **92**, 2055–2064 (1988).
17. G. Engelhardt, *Trends Analyt. Chem.* **8**, 343–347 (1989).
18. R. M. Barrer, *Hydrothermal chemistry of zeolites*, Academic Press, London, pp. 11 (1982).
19. J. V. Smith, *Chem. Rev.* **88**, 149–182 (1988).
20. H. Lee *et al.*, *Science* **373**, 104–107 (2021).
21. G. O. Brunner, W. M. Meier, *Nature* **337**, 146–147 (1989).
22. M. Li *et al.*, *Atmos. Chem. Phys.* **19**, 8897–8913 (2019).
23. C. Yang *et al.*, *Chem. Eng. J.* **370**, 1128–1153 (2019).
24. S. K. P. Veerapandian, N. De Geyter, J.-M. Giraudon, J.-F. Lamonier, R. Morent, *Catalysts* **9**, 98 (2019).
25. T. Blasco *et al.*, *J. Phys. Chem. B* **102**, 75–88 (1998).
26. Z. Zhu, H. Xu, J. Jiang, H. Wu, P. Wu, *ACS Appl. Mater. Interfaces* **9**, 27273–27283 (2017).
27. J. H. Cavka *et al.*, *J. Am. Chem. Soc.* **130**, 13850–13851 (2008).
28. M. Mazur *et al.*, *Nat. Chem.* **8**, 58–62 (2016).
29. A. Rojas, M. A. Camblor, *Angew. Chem. Int. Ed.* **51**, 3854–3856 (2012).

ACKNOWLEDGMENTS

We thank C. Márquez (ICP-CSIC) for the Fourier tranform infrared measurements, R. Bai (JLU) for kind suggestions in the synthesis and characterization on the titanosilicate ZEO-3, M. J. de la Mata (SIdI-UAM) for expedited help in collecting MAS NMR spectra, and A. Ziolkowska (Umeå University) for help in TEM sample preparation by ultramicrotomy. **Funding:** This work was supported by the National Natural Science Foundation of China (grants 22288101, 21920102005, 21835002, 22271115, 21601004, 21621001, 21776312, and 22078364), the National Key Research and Development Program of China (grants 2022YFA1503600 and 2021YFA1501202), the 111 Project (grant B17020), the Natural Science Foundation of the Higher Education Institutions of Anhui Province, China (grant KJ2020A0585), and the Spanish Ministry of Science Innovation (grants PID2019-105479RB-I00 project, MCIN/AEI/10.13039/501100011033, and RYC2018-024561-I). The cRED data were collected at the Electron Microscopy Center (EMC), Department of Materials and Environmental Chemistry (MMK), at Stockholm University with the support of the Knut and Alice Wallenberg Foundation (KAW, 2012-0112) through the 3DEM-NATUR project. Use of the Advanced Photon Source at Argonne National Laboratory was supported by the US Department of Energy, Office of Science, Office of Basic Energy Sciences, under contract no. DE-AC02-06CH11357. Additional funding from the European Union's Horizon 2020 research and innovation program (grant 823717-ESTEEM3) and the regional government of Aragon (grant DGA E13_20R) is also acknowledged. **Author contributions:** F.-J.C. conceived the project. J.L., M.A.C., F.-J.C., and J.Y. supervised the project. Z.R.G., Q.-F.L., C.Liu, F.G., S.Z., and F.-J.C. performed the synthesis work. J.L. solved the structure and performed Rietveld refinement and framework energy calculations. Z.R.G. analyzed the topology. J.L., Z.R.G., C.Liu, C.Lin, W.F., S.L., X.C., and M.A.C. performed the physicochemical characterization. H.D. worked on the VOC application. A.M. performed the high-resolution STEM tests. M.A.C. prepared the initial draft of the manuscript. J.L., W.F., M.A.C., F.-J.C., and J.Y. organized the work and helped with the draft of the manuscript. All the authors discussed the results and revised the final manuscript. **Competing interests:** J.L., Q.-F.L., Z.R.G., C.Lin, and F.-J.C. have filed a patent on zeolites ZEO-2 and ZEO-3. J.L., Z.R.G., and C.Lin are affiliated with the company holding the rights on that patent. The remaining authors declare no competing interests. **Data and materials availability:** The datasets generated during and/or analyzed during the present study are available from the corresponding authors on reasonable request. Crystallographic parameters for the structure of ZEO-2 and ZEO-3 refined against SPXRD and cRED data are archived at the Cambridge Crystallographic Data Center (www.ccdc.cam.ac.uk/) under reference nos. CCDC 2125815-2125816 (SPXRD data) and CCDC 2125677-2125678 (cRED data). **License information:** Copyright © 2023 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. https://www.science.org/about/science-licenses-journal-article-reuse

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.ade1771
Materials and Methods
Supplementary Text
Figs. S1 to S23
Tables S1 to S14
References (*30*–*53*)

Submitted 7 September 2022; accepted 29 November 2022
10.1126/science.ade1771

SOLAR CELLS

Radical polymeric p-doping and grain modulation for stable, efficient perovskite solar modules

Shuai You[1,2]†, Haipeng Zeng[1]†, Yuhang Liu[2]†, Bing Han[3]†, Min Li[1], Lin Li[1], Xin Zheng[1], Rui Guo[1], Long Luo[1], Zhe Li[4], Chi Zhang[5], Ranran Liu[1], Yang Zhao[1], Shujing Zhang[1], Qi Peng[1], Ti Wang[4], Qi Chen[5], Felix T. Eickemeyer[2], Brian Carlsen[2], Shaik M. Zakeeruddin[2], Liqiang Mai[6,7], Yaoguang Rong[1,6]*, Michael Grätzel[2]*, Xiong Li[1]*

High-quality perovskite light harvesters and robust organic hole extraction layers are essential for achieving high-performing perovskite solar cells (PSCs). We introduce a phosphonic acid–functionalized fullerene derivative in mixed-cation perovskites as a grain boundary modulator to consolidate the crystal structure, which enhances the tolerance of the film against illumination, heat, and moisture. We also developed a redox-active radical polymer, poly(oxoammonium salt), that can effectively p-dope the hole-transporting material by hole injection and that also mitigates lithium ion diffusion. Power conversion efficiencies of 23.5% for 1-square-centimeter mixed–cation-anion PSCs and 21.4% for 17.1-square-centimeter minimodules were achieved. The PSCs retained 95.5% of their initial efficiencies after 3265 hours at maximum power point tracking under continuous 1-sun illumination at 70° ± 5°C.

Organic-inorganic hybrid perovskites have attracted extensive research interest for photovoltaic (PV) applications (*1*–*3*). Although the power conversion efficiency (PCE) of organic-inorganic hybrid perovskite solar cells (PSCs) can now exceed 25% (*4*–*7*), such high efficiencies have been obtained only with small-area PSCs (aperture areas < 0.1 cm^2). Advances reported for small laboratory cells do not necessarily transfer to perovskite solar modules (PSMs), because issues such as film inhomogeneity, resistive losses from photocurrent collection, and instability resulting from ion diffusion escalate steeply with device area (*8*). Thus, larger-area devices based on current materials and fabrication techniques often have substantial performance losses (*9*–*12*), so it is essential to deposit compact perovskite light-harvesting layers and charge-transporting layers with low defect densities over larger areas.

The antisolvent spin-coating process has been widely used to deposit perovskite thin films (*13*–*15*). However, this technique cannot be applied for casting large-area perovskite films. Commonly used approaches to scale up device size use "perovskite inks" that are compatible with processes such as slot-die coating, doctor-blade coating, and vacuum flash–assisted solution processing (VASP) (*16*–*19*). Besides morphology control, it is also necessary to passivate defects, suppress unfavorable ion diffusion, and promote electronic charge carrier transport within the perovskite films (*9*, *20*, *21*). Functional molecular modulators were introduced into the grain boundaries (GBs) to heal defects and prevent ion migration (*22*–*29*). Our group has introduced phosphonic acid derivatives [$-PO(OH)_2$] as more efficient passivation and cross-linking agents compared with carboxylic acids (*26*, *27*). However, prior work generally made use of aliphatic backbones—which are electrochemically inert and do not interact with perovskite—to connect the functional groups. Hence, designing a grain modulator that can facilitate intergrain electron transfer, while strengthening their binding, is essential to fabricating high-quality large-area perovskite films for PSMs.

The fabrication of high-performance PSMs will also require both shortening the processing time and increasing the conductivity and stability of the hole-transporting layer (HTL). For an n-*i*-p device, typical hole-transporting materials (HTMs) such as 2,2′,7,7′-tetrakis-(*N*,*N*-di-p-methoxyphenyl-amine)-9,9′-spirobifluorene (spiro-MeOTAD) or poly[bis(4-phenyl)(2,4,6-trimethylphenyl)amine] (PTAA) are deposited on top of the perovskite layer. Because of their low intrinsic conductivity, a p-type doping process with hygroscopic lithium salts, for example, lithium bis(trifluoromethylsulfonyl)imide (LiTFSI), is usually applied. The conventional doping process of spiro-MeOTAD typically takes hours and results in by-products (e.g., Li_xO_y) that will seriously degrade cell performance (*30*–*32*). Efforts have been devoted to developing doping strategies for HTMs (*33*) in PSCs. Kong *et al.* doped spiro-MeOTAD by bubbling carbon dioxide through HTM precursor solution under ultraviolet (UV) light, thereby decreasing the doping time (*34*). Zhang *et al.* developed a method that avoids post-oxidation by using stable organic radicals and ionic salts as dopants (*35*). Although proven effective for small-area laboratory cells (*36*, *37*), efficient HTM doping for larger-area PSMs has not been reported yet.

We introduce two cooperative strategies to address these issues and realize efficient PSMs that achieve an unprecedented operational stability and, at the same time, can be manufactured on a large scale at a competitive production rate. We used a redox-active polymeric dopant consisting of a polyoxyethylene backbone endowed with 2,2,6,6-tetramethylpiperidin-1-yl)oxyl, abbreviated as PPO-TEMPO (molecular structure shown in Fig. 1A), as an oxidant to afford p-doping of the HTLs through hole injection. This strategy not only shortened the doping time to a few seconds, which enables high production rates for PSMs, but also stabilized the device by suppressing the Li ion aggregation and diffusion into the adjacent layers. We also synthesized the fullerene derivative 4-(1′,5′- dihydro-1′-methyl-2′*H*-[5,6] fullereno-C60-I_h-[1,9-c]pyrrol-2′-yl) phenylphosphonic acid (CPPA; molecular structure shown in Fig. 1A) to mitigate GB defects of mixed-cation perovskites. CPPA features a conductive fullerene core functionalized by phosphonic acid anchoring groups that can interact with perovskites through hydrogen bonding, Pb coordination, and fullerene–halide radical interactions (*26*–*29*). By self-assembling at the GBs, CPPA enhanced the tolerance of the perovskite film against illumination, heat, and moisture and accelerated intergrain electronic charge carrier transport. The combined action of the two agents suppresses interfacial ion migration, contributing in a crucial manner to the high performance of our devices. These approaches allowed us to fabricate efficient PSCs and PSMs showing outstanding long-term operational stability.

PSM fabrication and performance

The PSCs and PSMs we fabricated have the layered architecture indium tin oxide (ITO)/tin(IV) oxide (SnO_2)/perovskite/HTL/Ag or Au (Fig. 1B). The triple-cation perovskite $CsFAMAPb(I_{1-x}Br_x)_3$ [where Cs is cesium, FA is $HC(NH_2)_2$, and MA is CH_3NH_3] (*38*, *39*) was deposited by the VASP method (see supplementary materials for device fabrication details). The VASP method enables fast removal and recuperation of solvent, which boosts the rapid crystallization of the dimethyl sulfoxide (DMSO)/lead(II) iodide (PbI_2) Lewis base adduct intermediate phase (*14*, *15*). Upon thermal annealing, highly oriented, crystalline perovskite films can be obtained (*19*, *40*). This

[1]Michael Grätzel Center for Mesoscopic Solar Cells, Wuhan National Laboratory for Optoelectronics, Huazhong University of Science and Technology, Wuhan 430074, Hubei, China. [2]Laboratory of Photonics and Interfaces, École polytechnique fédérale de Lausanne, 1015 Lausanne, Switzerland. [3]Department of Materials Science and Engineering, Southern University of Science and Technology, Shenzhen 518055, Guangdong, China. [4]School of Physics and Technology and Key Laboratory of Artificial Micro- and Nanostructures of Ministry of Education, Wuhan University, Wuhan 430072, Hubei, China. [5]*i*-Lab, CAS Key Laboratory of Nanophotonic Materials and Devices, Suzhou Institute of Nano-Tech and Nano-Bionics, Chinese Academy of Sciences (CAS), Suzhou 215123, Jiangsu, China. [6]State Key Laboratory of Advanced Technology for Materials Synthesis and Processing, Wuhan University of Technology, Wuhan 430070, Hubei, China. [7]Hubei Longzhong Laboratory, Wuhan University of Technology (Xiangyang Demonstration Zone), Xiangyang 441000, Hubei, China.
***Corresponding author. Email: ygrong2022@outlook.com (Y.R.); michael.graetzel@epfl.ch (M.G.); xiongli@hust.edu.cn (X.L.)**
†These authors contributed equally to this work.

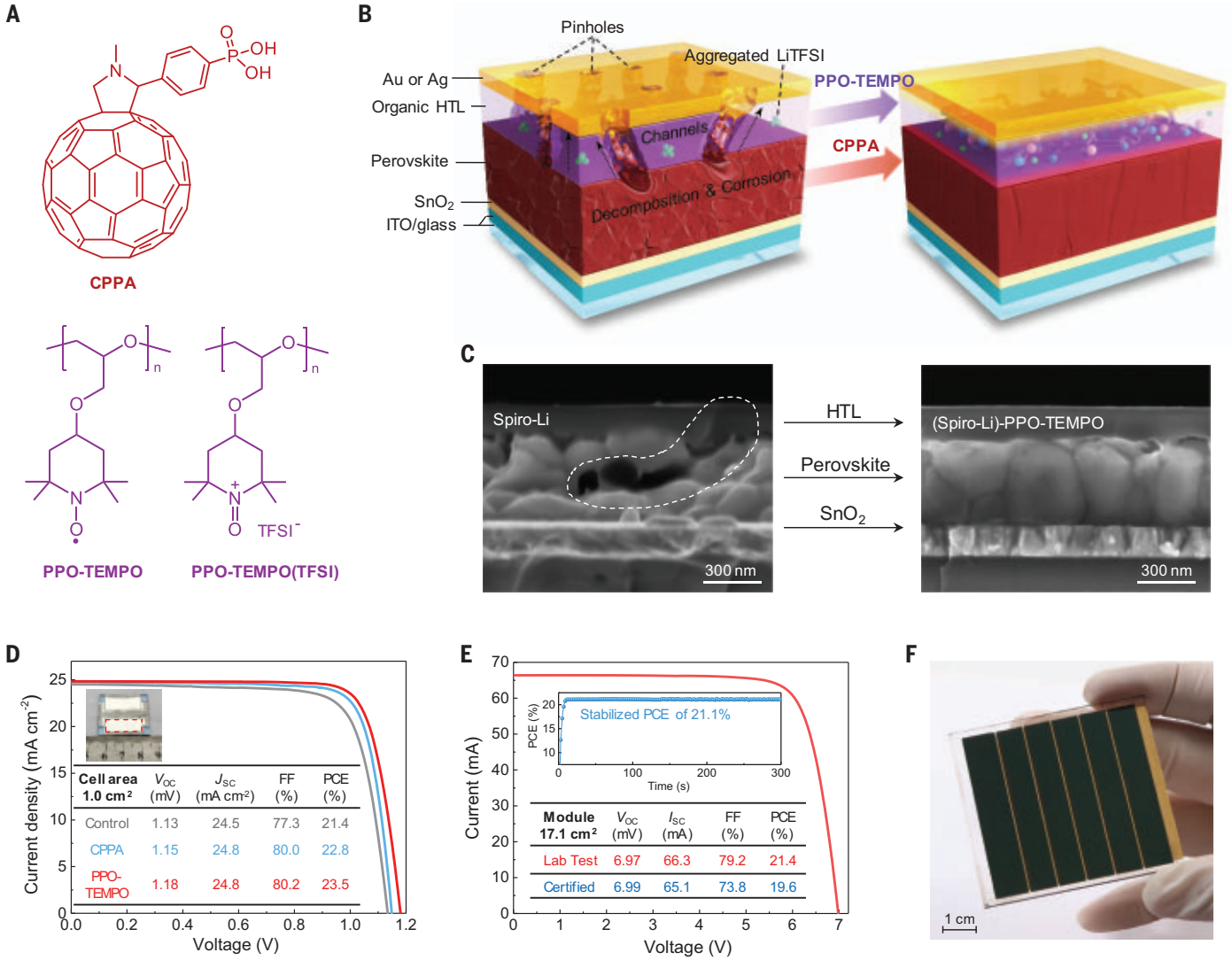


Fig. 1. Device configuration and photovoltaic performance. (**A**) Molecular structures of CPPA, PPO-TEMPO, and PPO-TEMPO(TFSI). (**B**) Schematic outline of the architecture of as-fabricated PSCs, and the stabilizing strategy for the perovskite layer and HTL. The perovskite layer had a composition $Cs_{0.05}FA_{0.85}MA_{0.10}Pb(I_{0.97}Br_{0.03})_3$. (**C**) Cross-sectional SEM images of perovskite/HTL/Ag aged under illumination at 75°C for 100 hours. The degraded area of the perovskite and HTL films is marked by the dashed line. In both images, Ag electrodes had been removed for conducting TOF-SIMS measurements of the HTL layer (fig. S1). (**D**) *J-V* curves and photovoltaic parameters of a representative cell that used CsFAMA and Spiro-Li (control), CsFAMA-CPPA and Spiro-Li (CPPA), and CsFAMA-CPPA and (Spiro-Li)-PPO-TEMPO (PPO-TEMPO). (**E**) *J-V* curves and stabilized output of the PSM recorded in the lab using a mask with an aperture area of 17.1 cm^2. The photovoltaic parameters certified by an accredited testing laboratory are also presented for comparison. (**F**) A photograph of the PSM.

method allows deposition of perovskite films on a large substrate area and can be applied to continuous production processes. Previously, we have demonstrated operationally stable PSCs using CsFAMA perovskite polycrystalline film and Spiro-MeOTAD/LiTFSI/4-*tert*-butyl pyridine (tBP) blend film (abbreviated as Spiro-Li) (*39*, *40*). However, the cells could only deliver stable electricity output at a temperature below 65°C. When the cells were stressed at a higher temperature, cell performance decayed severely owing to the degradation of the Spiro-Li layer. At elevated temperatures, small pinholes within the Spiro-Li layer were formed, enlarged, and ultimately caused severe morphological deformation that provided channels for the escape of volatile and corrosive ingredients from the perovskite layer (*41*–*44*). Upon introducing the PPO-TEMPO doping strategy to the Spiro-Li layer, the bulk morphology of the perovskite/HTL bilayer featuring well-defined interfaces was markedly stabilized under illumination at 75°C for 100 hours (Fig. 1C). Time-of-flight secondary ion mass spectrometry (TOF-SIMS) measurements showed that PPO-TEMPO doping could also minimize the interfacial diffusion of iodide from the perovskite layer to the HTL (fig. S1).

To evaluate the effect of the CPPA modulator and PPO-TEMPO doping on cell performance, we fabricated and tested a batch of 20 cells. The representative current density–voltage (*J-V*) curves are compared in Fig. 1D, and the statistical results are shown in table S1 and fig. S2. When CPPA was introduced into the perovskite layer, the fill factor (FF) increased from 77.3% to 80.0%, and the open-circuit voltage (V_{OC}) increased from 1.13 V to 1.15 V. We attribute the improved cell performance to enhanced intragrain charge transport and suppression of nonradiative charge recombination by CPPA molecules at the GBs (*45*). Meanwhile, when PPO-TEMPO doping was applied, the V_{OC} was further increased to 1.18 V, which we attribute to a more favorable

band alignment. For the cell using CsFAMA-CPPA as the absorbing layer and (Spiro-Li)-PPO-TEMPO as the HTL, we achieved a PCE of 23.5% with negligible hysteresis (fig. S3). We confirmed the measured short-circuit density (J_{SC}) of 24.8 mA cm^{-2} by recording incident photon–to-current efficiency (IPCE) spectra, which yield an integrated current density of 24.7 mA cm^{-2}. This value is in good agreement with the measured J_{SC} value and indicated that there was negligible spectral mismatch between our solar simulator and the standard global AM 1.5 solar emission (fig. S4).

To highlight the advantages of using the VASP method to deposit large-area perovskite films, we fabricated PSMs consisting of six subcells connected in series and obtained a PCE of 21.4% with an aperture area of 17.1 cm^2 (the stabilized PCE reached 21.1%) (Fig. 1E). We also sent one of the minimodules to a PV measurement laboratory, and the validated PCE reached 19.6% (fig. S5). A photograph of the PSM is presented in Fig. 1E. The short-circuit photocurrent (I_{SC}) of the module reached 66.3 mA, corresponding to a J_{SC} of 23.3 mA cm^{-2} for the subcells, which was 94% of the J_{SC} obtained with lab-scale cells. As confirmed by scanning electron microscopy (SEM), the perovskite film deposited by the VASP method showed uniform grain size and height at different locations over a 10-cm scale (fig. S6).

Grain modulation of the perovskite layer

Density functional theory calculations revealed that CPPA associated with lead halide perovskites through hydrogen bonding as well as coordinative bonding and fullerene–iodide interactions (see Fig. 2A, fig. S7, table S2, and supplementary note 1 in the supplementary text for details). Accordingly, we postulated that CPPA associated with GBs, thereby enhancing the stability of the perovskite film (*26*, *27*). To prove this, we used cryo–transmission electron microscopy (which minimizes the electron beam damage to the perovskites) to analyze the microstructures of aged perovskite films (*46*). The pristine and CPPA-modulated perovskite films were exposed to continuous illumination at 85°C. Representative high-resolution transmission electron microscopy (HRTEM) images of selected areas are shown in Fig. 2B and fig. S8. For the aged CsFAMA film, three distinct crystal structures were observed, and the interplanar spacings (ISs) and fast Fourier transform patterns (fig. S9) of a selected area were analyzed. The crystalline region with an IS of 3.16 Å was identified as the (200) domain of α-$FAPbI_3$, and the region with an IS of 6.90 Å corresponded to the hexagonal PbI_2 domain (*47*). For the aged CsFAMA-CPPA film, crystalline regions corresponded to (220) and (210) domains of α-$FAPbI_3$, and no PbI_2 domain was observed (fig. S10). Between the crystalline regions there was an amorphous region, which is considered beneficial for stress relaxation within crystalline perovskite films.

We analyzed the stabilization effect of CPPA on mixed-halide perovskites by measuring the photoluminescence (PL) properties of the CPPA-modulated perovskite films under blue-light irradiation. A typical mixed-halide perovskite composition of $MAPb(I_{0.5}Br_{0.5})_3$ was used as a model. When the films were exposed to high-intensity (13.0 W cm^{-2}) blue-light irradiation, significant phase evolution was observed (Fig. 2C). Both $MAPb(I_{0.5}Br_{0.5})_3$ and $MAPb(I_{0.5}Br_{0.5})_3$-CPPA showed PL peaks of an iodine-rich (I-rich) phase at ~747 nm after 1 min of irradiation, resulting from the regular photoinduced phase segregation (*48*–*51*). After 10 min of irradiation, a phase decomposition for the pristine $MAPb(I_{0.5}Br_{0.5})_3$ emerged that manifested itself by a weak PL peak at 537 nm, corresponding to the characteristic emission peak of $MAPbBr_3$ (*52*–*54*). This indicated that the I-rich phase decomposed under the intense light illumination, leaving $MAPbBr_3$ behind, which appeared to be less susceptible than $MAPbI_3$ to decomposition under the intense light beam. After 15 min, the intensity of the 537-nm peak became comparable to that of the I-rich phase with a peak at 737 nm. This indicated that $MAPbBr_3$ has become the dominant phase in the mixed-halide perovskite film. The decomposition of $MAPbI_3$ may be caused by either photolysis or thermolysis, according to the equation that follows (see also supplementary note 2)

$$MAPbI_3 = I_2 + Pb + MAI \quad (1)$$

In contrast, the CPPA-treated $MAPb(I_{0.5}Br_{0.5})_3$ resisted degradation under prolonged exposure to the intense blue light. Notably, the 747-nm PL peak remained unchanged within 15 min of irradiation. The stable PL emission indicated that the presence of CPPA prevented the decomposition of $MAPbI_3$ under high-intensity irradiation. The PL imaging results as a function of irradiation time are presented in Fig. 2D and fig. S11. For $MAPb(I_{0.5}Br_{0.5})_3$, it was found that green spots started to emerge at the illumination area after 10 min and became more intense after 15 min of irradiation. By contrast, the irradiated area of CPPA-treated $MAPb(I_{0.5}Br_{0.5})_3$ maintained its red emission during the whole irradiation process. This implied that CPPA could effectively enhance the photostability of mixed-halide perovskites (see supplementary note 2 for details).

The stabilization of $MAPbI_3$-rich grains formed during the phase segregation by CPPA clearly arises from its prevention of their photo- or thermally induced decomposition under blue-light irradiation. Our detailed analysis shows that $MAPbI_3$ decomposes into iodine, methylammonium iodide (MAI), and elemental lead. This process is triggered by the loss of iodine from the lattice, leaving MAI and Pb behind. The prevention of decomposition by CPPA indicates that its presence at the perovskite GBs impairs the escape of iodine. Such a stabilization effect may also contribute to the HTL, because the diffusion of iodide ions from the perovskite into the HTL can reduce the oxidized HTMs and cause dedoping issues (*55*).

We found that CPPA can also improve the hydrophobicity and moisture stability of the modulated perovskite film. We exposed the perovskite to ambient atmosphere with a high relative humidity of 70% and found that pristine CsFAMA rapidly degraded to δ-$FAPbI_3$ and PbI_2, whereas CsFAMA-CPPA maintained its original crystal structure (Fig. 2, E and F, and fig. S16). We performed conductive atomic force microscopy (c-AFM) measurements to investigate the charge transport properties of the perovskite films (fig. S17). Compared with pristine CsFAMA, CsFAMA-CPPA presented higher current flow through the film. In particular, the brighter contrast at GBs indicates a suppressed charge-transporting barrier across the GBs of CsFAMA-CPPA. Absolute spectral photon flux Φ was measured to derive the PL quantum yield (PLQY) of the perovskite films (Fig. 2G), which increased from 3.27% to 6.41% after doping the perovskite with CPPA. The quasi–Fermi level splitting $\Delta E_F/q$ (where E_F is the Fermi level and q is the elementary charge) increased from 1.176 to 1.192 V for CsFAMA versus CsFAMA-CPPA (table S3) (*56*, *57*) and reflected the suppression of nonradiative recombination with the CPPA additive.

Radical polymeric p-doping of the HTL

Effective and homogeneous doping and stabilization of the HTM is critically important for high-performance PSMs. In addition, it is necessary to develop effective HTM doping methods to facilitate the fabrication of PSMs. We oxidized the neutral radical polymer PPO-TEMPO to a poly(oxoammonium salt) of PPO-TEMPO(TFSI) (see the materials and methods section of the supplementary materials for synthesis details) and incorporated it as a p-dopant of spiro-MeOTAD. PPO-TEMPO/PPO-TEMPO(TFSI) had a redox potential of 0.925 V compared with a normal hydrogen electrode (NHE), as determined by electrochemical measurements (fig. S18), which is greater than that of spiro-MeOTAD/spiro-MeOTAD$^{\bullet+}$ (0.72 V compared with NHE) and provides a driving force of ~200 mV to oxidize spiro-MeOTAD through a one-electron transfer reaction (*58*). Upon the reduction of the oxoammonium, this cationic pendant group formed the stable radical polymer poly(4-glycidyloxy-2,2,6,6-tetramethylpiperidine-1-oxyl) (PPO-TEMPO), which can interact with Li^+ ions and stabilize the HTL (Fig. 3A) (*59*).

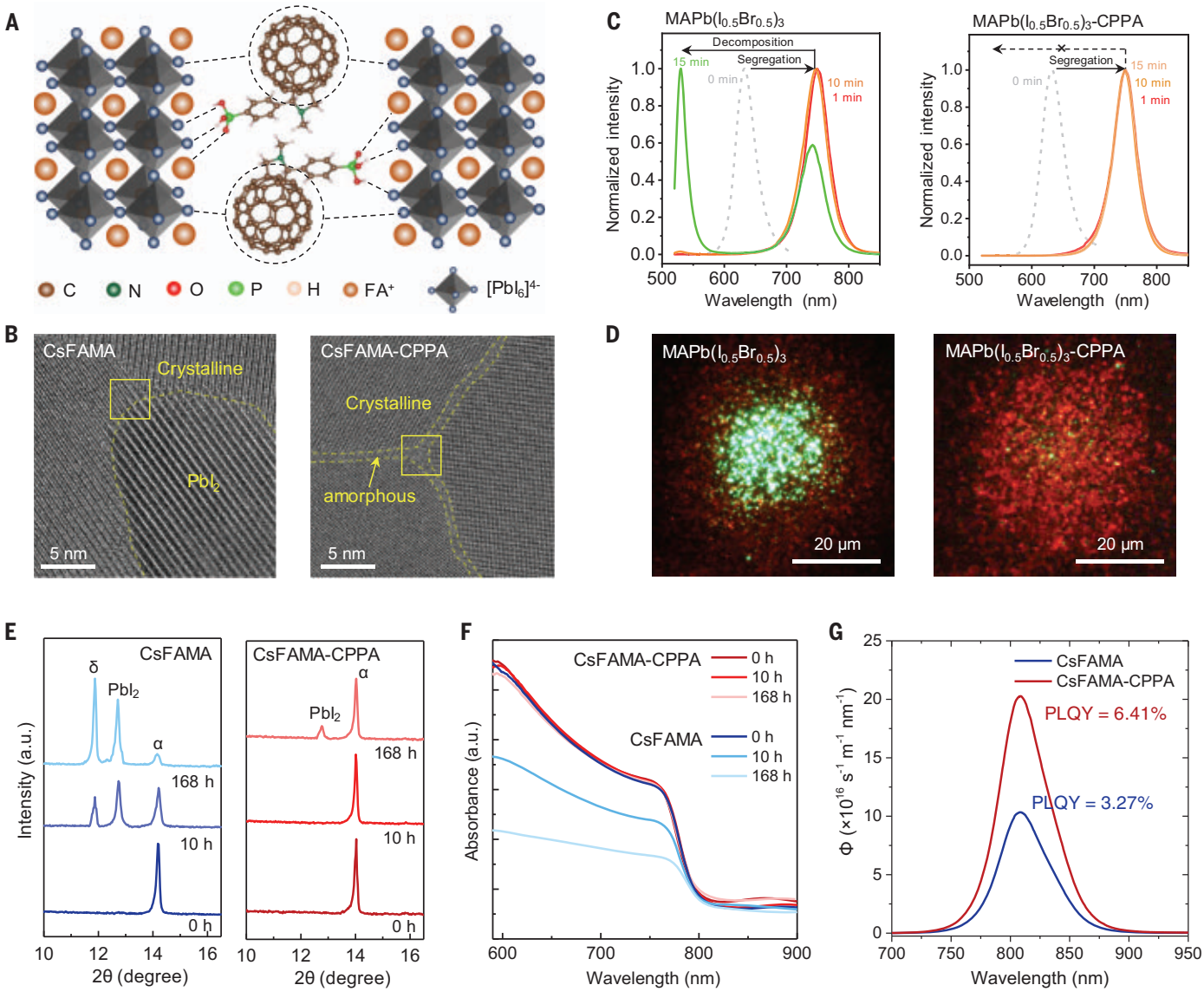


Fig. 2. Working mechanism of CPPA and luminescence enhancement in perovskite films. (**A**) The scheme of interaction modes between CPPA and perovskites. (**B**) Cryo-HRTEM images of perovskite films aged under continuous illumination at a temperature of 85°C for 7 days in nitrogen atmosphere. The magnifications of the boxed areas are shown in fig. S7. (**C**) The evolution of PL spectra of $MAPb(I_{0.5}Br_{0.5})_3$ and $MAPb(I_{0.5}Br_{0.5})_3$-CPPA films under continuous blue-light irradiation (mercury lamp; wavelength: 460 to 490 nm; intensity: 13.0 W cm^{-2}). (**D**) PL imaging results of $MAPb(I_{0.5}Br_{0.5})_3$ and $MAPb(I_{0.5}Br_{0.5})_3$-CPPA films after 15 min of blue-light irradiation. (**E**) X-ray diffraction (XRD) patterns and (**F**) UV-vis spectra of CsFAMA and CsFAMA-CPPA perovskite films aged in a humid environment (relative humidity: 70 to ~75%) for 0 to ~168 hours. (**G**) Photoluminescence spectra photon flux Φ of the perovskite films. a.u., arbitrary units.

The changes in the UV-visible (UV-vis) absorption spectrum of a spiro-MeOTAD solution in chlorobenzene upon PPO-TEMPO doping are shown in Fig. 3B. Increasing the amount of dopant gave rise to an absorption peak at ~520 nm from the monocation radical of spiro-MeOTAD$^{\bullet+}$ that increased linearly with dopant concentration (*30*). We estimated the doping efficacy, defined by the molar ratio of the oxidized spiro-MeOTAD to the added dopant, to be ~ 90%. Although spiro-MeOTAD was oxidized to the cation radical spiro-MeOTAD•+, PPO-TEMPO(TFSI) was reduced to its neutral polymeric radical state in the doped HTMs.

The doping effect was further investigated by conductivity and UV photoelectron spectroscopy (UPS) measurements. The conductivity of the Spiro-Li film was evaluated by measuring the *J-V* curves of a symmetrical device with Au/Spiro-Li/Au architecture. As shown in Fig. 3C, the undoped Spiro-Li film kept under N_2 showed quite a low electrical conductivity of 1.19×10^{-7} S cm^{-1}, but the Spiro-Li film with PPO-TEMPO doping had a high electrical conductivity of 3.32×10^{-5} S cm^{-1} without exposure to oxygen (see supplementary note 3 for details). UPS results confirmed that PPO-TEMPO doping improved the energy band alignment in the perovskite/HTL/Au stack (see figs. S19 and S20 and supplementary note 4 for details). The deepened Fermi level and increased conductivity confirmed efficient p-type doping with PPO-TEMPO/PPO-TEMPO(TFSI).

Compared with the gas-doping strategies with O_2 or CO_2, which usually involve photo-induced gas-liquid or gas-solid reactions (*34*), PPO-TEMPO doping in HTM solution was more efficient and would largely simplify the doping process. The Li^+/PPO-TEMPO radical interaction (*59*) can also prevent the migration of lithium cations under heat and voltage stress. To evaluate this effect, the Spiro-Li

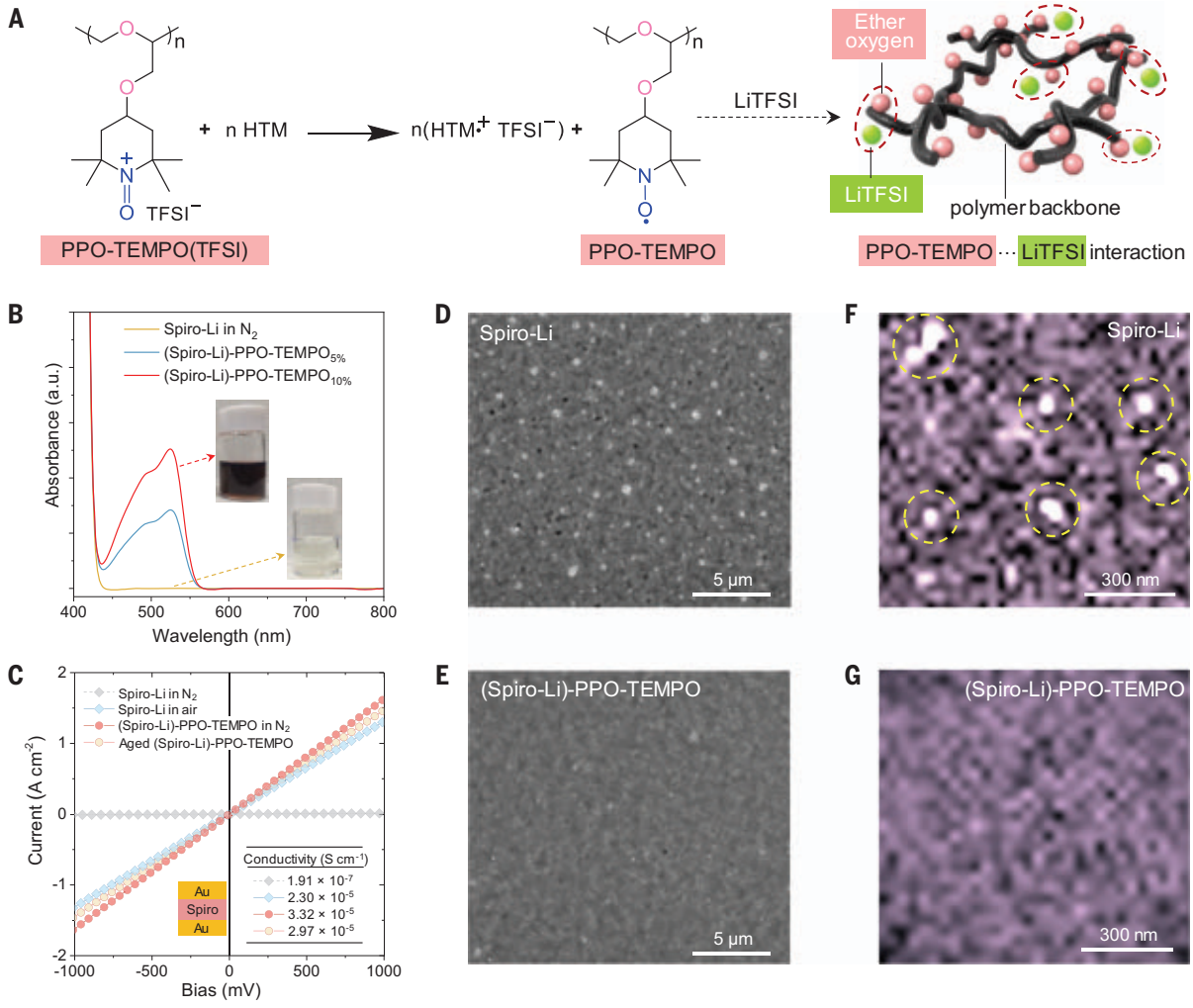


Fig. 3. Doping effect and stabilized bulk morphology. (**A**) Schematic illustration of the doping mechanism of PPO-TEMPO and PPO-TEMPO(TFSI) of organic HTMs, and the interaction mode between the reduction product PPO-TEMPO and LiTFSI additive. (**B**) The absorption spectra of Spiro-Li without doping and Spiro-Li with PPO-TEMPO doping of different ratios. (**C**) The *J-V* curves measured with symmetrical device of Au/Spiro-Li/Au for evaluating the conductivity of the HTL with and without PPO-TEMPO doping in air and nitrogen (the inset shows the extracted conductivity values). The aged samples were heated at 75°C for 30 min. (**D** and **E**) Surface SEM images of Spiro-Li films with and without PPO-TEMPO doping after aging at 75°C for 100 hours. (**F** and **G**) Li element EELS mapping images of aged Spiro-Li films with and without PPO-TEMPO doping.

films with and without PPO-TEMPO doping were aged under heat at 75°C in dry air. After 100 hours, voids, microdefects, and particles emerge on the film without PPO-TEMPO, whereas the film with PPO-TEMPO doping maintained its initial pinhole-free surface morphology (Fig. 3, D and E). The voids represent inhomogeneities within the HTL.

We correlated these morphological deformations with inhomogeneous distribution of Li^+ by analyzing the films with the EELS (electron energy loss spectroscopy) mapping technique (Fig. 3, F and G; see supplementary note 5 for details). In the HRTEM morphology images (fig. S21), several pits were observed, which was consistent with those in SEM images. For EELS mapping images, the relative color shift depicts the fluctuation in local Li concentration. In the control sample, the higher-contrast white spots highlighted by yellow cycles indicate a high Li content. These Li-rich domains of ~50-nm size were associated with the pits shown in the HRTEM morphology image. Thus, our direct mapping of Li in a control HTL film revealed nanoscale inhomogeneities (nonstoichiometric regions) associated with local phase separation. It appeared that PPO-TEMPO doping could improve the PSC stability by preventing such local Li^+ accumulation.

In addition, this polymeric p-doping strategy also provides an effective solution of HTM doping for the fabrication of PSMs. As shown in fig. S22, PSM using the regular Spiro-Li HTL required an ambient air exposure process of more than 24 hours. Such a process will unduly slow down the fabrication of PSMs. By contrast, the PSM with PPO-TEMPO doping exhibited its optimal photovoltaic performance with modular PCE > 20% right after fabrication. Our approach has provided effective HTM doping of the PSM and will facilitate future mass production of PSMs.

We investigated the operational stability of as-fabricated cells by conducting maximum power point (MPP) tracking under continuous illumination. The target cells treated with CPPA and PPO-TEMPO doping maintained 95.5% of their initial PCE after 3265 hours under continuous illumination at 70° ± 5°C, whereas the control cells without PPO-TEMPO doping lost >30% of their initial PCEs in just 1000 hours (Fig. 4A). The void-like structures and inhomogeneities within HTLs likely impeded the top-down charge transport during

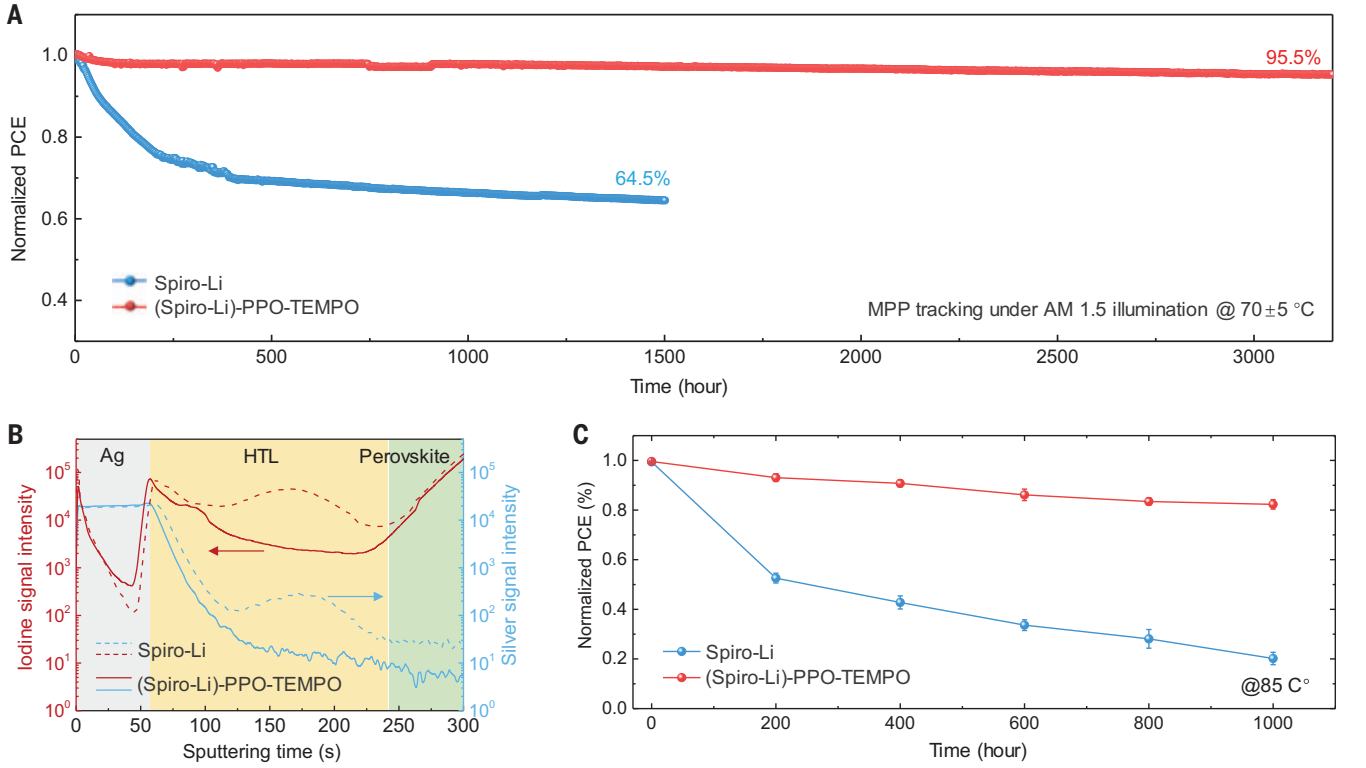


Fig. 4. Device stability and degradation analysis. (**A**) Evolution of the PCE over time, as measured by MPP tracking of encapsulated cells under light soaking with AM1.5G simulated illumination at a temperature of 70° ± 5°C. (**B**) TOF-SIMS depth profiles of the iodine and Ag in SnO_2/perovskite/HTL/Ag stacked layers. (**C**) Thermal stability test performed at 85°C for the cells using Spiro-Li as the HTL with and without PPO-TEMPO doping.

operation of PSCs and would lead to severe degradation of the devices by triggering the reactions between the perovskite and metal back contacts (*41*, *43*). TOF-SIMS results proved that PPO-TEMPO doping can effectively suppress the intrusion of iodide into the HTL (Fig. 4B) and prevent the reaction between the perovskite layer and metal electrode.

We also attributed the enhanced stability to the suppressed Li^+ diffusion in the HTL (*60*) and conducted TOF-SIMS depth profiling of Li^+ along the vertical direction through the stacks (fig. S23, and see supplementary note 6 for details). When PSCs were aged with a Spiro-Li layer, Li^+ was found at a higher concentration at the perovskite/SnO_2 interface. Li^+ intrusion into the perovskite active layer could also result in decomposition of the perovskite, leading to cell performance degradation (*61*). However, only a small amount of Li^+ diffused out from the aged PPO-TEMPO–doped HTL, which we attributed to the formation of Lewis acid-base interaction between Li^+ and the ether linkages in the PPO-TEMPO polymer backbone (*62*).

To investigate the stability of the cells under thermal stress, we tested PSCs under prolonged heat exposure at 85°C according to ISOS (International Summit on Organic Photovoltaic Stability) procedures (*63*). Given the glass-transition issues of Spiro-MeOTAD, the cells suffered a certain extent of initial losses (Fig. 4C). For the cells using Spiro-Li as the HTL, power output rapidly decayed by >40% after 200 hours, and by >70% after 1000 hours, of storage at 85°C. Conversely, after doping Spiro-Li with PPO-TEMPO, the cell power output decayed only by <20% after 1000 hours of storage at 85°C.

To show the universality of this polymeric redox couple doping strategy, we used PPO-TEMPO doping for two other representative polymeric hole conductors, PTAA and poly(3-hexylthiophene) (P3HT). Both of the doped HTLs exhibited higher conductivities than did pristine films (fig. S24), realizing more-efficient PSCs without any postfabrication treatments (see table S4, fig. S25, and supplementary note 7 for details). We infer that the doping and stabilizing effect of PPO-TEMPO doping on organic HTMs could be broadly applicable. The highly doped HTLs exhibited better conductivity, so the thickness of the HTL in PSCs can be increased to >150 nm. Integrated with the CPPA-enhanced perovskite layer, the holistic strategy for stabilizing the perovskite/HTL heterostructure provides promising technical routes for fabricating efficient and stable PSMs.

REFERENCES AND NOTES

1. A. Kojima, K. Teshima, Y. Shirai, T. Miyasaka, *J. Am. Chem. Soc.* **131**, 6050–6051 (2009).
2. Q. Dong *et al.*, *Science* **347**, 967–970 (2015).
3. G. E. Eperon, M. T. Hörantner, H. J. Snaith, *Nat. Rev. Chem.* **1**, 0095 (2017).
4. D. Bi *et al.*, *Nat. Energy* **1**, 16142 (2016).
5. J. J. Yoo *et al.*, *Nature* **590**, 587–593 (2021).
6. National Renewable Energy Laboratory, Photovoltaic Research: Best Research-Cell Efficiency Chart; https://www.nrel.gov/pv/cell-efficiency.html.
7. J. Jeong *et al.*, *Nature* **592**, 381–385 (2021).
8. E. Bi *et al.*, *Joule* **3**, 2748–2760 (2019).
9. Y. Deng *et al.*, *Nat. Energy* **6**, 633–641 (2021).
10. N.-G. Park, K. Zhu, *Nat. Rev. Mater.* **5**, 333–350 (2020).
11. Z. Yang *et al.*, *Sci. Adv.* **7**, eabg3749 (2021).
12. Y. Deng *et al.*, *Sci. Adv.* **5**, eaax7537 (2019).
13. N. J. Jeon *et al.*, *Nat. Mater.* **13**, 897–903 (2014).
14. Y. Rong *et al.*, *Nanoscale* **7**, 10595–10599 (2015).
15. N. Ahn *et al.*, *J. Am. Chem. Soc.* **137**, 8696–8699 (2015).
16. F. Di Giacomo *et al.*, *Sol. Energy Mater. Sol. Cells* **181**, 53–59 (2018).
17. Y. Deng *et al.*, *Nat. Energy* **3**, 560–566 (2018).
18. D.-N. Jeong *et al.*, *ACS Energy Lett.* **4**, 1189–1195 (2019).
19. X. Li *et al.*, *Science* **353**, 58–62 (2016).
20. J. Xue, R. Wang, Y. Yang, *Nat. Rev. Mater.* **5**, 809–827 (2020).
21. L. K. Ono, S. F. Liu, Y. Qi, *Angew. Chem. Int. Ed.* **59**, 6676–6698 (2020).
22. B. Chen, P. N. Rudd, S. Yang, Y. Yuan, J. Huang, *Chem. Soc. Rev.* **48**, 3842–3867 (2019).
23. S. Yang *et al.*, *J. Am. Chem. Soc.* **141**, 5781–5787 (2019).
24. M.-H. Choi *et al.*, *Nano Energy* **71**, 104639 (2020).
25. Q. Jiang *et al.*, *Nat. Photonics* **13**, 460–466 (2019).
26. X. Li *et al.*, *Nat. Chem.* **7**, 703–711 (2015).
27. H. Xie *et al.*, *Joule* **5**, 1246–1266 (2021).
28. J. Xu *et al.*, *Nat. Commun.* **6**, 7081 (2015).
29. Y. Shao, Z. Xiao, C. Bi, Y. Yuan, J. Huang, *Nat. Commun.* **5**, 5784 (2014).
30. A. Abate *et al.*, *Phys. Chem. Chem. Phys.* **15**, 2572–2579 (2013).
31. U. B. Cappel, T. Daeneke, U. Bach, *Nano Lett.* **12**, 4925–4931 (2012).
32. F. Lamberti *et al.*, *Chem* **5**, 1806–1817 (2019).
33. T. H. Schloemer, J. A. Christians, J. M. Luther, A. Sellinger, *Chem. Sci.* **10**, 1904–1935 (2019).

34. J. Kong *et al*., *Nature* **594**, 51–56 (2021).
35. T. Zhang *et al*., *Science* **377**, 495–501 (2022).
36. T. H. Schloemer *et al*., *ACS Energy Lett.* **4**, 473–482 (2019).
37. J. Urieta-Mora, I. García-Benito, A. Molina-Ontoria, N. Martín, *Chem. Soc. Rev.* **47**, 8541–8571 (2018).
38. M. Saliba *et al*., *Energy Environ. Sci.* **9**, 1989–1997 (2016).
39. S. You *et al*., *Adv. Mater.* **32**, 2003990 (2020).
40. D. Bi *et al*., *Nat. Commun.* **9**, 4482 (2018).
41. Y. Kato *et al*., *Adv. Mater. Interfaces* **2**, 1500195 (2015).
42. X. Niu *et al*., *J. Mater. Chem. A Mater. Energy Sustain.* **7**, 7338–7346 (2019).
43. K. Domanski *et al*., *ACS Nano* **10**, 6306–6314 (2016).
44. S. Wang, Y. Jiang, E. J. Juarez-Perez, L. K. Ono, Y. Qi, *Nat. Energy* **2**, 16195 (2016).
45. D.-Y. Son *et al*., *Nat. Energy* **1**, 16081 (2016).
46. M. U. Rothmann *et al*., *Adv. Mater.* **30**, e1800629 (2018).
47. B.-W. Park *et al*., *Nat. Energy* **6**, 419–428 (2021).
48. E. T. Hoke *et al*., *Chem. Sci.* **6**, 613–617 (2015).
49. C. G. Bischak *et al*., *Nano Lett.* **17**, 1028–1033 (2017).
50. M. C. Brennan, A. Ruth, P. V. Kamat, M. Kuno, *Trends Chem.* **2**, 282–301 (2020).
51. A. J. Knight, L. M. Herz, *Energy Environ. Sci.* **13**, 2024–2046 (2020).
52. J. T. DuBose, P. V. Kamat, *Acc. Mater. Res.* **3**, 761–771 (2022).
53. Z. Li *et al*., *Adv. Sci. (Weinh.)* **9**, e2103948 (2022).
54. G. F. Samu *et al*., *J. Am. Chem. Soc.* **141**, 10812–10820 (2019).
55. T. Wang *et al*., *Science* **377**, 1227–1232 (2022).
56. W. Tress *et al*., *Adv. Energy Mater.* **5**, 1400812 (2015).
57. H. Zhu *et al*., *Adv. Mater.* **32**, e1907757 (2020).
58. J. Burschka *et al*., *J. Am. Chem. Soc.* **133**, 18042–18045 (2011).
59. I. Yu, D. Jeon, B. Boudouris, Y. Joo, *Macromolecules* **53**, 4435–4441 (2020).
60. S.-G. Kim *et al*., *Adv. Mater.* **33**, e2007431 (2021).
61. J. A. Dawson *et al*., *ACS Energy Lett.* **2**, 1818–1824 (2017).
62. Y. Joo, V. Agarkar, S. H. Sung, B. M. Savoie, B. W. Boudouris, *Science* **359**, 1391–1395 (2018).
63. M. V. Khenkin *et al*., *Nat. Energy* **5**, 35–49 (2020).

ACKNOWLEDGMENTS

We thank the Analytical and Testing Center of Huazhong University of Science and Technology and the Center for Nanoscale Characterization and Devices at Wuhan National Laboratory for Optoelectronics for performing various characterizations and measurements; the Pico and Cryo-TEM Center at Southern University of Science and Technology (SUSTech) Core Research Facility; Y. Zou and Y. Deng at the Cryo-TEM Center of SUSTech for help with TEM measurements. H. Xu for help with PL measurements; Z. Xiao for providing VASP software copyright and server computing resources; and W. Tress for valuable discussions. **Funding:** This work was supported by the National Key Research and Development Program of China [2020YFA0715000 (L.M.) and 2022YFB4200305 (X.L. and Y.R.)]; the Chinese National Thousand Talents Plan (X.L.); the National Natural Science Foundation of China [21875081 and 22279039 (X.L.), 52172200 (Y.R.), and 52127816 (L.M.)]; the Science and Technology Department of Hubei Province [2021CFB315 (Y.R.)]; the Frontier Project of the Application Foundation of Wuhan Science and Technology Plan Project [2020010601012202 (X.L.)]; the Innovation Project of Optics Valley Laboratory OVL2021BG008 (X.L.); the Foundation of State Key Laboratory of New Textile Materials and Advanced Processing Technologies Grant FZ2021011 (X.L.); the Graduates' Innovation Fund of Huazhong University of Science and Technology 2021yjsCXCY037 (S.Y.); and the Günes Corporation (M.G.). **Author contributions:** Conceptualization: X.L., Y.R., and M.G. Investigation: S.Y., X.Z., Y.L., B.H., T.W., H.Z., C.Z., Q.C., R.L., Y.Z., L.Li, L.Luo, S.M.Z., R.G., Y.R., F.T.E., B.C., S.Z., Q.P., and L.M. Visualization: S.Y., Y.L., Y.R., and X.L. Funding acquisition: X.L., Y.R., and L.M. Project administration: X.L. and Y.R. Supervision: X.L., M.G., and Y.R. Writing – original draft: S.Y. and Y.R. Writing – review & editing: Y.R., Y.L., M.G., and X.L. X.L. and Y.R. conceived of and supervised the project. X.L., Y.R., and M.G. designed this project. S.Y., M.L., R.L., and S.Z. fabricated the PSCs and PSMs. H.Z. synthesized the chemicals and conducted UV-vis measurements. B.H. conducted cryo-HRTEM measurements and analyzed the data. S.Y. and Q.P. conducted SEM measurements. Z.L. and T.W. conducted PL imaging measurements. X.Z. and L.L. conducted the DFT calculations. L.L. and Y.Z. conducted TOF-SIMS and UPS measurements. C.Z. and Q.C. conducted c-AFM measurements. L.M. contributed to PL, XRD, and x-ray photoelectron spectroscopy measurements. B.C. and S.M.Z. performed PLQY measurements. Y.R., Y.L., R.G., F.T.E., and L.M. performed the investigations and analyzed the data. Y.R. and S.Y. wrote the original draft. Y.R., Y.L., M.G., and X.L. contributed to the editing of the manuscript. All authors discussed the results and reviewed the manuscript. **Competing interests:** X.L. and H.Z. have filed a patent based on the concept developed in this manuscript. The other authors declare no competing interests. **Data and materials availability:** All data are available in the main text or the supplementary materials. **License information:** Copyright © 2023 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. https://www.science.org/about/science-licenses-journal-article-reuse

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.add8786
Materials and Methods
Supplementary Text
Figs. S1 to S25
Tables S1 to S4
References (*64*–*70*)

Submitted 11 July 2022; accepted 15 December 2022
10.1126/science.add8786

OPTICS

Breaking the limitation of polarization multiplexing in optical metasurfaces with engineered noise

Bo Xiong[1]†, Yu Liu[1]†, Yihao Xu[2]†, Lin Deng[3], Chao-Wei Chen[1], Jia-Nan Wang[1], Ruwen Peng[1]*, Yun Lai[1], Yongmin Liu[2,3]*, Mu Wang[1,4]*

Noise is usually undesired yet inevitable in science and engineering. However, by introducing the engineered noise to the precise solution of Jones matrix elements, we break the fundamental limit of polarization multiplexing capacity of metasurfaces that roots from the dimension constraints of the Jones matrix. We experimentally demonstrate up to 11 independent holographic images using a single metasurface illuminated by visible light with different polarizations. To the best of our knowledge, it is the highest capacity reported for polarization multiplexing. Combining the position multiplexing scheme, the metasurface can generate 36 distinct images, forming a holographic keyboard pattern. This discovery implies a new paradigm for high-capacity optical display, information encryption, and data storage.

Metasurfaces (*1*–*3*) provide rich degrees of freedom (DOFs) to modulate the phase (*4*, *5*), amplitude (*6*, *7*), and polarization (*8*, *9*) of light based on subwavelength structures. Compared with traditional optical devices, metasurfaces possess distinct advantages of the ultrathin thickness (*10*, *11*), high efficiency (*12*–*14*), and broadband responses (*15*–*17*). In recent years, metasurfaces have been demonstrated for wavelength multiplexing (*18*, *19*), angle multiplexing (*20*), and orbital angular momentum (OAM) multiplexing (*21*), as well as reconfigurable metadevices (*22*, *23*).

Polarization (*24*–*27*) has been explored in metasurfaces for large-capacity multiplexing technology, which can transmit information through independent channels to different desired targets. However, there exists an upper limit in the capacity of polarization multiplexing. Traditionally, two orthogonal linear polarizations (e.g., x- and y-polarized components) operate dual independent functionalities without cross-talk in a two-dimensional (2D) planar optical platform such as a metasurface (*28*, *29*). This approach relies on the diagonal elements of the Jones matrix, which is a 2 × 2 matrix representing the optical response of a metasurface. Amplitude modulation has also been introduced in polarization multiplexing. Applying Malus's law has led to the design of metadevices to encode nanoprinting and holography simultaneously in the near- and far-field, respectively (*30*, *31*). More recently, the off-diagonal elements of the Jones matrix were used as an extra degree of freedom, realizing three independent polarization channels (*32*) and achieving triple sets of printing-hologram integrations (*33*). This approach reaches the theoretical upper limit (≤3 independent channels) of polarization multiplexing capacity in 2D planar metasurfaces. This restriction applies not only to linear polarizations, but also to circular polarizations [left circular polarization (LCP) and right circular polarization (RCP)] (*34*, *35*), and even arbitrary polarizations (*36*). Although polarization multiplexing has been substantially advanced through physics-guided forward design (*37*, *38*) and data-driven inverse design based on machine learning (*39*, *40*), the fact that the number of independent channels is limited to three is a notable constraint in practical applications. Breaking this upper limit will be crucial for developing high-capacity optical displays and data storage.

We develop an approach based on noise engineering that breaks this fundamental limit.

[1]National Laboratory of Solid State Microstructures, School of Physics, and Collaborative Innovation Center of Advanced Microstructures, Nanjing University, Nanjing, 210093, China. [2]Department of Mechanical and Industrial Engineering, Northeastern University, Boston, MA 02115, USA. [3]Department of Electrical and Computer Engineering, Northeastern University, Boston, MA 02115, USA. [4]American Physical Society, 100 Motor Pkwy, Hauppauge, NY 11788, USA.
***Corresponding author. Email: rwpeng@nju.edu.cn (R.P.); y.liu@northeastern.edu (Y.Liu); muwang@nju.edu.cn (M.W.)**
†These authors contributed equally to this work.

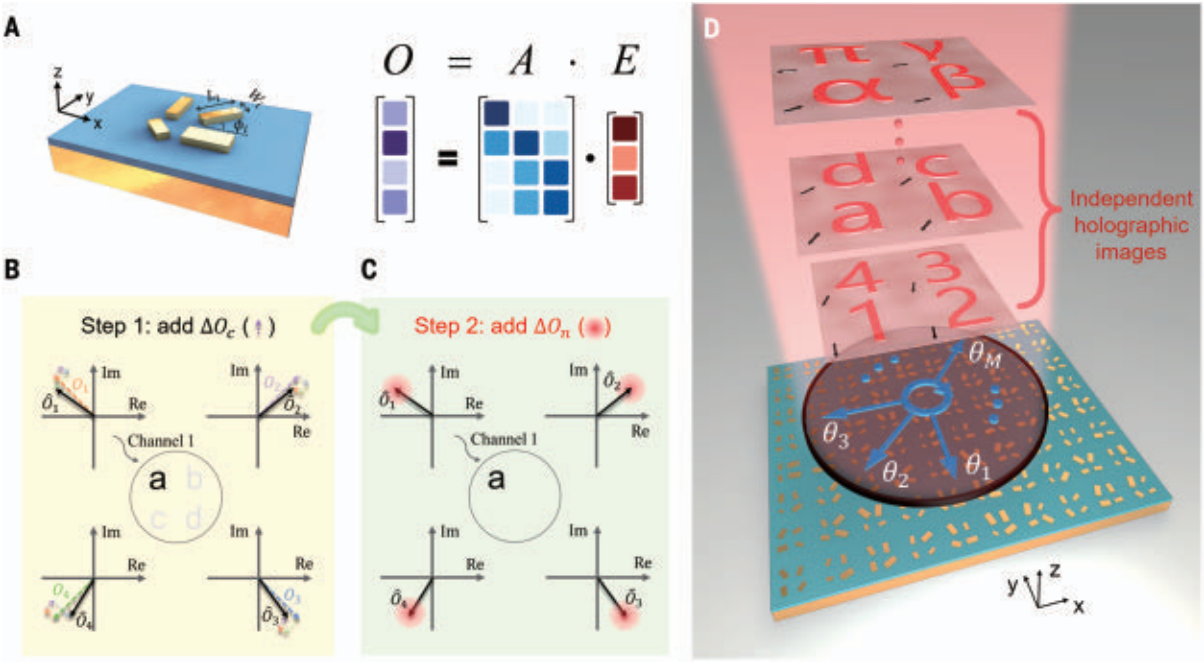


Fig. 1. Schematic of the design of polarization-multiplexed metasurface. (**A**) Designed unit cell and the corresponding response matrix. (**B**) In step 1, the least-squares solutions (solid black arrows) with the correlated noise (dashed short arrows) are introduced to create new polarization channels, where the target signals (dashed long arrows) are entirely distinct from each other. However, as shown in the inset, the solved complex outputs still have some correlations, resulting in the moderate cross-talk of holographic images. (**C**) In step 2, the random noncorrelated noise (red shadow circles) is introduced, which breaks the correlation between different channels and eliminates the cross-talk. (**D**) Schematic of the polarization-multiplexed metasurface. By changing the polarization direction of the incident light, the metasurface can generate multiple independent holographic images.

By introducing the correlated noise with least-squares estimation, new polarization channels can be created with only moderate cross-talk. To reduce the cross-talk, we further introduce the noncorrelated noise with a random distribution to push the limit of polarization multiplexing. Illuminated by the incidence with different linear polarization states, the metasurface can generate up to 11 independent holographic images at the corresponding channels. Simulation shows that the capacity could be expanded further by combining our methodology with position multiplexing or OAM multiplexing. As a proof of concept, we demonstrate 36 distinct holographic images, forming a typical keyboard pattern by simultaneously using polarization multiplexing and position multiplexing.

Design strategy for engineered noise

Our design strategy is based on approximate solutions of the Jones matrix elements with sufficient engineered noise, allowing us to overcome the constraint in the polarization multiplexing capacity of metasurfaces (Fig. 1). The unit cell of the metasurface (Fig. 1A) working in the reflective configuration consists of multiple rectangular resonators with distinct lengths (L), widths (W), and rotation angles (ϕ), each of which supports two eigen dipoles along the short and long axes (see details in supplementary note 1). The Jones matrix of the unit cell is given by $\begin{bmatrix} E_x & E_{xy} \\ E_{xy} & E_y \end{bmatrix}$, where E_x, E_{xy}, and E_y are complex numbers and represent the effective Jones matrix elements. They are determined by the parameter group $\{L_i, W_i, \phi_i\}$, where i indicates the serial number of rectangular resonators. In the presence of M incident light beams with different linear polarizations, the corresponding optical responses of the unit cell can be generalized to Eq. 1 (details in supplementary note 2)

$$\begin{bmatrix} O_1 \\ O_2 \\ O_3 \\ \cdots \\ O_M \end{bmatrix} = \begin{bmatrix} \cos^2\theta_1 & \sin 2\theta_1 & \sin^2\theta_1 \\ \cos^2\theta_2 & \sin 2\theta_2 & \sin^2\theta_2 \\ \cos^2\theta_3 & \sin 2\theta_3 & \sin^2\theta_3 \\ & \cdots & \\ \cos^2\theta_M & \sin 2\theta_M & \sin^2\theta_M \end{bmatrix} \cdot \begin{bmatrix} E_x \\ E_{xy} \\ E_y \end{bmatrix} \quad (1)$$

Here O_M denotes the complex reflected field at the M-th linear polarization channel along the polarization angle θ_M (with respect to the x axis). Equation 1 is a collection of nonhomogeneous linear equations that can be written as $O = A \cdot E$. According to linear algebra (*41*), Eq. 1 has a precise solution for the required Jones matrix elements only if $M \leq 3$. Therefore, the independent channels for polarization multiplexing have an upper limit of 3, consistent with the conclusion of several previous works (*33*–*36*) based on the analysis of DOFs. When $M = 4$, the fourth channel is the superposition of the first three polarization channels (see details in supplementary note 3).

We deliberately introduce noise into our system to break the polarization multiplexing limit of metasurfaces, offering new DOFs in the design space. In the first step, the introduction of correlated noise creates additional polarization channels by considering the least-squares solution (*41*) instead of the precise solution of Eq. 1. By solving such overdetermined linear equations, the correlated noise is automatically introduced. The new solution can be written as $\hat{E} = (A^T A)^{-1} A^T O$ (see details in supplementary note 4), and thus the complex outputs $\hat{O} = A \cdot \hat{E}$ (e.g., $\hat{O}_1$, $\hat{O}_2$, $\hat{O}_3$, $\hat{O}_4$ when $M = 4$) can be expressed by the solid black arrows in Fig. 1B. The dashed long arrows represent the target signals at the corresponding polarization channels, whereas the dashed short ones represent the correlated noise (defined as $\Delta O_c = |A \cdot \hat{E} - O|$). It follows that the target signal of the fourth channel is entirely distinct from the other three, indicating the generation of new functionality. However, we emphasize that the correlated noise can be theoretically decomposed into the perturbations from the other channels (see details in supplementary note 4). The outputs at different polarization channels

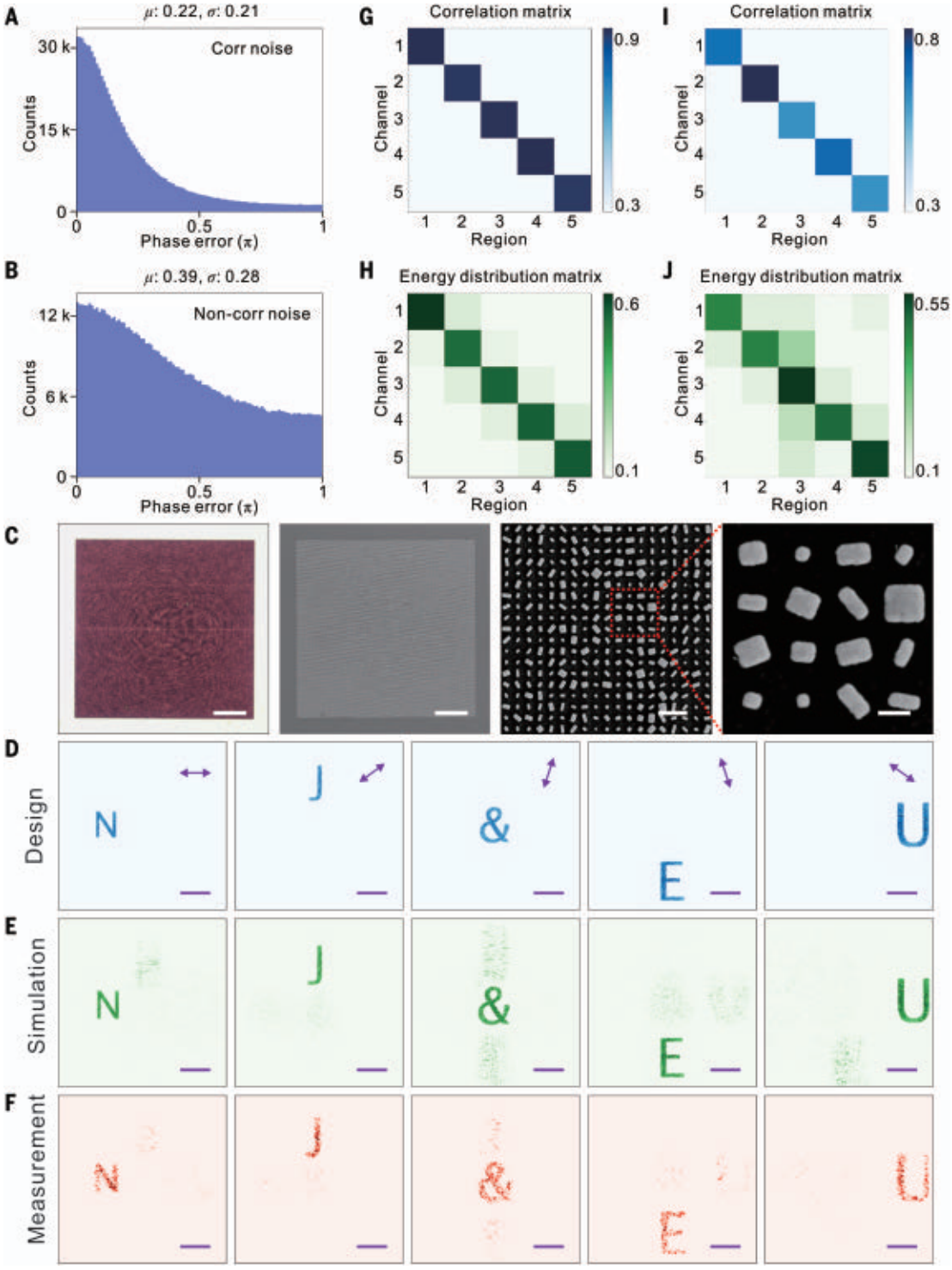


Fig. 2. Design and experimental demonstration of the five-channel polarization-multiplexed metasurface. (**A**) Calculated correlated noise distribution based on Eq. 1, which has a mean value of ~0.22π and a standard deviation of ~0.21π. (**B**) Added noncorrelated noise distribution with a mean value of 0.39π and a standard deviation of 0.28π. (**C**) Optical and SEM micrographs of the fabricated sample with different magnifications. The scale bars from left to right are 40 μm, 40 μm, 0.5 μm, and 250 nm, respectively. (**D**) Designed, (**E**) simulated, and (**F**) measured holographic images with different polarization incidences, which are marked by the purple arrows. All scale bars are 40 μm. When the incident light is polarized along 0°, 36°, 72°, 108°, and 144°, the metasurface will generate five independent holographic images, including "N", "J", "&", "E", and "U". (**G**) Calculated correlation coefficients matrix and (**H**) energy distribution matrix based on the simulation results in (E). (**I**) Measured correlation coefficients matrix and (**J**) energy distribution matrix based on the experimental results in (F), which agree well with the simulations.

are linearly correlated, resulting in the moderate cross-talk of holographic images between different polarization channels (Fig. 1B, insets).

To eliminate such cross-talk, we proceed to the second step and add the noncorrelated noise (details shown in supplementary note 5), represented by the red shadow circles in Fig. 1C. In this case, the new solution is expressed as $\hat{E}_n = (A^T A)^{-1} A^T O + \delta E$, where each component of δE is randomly sampled from a 2D complex normal distribution $r \times N(0, I)$. Here r indicates the amplitude of the random noise, and I is the identity matrix. Then the noncorrelated noise can be written as $\Delta O_n = |A \bullet \delta E|$. Such noise is similar to the additive Gaussian noise in information theory (*42*), which mimics the effects of many random processes occurring in nature, including black body radiation from Earth and other warm objects, thermal vibrations of atoms in conductors (referred to as thermal noise), and random manufacturing error in experiments. According to the central limit theorem (*43*), the probability distribution of the mean noise produced by vast random events actually approaches a Gaussian distribution. They are typically undesirable but ultimately are beneficial in our design. As illustrated in fig. S5, when the value of r increases from 0 to 1, the correlation between the neighboring channels is gradually broken down by the noncorrelated noise, although the intensity in the target channels becomes lower. Consequently, the cross-talk between different polarization channels diminishes by selecting an appropriate value of r, realizing polarization multiplexing with a single metasurface (Fig. 1D). This result highlights that correlated and noncorrelated noise have distinct physical origins and implications. The correlated noise can be passively introduced by solving the overdetermined linear equations, which creates additional information channels with certain cross-talk. The intentionally introduced noncorrelated noise with a Gaussian noise formula substantially reduces the cross-talk,

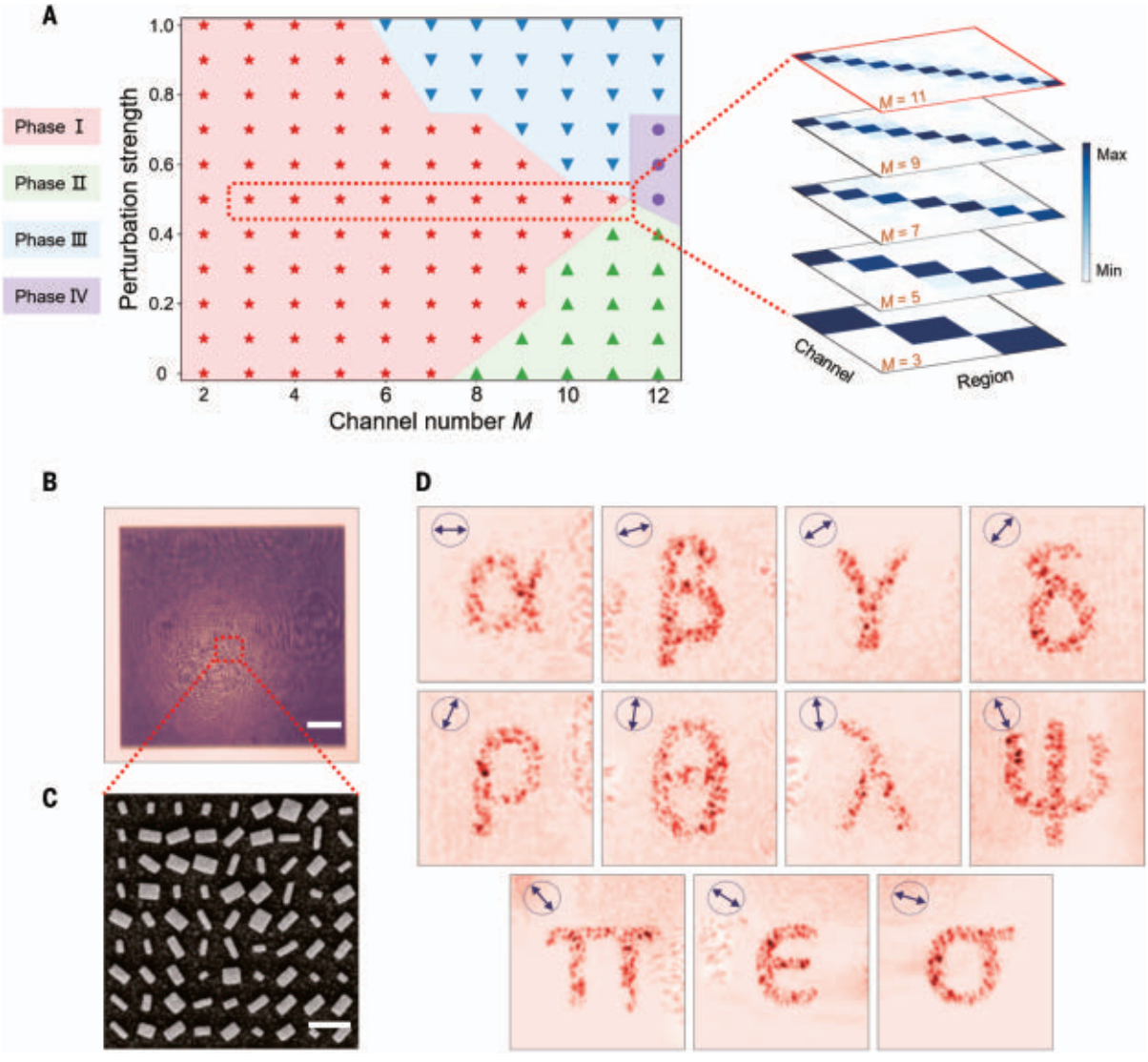


Fig. 3. Simulation and experimental demonstration of the new upper limit. (**A**) Phase diagram for polarization multiplexing in our design. Phase I represents the conditions that satisfy the threshold requirements. Phase II represents the cases in which the noncorrelated noise is so small that the nondiagonal term is more significant than 0.2, indicating that the cross-talk is rather substantial. Phase III represents the cases in which the noncorrelated noise is so large that the diagonal term is larger than 0.8, indicating that the energy in the design channel is rather weak. Phase IV indicates the scenarios where all the threshold conditions are not satisfied. The inset in the right panel shows the corresponding correlation coefficient matrices for different channel numbers from $M = 3$ to 11, all of which satisfy the threshold requirements. (**B**) Optical and (**C**) SEM micrograph of the fabricated sample for $M = 11$. The scale bars are 50 μm and 0.4 μm, respectively. (**D**) Measured holographic images at 11 linear polarization states. The arrows denote the polarization orientations.

despite sacrificing the signal-to-noise ratio in the primary channels.

Testing the design strategy

We first take $M = 5$ to test our design. Figure 2A shows the calculated histogram phase distribution of ΔO_c with a mean value of ~0.22π and a standard deviation of ~0.21π. Such a noise value is relatively low in the metasurface design, close to the noise amplitude in four-level phase discretization (~0.25π). It follows that five holographic images can be generated at five polarization states (fig. S6). However, the created holographic images still contain some cross-talk from the other channels. To suppress the cross-talk, we introduce an appropriate noncorrelated noise ($r = 0.3$) with a mean value of ~0.39π and a standard deviation of ~0.28π (Fig. 2B), which is similar to the case of binary phase design. These results demonstrate that the engineered noise is within the margin of the design error. Then we design a device with multiplexed holographic images (Fig. 2D), following the iterative Fourier transform algorithm (*44*) and genetic algorithm (*45*). The final approximate solutions with engineered correlated and noncorrelated noise are fed into the genetic algorithm as complex numbers, which can search the optimized geometrical parameters for each pixel with the smallest mean square error. The genetic algorithm is critical to the design of the metasurface. By mimicking the natural processes of selection, reproduction, mutation, and crossover, the genetic algorithm is a metaheuristic algorithm that can solve global optimization problems and effectively explore the parameter space (details shown in supplementary note 6). Figure 2E presents the simulation result in which five independent holographic images can be generated from the same metasurface when the polarization of incident light varies. Figure 2C shows the optical and scanning electron microscopy (SEM) micrographs of the sample (220 μm by 220 μm in size) fabricated by standard electron beam lithography. The working wavelength of the metasurface is 700 nm. From the measurement results presented in Fig. 2F, we can observe that the same metasurface generates five independent holographic images—"N", "J", "&", "E", and "U"—at the corresponding polarization states (0°, 36°, 72°, 108°, 144°) (details shown in supplementary note 7).

To quantify the cross-talk between the images at different channels, we introduce the correlation coefficient, which is defined as (*46*, *47*)

$$\rho(T,R) = \frac{\mathrm{COV}(T,R)}{\sqrt{D(T)\cdot D(R)}} \qquad (2)$$

Here T and R represent the intensity distribution of the target object and reconstructed object, $\mathrm{COV}(T,R)$ is the covariance of T and R, and $D(T)$ and $D(R)$ represent the variances of T and R, respectively. According to this definition, if we consider M polarization channels, we can calculate the correlation matrix with a size of $M \times M$, in which the diagonal terms represent the target signal, whereas the nondiagonal values indicate the cross-talk. The simulated correlation coefficients of the nondiagonal terms are all below 0.3 (Fig. 2G), confirming that these polarization channels are relatively uncorrelated and hence independent. Moreover, up to 60% of the power is directed to the targeted channel (Fig. 2H). The measured



Fig. 4. Experimental demonstration of the holographic keyboard. (**A**) Optical and (**B**) SEM micrograph of the fabricated sample with simultaneously polarization- and position-multiplexed design. The scale bars are 50 μm and 0.4 μm, respectively. (**C**) Schematic and (**D**) measured results for the performance of the holographic keyboard pattern. Here we choose nine different polarization channels covering from 0° to 160°. Four holographic images are simultaneously generated at different planes away from the metasurface for each polarization channel. The locations of the planes are z_1 = 330 μm, z_2 = 430 μm, z_3 = 570 μm, and z_4 = 830 μm. As a result, 36 holographic images can be generated to form a typical keyboard pattern.

correlation matrix and energy distribution matrix in Fig. 2, I and J, also agree well with the simulations. Here the fabrication error acts as an additional noncorrelated noise component that further reduces the cross-talk between the channels. These results demonstrate that by combining ΔO_c and ΔO_n, we can successfully break the constraint in the polarization multiplexing capacity of metasurfaces and realize the polarization-multiplexed metasurface with an independent channel number larger than 3.

The upper limit for channel number

We calculate the phase diagram for polarization multiplexing (Fig. 3). The correlated noise distributions from $M = 2$ to $M = 12$ are first solved (fig. S9). The mean value of ΔO_c increases slowly from 0 to ~ 0.33π as the channel number increases. This is because the least-squares solution requires more perturbations to fulfill Eq. 1 when M becomes larger. We have also swept the value of r from 0 to 1, and the calculated correlation matrices for each scenario are shown in fig. S11. It is noteworthy that because of the random nature of ΔO_n, the correlation matrices vary from time to time, even when the same level of ΔO_n is applied. To prevent fluctuation in the results, we take the statistical average of 10 correlation matrices. Lastly, to evaluate the multiplexing capacity, we follow a threshold value of 0.8 (*47*) for the correlation coefficient, resulting in the phase diagram shown in Fig. 3A. The examples that meet the threshold requirements (diagonal terms larger than 0.8 and nondiagonal terms less than 0.2 in this study) are marked in red (Phase I). By contrast, the rest (Phase II, Phase III, and Phase IV) cannot meet the requirements. When the noncorrelated noise is too small, the green region (Phase II, nondiagonal terms larger than 0.2) shows that the cross-talk is too strong. When the noncorrelated noise is too strong, the blue area (Phase III, diagonal terms less than 0.8) indicates that the patterns in the targeted channels have low intensity. We can see that the new upper limit is ~11 with $r = 0.5$. The upper limit might be increased further by relieving the threshold conditions (fig. S12). Figure 3, B and C, shows the optical and SEM micrographs of the sample for $M = 11$. Figure 3D presents the multiplexed holographic images measured in our experiment. By changing the polarization angles of incident light from 0° to 163.63° in the step of 16.36° each, 11 different holographic images are generated from the same metasurface. Despite some background noise, we can recognize all the designed Greek letters, including α, β, γ, δ, ρ, θ, λ, ψ, π, ε, and σ, demonstrating excellent performance of polarization multiplexing design in the new upper limit. The experimental results agree very well with the simulations shown in fig. S13. In addition to holograms, we expect that our design methodology can be applied to other metadevices such as metalenses. The key is to choose the appropriate parameters as the figure of merit and find the phase diagram for the specific applications (details shown in supplementary note 8).

Furthermore, our polarization multiplexing strategy based on noise engineering can be combined with other multiplexing methods, such as position multiplexing (fig. S14) and OAM multiplexing (fig. S15), to boost the information-encoding capacity. As in the last example, nine linear polarization channels (0°, 20°, 40°, 60°, 80°, 100°, 120°, 140°, 160°) are chosen, and each channel contains four different holographic images at the planes 330, 430, 570, and 830 μm away from the metasurface (Fig. 4). The device is 330 μm by 330 μm in size. The optical and SEM micrographs of the metasurface are shown in Fig. 4, A and B. The target holographic images for the 36 channels at the wavelength of 700 nm are 26 letters and 10 Arabic digits, resulting in a typical keyboard pattern, as seen in the dashed region of Fig. 4C. Figure 4D depicts the measured results from the constructed metasurface, which create a total of 36 separate holographic images and match the design well. To the best of our knowledge, this is the largest capacity reported in metasurfaces with polarization and position multiplexing design, and it can be further improved by OAM multiplexing, wavelength multiplexing, angle multiplexing, and other techniques. The design combining noise engineering and OAM multiplexing is presented in supplementary note 9.

Concluding remarks

To summarize, we break the polarization multiplexing limit of metasurfaces with engineered noise in the visible region. It is expected that the capacity of traditional optical communication systems (e.g., wavelength multiplexing) and information storage systems (e.g., polarization multiplexing) can be further boosted with more parallel channels by using our approach. Furthermore, although we focus on metadevices operating in the visible frequencies, this design strategy can be readily applied to infrared, terahertz, or microwave frequencies. It opens a new pathway to enhance parallel processing, reduce cross-talk between channels, and increase the storage capacity in optical communication, optical encryption, and information storage systems.

REFERENCES AND NOTES

1. N. Yu *et al*., *Science* **334**, 333–337 (2011).
2. X. Ni, N. K. Emani, A. V. Kildishev, A. Boltasseva, V. M. Shalaev, *Science* **335**, 427–427 (2012).
3. S. Sun *et al*., *Nat. Mater.* **11**, 426–431 (2012).
4. L. Huang *et al*., *Nano Lett.* **12**, 5750–5755 (2012).
5. S. Chen, Z. Li, W. Liu, H. Cheng, J. Tian, *Adv. Mater.* **31**, e1802458 (2019).
6. Q. Wang *et al*., *ACS Photonics* **5**, 599–606 (2017).
7. J. N. Wang *et al*., *Small* **17**, e2101282 (2021).
8. Z. L. Deng *et al*., *Nano Lett.* **18**, 2885–2892 (2018).
9. Z. J. Wang *et al*., *ACS Photonics* **3**, 2096–2101 (2016).
10. Y. Zhao, M. A. Belkin, A. Alù, *Nat. Commun.* **3**, 870 (2012).
11. D. Lin, P. Fan, E. Hasman, M. L. Brongersma, *Science* **345**, 298–302 (2014).
12. M. I. Shalaev *et al*., *Nano Lett.* **15**, 6261–6266 (2015).
13. G. Zheng *et al*., *Nat. Nanotechnol.* **10**, 308–312 (2015).
14. M. Khorasaninejad *et al*., *Science* **352**, 1190–1194 (2016).
15. S. Wang *et al*., *Nat. Nanotechnol.* **13**, 227–232 (2018).
16. W. T. Chen, A. Y. Zhu, J. Sisler, Z. Bharwani, F. Capasso, *Nat. Commun.* **10**, 355 (2019).
17. R. H. Fan, B. Xiong, R. W. Peng, M. Wang, *Adv. Mater.* **32**, e1904646 (2020).
18. B. Wang *et al*., *Nano Lett.* **16**, 5235–5240 (2016).
19. B. Xiong *et al*., *Adv. Mater.* **33**, e2005864 (2021).
20. S. M. Kamali *et al*., *Phys. Rev. X* **7**, 041056 (2017).
21. X. Y. Fang, H. R. Ren, M. Gu, *Nat. Photonics* **14**, 102–108 (2020).
22. F. Z. Shu *et al*., *Laser Photonics Rev.* **15**, 2100155 (2021).
23. L. Li *et al*., *Nat. Commun.* **8**, 197 (2017).
24. M. Born, E. Wolf, *Principles of Optics* (Pergamon, ed. 6, 1980).
25. N. K. Grady *et al*., *Science* **340**, 1304–1307 (2013).
26. Y. J. Gao *et al*., *Phys. Rev. X* **10**, 031035 (2020).
27. E. Arbabi, S. M. Kamali, A. Arbabi, A. Faraon, *ACS Photonics* **6**, 2712–2718 (2019).
28. A. Arbabi, Y. Horie, M. Bagheri, A. Faraon, *Nat. Nanotechnol.* **10**, 937–943 (2015).
29. F. Ding, R. Deshpande, S. I. Bozhevolnyi, *Light Sci. Appl.* **7**, 17178 (2018).
30. J. Li *et al*., *Adv. Mater.* **33**, e2007507 (2021).
31. Z. L. Li *et al*., *Laser Photonics Rev.* **14**, 2000032 (2020).
32. Z. Li *et al*., *Opt. Express* **29**, 134–144 (2021).
33. Y. Bao, L. Wen, Q. Chen, C. W. Qiu, B. Li, *Sci. Adv.* **7**, eabh0365 (2021).
34. D. Wen *et al*., *Nat. Commun.* **6**, 8241 (2015).
35. F. Zhang, M. B. Pu, J. Luo, H. L. Yu, X. G. Luo, *Opto-Electron. Eng.* **44**, 319–325 (2017).
36. J. P. Balthasar Mueller, N. A. Rubin, R. C. Devlin, B. Groever, F. Capasso, *Phys. Rev. Lett.* **118**, 113901 (2017).
37. R. Zhao *et al*., *Light Sci. Appl.* **7**, 95 (2018).
38. H. Zhou *et al*., *ACS Nano* **14**, 5553–5559 (2020).
39. D. Zhu, Z. Liu, L. Raju, A. S. Kim, W. Cai, *ACS Nano* **15**, 2318–2326 (2021).
40. W. Ma *et al*., *Adv. Mater.* **34**, e2110022 (2022).
41. K. Hoffman, R. Kunze, *Linear Algebra*, Prentice-Hall, ed. 2 (1971).
42. A. E. Gamal, Y. H. Kim, *Network Information Theory*, Cambridge University Press (2011).
43. S. Ihara, *Information Theory for Continuous Systems*, World Scientific (1993).
44. Z. Shi *et al*., *Nano Lett.* **18**, 2420–2427 (2018).
45. M. D. Huntington, L. J. Lauhon, T. W. Odom, *Nano Lett.* **14**, 7195–7200 (2014).
46. G. Situ, J. J. Zhang, *J. Opt. A, Pure Appl. Opt.* **8**, 391–397 (2006).
47. Q. S. Wei, L. L. Huang, X. W. Li, J. Liu, Y. T. Wang, *Adv. Opt. Mater.* **5**, 1700434 (2017).

ACKNOWLEDGMENTS

Funding: R.P. and M.W. acknowledge financial support from the National Key R&D Program of China (grant nos. 2022YFA1404303 and 2020YFA0211300) and the National Natural Science Foundation of China (grants nos.12234010, 61975078, and 11974177). Yongmin Liu acknowledges financial support from the National Science Foundation (ECCS-1916839, ECCS-2136168, and DMR-2202268).
Author contributions: R.P., B.X., Yongmin Liu, Y.X., and M.W. conceived this work. B.X., Y.X., and L.D. contributed to the theoretical design; B.X., Yu Liu, and C.-W.C. fabricated the samples with the assistance of J.-N.W.; B.X. and Yu Liu performed the optical experiments; B.X., Y.X., L.D., R.P., Yongmin Liu, M.W., and Y. Lai contributed to the data analysis and discussion. R.P., Yongmin Liu, and M.W. directed and supervised the research. B.X., Y.X., L.D., Yongmin Liu, R.P., and M.W. drafted and finalized the manuscript. All authors revised the manuscript. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the main text and supplementary materials.

License information: Copyright © 2023 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. https://www.science.org/about/science-licenses-journal-article-reuse

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.ade5140
Materials and Methods
Supplementary Text
Figs. S1 to S15
References (*48*)

Submitted 21 August 2022; accepted 16 December 2022
10.1126/science.ade5140

STELLAR ASTROPHYSICS

Spin-down by dynamo action in simulated radiative stellar layers

Ludovic Petitdemange[1]*†, Florence Marcotte[2]*†, Christophe Gissinger[3,4]*†

The evolution of a star is influenced by its internal rotation dynamics through transport and mixing mechanisms, which are poorly understood. Magnetic fields can play a role in transporting angular momentum and chemical elements, but the origin of magnetism in radiative stellar layers is unclear. Using global numerical simulations, we identify a subcritical transition from laminar flow to turbulence caused by the generation of a magnetic dynamo. Our results have many properties of the theoretically proposed Tayler-Spruit dynamo mechanism, which strongly enhances transport of angular momentum in radiative zones. The dynamo generates deep toroidal fields that are screened by the stellar outer layers. This mechanism could produce strong magnetic fields inside radiative stars without an observable field on their surface.

As young stars form through accretion, or after aging stars have burned through their hydrogen fuel, the star's core contracts. Conservation of angular momentum causes the core to spin up (rotate at a faster rate), producing strong gradients in the angular velocity as a function of radius (the rotation profile), a situation known as differential rotation. Measurements with stellar pulsations (asteroseismology) have shown that stars at various stages of their evolution have internal rotation profiles that are flatter than expected from stellar evolution models, especially across radiative zones—the layers in which the outward transport of energy occurs through radiative diffusion, rather than convection (*1*–*3*). This discrepancy could be resolved if there were an unidentified mechanism that extracts angular momentum from the stellar core and suppresses differential rotation as the star evolves (*4*).

A potential mechanism for enhanced angular momentum transport is stellar magnetism (*5*–*7*). Magnetic fields on stellar surfaces can collimate jets of plasma (*8*) or power flares (*9*). Theoretical treatments of magnetic fields in radiative zone models have shown that they modify the predicted dynamics of stars (*10*–*14*). However, these predictions are limited by two theoretical problems. First, magnetic fields are difficult to observe in deep stellar layers, including the radiative cores of stars with less than ~1.3 solar masses. Even in stars for which the radiative zone is located in the envelope (outside the core), in most cases the amplitude of potential magnetic fields falls below the detection limits of direct observations of stellar surfaces—except for 10% of these stars, for which strong dipolar magnetic fields have been measured (*15*). Secondly, the mechanism by which a dynamo magnetic field can be generated inside a radiative stellar layer remains unclear.

Theoretical predictions

Dynamo instability is the spontaneous development of an amplification loop by which a magnetic field directed toward the poles (poloidal field) is converted into a magnetic field parallel to lines of latitudes (toroidal field), and vice versa (*16*). This conversion is mediated by plasma motions, which must be sufficiently powerful for an initially weak magnetic field to undergo self-amplification. In convective stellar layers, the required flow complexity can be provided by turbulent buoyant plumes (*17*, *18*). However, in stably stratified, radiative

[1]Laboratoire d'Etudes du Rayonnement et de la Matière en Astrophysique et Atmosphères (LERMA), Observatoire de Paris, Paris Sciences & Lettres (PSL) Research University, French National Centre for Scientific Research (CNRS), Sorbonne Université, Paris, France. [2]Université Côte d'Azur, National Institute for Research in Digital Science and Technology (Inria), CNRS, Laboratoire J.A. Dieudonné, Nice, France. [3]Laboratoire de Physique de l'Ecole Normale Supérieure (ENS), Paris Sciences & Lettres (PSL) Research University, CNRS, Paris, France. [4]Institut Universitaire de France, Paris, France.
***Corresponding author. Email: ludovic.petitdemange@upmc.fr (L.P.); florence.marcotte@inria.fr (F.M.); christophe.gissinger@phys.ens.fr (C.G.)**
†These authors contributed equally to this work

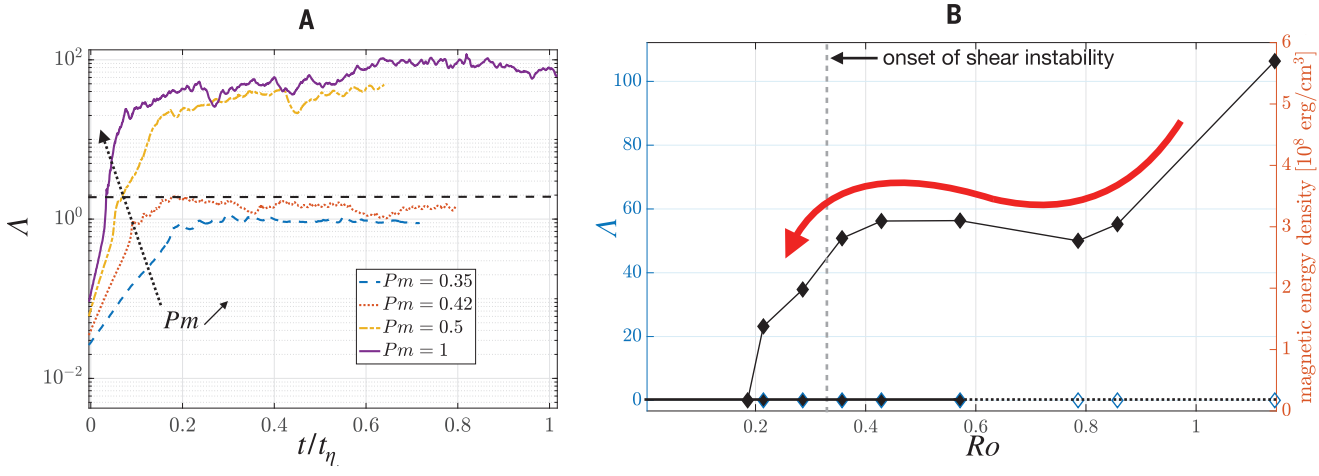


Fig. 1. Subcritical dynamo bifurcation. (**A**) Time series of the magnetic energy (quantified by the dimensionless Elsasser number Λ) at fixed rotation, stratification, and [specifically $E = 10^{-5}$, $N/\Omega = 1.24$, $Pr = 0.1$, $Ro = 0.78$ (*25*)], but varying magnetic Prandtl number {Pm = [0.35 (blue); 0.42 (red); 0.5 (yellow); 1 (purple)]}, ordered as indicated by the dashed black arrow. Times (t) are normalized by the Ohmic diffusion time t_η (*25*). A secondary instability occurs when the magnetic energy exceeds a typical value $\Lambda_* \sim 1$ (dashed black line). (**B**) Time-averaged magnetic energy density of the saturated dynamo as a function of shear rate (quantified by the Rossby number Ro), for $Pm = 1$) [other parameters the same as in (A)]. Empty diamonds indicate linearly unstable solutions, from which the magnetic field grows exponentially. Solid diamonds illustrate the hysteresis cycle (black arrows) between a nonmagnetic solution ($\Lambda = 0$) and a strong, toroidal dynamo solution. The scenario we propose follows the red arrow: the dynamo arises for initially large differential rotation, suppresses the shear as the star evolves, and maintains the magnetic field below the stability threshold.

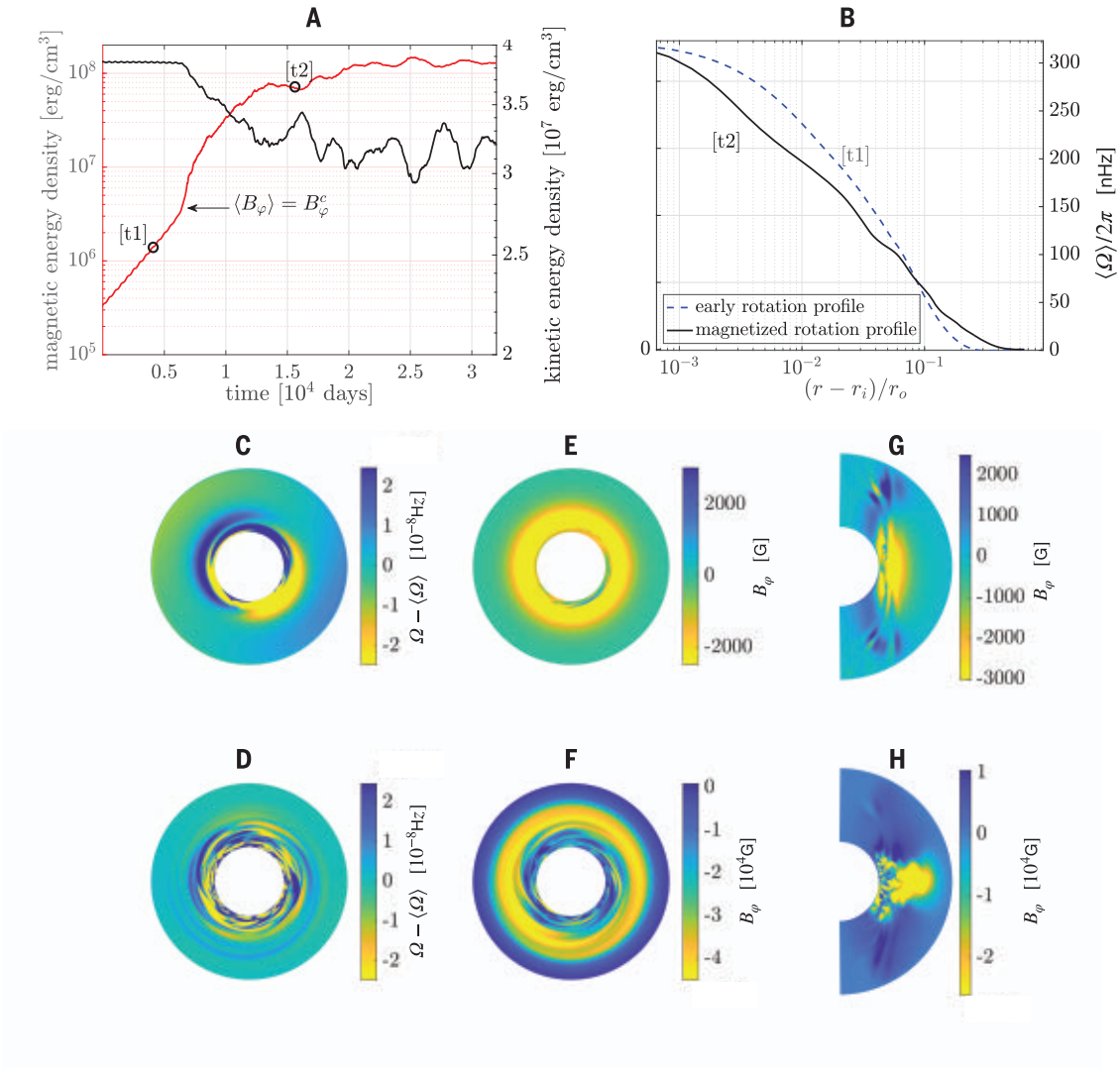


Fig. 2. Transition to turbulence. (**A**) Time series of the total kinetic and magnetic energy densities. The arrow marks the time where the amplitude of the axisymmetric component of the azimuthal magnetic field $\langle B_\varphi \rangle$ locally exceeds the prediction of the Tayler instability, B_φ^C (*21*). Angle brackets denote spatial averaging in the azimuthal direction (φ coordinate). (**B**) Radial profiles of the azimuthally averaged angular velocity $\langle\Omega\rangle$ in the equatorial plane for two distinct times, labeled [t1] and [t2] in (A), with r_i and r_o as the radii of the inner and outer shells, respectively. The onset of the instability causes the rotation profile to flatten between [t1] and [t2]. Snapshots below are of (**C** and **D**) the non-axisymmetric angular velocity in the equatorial plane, (**E** and **F**) the azimuthal magnetic field in the equatorial plane and (**G** and **H**) the same quantity in the meridional plane. Results are shown for the times [t1] (panels C, E, and G) and [t2] (panels D, F, and H), before and after the onset of secondary (Tayler) instability. This simulation has parameters $E = 10^{-5}$, $N/\Omega = 1.24$, $Pr = 0.1$, $Ro = 0.78$, and $Pm = 1$.

layers, dynamo action (and thus magnetic braking) requires a different source of hydrodynamic turbulence.

Several models have been considered to provide angular momentum transport, including internal waves (*19*) or magnetic instabilities (*20*). An example of the latter approach is the Tayler-Spruit (TS) dynamo model (*21*, *22*). In this model, magnetic field generation in radiative layers relies on (i) the winding of a poloidal field into a toroidal field by differential rotation [the Ω-effect (*16*)] and (ii) the destabilization of the resulting strong, toroidal, and axisymmetric magnetic field by the Tayler instability (*23*), which regenerates a poloidal field and thus (in theory) closes the dynamo loop initiated by differential rotation. However, global numerical simulations have not produced TS dynamos, casting doubt upon whether the simplifications made in the theory are valid when the plasma is turbulent (*24*). It is therefore unclear whether a dynamo mechanism could operate in a stably stratified stellar layer. We sought to numerically investigate whether a magnetic field can build up through dynamo instability, trigger magnetohydrodynamic turbulence, and achieve efficient angular momentum transport in a radiative star.

Global numerical simulations

We model a radiative stellar layer by considering the swirling flow of a stratified, nonideal, electrically conducting fluid between two coaxial, spherical shells spinning at different rates. The intensity of the differential rotation is controlled by the dimensionless Rossby number $Ro \equiv \Delta\Omega/\Omega$, where Ω and $\Omega + \Delta\Omega$ are the angular velocities of the outer and inner shell, respectively, and the strength of stratification is quantified by the buoyancy frequency N (*25*).

When the differential rotation across the radiative zone is weak (compared with overall rotation), the flow is stable and axisymmetric and all velocity perturbations decay away rapidly. For steeper rotation profiles, the rotational invariance of the flow is broken by the destabilization of a free shear layer. In this case, Fig. 1A shows the resulting magnetic energy. Initially weak magnetic fields undergo exponential amplification and saturate to a magnetic state, whose final amplitudes can be tuned in our simulations by varying the fluid's ratio of molecular to magnetic diffusivities (the magnetic Prandtl number Pm). For large values of the plasma magnetic diffusivity ($Pm < 0.5$), the shear instability amplifies a laminar and mostly axisymmetric toroidal dynamo which saturates at weak magnetic energies [the dimensionless Elsasser number Λ (*25*) is always less than one]. However, if the toroidal energy exceeds a transition value $\Lambda_* \sim 1$, a secondary

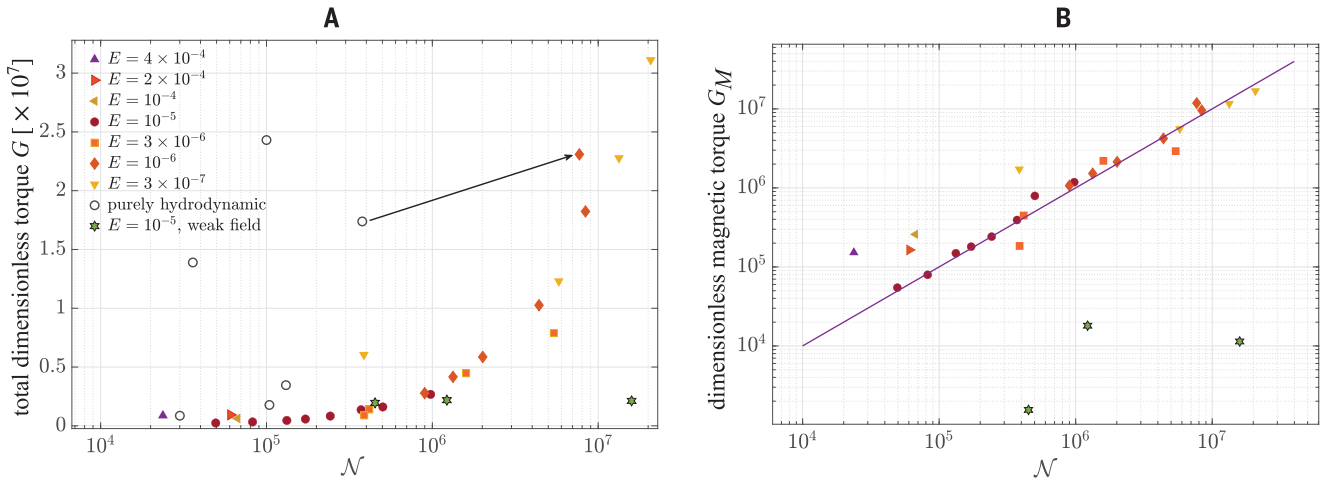


Fig. 3. Angular momentum transport. (**A**) Total dimensionless torque (*25*) exerted on the inner sphere as a function of the dimensionless quantity $N \equiv r_i^{5/2}(U_0\Omega)^{3/2}/N\nu^2$ (*25*). Data points are for simulations that result in non-magnetic flows (empty circles), weak-dynamos (green stars), and strong-dynamos [red, purple, orange, and yellow symbols (see legend)]. The arrow connects two simulations with identical parameters, with and without magnetic fields. (**B**) Results for the magnetic torque only. The purple line is the theoretical prediction for the TS dynamo (*21*).

instability is triggered: the exponential growth steepens, the system becomes turbulent, and the magnetic energy saturates at values nearly two orders of magnitude greater. We refer to these two dynamo solutions as the weak and strong dynamos, by analogy with Earth's dynamo (*26*). Figure 1B shows a bifurcation diagram that illustrates the subcritical behavior of the strong dynamo branch: once the dynamo has been triggered for steep rotation profiles, it can be maintained even when the shear rate is decreased well below that of the onset of hydrodynamic instability.

We propose that, when stars exhaust their fuel, the differential rotation across the radiative zone is large enough (high *Ro*) for extremely weak initial magnetic fields to be amplified by dynamo action. This induces enhanced outward transport of angular momentum, gradually flattening the rotation profile (decreasing the effective *Ro*) as the star evolves. The magnetic field then dynamically adjusts to the smoother rotation profile and sustains the turbulent motions on which it feeds, thus maintaining the magnetic field (Fig. 1B).

Figure 2 shows snapshots of our simulations before and after the steepening of the exponential growth to illustrate how the strong dynamo causes the subcritical transition to hydrodynamic turbulence. Steeper magnetic growth occurs at the same time as flow destabilization; after this time, the magnetic field has become strongly chaotic and exhibits fluctuations at small scale. The corresponding velocity field becomes highly turbulent, especially in the inner regions of the star. This transition to turbulent flow motions suppresses differential rotation across the fluid, causing the rotation to flatten (Fig. 2B) as the strong dynamo builds up (movie S1).

Comparison with the Tayler-Spruit mechanism

The strong dynamo that appears in our simulations shares several properties with the TS mechanism (*21*). First, this dynamo feeds on the interaction between a large-scale, toroidal magnetic field and differential rotation. The resulting fields have a dominant axisymmetric, toroidal component, containing more than 80% of the magnetic energy. Second, the spatial structure displayed by the magnetic field has a smaller length scale in the radial direction than in the azimuthal direction. Third, our simulations show that strong dynamos arise when the maximum amplitude of the axisymmetric component of the azimuthal magnetic field exceeds the local stability threshold of the Tayler instability, in agreement with theoretical predictions (*21*).

However, the TS dynamo loop in our simulations is initiated differently from the theoretical prediction (*21*), which may explain why TS dynamos have not appeared in previous numerical simulations (supplementary text). In theoretical models, the finite amplitude toroidal field required to trigger the Tayler instability is assumed to have grown (linearly) out of an infinitesimal poloidal field, wound up by differential rotation. In our simulations, this initial step is instead provided by the weak dynamo instability, which (exponentially) amplifies the toroidal field and kickstarts a subcritical TS dynamo. Previous numerical simulations have shown that several mechanisms could perform this role in stellar interiors (supplementary text). The dynamo transition we identify is turbulent, with many unstable modes excited simultaneously, which is consistent with laboratory analogs of the Tayler instability (supplementary text). This indicates that a fluctuations-based TS dynamo occurs in our simulations, in which the axisymmetric field is replenished by the mean electromotive force, so magnetic fluctuations influence the saturation mechanism (*27*). The turbulent nature of the transition resolves an issue with the predicted TS dynamo model: that the dynamo loop cannot be closed with a single nonaxisymmetric mode becoming unstable at the onset of the Tayler instability, because winding up the latter would not (alone) replenish the required axisymmetric toroidal field (*24*) (supplementary text).

Angular momentum transport

We quantify the enhanced transport of angular momentum to the outer regions of the star by measuring the total (G) and magnetic (G_M) torques exerted on the swirling fluid. We performed a range of simulations (table S1) that systematically varied the overall rotation rate and the stratification of the radiative layers. Figure 3 shows the resulting torques for strong dynamo action, which follow a transposed version of the theoretically predicted power law $G_M \sim N \equiv \beta r_i^{5/2} \frac{(U_0\Omega)^{3/2}}{N\nu^2}$ (*21*), where r_i is the inner shell radius, ν the plasma viscosity, U_0 the local azimuthal velocity measured in the dynamo region, and β is an adjustable parameter (*25*). Weak-field dynamos do not follow this relation (Fig. 3).

Implications for stellar structure

Our simulations produce a turbulent radiative dynamo that shares many features with the TS model. A strong magnetic field can be

sustained by dynamo action inside a stably stratified radiative zone, suppressing differential rotation and causing spin-down of the stellar core. As such, this dynamo provides a plausible mechanism to account for the enhanced transport of chemical elements and angular momentum in nonconvective stellar layers. In particular, it also provides a potential additional transport mechanism for the radiative layers of solar-type stars (*28*). Helioseismology has shown that the Sun has a flat rotation profile in its radiative zone (*29*).

The poloidal component of the dynamo mechanism we identify is extremely weak, so the resulting magnetic fields are almost entirely toroidal. They are also embedded deep in the star's internal layers, where intense differential rotation takes place. The thick stellar outer layers screen the magnetic field from the surface, preventing direct observations—but might potentially be inferred by using asteroseismology (*30*). Our results therefore provide a physical mechanism for enhanced transport of angular momentum in stellar interiors through a dynamo action that produces no surface magnetic field.

REFERENCES AND NOTES

1. T. Ceillier, P. Eggenberger, R. A. Garcia, S. Mathis, *Astron. Astrophys.* **555**, A54 (2013).
2. S. Deheuvels *et al*., *Astron. Astrophys.* **564**, A27 (2014).
3. T. Van Reeth *et al*., *Astron. Astrophys.* **618**, A24 (2018).
4. M. Cantiello, C. Mankovich, L. Bildsten, J. Christensen-Dalsgaard, B. Paxton, *Astrophys. J.* **788**, 93 (2014).
5. D. Moss, *Mon. Not. R. Astron. Soc.* **257**, 593–601 (1992).
6. H. C. Spruit, *Astron. Astrophys.* **349**, 189–202 (1999).
7. J. Braithwaite, H. C. Spruit, *R. Soc. Open Sci.* **4**, 160271 (2017).
8. B. Albertazzi *et al*., *Science* **346**, 325–328 (2014).
9. G. D. Fleishman *et al*., *Science* **367**, 278–280 (2020).
10. J.-P. Zahn, *Astron. Astrophys.* **265**, 115–132 (1992).
11. A. Heger, S. E. Woosley, N. Langer, H. C. Spruit, in *Stellar Rotation*, Proceedings of IAU Symposium No. 215, A. Maeder, P. Eenens, Eds. (2004), p. 591.
12. A. Maeder, G. Meynet, *Astron. Astrophys.* **411**, 543–552 (2003).
13. V. Petit *et al*., *Mon. Not. R. Astron. Soc.* **466**, 1052–1060 (2017).
14. A. T. Potter, S. M. Chitre, C. A. Tout, *Mon. Not. R. Astron. Soc.* **424**, 2358-2370 (2012).
15. G. A. Wade *et al*., *Mon. Not. R. Astron. Soc.* **456**, 2–22 (2016).
16. H. K. Moffatt, in *Cambridge Monographs on Mechanics and Applied Mathematics* (Cambridge Univ. Press, 1978).
17. P. Charbonneau, *Annu. Rev. Astron. Astrophys.* **52**, 251–290 (2014).
18. A. S. Brun, M. K. Browning, *Living Rev.* Sol. Phys. **14**, 4 (2017).
19. T. M. Rogers, D. N. C. Lin, J. N. McElwaine, H. H. B. Lau, *Astrophys. J.* **772**, 21 (2013).
20. L. Jouve, T. Gastine, F. Lignières, *Astron. Astrophys.* **575**, A106 (2015).
21. H. C. Spruit, *Astron. Astrophys.* **381**, 923–932 (2002).
22. A. Maeder, G. Meynet, *Astron. Astrophys.* **422**, 225–237 (2004).
23. R. J. Tayler, *Mon. Not. R. Astron. Soc.* **161**, 365–380 (1973).
24. J.-P. Zahn, A. S. Brun, S. Mathis, *Astron. Astrophys.* **474**, 145–154 (2007).
25. Materials and methods are available as supplementary materials.
26. P. H. Roberts, in *Rotating Fluids in Geophysics*, P. H. Roberts, A. M. Soward, Eds. (Academic Press, 1978), pp. 421–435.
27. J. Fuller, A. L. Piro, A. S. Jermyn, *Mon. Not. R. Astron. Soc.* **485**, 3661–3680 (2019).
28. P. Eggenberger, A. Maeder, G. Meynet, *Astron. Astrophys.* **440**, L9–L12 (2005).
29. R. Howe, *Living Rev. Sol. Phys.* **6**, 1 (2009).
30. S. Mathis *et al*., *Astron. Astrophys.* **647**, A122 (2021).

ACKNOWLEDGMENTS

This study used the HPC resources of MesoPSL financed by the Région île-de-France and the project EquipMeso (reference ANR-10-EQPX-29-01) of the program Investissements d'Avenir, supervised by the Agence Nationale pour la Recherche. Numerical simulations were also carried out at the CINES Occigen computing centers (GENCI project A001046698). **Funding:** L.P. acknowledges financial support from "Programme National de Physique Stellaire" (PNPS) of CNRS/INSU, France. F.M. acknowledges financial support from the French program "T-ERC" managed by Agence Nationale de la Recherche (Grant ANR-19-ERC7-0008-01). C.G. acknowledges financial support from the French program "JCJC" managed by Agence Nationale de la Recherche (Grant ANR-19-CE30-0025-01). **Author contributions:** L.P., F.M., and C.G. jointly conceptualized the project, defined the methodology, carried out the study, analyzed the data, and wrote the manuscript. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** We used the Parody-JA code (*31*, *32*), which is available from its author at www.ipgp.fr/~aubert/tools.html, combined with the SHTns library (*33*), available at https://bitbucket.org/nschaeff/shtns/src/master/. All the scripts, input parameters, and output data required to reproduce our simulations are archived on Zenodo (*34*). **License information:** Copyright © 2023 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. https://www.science.org/about/science-licenses-journal-article-reuse

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abk2169
Materials and Methods
Supplementary Text
Figs. S1 to S4
Table S1
References (*31*–*43*)
Movie S1

Submitted 29 June 2021; accepted 19 December 2022
10.1126/science.abk2169

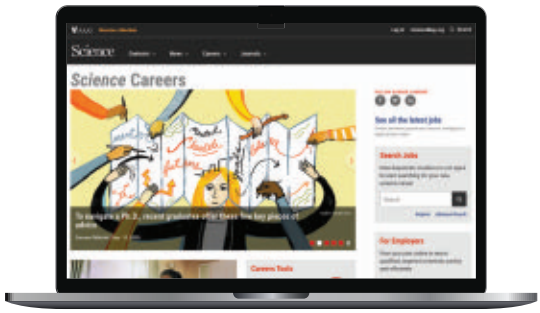
ScienceCareers online @sciencecareers.org



Science Careers helps you advance your career. Learn how!

- Register for a free online account on **ScienceCareers.org**.
- Search hundreds of job postings and find your perfect job.
- Sign up to receive e-mail alerts about job postings that match your criteria.
- Upload your resume into our database and connect with employers.
- Watch one of our many webinars on different career topics such as job searching, networking, and more.
- Download our career booklets, including Career Basics, Careers Beyond the Bench, and Developing Your Skills.
- Complete an interactive, personalized career plan at "my IDP."
- Visit our Employer Profiles to learn more about prospective employers.
- Read relevant career advice articles from our library of thousands.

Visit **ScienceCareers.org** today — all resources are free



SCIENCECAREERS.ORG

ScienceCareers
FROM THE JOURNAL SCIENCE AAAS

63 north
KNOWLEDGE FOR A BETTER WORLD



THE ONSAGER FELLOWSHIPS

16 tenure-track positions available at NTNU

The Norwegian University of Science and Technology (NTNU) is Norway's primary institution for educating the future's engineers and scientists. The university also has strong programmes in the social sciences, teacher education, the arts and humanitites, medicine, architecture and fine art. NTNU's cross-disciplinary research delivers creative innovations that have far-reaching social and economic impact and that help contribute to a better world.

The Onsager Fellowship programme at NTNU is designed to attract the most talented scholars with an established reputation for high quality research and a commitment to learning and teaching at the university level.

APPLY FOR A TENURE-TRACK POSITION AS AN ASSOCIATE PROFESSOR IN:

- Quantum chemistry
- Computational Porous Media Physics
- Socially-innovative cities
- Circular Building Design
- Experimental Environmental Psychology
- Political Behaviour
- Solid Mechanics
- Marine structures
- European Studies
- STS and sustainability transitions
- Post-quantum Cryptography
- Experimental Quantum Communication
- Artificial Intelligence (AI) and Health
- Brain MRS
- Public and Political Economics
- Biosystematics

More info at: www.ntnu.edu/onsager-fellowship/
Closing date: 1 February 2023

NTNU | Norwegian University of Science and Technology

TOP RESEARCH DEMANDS BRILLIANT MINDS – WE'RE ALWAYS LOOKING FOR THE BEST

DARTMOUTH

Center for Technology and Behavioral Health

FACULTY POSITIONS IN RESEARCH METHODS AND DATA ANALYTICS IN DIGITAL HEALTH AS APPLIED TO BEHAVIORAL HEALTH

The Department of Biomedical Data Science at the Geisel School of Medicine at Dartmouth seeks two outstanding individuals to join our faculty within the Center for Technology and Behavioral Health (CTBH), an interdisciplinary research group whose mission is to inform the optimal development, scientific evaluation, and sustainable implementation of digital therapeutics for behavioral health. Candidates who currently hold or have experience commensurate with the rank of **Assistant Professor** are preferred.

The successful candidates will be pursuing a vigorous research program focused on the use of novel methodologies and/or data analytics in digital health as applied to behavioral health. Research focused on clinical effectiveness and/or implementation of digital health behavior-change interventions or research designed to understand mechanisms of behavior change with digital health interventions is also of interest, as are interventions focusing on addictive and/or cancer-causing behaviors.

Applicants for all positions must have earned a PhD and/or MD (or equivalent), have advanced training in fields described above, a successful track record of peer-reviewed publications, and a history of extramural funding. Expectations for teaching and mentoring are a critical part of this role, and individuals will be provided with such opportunities through membership in relevant PhD graduate programs.

Dartmouth College is in the picturesque Upper Connecticut River Valley on the NH and VT border; a vibrant, academic and professional community offering excellent schools, lively arts, and unmatched quality of life in a beautiful setting.

Dartmouth College is an equal opportunity/affirmative action employer with a strong commitment to diversity and inclusion. We prohibit discrimination on the basis of race, color, religion, sex, age, national origin, sexual orientation, gender identity or expression, disability, veteran status, marital status, or any other legally protected status. Applications by members of all underrepresented groups are encouraged.

Using Interfolio (**http://apply.interfolio.com/115050**) applicants should submit a cover letter addressed to Chair of the Search Committee, Dr. Nicholas Jacobson, a curriculum vitae, a 2-3 paged description of their research interests and future research plans, and a statement on teaching interests. Applicants should also arrange for submission of 3 letters of recommendation. Review of applications will continue until position is filled.

Cancer Center at Illinois

SEEKING SENIOR CANCER RESEARCHERS for Phil and Ann Sharp Endowed Faculty Positions

The Cancer Center at Illinois (CCIL) is leading the University of Illinois Urbana-Champaign's (UIUC) campus-wide recruitment of mid-career and senior level cancer-focused researchers. We are seeking several investigators at the **Associate Professor** and **Professor** levels who possess a demonstrated track record of cancer research productivity and interdisciplinary collaboration.

Selected researchers may be awarded the title of Phil and Ann Sharp Scholar in Cancer Research, and exceptionally qualified candidates may be considered for appointment as the Phil and Ann Sharp Endowed Chair in Cancer Research.

Candidates may come from a broad range of disciplines and will be expected to hold tenured positions in one or more UIUC academic departments. In addition to their activities within their home unit, they would also become CCIL members, and potentially assume leadership positions at the program leader or associate director level.

The CCIL emphasizes collaborative technology-biology-drug discovery approaches to cancer research, uniting world-class faculty, resources, and students to initiate and harness the combined power of engineering, technology, and the basic sciences to transform cancer research, detection, and treatment. The current 100+ Cancer Center members represent over 25 departments across campus.

The CCIL features two research programs, Cancer Measurement Technology and Data Science and Cancer Discovery Platforms Bridging the Engineering-Biology Continuum, along with a wide variety of educational and training activities. Candidates will contribute to and collaborate within these research programs and will play an integral role in leading the CCIL to achieve its strategic objectives.

Rank and salary will be commensurate with qualifications. Candidates are expected to demonstrate a commitment to diversity, equity, and inclusion through research, teaching, and/or service.

Please submit inquiries to **cancercenter@illinois.edu.** To learn more about the CCIL, visit: **cancer.illinois.edu**.

The University of Illinois is an Equal Opportunity, Affirmative Action employer that recruits and hires qualified candidates without regard to race, color, religion, sex, sexual orientation, gender identity, age, national origin, disability, or veteran status. For more information, visit http://go.illinois.edu/EEO.

By **Sylvia Dee**

Changing the climate

"Sir, I am a climate scientist. Your remarks are unequivocally false." I listened in admiration as a fellow scientist discredited a climate denier in her audience who had spouted false information during a Q&A session. As a climate scientist who had recently moved to Texas to start a faculty position, I was nervous about having to field similar questions during talks to the general public. I had given those kinds of talks in the past, but they were mostly to groups in more liberal states, where I felt I was preaching to the choir. Texas was a different story.

I have always viewed science communication as one of the most important and urgent parts of my job—and a main reason I wanted to become an expert in climate science in the first place. As a Ph.D. student in California, I leveraged my training by preparing climate change summaries for a local congressman. I coached elementary school teachers in how to teach students about climate. And I worked with young women to empower them to work on climate solutions.

It felt good to step outside academia and share my expertise on issues that were important to me. But at the same time, I sometimes wondered: Am I really making a difference? When I spoke, I was almost always met with knowing nods and immediate moves to action. In Rhode Island, too, where I did my postdoc, my audiences were already on board.

Then, I moved to Texas. My partner had landed a job there, and I was able to secure a postdoc nearby followed by a faculty position. I was excited at the prospect of engaging in outreach with people who weren't yet concerned about the climate. But I wasn't sure what to expect—or how I'd deal with audiences that were skeptical of what I had to say.

I started by doing what I'd always done in the past—calling local organizations and offering to work with them. But this time, it wasn't always smooth sailing. I spoke on panels where I was insulted, threatened by elected officials, and talked down to. Unaccustomed to those kinds of attacks, I became flustered and angry, and sometimes it showed.

Those experiences were stressful and emotionally taxing. But I had to learn to deal with them if I wanted to make an impact. Luckily, I found that the state's network of climate scientists was close-knit. I immediately had colleagues from other universities introducing themselves and inviting me to give talks. They had advice about how to deal with hostility and misinformation. "Don't get visibly frustrated," one told me. "Stay as professional as possible. Be polite but be firm in correcting inaccurate statements," another said.



"I explained the state of the science and gently but directly corrected falsehoods."

During a panel discussion a few months later, I was able to call on their advice. When a fellow panelist made erroneous statements, I explained the state of the science and gently but directly corrected falsehoods. Zoom worked to my advantage: I placed figures in the chat to fact-check misstatements. When I sensed the conversation moving in a bad direction, I didn't become confrontational—I simply spoke when it was my turn and went back to the facts. Although I might not have changed the other panelist's mind, I hope I got through to at least some audience members.

As time wore on, I became more comfortable dealing with those kinds of situations. And I started to truly feel that I was making a difference. During a Girl Scout event, parents came up to me to say they'd never learned about climate science before, with one remarking, "I know you are here to teach the girls, but it was really valuable for the parents, too." A nonprofit organization founded by Republican women also reached out to me, and I've been informally advising one of their members—a former state representative—who is preparing to speak with Republican voters about climate change.

Speaking to audiences in California and Rhode Island may have been more comfortable. But I am glad that I now have the opportunity to do this work where it really matters—in places where I am not preaching to the choir. I finally have an opportunity to help mend a harmful divide. We all need clean air and a stable climate. And so, I encourage my fellow scientists who are interested in public outreach to run toward, rather than away from, these kinds of opportunities. Talk to everyone you can find. People are listening. ■

Sylvia Dee is an assistant professor at Rice University.

ILLUSTRATION: ROBERT NEUBECKER

science.org/journal/sciimmunol



READY TO PUT THE SPOTLIGHT ON YOUR RESEARCH?

Submit your research:
cts.ScienceMag.org

Science Immunology
AAAS

Twitter: @SciImmunology
Facebook: @ScienceImmunology

TRILLIONS OF MICROBES ONE ESSAY

The NOSTER *Science* Microbiome Prize is an international prize that rewards innovative research by investigators who have completed their terminal degree in the last 10 years and are working on the functional attributes of microbiota. The research can include any organism with the potential to contribute to our understanding of human or veterinary health and disease, or to guide therapeutic interventions. The winner and finalists will be chosen by a committee of independent scientists chaired by a senior editor of *Science*. The top prize includes a complimentary membership to AAAS, an online subscription to *Science*, and USD 25,000. Submit your research essay today.



Jennifer Hill, Ph.D.
2022 Winner

NOSTER | Science
MICROBIOME
PRIZE

Apply by 24 January 2023 at www.science.org/noster

Sponsored by Noster, Inc